



Health Technology Assessment

Volume 27 • Issue 31 • December 2023

ISSN 1366-5278

Tonsillectomy compared with conservative management in patients over 16 years with recurrent sore throat: the NATTINA RCT and economic evaluation

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



Tonsillectomy compared with conservative management in patients over 16 years with recurrent sore throat: the NATTINA RCT and economic evaluation

Janet A Wilson^{1*}, Tony Fouweather¹,
Deborah D Stocken², Tara Homer¹,
Catherine Haighton³, Nikki Rousseau²,
James O'Hara⁴, Luke Vale¹,
Rebecca Wilson⁵, Sonya Carnell⁵,
Scott Wilkes⁶, Jill Morrison⁷, Kim Ah-See⁸,
Sean Carrie⁴, Claire Hopkins⁹, Nicola Howe⁵,
Musheer Hussain¹⁰, Lyndsay Lindley¹¹,
Kenneth MacKenzie¹², Lorraine McSweeney¹,
Hisham Mehanna¹³, Christopher Raine¹⁴,
Ruby Smith Whelan⁵, Frank Sullivan¹⁵,
Alexander von Wilamowitz-Moellendorff⁵ and
Dawn Teare¹

¹Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK

²Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, UK

³Department of Social Work, Education and Community Wellbeing, Northumbria University, Newcastle upon Tyne, UK

⁴Ear, Nose and Throat Department, Newcastle Upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK

⁵Newcastle Clinical Trials Unit, Newcastle University, Newcastle upon Tyne, UK

⁶School of Medicine, Faculty of Health Sciences and Wellbeing, University of Sunderland, Sunderland, UK

⁷Senate Office, University of Glasgow, Glasgow, UK

⁸Department of Otolaryngology Head and Neck Surgery, NHS Grampian, Aberdeen, UK

⁹Ear, Nose and Throat and Head and Neck Department, Guy's and St Thomas' NHS Foundation Trust, London, UK

¹⁰School of Medicine, University of Dundee, Dundee, UK

¹¹Social Policy Research Unit, University of York, York, UK

¹²Department of Ear, Nose and Throat Surgery, NHS Greater Glasgow and Clyde, Glasgow, UK

¹³Institute of Head and Neck Studies and Education, University of Birmingham, Birmingham, UK

¹⁴Ear, Nose and Throat Department, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, UK

¹⁵Population and Behavioural Science Division, School of Medicine, University of St Andrews, St Andrews, UK

*Corresponding author

Disclosure of interests

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at <https://doi.org/10.3310/YKUR3660>.

Primary conflicts of interest: Catherine Haighton reports being a member of the College of Experts National Institute for Health and Care Research (NIHR) call for COVID Recovery and Learning Research (2020–present), the NIHR Programme Grants for Applied Research Sub Committee (2019–present) and the NIHR Research for Patient Benefit Programme North East Regional Funding Committee (2010–14). Luke Vale reports being a member of the NIHR Health Technology Assessment (HTA) Programme Clinical Trials and Evaluation Panel (2015–18). Musheer Hussain reports that he was chairperson of the most recent Scottish Intercollegiate Guidelines Network guideline on ‘Management of sore throat and indications for Tonsillectomy’ (2010) and chairperson of the Scottish Otolaryngology Society (ENT-Scotland) committee on ‘Should Reusable equipment for Tonsillectomy be abandoned?’ (2014). Hisham Mehanna reports personal fees from Merck Sharp & Dohme Corporation (Kenilworth, NJ, USA), Sanofi Pasteur (Lyon, France) and Merck (Darmstadt, Germany); grants from GlaxoSmithKline Biologicals (Brentford, UK), AstraZeneca (Cambridge, UK) and GSK PLC; directorship and employment from Warwickshire Head Neck Clinic Ltd; and travel and accommodation expenses from MSD, outside the submitted work. Professor Mehanna also reports being a member of the HTA Technology Assessment Clinical Evaluation and Trials Board (2013–18) and a member of the HTA Unit Interventional Technologies Panel (2009–18). Frank Sullivan reports being a member of the Efficacy and Mechanism Evaluation (EME) Strategy Advisory Committee (2018–19), the EME Funding Committee (2016–19) and the EME Funding Committee Sub-group Remit and Comp Check (2018–19).

Published December 2023
DOI: 10.3310/YKUR3660

This report should be referenced as follows:

Wilson JA, Fouweather T, Stocken DD, Homer T, Haighton C, Rousseau N, *et al.* Tonsillectomy compared with conservative management in patients over 16 years with recurrent sore throat: the NATTINA RCT and economic evaluation. *Health Technol Assess* 2023;27(31). <https://doi.org/10.3310/YKUR3660>

Health Technology Assessment

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.6

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.6 and is ranked 32nd (out of 105 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2021 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), Embase (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta.

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 12/146/06. The contractual start date was in July 2014. The draft report began editorial review in April 2021 and was accepted for publication in November 2021. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care.

Copyright © 2023 Wilson *et al.* This work was produced by Wilson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: <https://creativecommons.org/licenses/by/4.0/>. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland, and final files produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).

NIHR Journals Library Editor-in-Chief

Dr Cat Chatfield Director of Health Services Research UK

NIHR Journals Library Editors

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HSDR, PGfAR, PHR journals) and Editor-in-Chief of HSDR, PGfAR, PHR journals

Dr Peter Davidson Interim Chair of HTA and EME Editorial Board, Consultant Advisor, School of Healthcare Enterprise and Innovation, University of Southampton, UK

Professor Matthias Beck Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Consultant in Public Health, Delta Public Health Consulting Ltd, UK

Ms Tara Lamont Senior Adviser, School of Healthcare Enterprise and Innovation, University of Southampton, UK

Dr Catriona McDaid Reader in Trials, Department of Health Sciences, University of York, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Emeritus Professor of Wellbeing Research, University of Winchester, UK

Professor James Raftery Professor of Health Technology Assessment, School of Healthcare Enterprise and Innovation, University of Southampton, UK

Dr Rob Riemsma Consultant Advisor, School of Healthcare Enterprise and Innovation, University of Southampton, UK

Professor Helen Roberts Professor of Child Health Research, Child and Adolescent Mental Health, Palliative Care and Paediatrics Unit, Population Policy and Practice Programme, UCL Great Ormond Street Institute of Child Health, London, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Please visit the website for a list of editors: www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: journals.library@nihr.ac.uk

Abstract

Tonsillectomy compared with conservative management in patients over 16 years with recurrent sore throat: the NATTINA RCT and economic evaluation

Janet A Wilson^{1*}, Tony Fouweather¹, Deborah D Stocken², Tara Homer¹, Catherine Houghton³, Nikki Rousseau², James O'Hara⁴, Luke Vale¹, Rebecca Wilson⁵, Sonya Carnell⁵, Scott Wilkes⁶, Jill Morrison⁷, Kim Ah-See⁸, Sean Carrie⁴, Claire Hopkins⁹, Nicola Howe⁵, Musheer Hussain¹⁰, Lyndsay Lindley¹¹, Kenneth MacKenzie¹², Lorraine McSweeney¹, Hisham Mehanna¹³, Christopher Raine¹⁴, Ruby Smith Whelan⁵, Frank Sullivan¹⁵, Alexander von Wilamowitz-Moellendorff⁵ and Dawn Teare¹

¹Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK

²Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, UK

³Department of Social Work, Education and Community Wellbeing, Northumbria University, Newcastle upon Tyne, UK

⁴Ear, Nose and Throat Department, Newcastle Upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK

⁵Newcastle Clinical Trials Unit, Newcastle University, Newcastle upon Tyne, UK

⁶School of Medicine, Faculty of Health Sciences and Wellbeing, University of Sunderland, Sunderland, UK

⁷Senate Office, University of Glasgow, Glasgow, UK

⁸Department of Otolaryngology Head and Neck Surgery, NHS Grampian, Aberdeen, UK

⁹Ear, Nose and Throat and Head and Neck Department, Guy's and St Thomas' NHS Foundation Trust, London, UK

¹⁰School of Medicine, University of Dundee, Dundee, UK

¹¹Social Policy Research Unit, University of York, York, UK

¹²Department of Ear, Nose and Throat Surgery, NHS Greater Glasgow and Clyde, Glasgow, UK

¹³Institute of Head and Neck Studies and Education, University of Birmingham, Birmingham, UK

¹⁴Ear, Nose and Throat Department, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, UK

¹⁵Population and Behavioural Science Division, School of Medicine, University of St Andrews, St Andrews, UK

*Corresponding author janet.wilson@newcastle.ac.uk

Background: The place of tonsillectomy in the management of sore throat in adults remains uncertain.

Objectives: To establish the clinical effectiveness and cost-effectiveness of tonsillectomy, compared with conservative management, for tonsillitis in adults, and to evaluate the impact of alternative sore throat patient pathways.

Design: This was a multicentre, randomised controlled trial comparing tonsillectomy with conservative management. The trial included a qualitative process evaluation and an economic evaluation.

Setting: The study took place at 27 NHS secondary care hospitals in Great Britain.

Participants: A total of 453 eligible participants with recurrent sore throats were recruited to the main trial.

Interventions: Patients were randomised on a 1:1 basis between tonsil dissection and conservative management (i.e. deferred surgery) using a variable block-stratified design, stratified by (1) centre and (2) severity.

Main outcome measures: The primary outcome measure was the total number of sore throat days over 24 months following randomisation. The secondary outcome measures were the number of sore throat episodes and five characteristics from Sore Throat Alert Return, describing severity of the sore throat, use of medications, time away from usual activities and the Short Form questionnaire-12 items. Additional secondary outcomes were the Tonsil Outcome Inventory-14 total and subscales and Short Form questionnaire-12 items 6 monthly. Evaluation of the impact of alternative sore throat patient pathways by observation and statistical modelling of outcomes against baseline severity, as assessed by Tonsil Outcome Inventory-14 score at recruitment. The incremental cost per sore throat day avoided, the incremental cost per quality-adjusted life-year gained based on responses to the Short Form questionnaire-12 items and the incremental net benefit based on costs and responses to a contingent valuation exercise. A qualitative process evaluation examined acceptability of trial processes and randomised arms.

Results: There was a median of 27 (interquartile range 12–52) sore throats over the 24-month follow-up. A smaller number of sore throats was reported in the tonsillectomy arm [median 23 (interquartile range 11–46)] than in the conservative management arm [median 30 (interquartile range 14–65)]. On an intention-to-treat basis, there were fewer sore throats in the tonsillectomy arm (incident rate ratio 0.53, 95% confidence interval 0.43 to 0.65). Sensitivity analyses confirmed this, as did the secondary outcomes. There were 52 episodes of post-operative haemorrhage reported in 231 participants undergoing tonsillectomy (22.5%). There were 47 re-admissions following tonsillectomy (20.3%), 35 relating to haemorrhage. On average, tonsillectomy was more costly and more effective in terms of both sore throat days avoided and quality-adjusted life-years gained. Tonsillectomy had a 100% probability of being considered cost-effective if the threshold for an additional quality-adjusted life year was £20,000. Tonsillectomy had a 69% probability of having a higher net benefit than conservative management. Trial processes were deemed to be acceptable. Patients who received surgery were unanimous in reporting to be happy to have received it.

Limitations: The decliners who provided data tended to have higher Tonsillectomy Outcome Inventory-14 scores than those willing to be randomised implying that patients with a higher burden of tonsillitis symptoms may have declined entry into the trial.

Conclusions: The tonsillectomy arm had fewer sore throat days over 24 months than the conservative management arm, and had a high probability of being considered cost-effective over the ranges considered. Further work should focus on when tonsillectomy should be offered. National Trial of Tonsillectomy IN Adults has assessed the effectiveness of tonsillectomy when offered for the current UK threshold of disease burden. Further research is required to define the minimum disease burden at which tonsillectomy becomes clinically effective and cost-effective.

Trial registration: This trial is registered as ISRCTN55284102.

Funding: This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 12/146/06) and is published in full in *Health Technology Assessment*; Vol. 27, No. 31. See the NIHR Funding and Awards website for further award information.

Contents

List of tables	xiii
List of figures	xvii
List of supplementary material	xxi
List of abbreviations	xxiii
Plain language summary	xxv
Scientific summary	xxvii
Chapter 1 Introduction	1
Scientific background	1
<i>Economic burden of adult sore throat disease</i>	1
<i>Association over time between the rate of surgical tonsil intervention and admission to hospital with pharyngotonsillar inflammation</i>	2
<i>Sore throats and recurrent tonsillitis</i>	2
<i>Role of tonsillectomy in the management of recurrent sore throats</i>	2
<i>The risks of tonsillectomy</i>	3
<i>Conservative therapy for recurrent sore throats</i>	4
<i>Assessment of baseline sore throat severity</i>	4
Rationale	5
Context	5
Aims and objectives	6
<i>Feasibility study aim and objectives</i>	6
<i>Internal pilot objectives</i>	7
<i>Main trial objectives</i>	7
Chapter 2 Methods	9
Overview of the trial design	9
Trial registration and protocol availability	9
Ethics and governance	9
Setting	9
Participants	10
<i>Inclusion criteria</i>	10
<i>Exclusion criteria</i>	11
Intervention	11
<i>Tonsillectomy</i>	11
<i>Conservative management</i>	11
<i>Funding of the trial intervention</i>	11
Outcome measurement	12
<i>Primary outcome</i>	12
<i>Secondary outcomes</i>	13
Economic analysis	15
Qualitative analysis	15
Participant timeline	16

CONTENTS

<i>Feasibility study</i>	16
<i>Internal pilot</i>	16
<i>Identification, screening and recruitment of participants</i>	17
<i>Declining patients</i>	18
<i>Consent procedure</i>	18
<i>Randomisation</i>	18
<i>Treatment, crossover and withdrawals</i>	18
<i>Schedule of events</i>	19
<i>Safety</i>	19
<i>Definition of the end of the trial</i>	20
Participant expenses	20
Patient and public involvement	20
Statistical methods	20
<i>Sample size calculation</i>	20
<i>Statistical analysis</i>	21
<i>Primary analyses</i>	22
<i>Sensitivity analysis</i>	23
<i>Additional analyses of the primary outcome</i>	23
<i>Secondary outcome measures</i>	24
Data monitoring, quality control and assurance	26
Trial management and oversight	26
<i>Trial Management Group</i>	27
<i>Oversight committees</i>	27
Chapter 3 Results	29
Recruitment	29
<i>Randomisation and stratification factors</i>	29
<i>Consolidated Standards of Reporting Trials flow diagram</i>	29
<i>Characteristics of those declining to participate</i>	30
<i>Baseline participant characteristics</i>	32
<i>Data quality and completeness</i>	33
<i>Summary of primary outcome data: Sore Throat Alert Return</i>	33
<i>Assessment of compliance with face-to-face and postal completion of questionnaires</i>	33
Trial populations subjected to analysis	34
<i>Treatment received</i>	34
<i>Time from randomisation to tonsillectomy</i>	37
<i>Type of tonsillectomy received</i>	38
Safety	38
<i>Serious adverse events</i>	39
Descriptive analysis of the primary outcome measure	40
<i>Retention flow diagram</i>	40
<i>Primary outcome measure</i>	40
Efficacy analysis: primary outcome	41
<i>Primary analysis</i>	41
<i>Negative binomial regression of the primary outcome</i>	42
<i>Exploration of Sore Throat Alert Return profile by randomised arm</i>	44
<i>over the 105 weeks</i>	44
<i>Addressing missing data</i>	46
Sensitivity analyses	48
<i>Per-protocol analysis</i>	50
<i>Addressing crossover (per-treatment analysis)</i>	50
Secondary outcome measures	54

<i>Missing secondary outcome data</i>	56
<i>Tonsillectomy Outcome Inventory-14</i>	56
<i>Quality-of-life analysis: Short Form questionnaire-12 items</i>	58
Summary	62
Chapter 4 Economic evaluation	63
Introduction	63
Methods	63
<i>Cost data collection</i>	63
<i>Derivation of costs</i>	65
<i>General practitioner linkage</i>	67
<i>Effectiveness measures</i>	67
<i>Comparative incremental analyses of costs and effects</i>	68
<i>Adjusted analysis: seemingly unrelated regression</i>	69
<i>Sensitivity analysis</i>	69
Results	70
<i>Data validity and completeness</i>	70
<i>Resource use and costs</i>	70
<i>Effectiveness outcomes</i>	72
<i>Economic evaluation</i>	74
<i>Sensitivity analysis</i>	77
Chapter 5 Qualitative study	79
Introduction	79
Methods	79
<i>Design</i>	79
<i>Setting and sample</i>	79
<i>Recruitment</i>	79
<i>Interviews</i>	80
<i>Data management and analyses</i>	80
Results	80
<i>Findings</i>	80
Discussion	93
Chapter 6 Discussion	97
Statement and interpretation of main results	97
<i>Crossover</i>	99
<i>Clinical effectiveness</i>	99
<i>Secondary outcome measures</i>	99
<i>Safety</i>	101
Qualitative study	103
Strengths and limitations	103
<i>Strengths</i>	103
<i>Limitations</i>	104
Conclusions	105
Implications for practitioners and health services	106
Implications for patients and the public	107
Implications for research	107
Acknowledgements	109
References	113

Appendix 1 Substantial amendments	123
Appendix 2 Reasons for requesting to switch randomised arms	127
Appendix 3 Additional information for eligible participants who declined to be randomised	129
Appendix 4 Stratification factors	131
Appendix 5 Recoding of employment status and education levels	133
Appendix 6 Tables and histograms showing 6- and 18-month postal return and 12- and 24-month visit compliance	135
Appendix 7 Further details of sensitivity analyses	139
Appendix 8 Withdrawals	141
Appendix 9 Time from randomisation to tonsillectomy	149
Appendix 10 Additional safety information	151
Appendix 11 Measuring Sore Throat Alert Return response	157
Appendix 12 Summary statistics for crossing or remaining in randomised arm and tonsillectomy versus no tonsillectomy, with rates	159
Appendix 13 Instrumental variables	161
Appendix 14 Secondary analyses: Tonsillectomy Outcome Inventory-14	163
Appendix 15 Addressing missing data	175
Appendix 16 Secondary analyses: quality of life (Short Form questionnaire-12 items)	177
Appendix 17 Economic evaluation	179

List of tables

TABLE 1 Details of questionnaires included in the packages at each follow-up contact	15
TABLE 2 Distribution of baseline severity by randomised arm for the ITT population	29
TABLE 3 Demographics and imputed TOI-14 scores of the accepting and declining patients	31
TABLE 4 Baseline demographic characteristics of 453 eligible participants by randomised arm	32
TABLE 5 Participant activities completed by time point	34
TABLE 6 Compliance with follow-up visits and postal questionnaire completion	35
TABLE 7 Intervention received: immediate tonsillectomy ($n = 233$) and conservative management ($n = 220$)	37
TABLE 8 Characteristics of the 191 AEs, with severity category	39
TABLE 9 Characteristics of the 51 SAEs with severity category	39
TABLE 10 Primary outcome measure: total sore throats over 24 months and the average number of sore throat days per week, by randomised arm and overall, for the ITT population	41
TABLE 11 Unadjusted NBR showing IRR and 95% CIs, with exposure variable of proportion of weekly STAR responses for difference in total number of sore throat days between randomised arms: ITT population	41
TABLE 12 Results of multilevel mixed-effect NBR (model 1): ITT population	42
TABLE 13 Summary statistics of the raw total number of sore throats reported by arm and baseline severity	43
TABLE 14 Cumulative percentage of weekly STAR responses 24 months following randomisation	46
TABLE 15 Proportion of participants responding to weekly STAR responses split by observed complete, partial and no response at 6-monthly intervals post randomisation	49
TABLE 16 The SF-12 and TOI-14 complete, partial or missing over time in the ITT population	55
TABLE 17 The TOI-14 overall scores at time points (% is the proportion completed out of eligible randomised)	57
TABLE 18 Test difference in TOI-14 means by arm over time with 95% CIs	58

TABLE 19 The SF-12 components (SF-12 MCS and PCS) summary statistics at time points, ITT population; lower scores = poorer QoL	61
TABLE 20 Response rates for data collection tools used to inform the economic analysis	71
TABLE 21 Average total costs for each cost category by randomised arm (ITT)	72
TABLE 22 Summaries of the outcome measures used in the CEA and CUA by randomised arm	73
TABLE 23 Cost-effectiveness analysis results: outcome – sore throat days, results (multiple imputation)	74
TABLE 24 Cost-utility analysis results: outcome – QALYs, results (multiple imputation)	76
TABLE 25 Cost-benefit analysis results: outcome – willingness to pay and number of sore throat days: results (multiple imputation)	76
TABLE 26 Interviewees (NATTINA trial participants/decliners)	81
TABLE 27 Interviewees (staff)	81
TABLE 28 Changes to the protocol requiring regulatory approval	123
TABLE 29 Summary of the 134 reasons that the 73 participants randomised to the conservative management arm gave for crossing over to receive tonsillectomy	127
TABLE 30 Reasons for patients randomised to tonsillectomy for crossing over to conservative management	128
TABLE 31 Reasons given at screening stage for declining to participate	129
TABLE 32 Completeness of TOI-14 for declining patients	129
TABLE 33 Distribution of participants by recruiting centre ($n = 27$)	131
TABLE 34 Number of participants mis-stratified based on TOI-14 severity categories	132
TABLE 35 Summary statistics of time from randomisation to visits/postal returns of secondary outcome questionnaires	135
TABLE 36 Summary statistics for time from randomisation to withdrawal in weeks by arm	141
TABLE 37 Participants who withdrew from the trial ($n = 43$)	142
TABLE 38 Summary statistics of time to tonsillectomy by arm and overall	149
TABLE 39 Type of tonsillectomy delivered in NATTINA	149
TABLE 40 Further details of complications with tonsillectomy reported ($n = 10$)	150

TABLE 41 One-hundred and ninety-one reported AEs for 90 participants undergoing tonsillectomy (SAEs removed)	151
TABLE 42 Details of SAEs deemed to be caused by the tonsillectomy ranked by severity: severe, moderate and mild ($n = 51$)	151
TABLE 43 Severity and causality of the SAEs owing to the intervention reported	156
TABLE 44 Summary statistics for crossing or remaining in randomised arm and tonsillectomy vs. no tonsillectomy, with rates [average sore throat days per return (week)]: per-treatment	159
TABLE 45 Results of the CACE analyses (unadjusted and adjusted) with each arm set as the reference category	161
TABLE 46 Tonsillectomy Outcome Inventory-14 throat discomfort score completion rate and scores at time points	163
TABLE 47 Mixed-model repeated measures for TOI-14	164
TABLE 48 Tonsillectomy Outcome Inventory-14 general health score completion rate and scores at time points	165
TABLE 49 Tonsillectomy Outcome Inventory-14 resource impact score completion rate and scores at time points	166
TABLE 50 Tonsillectomy Outcome Inventory-14 social/psychological score completion rate and scores at time points	167
TABLE 51 Details of partially missing TOI-14 questionnaires at trial visits and postal returns	168
TABLE 52 Missing data examination: summarise exposure variable by baseline severity	175
TABLE 53 Summary statistics for overall return rate split by baseline severity to see whether or not patterns are different for return of the TOI-14 (secondary outcome)	176
TABLE 54 Crosstab by those who complete at least 80% STAR data by 80% TOI-14	176
TABLE 55 Mixed-model repeated measures for SF-12 PCS	177
TABLE 56 Mixed-model repeated measures for SF-12 MCS	178
TABLE 57 Summary of post-surgical AE treatments, as recorded in the eCRF	179
TABLE 58 Combined STAR/STARLET summaries by severity	179
TABLE 59 Summary of healthcare resource use, as reported in the health utilisation questionnaire	180
TABLE 60 Summary of tonsillectomy data, as recorded in the eCRF	184

TABLE 61 Cost-utility analysis sensitivity analysis: utility decrements associated with a sore throat episode are included in the QALY equation	186
TABLE 62 Cost-utility analysis sensitivity analysis: tonsillectomy costs estimated using microcosting	189
TABLE 63 Cost-utility analysis sensitivity analysis: participant costs	191
TABLE 64 Cost-utility analysis sensitivity analysis: healthcare resource use costs estimated based on GP linkage data	193

List of figures

FIGURE 1 Study pathway	14
FIGURE 2 The CONSORT flow diagram	30
FIGURE 3 Box plots showing time from randomisation to withdrawal by arm ($n = 43$)	36
FIGURE 4 Histograms showing distribution of time from randomisation to tonsillectomy by arm	38
FIGURE 5 Flow chart showing primary outcome data	40
FIGURE 6 Weekly average sore throat days with histograms showing point of tonsillectomy by arm	44
FIGURE 7 Average weekly sore throat rate for participants not receiving tonsillectomy and those who received tonsillectomy	45
FIGURE 8 Bar chart showing observed number of STARs by categories (complete, partial and none) by arm at time points (6, 12, 18 and 24 months)	47
FIGURE 9 Box plots showing distribution of time from randomisation to last STAR response in the 24-month follow-up period by arm	50
FIGURE 10 Forest plot showing IRRs and 95% CIs for primary ITT analysis and sensitivity analyses on ITT population	51
FIGURE 11 Forest plot with sensitivity analyses restricting the ITT population to per-protocol and per-treatment groups	52
FIGURE 12 Weekly average sore throat days by randomised arm and occurrence of tonsillectomy	54
FIGURE 13 Box plots showing percentage of total STARs by four category per-treatment groups ($n = 429$)	55
FIGURE 14 The TOI-14 at 6, 12, 18 and 24 months: ITT population	56
FIGURE 15 Box plots of normalised mental health scores (SF-12 MCS) at five time points by randomised arm	60
FIGURE 16 Box plots of normalised physical health scores (SF-12 PCS) at five time points by randomised arm	60
FIGURE 17 Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CEA multiple imputation results	75
FIGURE 18 Cost-effectiveness acceptability curves for tonsillectomy vs. conservative management using the adjusted bootstrapped CEA multiple imputation results	75

FIGURE 19 Box plots of baseline TOI-14 scores for participants accepting and declining to be randomised into NATTINA	130
FIGURE 20 The 6-month postal returns	136
FIGURE 21 The 12-month visit	137
FIGURE 22 The 18-month postal returns	137
FIGURE 23 The 24-month visit	138
FIGURE 24 The TOI-14 subscale throat discomfort	169
FIGURE 25 The TOI-14 subscale general health	170
FIGURE 26 The TOI-14 subscale resource impact	170
FIGURE 27 The TOI-14 subscale social psychological	171
FIGURE 28 The TOI-14 at five time points in the per-protocol population: baseline and 6, 12, 18 and 24 months	172
FIGURE 29 TOI-14 at five time points in the per-treatment population: baseline and 6, 12, 18 and 24 months	173
FIGURE 30 Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA multiple imputation results	187
FIGURE 31 Cost-effectiveness acceptability curves for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA multiple imputation results	187
FIGURE 32 Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CBA multiple imputation results	188
FIGURE 33 Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (tonsillectomy costs were estimated using microcosting)	190
FIGURE 34 Cost-effectiveness acceptability curves for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (tonsillectomy costs were estimated using microcosting)	190
FIGURE 35 Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (costs and QALYs were estimated using multiple imputation and participant costs have been included with total costs)	192
FIGURE 36 Cost-effectiveness acceptability curves for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (costs and QALYs were estimated using multiple imputation and participant costs have been included with total costs)	192

- FIGURE 37** Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (QALYs were estimated using multiple imputation and healthcare utilisation costs were estimated using GP linkage data) 194
- FIGURE 38** Cost-effectiveness acceptability curves for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (QALYs were estimated using multiple imputation and healthcare utilisation costs were estimated using GP linkage data) 194
- FIGURE 39** Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (QALYs were estimated using multiple imputation and include utility decrements associated with a sore throat episode) 195
- FIGURE 40** Cost-effectiveness acceptability curves for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (QALYs were estimated using multiple imputation and include utility decrements associated with a sore throat episode) 195

List of supplementary material

Report Supplementary Material 1 Economic evaluation supplementary tables and figures

Report Supplementary Material 2 Are adults willing to pay to avoid a sore throat day?
A contingent valuation study

Report Supplementary Material 3 NATTINA qualitative process evaluation topic guide

Supplementary material can be found on the NIHR Journals Library report page (<https://doi.org/10.3310/YKUR3660>).

Supplementary material has been provided by the authors to support the report and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer reviewed.

List of abbreviations

A&E	accident and emergency	NBR	negative binomial regression
AE	adverse event	NCTU	Newcastle Clinical Trials Unit
CACE	complier-average causal effect	NESSTAC	North of England and Scotland Study of Tonsillectomy and Adenotonsillectomy in Children
CBA	cost-benefit analysis	NICE	National Institute for Health and Care Excellence
CCG	Clinical Commissioning Group	NIHR	National Institute for Health and Care Research
CEA	cost-effectiveness analysis	NVQ	National Vocational Qualification
CEAC	cost-effectiveness acceptability curve	ODP	operating department practitioner
CI	confidence interval	PCS	physical component score
CONSORT	Consolidated Standards of Reporting Trials	PI	principal investigator
CUA	cost-utility analysis	PIS	participant information sheet
DMC	Data Monitoring Committee	PPI	patient and public involvement
DVD	digital versatile disc	PRO	patient-reported outcome
eCRF	electronic case report form	PSS	Personal and Social Services
ENT	ear, nose and throat	QALY	quality-adjusted life-year
GCSE	General Certificate of Secondary Education	QoL	quality of life
GDPR	General Data Protection Regulation	RCT	randomised controlled trial
GP	general practitioner	REC	Research Ethics Committee
HND	Higher National Diploma	SAE	serious adverse event
HTA	Health Technology Assessment	SAP	statistical analysis plan
ICER	incremental cost-effectiveness ratio	SD	standard deviation
ICF	informed consent form	SE	standard error
IQR	interquartile range	SF-12	Short Form questionnaire-12 items
IRR	incident rate ratio	SF-6D	Short Form questionnaire-6 Dimensions
ITT	intention to treat	SIGN	Scottish Intercollegiate Guidelines Network
IVR	interactive voice response	STAR	Sore Throat Alert Return
MCS	mental component score		
MOD	Ministry of Defence		
NATTINA	NATional Trial of Tonsillectomy IN Adults		

LIST OF ABBREVIATIONS

SUR	seemingly unrelated regression	TOI-14	Tonsillectomy Outcome Inventory-14
TMG	Trial Management Group	TSC	Trial Steering Committee
		WTP	willingness to pay

Plain language summary

Tonsillectomy is an operation to take out the pair of tonsil glands at the back of the throat. It is an option for adults who suffer from repeated, severe sore throats. Adults who have a tonsillectomy say that they get fewer sore throats afterwards, but it is not clear whether or not they would have got better over time without the operation. There is pressure on doctors to limit the number of tonsillectomies carried out. At the same time, emergency hospital admissions for adults with severe throat infections have been increasing. NAtional Trial of Tonsillectomy IN Adults aimed to find out whether tonsillectomy is an effective and worthwhile treatment for repeated severe sore throats or whether patients would be better off treated without an operation.

A total of 453 patients from 27 hospitals in Great Britain took part in the study. Patients were assigned at random to receive either tonsillectomy or conservative management (treatment as needed from their general practitioner). We measured how many sore throats patients had in the next 2 years by sending them text messages every week. We asked about the impact of their sore throats on their quality of life and time off work, and looked at the costs of treatment. We also interviewed 47 patients, general practitioners and hospital staff about their experiences of tonsillectomy and NAtional Trial of Tonsillectomy IN Adults. The typical patient in the tonsillectomy arm had 23 days of sore throat compared with 30 days of sore throat in the conservative management arm. Tonsillectomy resulted in higher quality of life. We looked to see whether or not it was only those with the most severe sore throats who benefited from tonsillectomy, but we found that patients with more or less severe sore throats at the start all did better with tonsillectomy. Patients who had a tonsillectomy were happy to have undertaken this. Our findings suggest a clear benefit of tonsillectomy using modest additional NHS resources for adults with repeated severe sore throats.

Scientific summary

Background

Sore throats cost the NHS over £120M per annum, including ≈ £60M for general practitioner (GP) consultations and medical therapy. The place of tonsillectomy in the management of sore throat remains uncertain.

Objectives

Clinical

- To establish the clinical effectiveness of tonsillectomy, compared with conservative management, for tonsillitis in adults.
- To report the number of adverse events (AEs), visits to the GP/walk-in clinic/accident and emergency, prescriptions issued and additional interventions required to manage sore throats and related events through weekly Sore Throat Alert Return (STAR) response data and primary care patient records.
- To adjust the estimate of effectiveness in the light of other baseline covariates, including severity of tonsillitis.
- To evaluate the impact of alternative sore throat patient pathways by observation and statistical modelling of outcomes.
- To assess to what extent trial participants were representative of the total population of sore throat patients referred to ear, nose and throat (ENT) clinics.
- To inform future research.

Qualitative process evaluation

- To examine the acceptability of the trial, treatments and unforeseen consequences from the perspective of participants and stakeholders, including ENT staff and GPs.

Economic evaluation

- To compare costs incurred by the NHS, Personal and Social Services (PSS) and participants to manage recurrent sore throats in adults.
- To compare quality-adjusted life-years (QALYs) using the area-under-the-curve method based on Short Form questionnaire-6 Dimension (SF-6D) scores derived from longitudinal Short Form questionnaire-12 items (SF-12) responses throughout the study and during self-reported sore throat episodes.
- To compare the cost-effectiveness measured in terms of the incremental:
 - cost per sore throat day avoided
 - cost per QALY gained
 - net benefit – estimated based on self-reported number of sore throat days and responses to a contingent valuation questionnaire administered at baseline asking participants' willingness to pay (WTP) to avoid a sore throat day.

Methods

Design

A multicentre, randomised controlled trial comparing the effectiveness of tonsillectomy for recurrent sore throat with that of non-surgical conservative management in a 1:1 ratio over a 24-month period.

Surgeons, participants and site staff could not be blinded to the allocated procedure. The main trial began following the completion of a feasibility study and included a qualitative process evaluation, as well as an economic evaluation. The design, conduct and reporting of the trial were informed by patients. Crossover of participants between arms was permitted.

Interventions

- Tonsil dissection: dissection of the palatine tonsils preferably within 6 weeks, and no more than 8 weeks, following randomisation (dissection method at the discretion of the participating centres).
- Conservative (non-surgical) management (i.e. deferred surgery): participants entering the conservative management arm were asked to defer surgery for up to 24 months on the understanding that they would be reviewed at 12 months.

Setting and participants

A total of 453 patients were recruited to the main trial from 27 NHS secondary care hospitals in Great Britain. Eligible patients were identified via general ENT referrals and established sore throat-specific referral pathways, some of which were run by ENT nurse practitioners.

Inclusion criteria

- Aged ≥ 16 years.
- Recurrent sore throats that fulfil current Scottish Intercollegiate Guidelines Network (SIGN) guidance for elective tonsillectomy.

Exclusion criteria

- Previous tonsillectomy.
- Listed directly (i.e. added to waiting list without prior elective ENT outpatient appointment) during emergency admission (e.g. owing to peritonsillar abscess/quinsy).
- Primary sleep breathing disorder.
- Suspected malignancy.
- Tonsilloliths (as primary referral).
- Pregnant or breastfeeding.
- Bleeding diathesis (including haemophilia, sickle cell disease and platelet dysfunction).
- Therapeutic anticoagulation.
- Inability to complete self-reported questionnaires and sore throat returns.

Main outcome measures

Primary outcome

The primary outcome measure was the total number of sore throat days over the 24 months following randomisation.

The number of sore throat days was collected by a database that was designed for use in the trial (the STAR database). This database allowed participants to respond weekly to alerts by providing the number of sore throats that they had experienced in the previous 7 days (between 0 and 7 sore throat days). From the beginning of the trial, participants were able to choose their preferred method for receiving these alerts: e-mail, text message or interactive voice response (IVR) via telephone.

Secondary outcomes

A participant who had suffered from a sore throat in the past week (i.e. sore throat days >0) was asked to provide information on the severity of the sore throat(s) and additional data for health economics and other secondary outcomes in a STAR questionnaire.

Quality-of-life data, reported as the SF-12 physical component score (PCS) and mental component score (MCS) and additional symptoms of tonsillitis, were collected every 6 months.

The impact of alternative NHS sore throat pathways was measured by observation and statistical modelling of outcomes. The extent to which trial participants represented the total population of sore throat patients referred to ENT clinics was assessed through analysis of site screening logs.

Adverse events

Information regarding the AEs related to the trial intervention was collected during telephone calls at 1 and 2 weeks post tonsillectomy; all tonsillectomy arm serious adverse events (SAEs) were recorded throughout the duration of the trial for all participants.

Economic evaluation

The cost-effectiveness of tonsillectomy compared with conservative management was evaluated by estimating the total costs incurred by the NHS and PSS, and averaging these costs across participants in each trial arm. Three different analyses were undertaken: (1) a cost-effectiveness analysis (CEA), (2) a cost-utility analysis (CUA) and (3) a cost-benefit analysis (CBA). All three analyses measured costs using the same methodology but differed in their measure of effectiveness. An incremental cost-effectiveness ratio was estimated for the CEA and the CUA by dividing the difference in average total costs by the difference in average total effects. The CEA estimated the incremental cost per sore throat day avoided. The number of sore throat days was derived from the primary outcome data. The CUA estimated the incremental cost per QALY gained. QALYs were derived using the SF-6D algorithm, which estimates utility values based on responses to the SF-12, which was administered at baseline and 6, 12, 18 and 24 months post randomisation and at the time of a sore throat episode. The CBA estimated the incremental net benefit, comparing costs and benefits in monetary terms. Participants' WTP values, estimated from the contingent valuation study, were multiplied by participants' self-reported number of sore throat days to estimate the reduction in patient benefits in monetary terms; from this costs were subtracted to give the net benefit.

Qualitative process evaluation

Qualitative and cognitive interviews were carried out by researchers from Newcastle University for the feasibility study, pilot and main trial. Interviews were held with adult patients with acute tonsillitis who had been referred to ENT outpatient clinics for recurrent sore throats, ENT staff who were working at a National Trial of Tonsillectomy IN Adults (NATTINA) trial site and GPs. These interviews addressed the acceptability of the trial/treatments, unforeseen consequences from the perspective of participants and how patient experience may shape future research.

Statistical analysis

The primary statistical analysis was carried out on an intention-to-treat (ITT) basis, retaining patients in their randomised arms and including protocol violator and ineligible patients. Patients randomised to conservative management were asked to commit to 'deferred surgery'. We anticipated that a number of patients would take the opportunity to switch to surgery. We also undertook sensitivity analyses, including a per-treatment (as treated) and a per-protocol analysis. The cumulative total number of sore throat days reported on a patient level was provided separately for each randomised arm. Negative binomial regression was used to compare these, adjusting for stratification variables (baseline severity as a fixed effect and site as a random effect). The summary comparative statistic reported is the incident rate ratio (IRR). The analysis took account of incomplete weekly returns by use of an exposure variable included in the model. The analysis of secondary outcomes followed a broadly similar strategy to the primary outcome. Analyses were adjusted for baseline severity and site, and repeated measures were analysed using random-effects models with appropriate error structure.

Results

Primary outcome

The participants in this trial reported a median of 27 [interquartile range (IQR) 12–52] sore throats over the full 24 months of follow-up. Fewer sore throats were reported in the tonsillectomy arm (median 23, IQR 11–46) than in the conservative management arm (median 30, IQR 14–65). When the primary outcome (total sore throats) was compared between the two randomised arms on an ITT basis, a reduction in sore throats was seen in the tonsillectomy arm. The tonsillectomy arm had 0.53 [IRR, 95% confidence interval (CI) 0.43 to 0.65] times the total sore throat days over the 24-month period than those in the conservative management arm. Sensitivity analyses on the ITT population confirmed this result. The analysis of the per-protocol population of 224 participants had, on average, a greater reduction in the number of sore throats, with patients in the tonsillectomy arm having 0.42 (IRR 95% CI 0.31 to 0.55) times the total score throats compared with the conservative management arm. Around 25% of participants did not receive the treatment that they were randomised to, which meant that some opted not to receive a tonsillectomy and some opted to cross to tonsillectomy. There is evidence to suggest that those with larger numbers of sore throats following randomisation were more likely to either opt for or remain in the tonsillectomy arm. Conversely, those with slightly smaller numbers of sore throats following randomisation were more likely to remain in conservative management or opt out of the tonsillectomy arm. Despite these crossovers, the ITT, per-protocol and per-treated analyses all confirmed that there was a significant reduction in total sore throats for those randomised to tonsillectomy.

Secondary outcome measures

The benefits of tonsillectomy were also seen in the secondary outcome measures. Tonsillectomy Outcome Inventory-14 (TOI-14) scores improved in both arms, but show a greater improvement in the tonsillectomy arm than the conservative management arm, where at 12 months the difference between the mean scores was –13.17 units (95% CI –17.41 to –8.92 units), indicating a reduction in symptoms. The SF-12 MCS and SF-12 PCS also show significant and beneficial differences in favour of the tonsillectomy arm over time [SF-12 MCS scores 3.71 units higher (95% CI 2.10 to 5.47) and SF-12 PCS 2.77 units higher (95% CI 0.30 to 5.23) in the tonsillectomy arm than the conservative management arm].

Adverse events

There were 52 episodes of post-operative haemorrhage reported in 231 participants undergoing tonsillectomy (22.5%). Of these episodes, 35 were reported as SAEs: 8 as mild events, 22 as moderate events and 5 as severe events. No deaths were reported. Seventeen episodes were recorded as AEs, for which patients did not attend hospital. All episodes of bleeding were managed conservatively with no returns to theatre.

Economic evaluation

On average, tonsillectomy was more costly and more effective than conservative management. In the CEA, the incremental cost to avoid a sore throat episode was £24 per sore throat day. From the contingent valuation, the mean WTP to avoid a sore throat day was £43 (95% CI £2 to £100). In the CUA, tonsillectomy had an 87% probability of being considered cost-effective at a £5000 threshold for an additional QALY; this increased to 100% as the threshold values for an additional QALY increased. In the CBA, tonsillectomy had a 69% probability of having a higher net benefit than conservative management.

Qualitative study

Trial processes were deemed as acceptable, with only a few sites experiencing barriers to treatment. The use of technology to collect data was particularly well received. However, there were some challenges with recruitment, particularly for staff who lacked equipoise. ENT staff alluded to having to negotiate surgery dates with patients, which, at times, meant that they had to deviate from trial protocols. Some patients did not fully understand the process of randomisation. Patients who had received surgery were unanimous in reporting to be happy to have undertaken this, despite the challenging recovery period.

Limitations

There was some evidence that those with the most severe extent of disease were reluctant to enter the study (around 5 points higher scores overall on the TOI-14). Some symptoms may not be a result of tonsillitis. Not all patients were offered, or chose, to watch the trial recruitment video. The ITT analysis is likely to offer a conservative underestimate of the true impact of tonsillectomy in reducing sore throat days, as a result of patients crossing over to receive tonsillectomy. A challenge of the economic evaluation was the progressive loss of data over the 24-month follow-up.

Conclusions

Tonsillectomy in adults is a clinically effective intervention. It was more costly but had a high probability of being considered cost-effective over the range of analyses conducted. Participants with recurrent tonsillitis, who met current UK NHS guidelines to undergo tonsillectomy, suffered significantly fewer sore throat days over 24 months than similar participants treated conservatively.

How should health services react?

Pre-NATTINA, UK guidelines were a translation of level 1 evidence in children, applied to adults. Access to tonsillectomy in the UK was governed by application of national guidance, which is predicated on a qualifying number of episodes of tonsillitis. NATTINA participants in the tonsillectomy arm reported, on average, fewer healthcare contacts, fewer sore throat days and higher QALYs than those in the conservative management arm. Within the UK, tonsillectomy is listed as a 'procedure of limited clinical value' (The Royal College of Surgeons of England. *Procedures of Limited Clinical Value: Royal College of Surgeons Briefing*. London: The Royal College of Surgeons of England; 2011). To the best of our knowledge, NATTINA is the first definitive trial to demonstrate that tonsillectomy performed according to the current UK national guidelines is effective for patients, and the probability of it being considered cost-effective is high. Guideline reassessment, in particular how guidelines translate into healthcare commissioning, is called for.

How should practitioners and patients respond to these findings?

We have identified communication issues at the primary–secondary care interface in our qualitative work. There is, therefore, a need to convert the findings of NATTINA into a practical decision support tool for patients and surgeons.

Implications for research (in priority order)

The top research priority to emerge from NATTINA is to determine the optimum timing of tonsillectomy in adults with recurrent acute tonsillitis. Work is required also to optimise metrics for disease burden severity, and to exploit the novel real-time data collection methods elaborated in NATTINA. There is also a need to better understand optimum treatment strategies, including oral steroids, for tonsillitis in primary care.

Trial registration

This trial is registered as ISRCTN55284102.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 12/146/06) and is published in full in *Health Technology Assessment*; Vol. 27, No. 31. See the NIHR Funding and Awards website for further award information.

Chapter 1 Introduction

Scientific background

Around 12% of the population of Great Britain experience recurrent tonsillitis, with significant associated morbidity and impacts on work, education, family and social life.¹⁻³ For those patients experiencing frequent, disabling tonsillitis over time, tonsillectomy (surgical removal of the tonsils) is a treatment option. Although most of those undergoing tonsillectomy report benefit, it has been unclear to what extent some of the improvements seen following tonsillectomy would have occurred without surgical intervention. This uncertainty, together with the cost and morbidity associated with surgery, has led some to question whether or not tonsillectomy should be considered a 'low-value surgical procedure'.⁴ However, in recent years, emergency hospital admissions for adults with pharyngotonsillitis have risen to 30,000 per year in England: twice as many as adults who undergo tonsillectomy. With this context, the National Trial of Tonsillectomy IN Adults (NATTINA) set out to provide much-needed evidence regarding the place of tonsillectomy in the treatment of adults with tonsillitis.

Economic burden of adult sore throat disease

There have been several attempts to quantify the economic impacts of sore throat. One recent study initially estimated the annual cost of consultations and lost productivity to be £40M, but concluded that the actual all-encompassing bill might be as high as £2.35B.⁵ A more current estimate of primary care costs can be drawn from the pilot data of an NHS study: Sore Throat Test and Treat.⁶ Extrapolation of the Sore Throat Test and Treat pilot figures to the estimated 42,000 full-time equivalent general practitioners (GPs) in England and Scotland would translate to 5.9 million consultations per annum. Taking the cost of a GP appointment from the unit costs of health and social care produced by the Personal Social Services Research Unit⁷ to be £34, these consultations equate to a nationwide bill of over £201M per annum. In addition, if there are antibiotics prescribed in approximately three-fifths of those attending primary care with sore throats, the total cost even of simple penicillin V at £1.25 for 28 500-mg tablets four times per day for 1 week is £19.2M.⁶ The cumulative total of consultations and antibiotic prescriptions is, therefore, in the order of £219M per annum in England and Scotland. The latest available NHS Healthcare Resource Group data give an expenditure of over £70M for tonsil surgery per year in England, in all ages. Between 2013 and 2017, there were 4600 tonsillectomies per year on average, in Scotland,⁸ costing an additional £7.5M, although the emergence of the Scottish Intercollegiate Guidelines Network (SIGN) guidance has been reported to have impacted differently in England and Wales than in Scotland.⁹

Numerous additional interventions take place in the independent sector that are not recorded in NHS digital data. The cost of pharyngeal inflammation emergency admissions, as outlined below, exceeds £20M. Adding these secondary care episodes to those in the community gives an overall, all-age cost of £309M.

Estimated cost of adult pharyngotonsillitis in secondary care

Around 60% of tonsil disease occurs in childhood; the present trial concerns optimum treatment strategies for tonsil infections in adults aged ≥ 16 years. In adults, the total cost of tonsil surgery in 2018–19 in England was over £29.4M.¹⁰ At a conservative estimate, over 30,000 adults are admitted with pharyngitis and tonsillitis diagnoses as emergencies in England per annum, as outlined in the following section.¹¹ The cost of the associated bed occupancy alone exceeds £10M,¹⁰ a total secondary care expenditure around £40M.

Association over time between the rate of surgical tonsil intervention and admission to hospital with pharyngotonsillar inflammation

Several reports¹² indicate that as the UK rate of tonsillectomy has fallen, the number of acute admissions has risen, although no causality has been proven in this association. For example, looking at English Hospital Episode Statistics data,¹³ in 2009–10 there were 21,540 tonsillectomy operations in adults aged ≥ 15 years, and around 15,000 emergency adult pharyngotonsillitis admissions. In 2019–20, there were only 16,000 adult tonsillectomy operations, whereas emergency adult admissions in England with a main diagnosis of pharyngitis and tonsillitis had risen to over 30,000.¹¹ A similar reduction in surgical intervention was seen in recent years in Germany, where the tonsillectomy rate fell by 46% in males and 40% in females from 2005 to 2017.¹⁴

Sore throats and recurrent tonsillitis

'Sore throat' refers to pain in the pharyngotonsillar area and can be caused by viral or bacterial infection, or by non-infective irritants, such as smoke and hay fever. 'Tonsillitis' is a subset of sore throats in which there is inflammation of the tonsils, and this can range from relatively mild symptoms associated with viral infections to bacterial tonsillitis and quinsy (peritonsillar abscess). The pathway through which patients are offered tonsillectomy is inevitably predicated on patient complaints, not only sore throat but also, as our own NATTINA patient and public involvement (PPI) group advised, the concomitant experience of general illness and malaise. It is the systemic nature of the illness and the impact on the patients' quality of life (QoL) that leads patients to seek treatment¹⁵ and, eventually, to be referred to secondary care. Our PPI group reported varied access to antibiotic therapy for tonsillitis in primary care.

In the USA, the term 'strep throat' is widely used to distinguish between bacterial tonsillitis caused by group A *Streptococcus* and viral (usually milder) tonsillitis, but this distinction is less commonly made in the UK, particularly in primary care. Although anyone (with tonsils) can experience tonsillitis, a subset of people experience multiple episodes of acute tonsillitis over an extended period of time: recurrent tonsillitis. Recurrent tonsillitis is more common in females, and studies in twins demonstrate a strong hereditary element.¹⁶ A failure to differentiate between different forms of sore throat may have contributed to an inadequate understanding of the features and impacts of recurrent tonsillitis and limited the development of effective alternatives to antibiotics and tonsillectomy, in spite of the high prevalence and morbidity associated with recurrent tonsillitis. Recent work has identified specific genetic variations in the immune system that are associated with a risk of recurrent tonsillitis, suggesting potential future treatment targets.¹⁷ In NATTINA, consistent with SIGN¹⁸ and National Institute for Health and Care Excellence (NICE) guidance,¹⁹ we refer throughout to 'recurrent sore throats'.

Role of tonsillectomy in the management of recurrent sore throats

The optimum role of tonsillectomy in the management of adults with recurrent sore throat remains uncertain. Audits consistently report over 95% of patients gaining symptom relief from surgery,^{20,21} but a lack of level I evidence^{22,23} contributes to ongoing UK regional variation in tonsillectomy rates,²⁴ even where compliance to SIGN guidance¹⁸ approaches 90%.²⁵ The 2014 Cochrane review²² identified only one evaluable, small, adult trial comprising 70 participants²⁶ with only 90 days' follow-up. It concluded that reasonable levels of evidence were available for tonsillectomy in children only. Since then, there has been a further small (≈ 40 per arm) Finnish tonsillectomy trial,²⁷ which demonstrated no impact on the primary outcome, that is the proportion of patients with severe inflammation of the pharynx within 5 months. However, there were 10 times as many consultations for pharyngitis in the conservative management arm.

Except in rare situations (when tonsillar material remains or regrows), tonsillectomy should remove the possibility of tonsillitis, but its impact on non-tonsillitis sore throats, if any, is unknown. The questions that patients, doctors and healthcare providers wish to answer relate to the relative costs, risks and benefits of tonsillectomy, which must be weighed against those of conservative alternative treatments for recurrent sore throat. Research may also usefully establish whether or not more refined surgical indications could maximise the benefit–risk ratio.

The risks of tonsillectomy

The most common complications of tonsillectomy are pain and bleeding. From 1995 to 2010, there were 40 claims of clinical negligence related to NHS tonsillectomy, most commonly bleeding related and followed by nasopharyngeal regurgitation.²⁸ Less common but intrusive complications include changes in taste²⁹ or tongue sensation.³⁰ In the Swedish 2012 audit,²¹ almost 14% of patients had an unplanned postoperative clinic visit. A Danish study of 614 outpatient tonsillectomy patients of all age groups identified unscheduled postoperative contacts in 23% of patients, about half because of pain. Contacts were more frequent when discharge occurred in under 4 hours.³¹ Comprehensive ascertainment of complications may be difficult precisely because patients are discharged promptly, increasingly on the day of operation; thus, not every complication will involve a hospital contact. Supplementing postal surveys by telephone calls can increase response rates on postoperative events following tonsillectomy by about 50%.³²

Pain

Tonsillectomy, the removal of the palatine tonsil tissue within the oropharynx, is a painful procedure³³ that requires an average of 14 days sick leave.^{34,35} The dissection of the tonsil can be carried out with 'cold steel' or using electrical (diathermy) or radiofrequency (coblation) energy.³⁶ In one small study,³⁷ the level of pain correlated to the amount of bipolar diathermy used in the procedure.³⁷ Many trials have addressed the optimum pain relief post tonsillectomy, with the recent emphasis on any potential benefit from the addition of dexamethasone, which may reduce pain, post-operative nausea and vomiting, and overall complications.³⁸ Steroids do not appear to reduce pain beyond the first post-operative day nor have any impact on post-operative bleeding.³⁹ A recent review⁴⁰ also emphasised the importance of the use of multiple painkillers on the day of surgery to maximise analgesic effect.

Bleeding

Primary post-operative haemorrhage occurs close to the time of the surgery. Its incidence was 1.3% in a Swedish audit of 54,696 tonsillectomies between 1997 and 2008.⁴¹ Predictors of bleeding rates included older age, male sex and inpatient as opposed to day case surgery.⁴¹ Post-tonsillectomy, delayed, secondary haemorrhage is generally much more common. Evans *et al.*⁴² attempted to identify overlooked instances of UK post-tonsillectomy haemorrhage by a telephone survey. Of 60 patients, 40% had experienced oral blood flow for more than 1 minute. Only 8% required re-admission and 3% required a return to theatre. The authors suggested that return to theatre rates were a more valuable metric than generic reporting of any post-operative haemorrhage. A recent Scottish review⁸ compared 27,819 patients undergoing tonsillectomy from 1998 to 2002 with 23,184 undergoing tonsillectomy from 2013 to 2017. There was a notable increase in the adult 30-day re-admission rate from 9.8% to 19.9% between the two time periods, and also a small increase in adults returning to theatre for arrest of haemorrhage from 3.6% to 4.9%. The change was possibly thought to be the result of an underestimation of previous rates and an alteration in patient demographics.⁸ Temporal changes in complication rates were also studied in a much larger German all-age series ($n = 1,452,637$) from 2005 to 2017.¹⁴ As in the Swedish study of primary haemorrhage,⁴¹ male sex and increasing age were significant risk factors for haemorrhage. The sex difference was the largest in 20- to 25-year-olds, in whom males had a 12% haemorrhage rate, twice that of females, but the overall risk of bleeding was stable throughout the time period.¹⁴ The Danish series from 2003 to 2005 showed that 4% of 614 patients of all ages were hospitalised because of bleeding, with 2% of adults returning to the theatre.³¹

Dysgeusia

Taste and throat sensation changes were assessed in >180 patients 2 weeks and 6 months after tonsillectomy. At 2 weeks, 32% of patients had taste disturbance, with 8% having persisting disturbance at 6 months, which was most often a metallic taste, while 20% of patients had a foreign body sensation.⁴³ Late persistence (up to 3 years) was observed in about 1% of patients.²⁹

Conservative therapy for recurrent sore throats

Antibiotics

When considering the use of antibiotics for recurrent sore throats, sore throats caused by viral inflammation will not respond to antibiotics. In practice, the distinction between tonsillitis and other forms of sore throat can be problematic. In primary care, inappropriate antibiotic prescribing remains common.⁴⁴ Patients who have three out of four of the Centor Clinical Prediction Rule criteria⁴⁵ (i.e. fever, tonsillar exudate, absence of cough and tender neck lymph glands) are more likely to have bacterial tonsillitis.⁴⁶ These criteria are widely advocated but lack specificity.⁴⁷ Nonetheless, the Centor criteria remain a key part of the current NICE decision-making guide for the use of antibiotics in treating sore throat¹⁹ because there is no good guidance on point-of-care detection of the typical bacterium, group A *Streptococcus*.⁴⁸ Antibiotic overuse in unselected community populations with viral/non-bacterial pharyngitis is costly for healthcare providers,⁴⁹ risks selection of resistant 'superbugs' and provides a rationale for the continual efforts to try to curtail antibiotic prescription in general practice.⁵⁰ A comparison of immediate, no and delayed antibiotic prescribing found that the main effect of antibiotic use was the promotion of medical consultation for sore throat.⁵¹ Patient-focused interventions, such as delayed prescription or information resources, did appear in a systematic review to affect antibiotic usage.⁵² However, those with more frequent and incapacitating episodes of sore throat require special consideration.¹⁸ It must be said, however, that the whole area is beset by different interpretations by patients, clinicians and researchers of the boundary between sore throat/non-specific pharyngitis and tonsillitis. Strictly speaking, of course, any pharyngitis is likely to involve some degree of inflammation of the tonsils.

A Dutch trial⁵³ of the efficacy of 7 days of penicillin compared with (1) 3 days of penicillin and (2) placebo recruited 561 patients. The 7-day course of penicillin shortened the duration of illness by 1.7–1.9 days (2.5 days in those with group A streptococcal infection).⁵³ These results were, however, in contrast to a larger earlier UK randomised controlled trial (RCT),⁵⁴ which found that antibiotics yielded only 1 fewer day of fever and a non-significant reduction in days of sore throat. Any such benefit of antibiotic treatment must, moreover, be balanced with the costs, side effects, risk of emergent antibacterial resistance and medicalisation of responses to less severe sore throats.^{47,51} Penicillin V remains the antibiotic of choice for streptococcal pharyngitis,^{55,56} but up to 25% of patients⁵⁷ fail to complete the recommended course of treatment.^{57,58} Traditionally, patients with streptococcal carriage have been given a larger number of antibiotic prescriptions than non-carriers,⁵⁹ and modern algorithms remain heavily predicated on protection from streptococcal infection,⁶⁰ where a delayed rather than an immediate prescription appears.⁶¹

Steroids have been considered as an alternative treatment for sore throat, but the UK Treatment Options without Antibiotics for Sore Throat (TOAST) study found insufficient evidence to support a single dose of dexamethasone.^{5,62} A recent update of a systematic review of the role of corticosteroids in sore throat comprised 1319 participants, including 950 adults.⁶³ Corticosteroids increased the likelihood of complete resolution of pain at 24 hours by 2.40 times. No differences were found in recurrence or in days lost from normal activities⁶³ and, in view of the potential adverse effects, the authors concluded that further research is required.

Assessment of baseline sore throat severity

Our previous North of England and Scotland Study of Tonsillectomy and Adenotonsillectomy in Children (NESSTAC)^{3,64} showed that the procedure reduced the prevalence of sore throats during the 24 months following tonsillectomy, but did not allow accurate mapping of outcomes to baseline severity. Baseline severity in tonsillitis, as in many other conditions, is an important predictor of outcome and may be the main determinant of patient satisfaction with tonsillectomy.⁶⁵ NESSTAC's inability to stratify outcomes according to disease severity or QoL impact significantly curtailed the trial's impact on clinical decision-making.

The assessment of baseline severity in tonsil disease has been simplified by the advent of the Tonsillectomy Outcome Inventory-14 (TOI-14) patient-reported outcome (PRO). The TOI-14 depicts the health-related QoL of patients with chronic tonsillitis and is validated for the German population.⁶⁶ The English TOI-14 paediatric version has also been shown to have good psychometric characteristics.⁶⁷ The TOI-14 adult version invites patients to reflect on the prior 6 months when responding. It is a straightforward PRO with appropriate face validity. We have shown the TOI-14 to have discriminant validity in a comparison of recurrent tonsillitis patients with healthy volunteers⁶⁸ and in peritonsillar abscess.^{69,70} Importantly, tonsillitis is a disease primarily of younger adults, usually with education or workplace activity profiles. Small descriptive series report substantial benefits in reduction in days lost from work or education following tonsillectomy.⁷¹

The NATTINA RCT was a mixed-methods study, incorporating a feasibility study, an internal pilot and a Phase III multicentre trial, randomising patients to immediate tonsillectomy or conservative management. Access to tonsillectomy in the UK is currently governed by application of SIGN/NICE guidance. This mandates that qualifying sore throats are well documented, clinically significant and adequately treated. The 'required' number of episodes for consideration of NHS tonsillectomy is seven in the previous year, five in each of the two consecutive previous years, or three in each of the three preceding years.¹⁸ Our PPI NATTINA preparatory work reported varied access to antibiotic therapy for tonsillitis in primary care, possibly as some clinicians' therapeutic research goal has been 'sore throat days saved'. In contrast to professionals, patients reported that they are often looking to minimise days of feeling ill rather than days of sore throat. Thus, any tonsillitis/pharyngitis treatment trial needs to reflect a wider range of outcomes than a simple count of sore throat days.⁵⁴

Research is required in the form of a modelling approach to identify what optimal treatment strategies might be and to inform future research. Patients in NATTINA were asked to identify at the time of a reported sore throat episode whether they regarded it as being (1) mild, (2) moderate or (3) severe. This distinction was supported by work in primary care showing that patients' own perception of sore throat severity was the single best predictor of streptococcal tonsillitis.⁷²

Rationale

Context

UK hospital admissions related to pharyngeal inflammation

The context of NATTINA was one of increasing admissions to hospital with acute tonsillitis and complications of tonsillitis,^{73,74} including sepsis,⁷⁵ and twice as many annual hospital admissions in England for throat infections as for tonsillectomy. The systemic features of tonsillitis in adults together with economic drivers may mean that early intervention is superior to delayed therapy.

Practice variation

Practice variation is a longstanding feature of tonsillectomy within and between countries. Paediatric tonsillectomy rates in England were noted some time ago to be seven times higher in some regions than in others.²⁴ More recently, in Scotland, considerable variation was found to persist, despite fairly consistent application of prevailing guidance across the country,²⁵ and in Germany inter-regional variation is described in all age tonsillectomy rates.⁷⁶

Access to tonsillectomy in 21st century UK

Current UK tonsillectomy practice is governed by SIGN guidance,¹⁸ outlined above, which was, perforce, drawn up without the benefit of level I or health economic evidence. By modelling the costs and consequences and setting these against the baseline TOI-14, patients, clinicians and health service

funders will be presented with a range of options as to what might be the optimum threshold for surgical intervention. Looking at the considerable excess of adult medical admissions with inflammation of tonsils over the number of adults with tonsillectomy, it may be that, for adults with recurrent tonsillitis, the present UK consensus-based surgical threshold¹⁸ has been set too high.¹¹

To our knowledge, there has been no previous attempt to map the current NHS referral criteria against any other metrics of severity. In other words, the characterisation of the population currently undergoing tonsillectomy is incomplete. NATTINA, which tracked the progress of over 450 adults across 27 UK centres, sought to address this. The modelling of outcomes in relation to baseline presentation generated a number of pathways for evaluation by patient, clinician and funding stakeholders. In addition, the trial design aimed to capture anonymised TOI-14 data from potentially eligible patients who declined trial recruitment, not only to check the representativeness of the trial sample but also to increase the bank of TOI-14 data captured nationally.

Aims and objectives

The overall aims of NATTINA were to:

- determine a design that will result in a trial achieving recruitment target in a timely fashion through a preliminary feasibility and internal pilot trial rehearsal
- establish, via the definitive, multicentre RCT, the clinical effectiveness and cost-effectiveness of tonsillectomy, compared with conservative management, for tonsillitis in adults
- evaluate the impact of alternative sore throat patient pathways by observation and statistical modelling of outcomes
- compare the costs and consequences (symptomatic and QoL) for patients and the NHS from each management strategy
- inform future research.

Feasibility study aim and objectives

The aim of the feasibility study was to address the key methodological issues raised via our PPI group and clinicians.

The objectives of the feasibility study were as follows:

- Evaluate otolaryngologists' and recruiting specialist nurse practitioners' willingness to randomise and randomly allocate patients to the randomised arms, and patients' willingness to be randomised, taking account of predicted variation in severity of sore throat.
- Identify barriers to and facilitators of, and propose strategies to address, the above items.
- Define clear-cut eligibility criteria for a trial acceptable to all stakeholders.
- Assess the acceptability of the usual-care randomised arm to patients whose primary care clinicians have referred them for specialist intervention. In addition, to review primary care clinicians' willingness to refer.
- Provide further information to maximise patients' equipoise when they present in the tonsil study secondary care clinic, including meetings/seminars with research participants from other sites.
- Investigate the feasibility of our proposed data collection methods and outcome measures.
- Explore illness features of concern to our PPI group to ensure that these were captured in trial outcome measures.
- Develop and test sore throat weekly data collection methods and storage, and Sore Throat Alert Return (STAR).
- Explore the processes of patient identification and recruitment.

- Develop patient recruitment materials, including the production of the standardised NATTINA randomisation online video-recording/digital versatile disc (DVD).⁷⁷

Patients who met the proposed inclusion criteria were identified and invited by study otolaryngologists to participate in the feasibility study (involving qualitative interviews, role play and cognitive interviews).

Internal pilot objectives

The NATTINA internal pilot target was 72 patients randomised in 6 months among six of our nine sites. In addition to assessing the ability to recruit, the internal pilot aimed to:

- ascertain whether or not all of the trial processes, including patient identification, eligibility criteria, randomisation and data collection processes, work as intended and cohesively
- gauge more precisely the number of potentially eligible patients identified in NATTINA screening clinics
- investigate referral, recruitment and acceptability across the baseline severity spectrum
- identify barriers to patient recruitment, suggest improvements to have an impact on recruitment rates and measure patient compliance in weekly submission of days of sore throat, plus STARs, during sore throat episodes
- identify any major emerging systematic differences between recruited patients and patients who declined to participate
- collate and report reasons for ineligibility/non-consent
- quantify missing data and measure attrition in sore throat data
- review activity against the go criterion – six screening clinics established, with target throughput of 396 eligible patients screened in 6 months and a minimum target of 72 randomised patients.

The structure of the internal pilot was, however, eventually abandoned, as described in [Chapter 2](#), owing to unexpectedly severe barriers to recruitment, particularly in certain regions of England where Clinical Commissioning Groups (CCGs) appeared to be following an audit commission health briefing, which labelled tonsillectomy among a number of ‘relatively ineffective procedures’.⁴

Main trial objectives

Primary objective

- To compare the clinical effectiveness (as number of sore throat days) and cost-effectiveness, over 24 months, of tonsillectomy with conservative management in adults with recurrent acute tonsillitis who have been referred to otolaryngology outpatient clinics.

Secondary objectives

Clinical effectiveness

- To provide group and subgroup comparisons of:
 - other metrics of sore throat severity, including responses on the TOI-14 and STARs for any sore throat episodes experienced
 - QoL, as recorded by the Short Form questionnaire-12 items (SF-12), from baseline throughout the study and during any episodes of sore throat experienced.
- To report the number of adverse events (AEs) and additional interventions required, supported by linkage to community prescribing data and, in Scotland, by Information Services Division linkage to SMR01 (acute hospital discharge episodes).

INTRODUCTION

- To adjust the estimate of effectiveness in the light of other baseline covariates, including severity of tonsillitis.⁷⁸
- To evaluate the impact of alternative sore throat patient pathways by observation and statistical modelling of outcomes.
- To assess to what extent trial participants are representative of the total population of sore throat patients referred to ear, nose and throat (ENT) clinics.

Economic evaluation

- To compare cost-effectiveness measured in terms of the incremental cost per sore throat day avoided from the perspective of the NHS and Personal Social Services (PSS) over 24 months.
- To compare cost-effectiveness with outcomes reported as incremental cost per quality-adjusted life-year (QALY) gained from the perspective of the NHS and PSS over 24 months.
- To compare cost-effectiveness with outcomes reported as net monetary benefits from the perspective of the NHS and PSS over 24 months.
- To compare costs to the NHS and patients of the different randomised treatments over 24 months.
- To compare QALYs based on SF-6D values derived from responses to the SF-12 administered at baseline and 6, 12, 18 and 24 months post randomisation and when participants' experienced a sore throat.
- To assess participants' willingness to pay (WTP) to avoid a day with sore throat derived from a contingent exercise administered at baseline and to compare QoL valued in monetary terms (i.e. WTP to avoid a day with sore throat × number of days of sore throat experienced).

Qualitative process evaluation

- To document the views and experiences of patients and clinicians regarding tonsillectomy and conservative management and how patient experience may shape any future research required.

Dissemination

- To consider the role of the trial outputs in informing shared decision-making between adults with recurrent tonsillitis and their GPs.
- To inform decisions by healthcare professionals in secondary care and their clinical commissioners.

Chapter 2 Methods

This chapter includes information about the NATTINA trial participants, interventions, objectives, outcomes, sample size, randomisation and statistical methods, and a summary of any substantial changes to the protocol during the trial.

Overview of the trial design

The NATTINA trial was a multicentre, randomised controlled, two-arm surgical trial that was set in NHS secondary care within the UK; this was funded by the National Institute for Health and Care Research (NIHR).

Adult patients aged ≥ 16 years who presented with recurrent acute tonsillitis and were referred to ENT outpatient clinics were identified and recruited.

The trial aimed to establish the effectiveness and efficiency of tonsillectomy for recurrent sore throat compared with non-surgical management over a 24-month period. Participants were randomised in a 1:1 ratio to receive immediate tonsillectomy or conservative management with surgery deferred.

The trial included a qualitative feasibility study,^{1,79} a 6-month internal pilot, a qualitative process evaluation (see [Chapter 5](#)), a full statistical evaluation (see [Chapter 3](#)) and an economic evaluation (see [Chapter 4](#)).

Trial registration and protocol availability

The trial was registered in the International Standard Randomised Controlled Trial Number registry on 4 August 2014: registration number ISRCTN55284102.

The latest version of the full protocol is available on the NIHR Health Technology Assessment (HTA) project web page (www.journalslibrary.nihr.ac.uk/programmes/hta/1214606/#/; accessed 7 October 2020).

Ethics and governance

The sponsor for the trial was the Newcastle upon Tyne Hospitals NHS Foundation Trust (reference 07065). Favourable ethics opinion was obtained on 10 November 2014 from the NHS Research Ethics Service Committee North East – Tyne and Wear South [Research Ethics Committee (REC) reference 14/NE/1144]. Subsequent approval was gained for the 11 substantial protocol amendments (see [Appendix 1, Table 28](#)).

Setting

The trial was conducted in the following 27 NHS secondary care hospitals in England, Scotland and Wales (the date of opening is presented in brackets):

- Aneurin Bevan University Health Board (12 July 2016)
- Bradford Teaching Hospitals NHS Foundation Trust (20 July 2015)

METHODS

- City Hospital Sunderland NHS Foundation Trust (now South Tyneside and Sunderland NHS Foundation Trust) (28 July 2015)
- Dorset County Hospital NHS Foundation Trust (16 February 2016)
- East and North Hertfordshire NHS Trust (2 March 2017)
- East Lancashire Hospitals NHS Trust (Royal Blackburn Hospital) (13 September 2016)
- East Lancashire Hospitals NHS Trust (Royal Preston Hospital) (30 September 2016)
- Frimley Health NHS Foundation Trust (11 March 2016)
- Guy's and St Thomas' NHS Foundation Trust (22 August 2014)
- James Paget University Hospital NHS Foundation Trust (5 October 2016)
- NHS Ayrshire and Arran (2 March 2017)
- NHS Grampian (30 April 2015)
- NHS Greater Glasgow and Clyde (30 April 2015)
- NHS Tayside (9 September 2014)
- North Cumbria University Hospitals NHS Trust – Cumberland Infirmary (now North Cumbria Integrated Care NHS Foundation Trust) (15 February 2016)
- Nottingham University Hospitals NHS Trust (18 July 2016)
- Oxford University Hospitals NHS Foundation Trust (7 April 2017)
- Plymouth Hospitals NHS Trust (29 September 2016)
- Salisbury NHS Foundation Trust (18 August 2016)
- Sheffield Teaching Hospitals NHS Foundation Trust (24 March 2016)
- South Tees Hospitals NHS Foundation Trust (13 October 2016)
- The Newcastle Upon Tyne Hospitals NHS Foundation Trust (15 April 2015)
- United Lincolnshire Hospitals NHS Trust (13 February 2017)
- University Hospitals Birmingham NHS Foundation Trust (23 October 2016)
- West Suffolk NHS Foundation Trust (26 September 2016)
- Worcestershire Acute Hospitals NHS Trust (26 August 2016)
- Wrightington, Wigan and Leigh NHS Foundation Trust (8 February 2016).

Participants

A total of 455 NATTINA participants were recruited and randomised following referral by their GP to an ENT clinic at one of the participating centres. Two individuals were randomised into the study in error; no data were collected for these individuals and they were withdrawn. All remaining individuals were assessed as eligible for tonsil dissection in accordance with prevailing NHS regulations. RCT participants were recruited from 22 of the 27 open sites.

Inclusion criteria

- Aged ≥ 16 years.
- Recurrent sore throats that fulfil current SIGN guidance¹⁸ for elective tonsillectomy. This guidance stipulates that to be eligible for tonsillectomy on the NHS sore throats should be a result of acute tonsillitis and that:
 - the episodes of sore throat should be disabling and prevent normal functioning
 - the patient has experienced seven or more well-documented, clinically significant, adequately treated sore throats in the preceding year, five or more such episodes in each of the preceding 2 years, or three or more such episodes in each of the preceding 3 years.
- The patient has provided written informed consent for participation in the trial prior to any trial-specific procedures.

Exclusion criteria

- Previous tonsillectomy.
- Listed directly (i.e. added to waiting list without prior elective ENT outpatient appointment) during emergency admission (e.g. owing to peritonsillar abscess/quinsy).
- Primary sleep breathing disorder.
- Suspected malignancy.
- Tonsilloliths (as primary referral).
- Pregnant or breastfeeding.
- Bleeding diathesis (including haemophilia, sickle cell disease and platelet dysfunction).
- Therapeutic anticoagulation.
- Inability to complete self-reported questionnaires and STARS.

Intervention

Current standard care, according to NICE and SIGN,^{18,19} is for adult patients with a qualifying frequency of recurrent sore throat to be referred by their GP to ENT outpatient clinics to be considered for tonsillectomy. Those confirmed by a member of the specialist team as being eligible for tonsillectomy, who consent to surgery and who do not have any pressing medical exclusions are then offered elective tonsil dissection. Non-surgical candidates revert to conservative management in primary care (analgesia with or without antibiotic therapy). NATTINA aimed to be a pragmatic trial, reflecting normal NHS practice; therefore, the process of referral to an ENT clinic by GPs was not altered. Eligible consenting participants were randomly allocated to one of two arms: tonsillectomy or conservative management.

Tonsillectomy

Participants randomised to the tonsillectomy arm received elective surgery to dissect the palatine tonsils, preferably within 6 weeks and no longer than 8 weeks following randomisation. The method of the surgical procedure was at the discretion of staff at the participating sites and was identical to that in standard care. Except for randomisation allocation, participants in both arms received standard non-surgical care, as per local policies. Participants who were randomised to undergo a tonsillectomy, but whose surgery was delayed owing to severe tonsillitis or other reasons, remained in the trial and continued to follow the surgery pathway.

Conservative management

Participants randomised to the conservative management arm had the option of deferred surgery for up to 24 months in addition to standard non-surgical care. A 12-month face-to-face review allowed confirmation of willingness to remain in the conservative management arm. This pathway was informed by the NATTINA PPI group and the qualitative feasibility study for NATTINA.^{80–82} A participant contemplating surgery was required to fulfil the SIGN guidance¹⁸ at the time point of review to be considered for tonsillectomy.

Participants in the conservative management arm received standard non-surgical care, which typically comprised self-administered analgesia plus or minus ad hoc primary care prescription of antibiotics, and attendance at walk-in centres or the accident and emergency (A&E) department for more severe episodes of sore throat. Participants were permitted to cross over to the tonsillectomy arm at any point in the trial on the condition that they continued to be eligible for surgery in accordance with the SIGN guidance.¹⁸

Funding of the trial intervention

Funding for the surgery was not required, as the procedure was a part of the usual standard-care pathway for this cohort of patients.

Outcome measurement

Primary outcome

The primary outcome was the number of sore throat days collected through weekly 'returns' from the participants over the 24-month period following randomisation. The data allowed comparison of tonsillectomy with conservative management to determine their effectiveness in treating recurrent sore throat.

Primary outcome data collection

A discrete database was designed for use in the trial to collect the number of sore throat days that participants experienced per week. The company who provided the software and who worked with the trial team to design the STAR database was Tay Dynamic, later rebranded as Intelem (Huddersfield, UK) and then as CI Data (but referred to throughout this report as Tay Dynamic). Tay Dynamic specialised in software for direct participant communication.

The weekly number of sore throat days was collected from 1 week after randomisation, throughout the follow-up period, to the 24-month follow-up visit.

Following recruitment, the Newcastle Clinical Trials Unit (NCTU) team liaised with sites to register the participant into the STAR system, indicating their preferred method of contact: e-mail, text message or IVR via telephone. Participants' identifiable data were stored securely on a Newcastle University (Newcastle, UK) server, and Tay Dynamic were provided with access to e-mail addresses and telephone numbers for the purpose of sending out the alerts and STAR questionnaires.

The participants then received weekly alerts asking them how many sore throat days they had experienced in the previous 7 days. Their response (a number from 0 to 7) was automatically recorded and a reminder was sent to complete a STAR questionnaire if the participant had experienced any sore throat days (score of >0). The STAR questionnaire collected the following secondary outcome data:

- the severity category of sore throat days (mild/moderate/severe)
- the name and frequency of over-the-counter and prescription medications used
- the nature of any professional healthcare advice sought (e.g. GP, walk-in clinic and pharmacist)
- the number of hours unable to undertake usual activities (including time off work and studies)
- the SF-12 questionnaire⁸³ (see also [Chapter 4](#)).

In the initial trial design, the STAR questionnaire was planned to be a paper form that would be completed and (free) posted to NCTU. Substantial amendment 1 (see [Appendix 1, Table 28](#)) was implemented prior to the recruitment start date, allowing the STAR questionnaire to be completed and submitted electronically if participants had access to a mobile telephone or computer. Those without such access submitted a paper copy.

Fifteen months into recruitment, it became apparent that the rate of electronic STAR responses considerably outnumbered the rate of paper STAR questionnaire returns. The STAR system was, therefore, modified (see substantial amendment 5 in [Appendix 1, Table 28](#)). The report of a sore throat day triggered a second STAR e-mail/text/call asking for a severity grade (1, mild; 2, moderate; 3, severe) (i.e. corresponding to the first item on the original the STAR questionnaire) (see [Chapter 4](#)). Receipt of the severity response generated a further text/e-mail request to submit an electronic or paper STAR questionnaire (one per 7-day reporting period).

The Tay Dynamic software identified non-responders. If participants did not return STAR questionnaires after several sore throat day reports, the trial management team issued at least one reminder letter.

Approximately 24 months after the recruitment of the first participant, the items on the STAR questionnaire used to inform the economic analysis were abbreviated as the STARLET questionnaire (see substantial amendment 7 in [Appendix 1, Table 28](#)):

- whether or not the participant had taken any over-the-counter and prescribed medications
- number of days unable to undertake work and usual activities
- SF-12.

The main weekly STAR text message/e-mail/IVR system continued to collect information on the number of sore throat days and the severity category of the sore throat episode. Report of at least 1 sore throat day in the previous week generated a request to complete the STARLET. Details of how the economic data from the STAR and STARLET were combined are provided in [Chapter 4](#).

Secondary outcomes

The secondary outcomes were:

- Clinical effectiveness. Intergroup comparison over the 24-month follow-up period of:
 - STAR severity scores of any sore throats experienced.
 - TOI-14^{66,68} scores collected 6-monthly. The TOI-14 is a 14-item questionnaire with a supplementary final open response item for additional symptoms that are related to tonsillitis over the preceding 6 months.⁸⁴ We have shown the TOI-14 to have discriminant validity in a comparison of recurrent tonsillitis patients with healthy volunteers⁸⁵ and in peritonsillar abscess.^{69,70}
 - An adjusted estimate of intervention effectiveness in the light of baseline covariates, including severity of tonsillitis, estimated from TOI-14 baseline scores and categories.
 - Intergroup comparison of QoL as recorded over 24 months by the SF-12.⁸³ This 12-item Short Form Survey is an abbreviated Short Form questionnaire-36 items generic QoL assessment tool, covering the same eight health domains of physical and mental health over the previous 4-week period.
 - The impact of alternative NHS sore throat patient pathways by observation and statistical modelling of outcomes.
 - To assess to what extent trial participants are representative of the total population of sore throat patients referred to ENT clinics through interrogation of site screening logs.
- Economic evaluation. To compare the following between the two arms over 24 months:
 - Costs. The average total cost per participant to the NHS and PSS to manage recurrent sore throats. Participant costs were considered in a sensitivity analysis.
 - QALYs using the area-under-the-curve method based on SF-6D scores derived from the SF-12 responses.⁸⁶
 - Willingness to pay to avoid a sore throat day. A contingent valuation questionnaire was administered at baseline to quantify the effect that a sore throat day has on individuals' QoL in monetary terms.
 - The number of AEs, visits to the GP/walk-in clinic/A&E, prescriptions issued and additional interventions required for sore throats and related events through STAR/STARLET data, and supported by data linkage to primary care patient records.
 - Cost-effectiveness analysis (CEA) – incremental cost per sore throat day avoided.
 - Cost-utility analysis (CUA) – incremental cost per QALY gained.
 - Cost-benefit analysis (CBA) – incremental net benefit.
- Qualitative process evaluation. To document the views and experiences of patients and clinicians regarding tonsillectomy and conservative management, and how patient experience may shape future research (see [Chapter 5](#)).
- Future research. To propose further research questions using newfound cost-benefit information to develop algorithms that guide and assess management of health services.

Secondary outcome data collection

The questionnaires used for secondary outcome data collection can be found on the project webpage (<https://fundingawards.nihr.ac.uk/award/12/146/06>; accessed 29 September 2021). Participant-reported outcomes were collated into questionnaire packages for each 6-month contact (Figure 1 and Table 1).

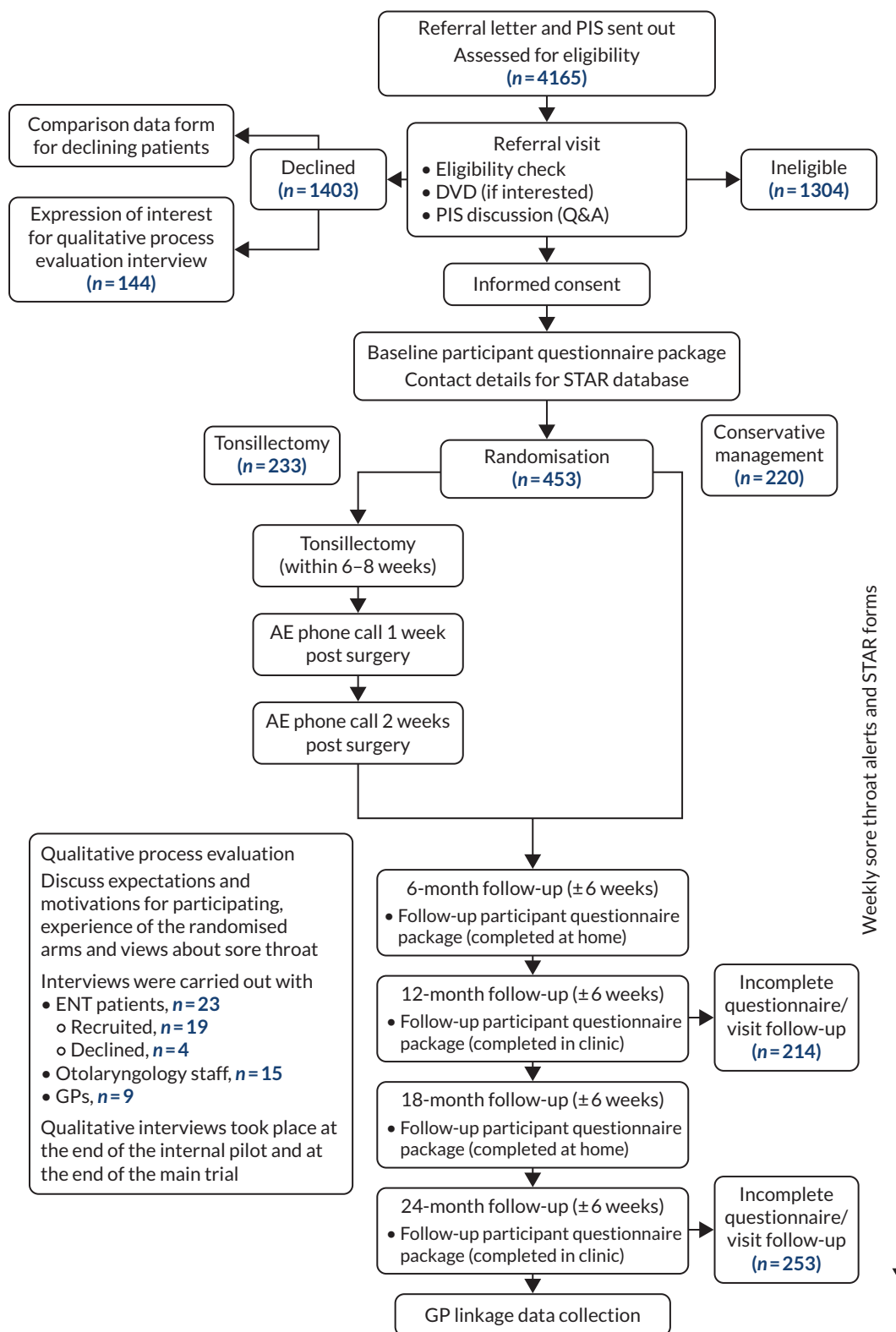


FIGURE 1 Study pathway. PIS, participant information sheet; Q&A, question and answer.

TABLE 1 Details of questionnaires included in the packages at each follow-up contact

Visit	Questionnaires collected at each time point					
	TOI-14	SF-12	Health Service Utilisation	About You	Value of Avoiding a Sore Throat Day	Participant Time and Travel
Baseline	X	X	X	X	X	
6 months	X	X	X			
12 months	X	X	X			
18 months	X	X	X			X ^a
24 months	X	X	X			

a At the 18-month visit, the Participant Time and Travel questionnaire was also completed, but not by all participants as it was removed in an attempt to increase the response rates of the postal 18-month questionnaire in substantial amendment 9. Schedule amendments are given in [Appendix 1, Table 28](#). Those undergoing surgery were telephoned at 1 and 2 weeks to record any related AEs or serious adverse events experienced.

The 'About You' set of questions collected data about participants' ethnic origin, education and employment. Primary healthcare service data were accessed from the participants' general practices following the completion of the 24-month follow-up visit, including the 12 months prior to randomisation. NCTU initially sent a GP linkage case report form for completion by the practice for each participant, along with a prepaid return envelope. A move to an electronic process achieved a higher response rate. The participant record report form defined the time window of enquiry – 12 months prior to 24 months post randomisation – on the following episodes for a sore throat or related event:

- attendance at a general practice (with a GP/nurse)
- telephone call with a GP/nurse from the practice over the telephone
- attended a walk-in clinic
- attended A&E
- attended an outpatient appointment
- admitted to hospital
- had any other intervention or contact with a healthcare service
- any medications prescribed.

The purpose of collecting these data was to capture participants' consultation rates, outcomes, prescribing information and additional interventions required for sore throat or related events throughout the period stipulated above. GP practices were paid £50 for each participant return.

Economic analysis

The NATTINA economic analysis aimed to determine the relative efficiency of tonsillectomy compared with conventional management, and is detailed in [Chapter 4](#).

Qualitative analysis

The NATTINA qualitative analysis aimed to document the views and experiences of patients and clinicians regarding tonsillectomy and conservative management, and how patient experience may shape any future research required, and is detailed in [Chapter 5](#).

Participant timeline

Feasibility study

The aim of the feasibility study was to assess the willingness of clinicians to refer and randomly allocate patients to the study arms, patients' willingness to be randomised and the acceptability of the deferred surgery randomised arm.

Qualitative and cognitive interviews were carried out by a researcher from Newcastle University. The study was planned to last for approximately 5 months, or until data saturation, and to take place over nine sites.

Fifteen ENT patient interviews were planned: these were with adult patients with acute tonsillitis who had been referred to ENT outpatient clinics for recurrent sore throat and who met the eligibility criteria (inclusion and exclusion criteria for the feasibility study were the same as for the main trial listed in [Participants](#)).

Twenty ENT staff interviews were planned for ENT staff members at one of the nine participating sites that were likely to be involved in the proposed NATTINA trial. Ten GP interviews were planned, with no eligibility criteria for this group of participants.

The outcome measures were:

- Patient identification and recruitment, eligibility criteria and ENT clinic staffs' willingness to randomise and randomly allocate patients to the randomised arms.
- Primary care clinicians' willingness to refer.
- Patients' willingness to be randomised and their acceptability of the conservative management arm for patients whose primary care clinicians have referred them for specialist intervention. Feedback on proposed data collection methods, including weekly sore throat alerts, STARs and questionnaires, was also collected.

In total, 15 ENT patients, 11 GPs and 22 ENT staff were recruited.

The success criteria for progression to the internal pilot study were met, with a plan to address potential barriers to recruitment and patient equipoise. Patient participants reported that recurrent sore throats severely affected their family, work and social life; ENT staff reported that patients faced increasing barriers to secondary care service access; and GPs were guided to reduce referrals for tonsillectomy owing to it being seen as having limited clinical value. This highlighted the need for further research regarding the effectiveness of tonsillectomy for this cohort of patients.

Internal pilot

NATTINA was originally planned to include a 6-month internal pilot with six of its nine sites that were planned to be opened to recruitment. The process was overseen by the Data Monitoring Committee (DMC) and the Trial Steering Committee (TSC) and was reviewed by the NIHR HTA programme to decide whether or not to continue with the main trial phase.

The proposed 'stop/go criteria' to be achieved in the 6-month pilot⁷⁹ were:

- six sites to be opened with screening clinics established
- 396 eligible patients screened over the six sites
- a minimum of 72 participants randomised.

Other aims were to:

- ascertain if all trial processes, including patient identification, eligibility criteria, randomisation and data collection, worked as intended and if the eligibility criteria were cohesively operational
- gauge more precisely the number of potentially eligible patients identified in NATTINA screening clinics
- investigate referral, recruitment and acceptability across the baseline severity spectrum
- develop participant recruitment materials, including the production of the standardised NATTINA randomisation online video-recording/DVD
- identify barriers to participant recruitment and suggest improvements to have an impact on recruitment rates
- measure participant compliance with the proposed weekly submission of number of sore throat days per week, plus STARs during sore throat episodes
- develop and test sore throat weekly data collection methods and storage, and STARs (participant interviews, cognitive interviews) with input from Tay Dynamic software
- explore illness features reported to us as of concern by our PPI group to ensure that they were captured in trial returns, for example chronic sore throat, night-time sore throat and disruption of ability to undertake not only usual activities but also leisure pursuits
- identify any major emerging systematic differences between recruited participants and individuals who declined to participate
- collate and report reasons for participation/ineligibility/decline
- quantify missing data and measure attrition in sore throat data.

The recruitment target of 72 participants was, however, not met in the pilot timescale. The Trial Management Group (TMG) and the NIHR HTA programme monitoring team met on 10 September 2015, 4 months from the first patient's first visit, and agreed that the proposed nine-site model would not succeed. The conversion rate from screening to recruitment had been lower than expected, and the embedded qualitative work^{81,82} showed that this was because patients' willingness to participate in the trial, particularly in certain regions, was much less than that anticipated. Some patients were having to wait so long to gain a referral to ENT clinics that they were unwilling to contemplate the prospect of further delay. The NIHR HTA programme team suggested a recovery plan with an increase in the number of sites from 9 to at least 20, as rapidly as was practicable. It also came to light that a draft sample size of 600 had been carried forward into the trial, but in fact applied to an earlier design with a third arm; therefore, the total number of participants required was only 510, not 600. The recovery plan was discussed at a second meeting on 23 May 2016 and was agreed in summer 2016.

Identification, screening and recruitment of participants

The clinical team identified potential participants who had been referred by a GP to be considered for tonsillectomy. A participant information sheet (PIS) along with an invitation letter and appointment letter, if appropriate, were posted out to the patients. The NATTINA PIS outlined details of the trial and signposted potential participants to the NATTINA information DVD/online video-recording on the trial website.⁷⁷ The recruitment DVD/online video-recording was developed for the study because it was found to be useful in the NESSTAC trial.^{3,64} Video information is more accessible to those who do not read extensive documents regularly and also ensured a degree of consistency of verbal information given across all of the trial sites.

As many patients as possible who were referred to ENT clinics with recurrent sore throat were screened for NATTINA eligibility. All site staff who were involved in screening and recruitment activities were fully trained on the trial protocol, had the necessary experience and qualifications to perform these tasks, and were delegated to do so on the site delegation log. Potential participants, who were previously posted a PIS, were shown the information DVD/online video-recording at their referral visit (unless already viewed online)⁷⁷ and were given the opportunity to discuss the trial with the designated member of the research team. Patients interested in NATTINA participation underwent an entry eligibility criteria check. Eligible patients were invited to join the trial. Potential participants at the clinic who did not

previously receive a PIS were given verbal information. Those with initial interest took away a PIS to read and consider, and if they were interested a further appointment was made for the consent visit, allowing a minimum intervening interval of 24 hours for their decision-making.

Screening and recruitment logs were completed at all sites to document all subjects who attended a referral visit, including their outcome status, reasons for ineligibility, if applicable and, where offered, reason for declining participation.

Declining patients

Patients who were eligible but decided not to take part were invited to provide anonymised baseline comparison data for the NATTINA database. These included age, sex, an estimate of number of sore throat days over the prior 6 months and the completion of a TOI-14 questionnaire. This allowed an analysis of the comparability of trial participants with the total pool of those referred at each of the sites. Patients who declined were also invited to take part in a qualitative interview with a researcher (see [Chapter 5](#)). Those who declined the trial will be the subject of a separate publication, incorporating factor analysis of the TOI-14 (in preparation). For the number of declining patients with data, see [Table 5](#). Information on reasons for declining to be randomised and the missing TOI-14 items for these patients are presented separately in [Appendix 3, Tables 31 and 32](#).

Consent procedure

Eligible patients wishing to take part were given the opportunity to discuss the trial and ask any questions before providing written informed consent on the informed consent form (ICF), which was witnessed, signed and dated by a member of the research team who had documented, delegated responsibility to do so. All consent forms were monitored by the NCTU for completeness and accuracy.

Randomisation

Randomisation was carried out via the NCTU secure web-based computer allocation system by delegated site research team members.

Participants were randomised to receive tonsillectomy or conservative management in a 1:1 ratio stratified by centre and severity using a blocked allocation (permuted random blocks of variable length) system.

The participant's severity category was determined by the total TOI-14 score: mild = 0–35, moderate = 36–48 or severe = 49–70.

Treatment, crossover and withdrawals

Participants randomised to receive conservative management were asked to defer surgery for up to 24 months and were informed that they would be assessed on their willingness to continue to delay surgery at the 12-month follow-up visit. However, these participants could decide to cross over to have the tonsillectomy at any point of the study, if they continued to meet the SIGN guidelines for tonsillectomy at that time point.¹⁸ Some participants who were randomised to receive tonsillectomy (see [Figure 2](#)) changed their minds prior to receiving the surgery and crossed over to the conservative management pathway, and some crossed over otherwise, for example some failed to receive surgery owing to reasons at the site level. The investigator could also withdraw a participant from the trial intervention if this was believed to be in the participant's best interests. Any reason for treatment withdrawal offered was recorded.

The number of withdrawals was recorded along with the time from randomisation (see [Figures 2 and 3](#), and corresponding summary statistics can be found in [Appendix 8, Table 36](#)). The number of crossovers was also recorded (see [Table 6](#)). The number of participants crossing over and reasons for crossover were tabulated (see [Appendix 2, Tables 29 and 30](#)). Those crossing over randomised arms could either continue with follow-up visits and data collection or withdraw completely and undertake no further follow-up visits or data submission if this was their wish.

Participants had the right to withdraw from the trial at any time without giving a reason. All data collected up until the point of withdrawal were included within the analysis.

Schedule of events

[Figure 1](#) details the participant flow through the trial, which was designed to maximise data accrual from the greatest available number of sources.

The baseline questionnaire package was completed by the participants at the screening/baseline visit after consent and before randomisation so that the total TOI-14 score was available to inform baseline stratification. Baseline, 12-month and 24-month contacts were in the clinic. Six-month and 18-month questionnaires were postal. [Table 1](#) lists the questionnaires completed at each time point.

At both clinic visits, conservative management participants were asked if they preferred to undergo surgery. The operation date for those in the surgery arm was cross-checked in clinic to facilitate identification of sore throat days directly attributable to surgery in the subsequent statistical modelling. After the final visit, participants' GPs were asked to provide information on participants' healthcare resource use for the previous 3 years via the GP linkage case report form (see [Chapter 4](#)).

Safety

Only AEs related to the trial intervention (tonsillectomy) were reported for this trial; these were captured and documented at the 1-week and 2-week post-surgery telephone calls. Participants were asked if they had experienced any AEs immediately after, or during recovery from, the tonsillectomy. Only participants who received the surgery were contacted.

All serious adverse events (SAEs) were recorded throughout the duration of the trial for all participants, regardless of whether or not participants underwent tonsillectomy; these did not include pre-planned hospitalisations, routine follow-up of the studied condition or treatment for pre-existing health conditions, not resulting in clinical deterioration. SAEs that were related to the trial intervention have been listed for the trial (see [Appendix 10, Table 42](#)).

All AEs that occurred in this trial, whether or not they were serious, were expected following the tonsillectomy procedure, with the exception of two SAEs that were related to the anaesthetic. The following AEs may be expected after receiving tonsillectomy:

- common AEs
 - post-operative pain
 - post-operative bleeding
 - temporary changes in taste/tongue sensation
 - difficulty swallowing
 - nausea
 - vomiting
 - infection

- uncommon AEs
 - long-term changes in taste/tongue sensation
 - chip/knock out of tooth

- very rare AEs
 - death.

Frequencies were estimated from the available data. Full requirements for the reporting of AEs and SAEs were described in the protocol, including guidance for the assessment of the relatedness and causality for SAEs in relation to the intervention.⁷⁹

Definition of the end of the trial

The end of the trial was the date when the final participant's 24-month follow-up data were obtained and when all SAEs were resolved.

Participant expenses

In recognition of the completion of the weekly returns and follow-up questionnaires, the participants received two £25 gift vouchers, one after the 12-month follow-up and the second after the date of the 24-month follow-up visit, totalling £50. Participants were reimbursed reasonable travel expenses incurred as a result of taking part in the NATTINA trial. Participants who gave a qualitative interview were given a £15 gift voucher.

Patient and public involvement

Patient and public involvement was incorporated from the funding application stage throughout the trial until the very end, including anticipated input into the trial dissemination phase.

Structured interviews were performed with PPI members and were used to develop the research proposal. Following this, PPI meetings were held in England and Scotland, at which PPI representatives provided input into the recruitment strategy, including the content of the recruitment DVD/online video-recording and the adequacy and accessibility of study documents, such as the study summary sheet, PIS, ICF and STAR form. Input was also given in relation to the weekly alert system and trial advertisement. Feedback was provided on a one-to-one basis on the health economics WTP to avoid a sore throat day question in the baseline questionnaire to provide further validation for the values used. Representatives attended TSC meetings throughout the trial, and towards the end of the trial PPI representatives reviewed the WTP paper and provided input into the dissemination of the final results.

Statistical methods

Analysis followed the full statistical analysis plan (SAP) written in accordance with published guidance⁸⁷ and approved by the chief investigator (CI), TMG and TSC. The fully signed-off SAP was in place prior to any comparative analysis and the final version was in place prior to the final data lock. Analysis was reported according to Consolidated Standards of Reporting Trials (CONSORT) recommendations⁸⁸ and the data analyses were conducted in the validated statistical software package Stata[®] v16 (StataCorp LP, College Station, TX, USA).

Sample size calculation

The planned target recruitment number was 510 participants, which included 72 participants in the internal pilot. This allowed for a total loss to follow-up of 25% over 24 months. In total, 382 participants (two groups of 191 participants providing complete data at 24 months) gave 90% power to detect an effect size of 0.33 [corresponding mean intergroup difference of 3.6 days of sore throat, based on a pooled estimated standard deviation (SD) of 10.8 days], assuming a type 1 error rate of 5%. The choice of this effect size was based on a number of considerations, including our prior experience with the NESSTAC study.^{64,89}

Statistical analysis

Definition of the primary outcome measure

The primary outcome measure was the total number of sore throat days during the 24 months following randomisation. Retention was defined as the time on the trial, which was calculated as the time from the date of randomisation to the date of the last weekly response. The average time in the trial for the participants who sent at least one weekly STAR response was calculated and is presented in box plots (see [Figure 9](#)); this allowed us to assess if retention was balanced by randomised arms.

Please note that the primary analyses were quality checked and validated (August 2020) by a member of the Biostatistics Research Group (Newcastle University) who was not otherwise involved in the NATTINA trial. This validation was for the univariate [unadjusted negative binomial regression (NBR)] and for the multivariate (adjusted for stratification variables of site and baseline severity) analyses. The validation results confirmed the primary analysis results.

Defining populations for analysis

The primary statistical analyses were carried out on an intention-to-treat (ITT) basis, retaining patients in their randomised arms. However, patients could switch over from conservative management to tonsillectomy and could also switch from tonsillectomy to conservative management and remain in the trial. In the NATTINA design in this situation, patients in the conservative management arm were asked to commit to 'deferred surgery'. The implication of such crossover, which typifies surgical trials, is that the ITT analysis may produce a conservative estimate of the effect of tonsillectomy. We, therefore, also carried out supplementary analyses to compare outcomes allowing for treatment switches.

The three analyses by population were as follows:

- ITT group – all eligible and ineligible participants as randomised were included in the analysis on an ITT basis, with participants kept in their randomised arm ($n = 429$).
- Per-treatment group – all randomised participants who started treatment were included in the analysis according to the treatment that they received ($n = 429$).
 - This is in effect a per-surgery analysis.
 - A second analysis of the per-treatment group (unplanned) further split into original randomised group with surgery status (four categories).
- Per-protocol as randomised – this was limited to patients who were randomised to receive surgery and received surgery within the 8-week window and those patients randomised to the conservative management arm who did not crossover and have tonsillectomy before end of follow-up ($n = 224$).
- Sensitivity analysis population (unplanned), including all participants who had returned $\geq 80\%$ of the expected STARs ($n = 263$).
 - This population was not specified in the SAP but we carried out this sensitivity analysis as part of our assessment of missing primary outcome data.

Weekly Sore Throat Alert Return response counts over the 24-month follow-up period

Counts of weekly STAR responses in both arms from baseline to week 105 were observed. The cumulative total number of sore throat days reported on a participant level has been provided separately for each randomised arm (for surgery split into pre surgery, surgery plus 2 weeks, and the remainder of the 24-month follow-up period) (see [Figure 7](#) and corresponding summary statistics in [Appendix 12, Table 44](#)). The STARs have been sub grouped and summarised for the never surgery, pre-surgery, immediate 2 weeks post-surgery and remainder of the post-surgery follow-up. The day of surgery was identified and the ensuing two STARs were taken to reflect the 2-week postoperative period.

Only data for the 24-month follow-up after randomisation were included in the statistical analysis for all participants, despite additional STAR responses being submitted by a small number of participants post 24-month follow-up.

The total number of sore throat days was calculated by summing all of the STARs for each participant. Negative binomial regression was used to analyse the primary outcome data, which allowed for an exposure variable to be included that takes account of the number of STARs received to make up this total. However, this means that viewing the raw totals themselves for the descriptive analysis may be misleading because they do not take account of the missing returns (unless it is assumed that missing returns mean there were no sore throats to report, which is not a reasonable assumption). We therefore transformed the data so that each participant's total sore throats was divided by the proportion of STARs returned.

Incident rate ratio

In NBR, coefficients are obtained that are interpreted as the difference between the log of expected counts, corresponding to a 1-unit change in the predictor variable. The parameter estimate can also be interpreted as the log of the ratio of expected counts, technically a rate. The response variable is the total number of sore throat days reported over follow-up, which by definition, is a rate. A rate is defined as the number of events per time, so can also interpret the regression coefficients as the log of the rate ratio.

Incident rate ratios (IRRs) are presented in [Primary analyses](#) and [Sensitivity analysis](#). Note that the IRR is obtained by exponentiating the coefficients of the regression model.

Differential loss to follow-up

Of 455 patients randomised, 26 participants did not complete any weekly STAR responses: 10 randomised to receive tonsillectomy and 16 randomised to receive conservative management. This includes the two ineligible participants and was balanced across arms (see [Figure 5](#)). The details of compliance to follow-up in terms of visits and postal questionnaire completion was also balanced across arms (see [Table 6](#)). The relatively low levels of engagement with follow-up visits and postal questionnaire completion did not affect the primary analysis of the trial.

Missing primary outcome data

Weekly STAR response was categorised at each time point (6, 12, 18 and 24 months post randomisation), as (1) complete; (2) partial, where some but not all expected weekly STAR responses were received; and (3) no returns (see [Table 15](#)).

In addition to assessing missing primary outcome data (STAR responses), we also assessed the exposure variable (i.e. a measure of missing data) in terms of the four per-treatment categories in terms of how this relates to the baseline severity (mild, moderate and severe). We also carried out a similar check regarding the secondary outcome measure of TOI-14 scores in a similar manner. Further details can be found in [Appendix 15, Tables 52](#) (exposure by baseline severity), [53](#) (summary statistics for overall return rate), [54](#) (crosstab showing numbers completing at least 80% STARs and 80% TOI-14 questionnaires at the follow-up time points).

Primary analyses

The primary hypothesis to be tested was H_0 : the number of sore throat days in 24 months is same for both randomised arms. A two-sided significance level of $p < 0.05$ was used throughout.

The primary outcome measure of the total number of sore throat days experienced over the 24 months of follow-up was analysed using NBR to compare the change between the NATTINA arms while adjusting for potential confounders, including the stratification variables: recruiting centre (as a random effect) and baseline severity (as a fixed effect). This analysis was undertaken on an ITT basis; however,

patients could switch over from conservative management to tonsillectomy. In the NATTINA design, patients were asked to commit to 'deferred surgery' (model 1).

Count data, by convention, often have an exposure variable indicating the duration of the observation period. Exposure variables for the NATTINA models were applied to the NBR model developed to account for differences in the proportion of the weekly returns completed by each participant (range from 0 to 1, with 1 representing a situation whereby all follow-up data were present).

The response variable is the total number of sore throat days over 24 months from randomisation to last follow-up. All participants were included in the analysis on an ITT basis with participants kept in their randomised arm, with the exception of the two participants randomised in error, as described in [Chapter 2](#).

Sensitivity analysis

The true effect of tonsillectomy is likely to lie between the estimate from the ITT analysis, which is the most parsimonious account owing to anticipated crossover into surgery, and the as-treated analyses, which tend to maximise the effect size of any surgical intervention. Outcome data analysis was at the end of the study and for DMC review, and followed the full SAP developed prior to the start of the analysis. Secondary analysis included estimation of the effects of tonsillectomy adjusted for potentially important clinical and demographical variables.

Sensitivity analyses on the ITT population were carried out to examine continuous baseline severity measure TOI-14 (model 2) and any other important baseline factors found to be potentially significant (model 3).

Further multivariable analyses considered other important baseline factors in the NBR model, including sex, age, ethnicity, education level, employment status, site and baseline levels. Non-linear continuous covariates were transformed when appropriate using simple first-degree polynomial transformations.

Additional analyses of the primary outcome

Addressing crossover (per-treatment analysis)

Participants in the conservative management arm of NATTINA were asked to commit to 'deferred surgery'. We anticipated that a number of participants would elect to cross over to surgery (see [Appendix 6](#)). The implication of such crossover, which typifies surgical trials, is that the ITT analysis would produce a very conservative estimate of the effect of tonsillectomy. We therefore also undertook sensitivity analyses for as treated, comparing those who received tonsillectomy with those who did not, as well as another four category per-treatment analysis that still considered randomised arm, but with each arm further divided into participants who received tonsillectomies and those who did not. A further unplanned sensitivity analysis included only participants who had returned $\geq 80\%$ of the STAR responses. A per-protocol sensitivity analysis was also carried out to confirm the results seen in the ITT primary analysis. This restricted the analysis to those participants who complied with the protocol and received tonsillectomy within 8 weeks after randomisation, after being initially randomised to the immediate tonsillectomy arm, compared with those randomised to the conservative management arm who did not cross over to receive a tonsillectomy. Non-compliance (including crossover) was addressed further using an 'as-treated' approach or complier-average causal effect (CACE) approach, given that the ITT analysis under non-compliance is biased when the intervention effect is large.⁹⁰ Alternative analysis can provide less biased estimates.⁹¹ The CACE approach results are reported in [Appendix 12](#) (instrumental variables).

Data with missing observations owing to loss to follow-up were examined to determine both its extent and whether it was missing at random or was informative. Owing to incomplete follow-up on

the primary outcome for some patients as anticipated, an appropriate exposure variable in the NBR model addressed this.

Ultimately, however, some of the crossovers from the conservative management arm did not actually undergo tonsillectomy. It was, therefore, decided to adopt 'per-treatment' to mean having surgery, and not those who crossed over in name/intention only. Therefore, the four groups in the per-treatment analysis were (1) randomised to surgery and surgery received, (2) randomised to surgery but no surgery received, (3) randomised to conservative management and no surgery received and (4) randomised to conservative management but received surgery.

Instrumental variables

The ITT principle kept the participants in the arm that they were randomised to regardless of the treatment that they actually received. Consequently, ITT analysis does not capture the full efficacy of surgery trials in which there is non-compliance with randomised arms originally randomised to. Instrumental variables rely on building a model that predicts the treatment actually received, accounting for unexpected behaviour between variables related to the treatment received. We have performed two types of instrumental variables analyses: CACE analysis and allowing for treatment compared with allocation interaction.

Complier-average causal effect analysis

An approach that better estimates the effect of the treatment than either per-protocol analysis, in which participants who comply with the allocated treatment are analysed, or per-treatment analysis, in which treatment actually received is analysed, is CACE.

Complier-average causal effect analysis allows unbiased assessment of treatment effect after grouping the intervention arm into compliers and non-compliers. CACE can be reduced for estimation to:

$$\text{CACE} = \text{ITT} / \text{proportion of compliers in randomised arm}, \quad (1)$$

where CACE (see [Appendix 12](#)) is defined as the mean of observed outcomes for participants who comply with allocated treatment.

Interactions with randomised arm allocation

In RCTs, the instruments include the randomised arm and the interactions between arm and baseline covariates. As an exploratory analysis, we modelled and reported parameter estimates of interaction term of randomised arm and baseline severity only.

Addressing 8-week surgery subgroup (per-protocol analysis)

For various reasons, surgery was delayed in a proportion of those in the tonsillectomy arm. Per-protocol analyses were restricted to those randomised to receive surgery and those receiving surgery within the specified maximum 8 weeks following randomisation, comparing them with those randomised to the conservative management arm who did not cross over to receive surgery.

Secondary outcome measures

The analysis of other secondary outcomes followed a similar strategy, with summary statistics tabulated and comparative box plots provided.

The analysis of secondary outcomes followed a broadly similar strategy: repeated measures were analysed using a random-effects model with an appropriate error structure. For example, QoL scores (SF-12) were calculated at baseline and 6, 12, 18 and 24 months post randomisation.

Scores were statistically analysed using longitudinal repeated-measure maximum-likelihood models developed for longitudinal data. The dependent variable was the overall QoL score [SF-12 mental

component score (MCS) and SF-12 physical component score (PCS)] and the analysis repeated with the dependent variable TOI-14 score for an individual participant at a particular time point. Both variation between participants and variation between responses nested within participants were modelled as random effects with a normal distribution. The results were adjusted for baseline scores and site, as previously. Differences between groups and changes over time were modelled as fixed effects. The analysis was adjusted for the treatment groups and stratification factor. The model assessed differences at follow-up time points beyond the initial 6-month time point, using the 6-month visit data as the reference value to compare the 12-, 18- and 24-month follow-up scores to.

Tonsillectomy Outcome Inventory-14

The TOI-14⁶⁶ is a validated disease-specific instrument for measuring health-related QoL. Our experience of using the TOI-14 in three centres pre and post tonsillectomy allowed us to (1) precisely estimate the effect size of tonsillectomy; (2) estimate the spectrum of baseline severity of those referred from primary care for consideration for surgery, hence generate stratification TOI-14 categories; and (3) evaluate the impact of alternative sore throat patient pathways by observation and statistical modelling of outcomes. The TOI-14 questionnaire data were collected at each follow-up visit for trial participants and also collected for declining patients at baseline.

Longitudinal TOI-14 scores were summarised using appropriate descriptive statistics along with 95% confidence intervals (CIs), but were not statistically tested across randomised arms. We tested the difference between TOI-14 means at each time point with CIs. The TOI-14 subscales of throat discomfort, general health, resource impact, social and psychological were dealt with in a similar way. After Q14, participants were asked to describe and score any other symptoms. These additional items were not incorporated into the TOI-14 scores and are described separately (see [Appendix 13](#)).

Sort Throat Alert Return questionnaire

The STAR questionnaire, later modified as STARLET (see above), yielded severity data for each registered sore throat episode, together with economic variables (see [Chapter 4](#)).

Short Form questionnaire-12 items

The analysis group for SF-12 was the ITT population. QoL scores based on the SF-12 were calculated in accordance with the scoring manual at baseline and 6, 12, 18 and 24 months post randomisation. Scores were described with summary statistics (mean, 95% CI about the mean, SD, median, and interquartile and overall ranges) and were graphically presented over time in the form of box plots.

Quality-of-life scores based on the SF-12 were processed using the Optum PRO CoRE software version 1.5 (Quality Metric Incorporated, Johnston, RI, USA). This software calculated composite scores for comparison with the normative data for the UK population. The SF-12 PCS and SF-12 MCS based on algorithms developed by the software owners at baseline and 6, 12, 18 and 24 months post randomisation were produced. In the normative data produced by the software, the mean score was transformed to be 50 (SD 10), adjusted for sex and age, with scores greater than 50 indicating better physical or mental health than the mean, and scores less than 50 indicating worse physical or mental health than the mean. The scores were analysed using models developed for longitudinal data and presented as summary statistics in a table and graphically by the use of box plots for each randomised arm. The dependent variable was the QoL score for an individual patient at a particular occasion. Both variation between patients and variation between occasions nested within patients were modelled as random effects with a normal distribution. Differences between groups and changes over time were modelled as fixed effects. The analysis was adjusted for the differences between strata.

Longitudinal repeated-measure maximum-likelihood models were used to analyse the dependent variable, that is the overall QoL score (both SF-12 MCS and SF-12 PCS), for an individual participant

at a particular occasion (transformed if non-normal). Both variation between participants and variation between responses nested within participants were modelled as random effects with a normal distribution. Differences between groups and changes over time were modelled as fixed effects. The analysis was adjusted for the treatment groups and stratification factors. Missing SF-12 data were assessed to decide whether or not to use complete-case analysis data and were presented.

Safety data

Safety data were not subject to statistical testing but are presented separately for AEs and SAEs categorised by severity, and further categorised by type of AE (see [Appendix 10](#)).

Sources of bias

Blinding

Because of the trial options, one surgical and one non-surgical, the trial was unblinded.

Response bias

For our primary outcome, sore throat days, we cannot eliminate surgical expectancy bias, but there is some evidence that accruing subjective data very regularly serves to minimise detection bias. We maximised capture of sore throat days by using patient-preferred methods of data return, principally electronic. The shorter the back-reference period, the less likely the risk of recall bias (telescoping or omission).⁹² NATTINA's substantial basket of secondary outcomes further mitigated any response bias and also maximised the opportunity to fulfil the commissioned goal to inform future research.

Data monitoring, quality control and assurance

Trial data were entered into and stored by authorised site staff in MACRO (MACRO Electronic Data Capture, InferMed, London, UK), a secure, web-based electronic case report form (eCRF) system operated from NCTU.⁹³ Data transferred from site to MACRO by remote access were encrypted and had restricted access. The database was maintained by NCTU staff. Data checks included reconciliation of data between MACRO and the NCTU randomisation system; identification of missing data; sense checks, for example on dates; checks for activation of in-built validations, such as warnings; and checks for completion rates of key variables, such as primary and secondary outcomes. Queries resulting from data checks were raised with sites. Any changes to the trial data were captured through an audit trail for data integrity.

Participants were identified by a unique participant ID that was allocated by the randomisation system, which was used on eCRFs and questionnaire front covers. Personal details (full name, address, e-mail address and telephone numbers) were stored on secure and restricted databases (separate to MACRO) at NCTU on the Newcastle University server for the purpose of sending out weekly sore throat alert prompts, STARs and follow-up questionnaires.

All interviews undertaken were audio-recorded and transcribed verbatim. Anonymous audio files and transcripts were stored electronically with other trial data. Data were handled, computerised and stored in accordance with the Data Protection Act 1998⁹⁴ and General Data Protection Regulation⁹⁵ (GDPR). Caldicott approval was sought during set up at each participating site to enable the collection and transfer of participant information for NATTINA. All trial data will be archived for at least the mandatory 5 years from the close of the trial.

Trial management and oversight

The trial was managed by NCTU on behalf of the sponsor. These responsibilities included obtaining initial and any required subsequent regulatory approvals; providing site training and set up; day-to-day

site support (including training through investigator meetings, site initiation visits and routine monitoring advice and guidance); trial monitoring; financial management; data management; liaising with, and reporting to, the trial funder and the REC; maintenance of the central trial master file; and site close down. The principal investigators (PIs) were responsible for the day-to-day trial conduct at the site. Quality control was maintained through adherence to the NCTU's standard operating procedures, the trial protocol, the principles of good clinical practice, research governance and clinical trial regulations.

Trial Management Group

The TMG ran the day-to-day NATTINA management and constituted the chief investigator, co-investigators, statisticians, health economists, qualitative researchers and the NCTU trial management team [senior trial manager, trial manager, data(base) manager, clinical trial administrator and trial secretary]. The TMG met monthly during recruitment and 6 weekly thereafter.

Oversight committees

A TSC was convened to oversee NATTINA, under an independent chairperson. Members were the chief investigator, independent clinicians, an independent statistician and PPI representatives. The committee met six times: prior to the start of the internal pilot and annually thereafter.

An independent DMC was established to undertake an independent review and to monitor efficacy and safety end points. The committee consisted of an independent chairperson, an independent clinician and an independent statistician. The DMC met seven times: first to discuss and advise on the inclusion of an interim analysis and possible adoption of a formal stopping rule for efficacy or safety, at the end of the internal pilot, three times in 2017 at the recovery plan and initial 12-month extension stage, and annually thereafter. Independent members of both the TSC and the DMC were approved by the funder prior to confirmation of their appointment.

Chapter 3 Results

The analysis presented here is reported according to the CONSORT flow diagram (see [Figure 2](#)), and is based on SAP version 2.0 (13 July 2020), with detailed clarification of the primary analysis SAP version 3.0 (August 2020) that provided guidelines for the analysis of the NATTINA trial based on protocol version 7.0 (11 September 2019) and the published protocol.⁷⁹ Any analyses that were not prespecified in the SAP are denoted as 'unplanned'.

Recruitment

NATTINA took place in 27 sites: five of the sites failed to recruit any participants to the trial, and one site that recruited one participant to the trial did not enter baseline data for this participant and was unresponsive when contact was attempted by the trial team (unresponsive site). The first site opened on 14 April 2014, the first participant was randomised on 11 May 2015, the last participant was randomised (before snapshot) on 30 April 2018 (at 15:07) and the date of data lock was 20 July 2020. A total of 455 participants were randomised before snapshot, of whom two were mistakenly entered into the NATTINA trial instead of different trials to which they had consented (randomised on 15 March 2016 and 26 May 2017); no further data were collected for these patients. These patients would have been included in the ITT analysis as ineligible participants, as per the definition of ITT analyses, but were not included given that there was no data return.

Randomisation and stratification factors

Participants were randomised to receive immediate tonsillectomy or conservative management. The allocation ratio was 1 : 1, stratified by recruiting site and baseline severity. The distribution of randomised participants by baseline severity and randomised arm is shown in [Table 2](#). The distribution by site is shown in [Appendix 4, Table 33](#).

Note that some participants were mis-stratified based on their baseline TOI-14 scores. In total, there were 36 such mis-stratifications, detailed in [Appendix 4, Table 34](#).

To mitigate for mis-stratification, we carried out a sensitivity analysis of the primary analysis using continuous baseline TOI-14 scores instead of the stratified categories (see [Sensitivity analyses](#)). The principal reason for this sensitivity analysis was, however, to utilise the full information from the continuous measure. The mitigation of mis-stratification was a secondary consideration.

Consolidated Standards of Reporting Trials flow diagram

In the CONSORT flow diagram ([Figure 2](#)), the numbers of returned TOI-14 and SF-12 questionnaires at the 6-monthly time points were different. If only one was returned/collected, the higher of the two numbers being returned is cited in the CONSORT flow diagram follow-up questionnaire boxes. Therefore, in the later secondary analyses sections, some numbers may differ from those presented in [Figure 2](#).

TABLE 2 Distribution of baseline severity by randomised arm for the ITT population

Baseline severity (by TOI-14 score)	Randomised arm, n (%)		Total, n (%)
	Immediate tonsillectomy	Conservative management	
Mild (0–35)	52 (22)	45 (20)	97 (21)
Moderate (36–48)	94 (40)	96 (43)	190 (42)
Severe (49–70)	88 (38)	80 (36)	168 (37)
Total	234 (100)	221 (100)	455 (100)

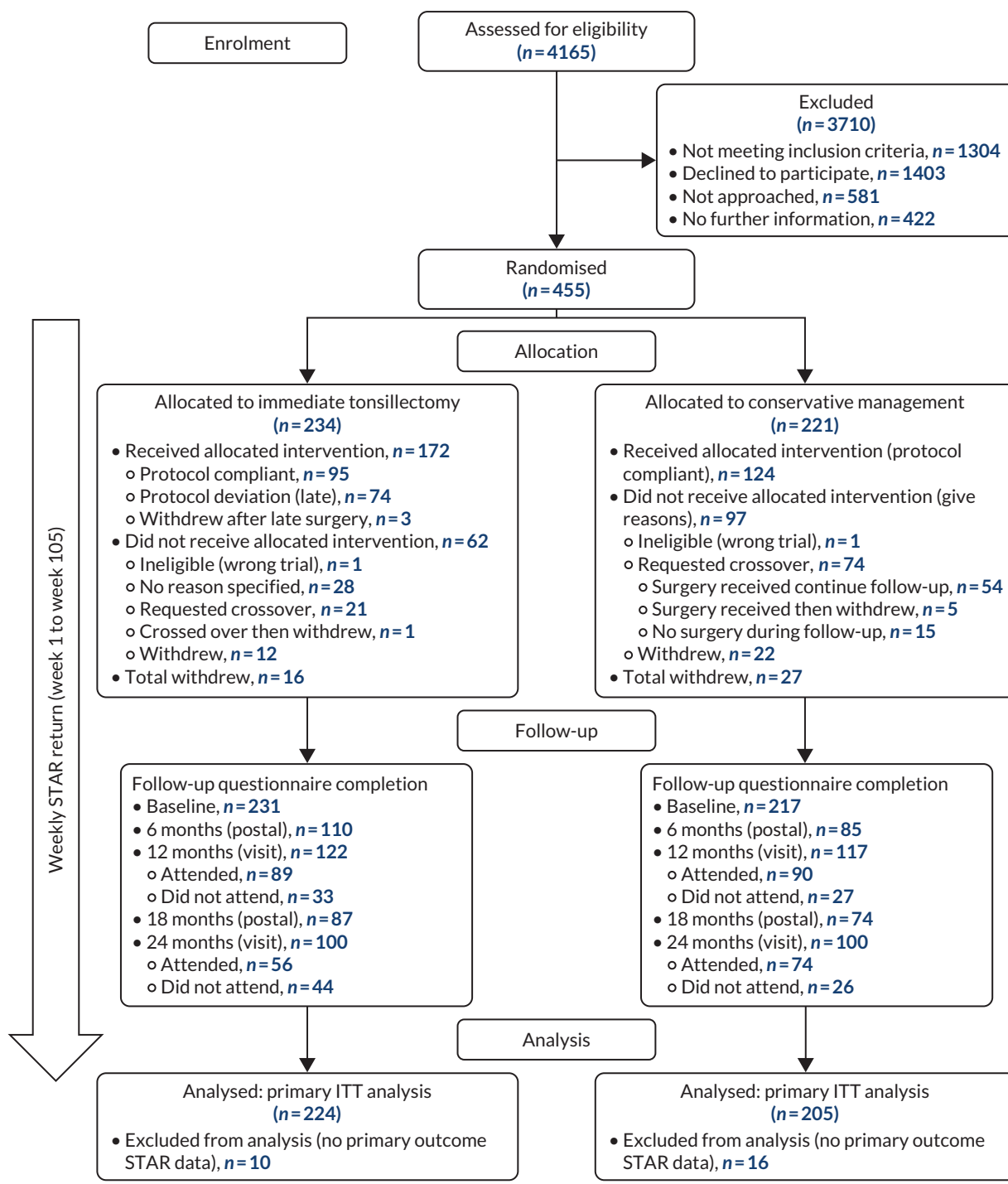


FIGURE 2 The CONSORT flow diagram.

Characteristics of those declining to participate

A total of 1403 patients declined to participate and 1002 (71%) of these patients gave the following reason: 'wanted tonsillectomy off study'. The full list of reasons given for declining to join the trial is summarised in [Appendix 3, Table 31](#). Patients who declined to participate were invited to provide anonymised baseline comparison data for the NATTINA database (age, sex, an estimate of number of sore throat days over the prior 6 months and a TOI-14 questionnaire) and 609 patients provided these data. Declining patients were also invited to participate in a qualitative interview with a researcher.

It can be seen from [Table 3](#) that the proportions of males (21%) and females (78%) who accepted or declined to be randomised into the trial were the same. There was a tendency for those who accepted to be randomised to be, on average, 2.5 years older (95% CI 1.7 to 3.3 years) and to have lower TOI-14 scores (mean difference -5.19, 95% CI -6.5 to -3.9) than those who declined to be randomised who provided data.

Differences between the distributions of TOI-14 scores between the two arms can be seen in the box plots in [Appendix 3, Figure 19](#).

TABLE 3 Demographics and imputed TOI-14 scores of the accepting and declining patients

Summary statistics	Decline (N = 609)	Accept (N = 453)
TOI-14		
n (%)	607 ^a (99.7)	448 (99)
Mean (SD)	48.7 (9.8)	43.5 (11.4)
95% CI	47.9 to 49.5	42.4 to 44.5
Median (IQR)	49 (43–56)	44 (36–51)
Min, max	17, 70	7, 69
Missing (%)	2 (<1)	5 (1)
Sex, n (%)		
Male	129 (21)	97 (21)
Female	476 (78)	355 (78)
Missing	4 (<1)	1 (<1)
Age (years)		
n (%)	605 (99)	453 (100)
Median (LQ, UQ)	21 (18, 25)	23 (19, 30)
Mean (SD)	22.7 (6.0)	25.2 (7.4)
95% CI	(22.2 to 23.2)	(24.5 to 25.9)
Range (min, max)	(0 ^b , 50)	(16, 59)
Missing (% of n)	4 (<1)	0 (0)
Sore throat days (previous 6 months)		
n (%)	581 (95)	Not collected for patients consenting to randomisation into trial
Mean (SD)	39.4 (32.3)	
95% CI	39.4 (36.8 to 42.0)	
Median (IQR)	30 (20–50)	
Min, max	0, 200 ^c	
Missing (%)	28 (5)	

IQR, interquartile range; LQ, lower quartile; max, maximum; min, minimum; UQ, upper quartile.

^a TOI-14 data missing for some participants who declined (see [Appendix 3, Table 32](#)).

^b One age given as 0 years.

^c Two gave estimates of sore throat days in previous 6 months greater than the number of days possible in that time (196 and 200 days).

Baseline participant characteristics

The population was ITT. Demographic and baseline characteristics were compared across randomised arms descriptively. Descriptive statistics were tabulated by randomised arm and overall (see [Table 4](#)):

- No baseline data were collected for the two participants who were randomised but were ineligible for the trial because the participants did not consent.
- No baseline data were collected for participants at the unresponsive site.
- The recoding of 'other' categories for employment ($n = 11$) and education ($n = 9$) was completed manually from the text provided by the participant, with the approval of the TMG. For a list of the free text provided for 'other' description, and how it was recoded, see [Appendix 5](#).

[Table 4](#) shows the baseline demographics for 453 eligible participants. It can be seen that the population recruited into the NATTINA trial was predominantly female (78%); white (90%); relatively young, with a median age of 23 years [interquartile range (IQR) 19–30 years]; and in paid employment (62%).

TABLE 4 Baseline demographic characteristics of 453 eligible participants by randomised arm

Variable	Randomised arm		Total
	Immediate tonsillectomy	Conservative management	
Sex, n (%)			
Female	175 (75)	180 (82)	355 (78)
Male	57 (24)	40 (18)	97 (21)
Missing ^a	1 (<1)	0 (0)	1 (<1)
Age (years)			
Total (n)	233	220	453
Median (IQR) (from randomisation log)	23 (19–30)	23 (19–30)	23 (19–30)
Min, max	16, 59	16, 56	16, 59
Ethnic origin, n (%)			
Total	232 (>99)	220 (100)	452 (>99)
White	211 (91)	196 (89)	407 (90)
Asian (Indian/Pakistani/Bangladeshi ancestry)	9 (4)	9 (4)	18 (4)
Other Asian	0 (0)	4 (2)	4 (<1)
Black or Afro-Caribbean (African or Caribbean ancestry)	6 (3)	6 (3)	12 (3)
Other ethnic origin	6 (3)	5 (2)	11 (2)
Missing ^a	1 (<1)	0 (0)	1 (<1)
Education level, n (%)			
Post graduate	20 (9)	23 (10)	43 (9)
Degree/Professional/Vocational (NVQ level 4)	48 (21)	70 (32)	118 (26)
Higher/A Level/National grade/vocational (HND)	100 (43)	74 (34)	174 (38)
O Level/O Grade/GCSE/Standard Grade/vocational (HNC)	57 (24)	47 (21)	104 (23)
No educational qualification	5 (2)	1 (<1)	6 (1)
Missing ^a	3 (1)	5 (2)	8 (2)

TABLE 4 Baseline demographic characteristics of 453 eligible participants by randomised arm (*continued*)

Variable	Randomised arm		Total
	Immediate tonsillectomy	Conservative management	
Employment status, n (%)	230 (99)	216 (98)	446 (98)
Self-employed	12 (5)	11 (5)	23 (5)
Paid employment (full or part time)	149 (64)	132 (60)	281 (62)
Unemployment (actively seeking work)	6 (3)	8 (4)	14 (3)
Retired	0 (0)	0 (0)	0 (0)
Maternity leave	1 (<1)	1 (<1)	2 (<1)
Looking after family or home	12 (5)	9 (4)	21 (5)
Full-time student/at school	48 (21)	52 (24)	100 (22)
Long-term sick or disabled	1 (<1)	3 (1)	4 (<1)
Government training scheme	1 (<1)	0 (0)	1 (<1)
Missing ^a	3 (1)	4 (2)	7 (2)

GCSE, General Certificate of Secondary Education; HNC, Higher National Certificate; HND, Higher National Diploma; NVQ, National Vocational Qualification.

a Missing data are for participants from unresponsive site. We have the age at randomisation for this participant, however, as date of birth and randomisation date were collected.

Data quality and completeness

[Table 5](#) shows the expected number of responses, actual number of responses and response rate for the trial case report forms. [Table 5](#) also includes the number of participants who returned at least one weekly STAR and counts of (1) the primary outcome measure 'sore throat days', collected via the weekly STAR response texts, and (2) the resultant STAR questionnaires when sore throat days were >0. Compliance for the two telephone calls scheduled for 1 and 2 weeks post tonsillectomy is also presented. AEs reported during these telephone calls are reported in [Safety](#).

Summary of primary outcome data: Sore Throat Alert Return

A total of 32,978 weekly STAR responses was received. Of these responses, 809 were for the weeks after the end of the follow-up period. Participants who withdrew from the trial were included in the analysis up to the point of withdrawal. This resulted in a further 117 STAR responses being omitted from the primary analysis data set. Therefore, 32,052 STARs were retained for the primary analysis.

Assessment of compliance with face-to-face and postal completion of questionnaires

[Table 6](#) shows participant engagement with the trial follow-up questionnaire at the follow-up visits and postal questionnaire returns points:

- In [Table 6](#), 'complied (± 6 weeks)' indicates participants who returned forms or attended visits within ± 6 weeks of each scheduled data collection point (6, 12, 18 and 24 months after the randomisation date).
- 'Early' represents forms returned/visits completed before the 6-week lower compliance window.
- 'Late' represents forms returned/visits completed after the 6-week upper compliance window.
- Further details, including a table of summary statistics of the time from randomisation to visits/postal returns of secondary outcome questionnaires and histograms showing compliance windows for the four follow-up data collection points, are shown in [Appendix 6](#), [Table 35](#), and [Figures 20–23](#).

TABLE 5 Participant activities completed by time point

eCRF	Number expected (n)	Actual number (n)	Return rate (%)
Baseline form (clinic visit 1)	453	452	99.8
Baseline (declined participants)	N/A	609	N/A
Participants with at least one weekly STAR	453	429	95
Weekly STAR responses (primary outcome) for 453 eligible participants making 105 weekly STARs	47,565	32,978	69
Weekly STAR responses (primary outcome) adjusted to include only weeks up to withdrawing	44,676	32,978	74
Weekly STAR responses (primary outcome) adjusted to include only weeks up to withdrawing and excluding surplus returns beyond 105 weeks	44,676	32,052	72
Reported STAR questionnaires (triggered by a weekly text response of >0)	5431	1395	26
Old paper STARs		721	
New electronic 'STARLETs' (date implemented 5 May 2017)		674	
Tonsillectomy form for tonsillectomies within follow-up	231	231	100
Crossover form	N/A	120	N/A
Actual crossovers requested		94	
Withdrawal form	N/A	71	N/A
Actual withdrawals		43	
Post-tonsillectomy telephone call 1 (n = 231 tonsillectomies carried out)	231	195	84
Post-tonsillectomy telephone call 2 (n = 231 tonsillectomies carried out)	231	184	80
6-month follow-up form (postal) (adjusted for withdrew: minus 18)	435	195	45
12-month follow-up (clinic visit 2) (adjusted for withdrew: minus 32)	421	239	57
18-month follow-up form (postal) (adjusted for withdrew: minus 38)	415	161	39
24-month follow-up (clinic visit 3) (adjusted for withdrew: minus 42 ^a)	411	200	49

N/A, not applicable.

a One withdrew after the end of 24-month follow-up.

Note that the relatively low levels of engagement with follow-up visits and postal questionnaire completion did not affect the primary analysis of the trial. The primary outcome measure was the number of sore throat days reported over the 24-month follow-up period, which was collected via the weekly STAR responses. These were independent of scheduled follow-up 6-monthly visits and postal returns, which primarily concerned the collection of secondary outcome data.

Trial populations subjected to analysis

The populations used for the statistical analysis are defined in [Chapter 2](#). The primary analyses results are presented in this chapter, with more details of sensitivity and unplanned analyses reported in [Appendix 10](#).

Treatment received

There were 43 withdrawals from the trial (see [Figure 2](#)). [Appendix 8, Table 36](#), reports the summary statistics for time from randomisation to withdrawal in weeks, by arm and overall. The box plots in

TABLE 6 Compliance with follow-up visits and postal questionnaire completion

Follow-up visit	Randomised arm, n (%)															
	Immediate tonsillectomy						Conservative management						Total, n (%)			
	Number expected ^a	Complied (±6 weeks)	Early <6 weeks	Late >6 weeks	Number expected ^a	Complied (±6 weeks)	Early <6 weeks	Late >6 weeks	Number expected ^a	Complied (±6 weeks)	Early <6 weeks	Late >6 weeks	Number expected ^a	Complied (±6 weeks)	Early <6 weeks	Late >6 weeks
6 months ^b	226	93 (41)	0 (0)	16 (7)	209	72 (34)	0 (0)	13 (6)	435	165 (38)	0 (0)	29 (7)	421	153 (36)	4 (<1)	81 (19)
12 months	221	80 (36)	3 (1)	38 (17)	200	73 (37)	1 (<1)	43 (22)	415	121 (29)	1 (<1)	40 (10)	411	143 (35)	3 (<1)	53 (13)
18 months ^b	220	59 (27)	1 (<1)	27 (12)	195	62 (32)	0 (0)	13 (7)	411	143 (35)	3 (<1)	53 (13)	411	143 (35)	3 (<1)	53 (13)
24 months	218	75 (34)	1 (<1)	23 (11)	193	68 (35)	2 (1)	30 (16)	411	143 (35)	3 (<1)	53 (13)	411	143 (35)	3 (<1)	53 (13)

^a Adjusted for withdrawals up to time points.

^b The 6- and 18-month collection is postal, not face to face.

RESULTS

[Figure 3](#) also illustrates the withdrawal information. Most withdrawals occurred in the first 12 months of the trial, with those randomised to receive conservative management tending to withdraw earlier than those randomised to receive tonsillectomy.

The preparatory PPI group work had advised that deferred tonsillectomy as the conservative management arm would facilitate recruitment. Therefore, the protocol permitted participants to opt out of their randomised treatment allocation and effectively cross over to the other randomised arm.

[Table 7](#) presents how the participants were randomised, how many received the allocated treatment, how many requested and received the treatment from the other arm, and how many withdrew (see [Figure 2](#)).

For participants randomised to the tonsillectomy arm (i.e. to receive immediate tonsillectomy), 172 out of 233 (74%) went on to receive the procedure. Of these participants, 95 (41%) received their tonsillectomy within 8 weeks of randomisation, in line with the protocol, and the remaining 77 (33%) received their tonsillectomy late. A total of 234 out of 453 (52%) eligible participants fully complied with the trial protocol (not crossing over or withdrawing and, in the case of the tonsillectomy arm, receive tonsillectomy within 8 weeks of randomisation). However, 10 of these participants did not make any STARs and, therefore, did not have primary outcome data and were not included in the per-protocol analysis, leaving 224 (49%) participants in this sensitivity analysis. A total of 139 out of 220 (63%) participants randomised to the conservative management arm went on to receive that treatment; however, of these participants, 15 had requested to receive tonsillectomy but had not received it by the end of follow-up, so were assumed to comply with conservative management (see [Table 7](#)).

A total of 94 out of 453 (21%) participants requested to stop the treatment that they were randomised to (shaded in [Table 7](#)): 21 out of 233 (9%) participants randomised to the immediate tonsillectomy arm requested to switch to conservative management, whereas 73 out of 220 (33%) participants randomised to the conservative management arm requested to receive tonsillectomy. There was no protocol provision for these participants to receive the tonsillectomy within a specified time compliance window, but 59 out of 73 (81%) of those requesting to cross over had received their tonsillectomy by the end of their 24-month follow-up period. This includes one participant who had tonsillectomy off trial, so did not complete a crossover form. Reasons given for requesting a crossover between randomised arms are tabulated in [Appendix 2](#), [Tables 29](#) and [30](#).

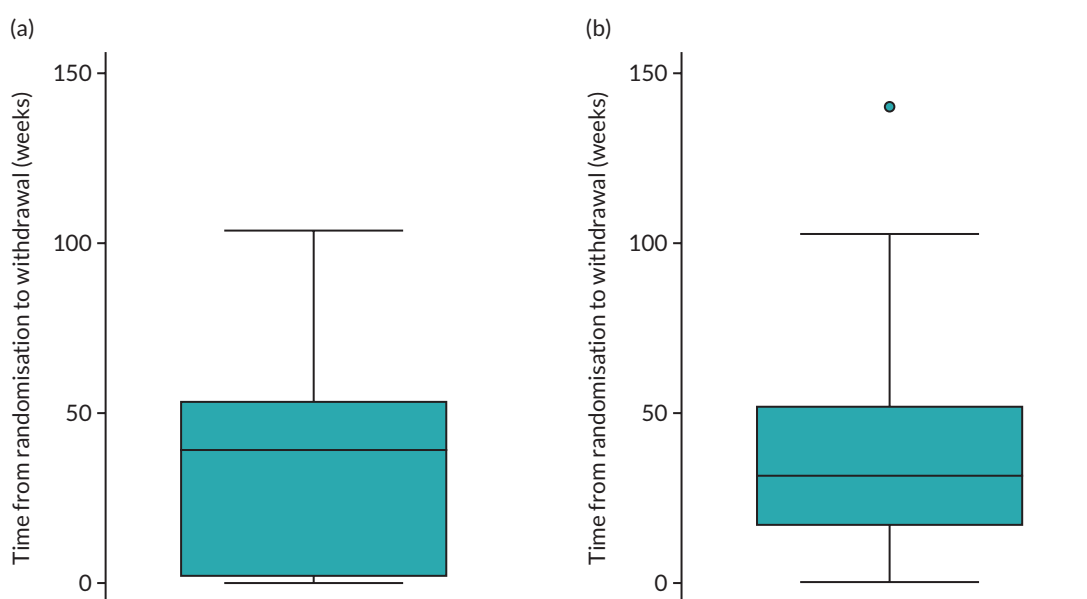


FIGURE 3 Box plots showing time from randomisation to withdrawal by arm ($n = 43$). (a) Conservative management ($n = 27$); and (b) immediate tonsillectomy ($n = 16$). The outlier in immediate tonsillectomy requested to withdraw before the end of follow-up.

TABLE 7 Intervention received: immediate tonsillectomy (n = 233) and conservative management (n = 220)

Randomised arm	Crossed over	Tonsillectomy received	Complied (tonsillectomy within 8 weeks)	Total (n)	Additional information
Tonsillectomy	No	Yes	Yes	95	Protocol complied. One withdrew but this was after the 24-month follow-up should have ceased ^a
Tonsillectomy	No	Yes	No (late)	77 ^a	Protocol deviation. Two withdrew before the end of follow-up ^a
Tonsillectomy	No	No	No	28	One was from the unresponsive site
Tonsillectomy	No	No	Withdrawn	12 ^a	Total of 16 ^a withdrew from the tonsillectomy arm but four are counted elsewhere
Tonsillectomy	Yes	No	N/A	21 ^a	Crossed over to conservative management. One crossed over to conservative management then withdrew ^a
Total: immediate tonsillectomy				233	
Conservative management	No	No	N/A	124	Protocol complied
Conservative management	No	No	Withdrawn	22 ^b	Total of 27 ^b withdrew from conservative management arm but five are counted elsewhere
Conservative management	Yes	No	N/A	15	Protocol complied. One had tonsillectomy beyond end of follow-up
Conservative management	Yes	Yes	N/A	59 ^b	Protocol deviation Five ^b had tonsillectomy then withdrew. One participant had tonsillectomy off trial so did not complete a crossover form
Total: conservative management				220	
Trial total				453	

N/A, not applicable.

a Refers to the total number of withdrawals in the immediate tonsillectomy arm. The total is 16, 12 of which are shown in the 'withdrawal row', the other four are shown elsewhere (one had tonsillectomy in the compliance window then withdrew, two had late tonsillectomy then withdrew and one crossed over before withdrawing).

b Similarly, for the conservative management arm, 27 withdrew: 22 are shown in the 'withdrew row' and the other five are shown elsewhere as they had tonsillectomy then withdrew.

Note

Shading indicates crossovers.

A total of 43 out of 453 (9%) participants officially withdrew their consent to further participate in the trial. All data collected up until withdrawal were retained for NATTINA research purposes, as no one requested that their prior data be withdrawn. More details of these 43 withdrawals are shown in [Appendix 8, Table 37](#). The summary statistics of time from randomisation to withdrawal are shown in [Appendix 8, Table 36](#).

Time from randomisation to tonsillectomy

[Figure 4](#) shows the distribution of time from randomisation to tonsillectomy for participants randomised to the immediate tonsillectomy arm (with a protocol compliance window of within 8 weeks of randomisation

RESULTS

shown for participants randomised to receive immediate tonsillectomy). The width of each bar is 8 weeks and, therefore, it can be seen that 95 participants complied with the protocol of receiving tonsillectomy within 8 weeks of randomisation. A further 55 participants received tonsillectomy in the 8-week period following protocol compliance. Summary statistics for the time from randomisation to tonsillectomy are shown in [Appendix 9, Table 38](#).

Type of tonsillectomy received

Of the 231 surgeries carried out, 122 (53%) were cold dissection, 91 (39%) were bipolar diathermy, 1 (<1%) was coblation, 1 (<1%) was laser, 8 (3%) were described as mixed (combination of cold and bipolar dissection) and 1 (<1%) was the tonsillectomy carried out off trial. For the remaining seven surgeries (3%), sites did not provide details of the tonsillectomy type received. [Appendix 9, Table 39](#), shows the breakdown of tonsillectomy type by the arm originally randomised to. Ten of these (4%) reported complications in the procedure (see [Safety](#)).

Safety

Safety data were not subject to comparative statistical analysis.

There were 47 re-admissions to hospital in the 231 participants undergoing tonsillectomy (20%). Three participants had a prolonged in-hospital stay following surgery. There were 54 episodes of post-operative haemorrhage reported in 44 participants. This equates to 44 out of 231 participants undergoing tonsillectomy (19%). Of these participants, 37 were reported as SAEs requiring re-admission: 9 were reported as mild events, 22 as moderate events and 6 as severe events. No deaths were reported. Seventeen were recorded as AEs, for which participants did not attend hospital. All episodes of bleeding were managed conservatively with no returns to theatre.

There were 10 complications from tonsillectomies reported. The details of the complications can be found in [Appendix 9, Table 40](#).

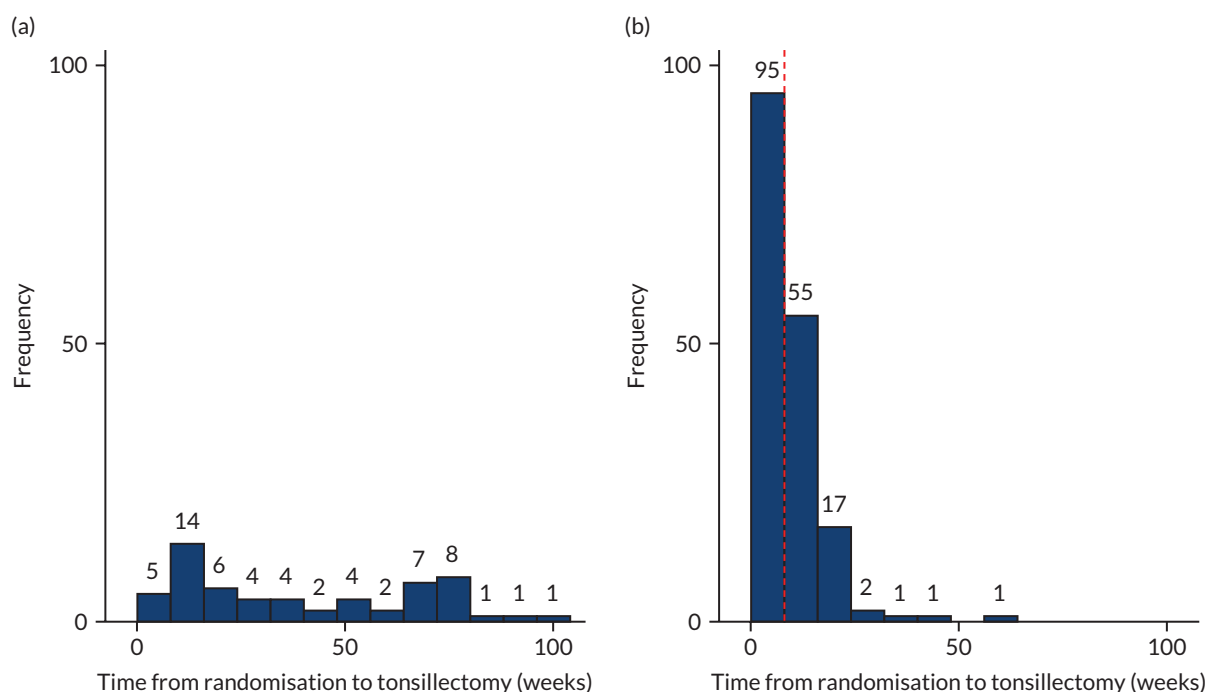


FIGURE 4 Histograms showing distribution of time from randomisation to tonsillectomy by arm. (a) Conservative management ($n = 59$); and (b) immediate tonsillectomy ($n = 172$). Red dashed line is 8 weeks after randomisation (protocol compliance for immediate tonsillectomy).

[Appendix 10, Table 41](#), gives the breakdown for the 191 AEs categorised as mild, moderate and severe, split by those randomised to receive the tonsillectomy and those who requested to cross over for tonsillectomy and received it within the trial follow-up period.

[Table 8](#) shows the 191 reported AEs recategorised based on the free text provided in the original MACRO entries:

- The most common AEs reported were throat/ear pain (35% of all AEs, approximately evenly split by severity).
- Summary statistics for time from tonsillectomy to first AE.
- Date available for all 90 participants. Median day AE was reported following the received intervention (lower quartile, upper quartile) is 1 day (0, 4).
- Maximum is 12 days for two participants.

Serious adverse events

Fifty-one SAEs, reported in 45 participants, were attributed to the intervention by the surgeon. These are reported in [Appendix 10, Table 43](#), where they are categorised as mild, moderate and severe and are split by randomised arm.

No fatal SAEs were reported in NATTINA, nor any life-threatening SAEs or SAEs that caused persistent or significant disability or incapacity.

[Table 9](#) shows the 51 reported SAEs recategorised based on the free text provided in the original MACRO entries.

TABLE 8 Characteristics of the 191 AEs, with severity category

AE category	Severity, n (%)			Total, n (%)	Participants (n)	Rate out of 231 tonsillectomies (%)
	Mild	Moderate	Severe			
Pain: throat or ear	22 (33)	21 (32)	23 (35)	66 (35)	52	23
Bleed	14 (82)	3 (18)	0 (0)	17 (9)	14	6
Infection/fever/temperature	14 (45)	16 (52)	1 (3)	31 (16)	29	13
Reduced diet/swallow difficulty	9 (45)	11 (55)	0 (0)	20 (10)	20	9
Nausea/vomiting	10 (56)	8 (44)	0 (0)	18 (9)	15	6
Tiredness/fatigue	4 (33)	8 (67)	0 (0)	12 (6)	11	5
Other	16 (59)	10 (37)	1 (4)	27 (11)	23	10
Total	89 (47)	77 (40)	25 (13)	191 (100)	90	39

TABLE 9 Characteristics of the 51 SAEs with severity category

SAE category	Severity, n (%)			Total, n (%)	Participants (n)	Rate out of 231 tonsillectomies (%)
	Mild	Moderate	Severe			
Bleed	9 (56)	22 (81)	6 (75)	37 (73)	34	15
Infection/fever/temperature	3 (19)	2 (7)	1 (13)	6 (12)	6	3
Pain	2 (13)	3 (11)	0 (0)	5 (10)	5	2
Prolonged hospital stay	2 (13)	0 (0)	1 (13)	3 (6)	2	<1
Total	16 (100)	27 (100)	8 (100)	51 (100)	45	19

The median time from tonsillectomy to the first SAE in these 45 participants was 5 days (IQR 2–7 days).

The SAE resulted in an extended stay in hospital in 50 out of the 51 (98%) SAEs judged to be causally related to surgery. All participants recovering from a SAE had fully recovered by the end of the trial follow-up. Further details are provided in [Appendix 10, Table 42](#).

Medication was the single most common remedial (32/51, 63%). Medication was combined with a blood/fluid transfusion in six cases (12%), surgery and medication were used in two cases (4%), surgery alone was used in one case (2%), blood/fluid transfusion was used in two cases (4%) and surgery/medication/transfusion was used in one case (2%). No action was required in four (8%) of the cases and information was not entered onto MACRO in three cases (6%).

Descriptive analysis of the primary outcome measure

Retention flow diagram

The flow chart in [Figure 5](#) shows the number of randomised participants and how they engaged with the weekly STAR response procedure.

Primary outcome measure

The primary outcome measure was the total number of sore throat days reported over the 24-month follow-up period. As [Table 5](#) shows, 72% of STARs were returned.

[Table 10](#) summarises the raw data of total sore throat days over the 24-month period. Participants in the tonsillectomy arm tended to report fewer sore throat days than those in the conservative management arm. The total reports of sore throats were difficult to compare formally owing to variability in the total number of returns per participant. To make a fairer comparison of the sore throat rates in the two arms, we have also presented the summary of the average weekly STAR for each participant, calculated as the total number of sore throat days reported in the 24-month (105-week) follow-up period divided by the number of returns made, by randomised arm and overall, for the ITT population. The data again show a positive skew: immediate tonsillectomy participants had median of 0.33 sore throats per week compared with 0.57 sore throats per week in the conservative management arm.

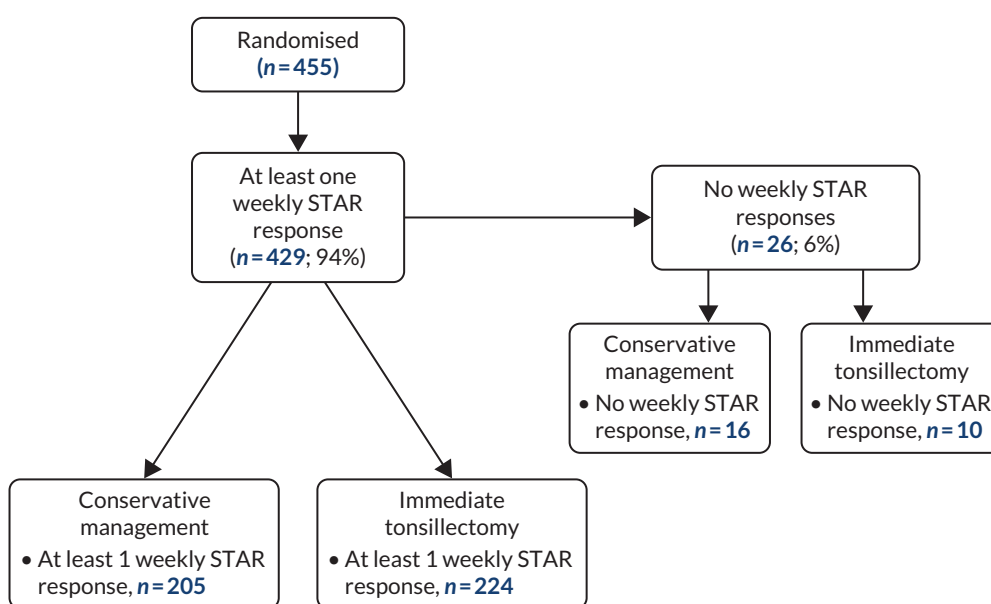


FIGURE 5 Flow chart showing primary outcome data.

TABLE 10 Primary outcome measure: total sore throats over 24 months and the average number of sore throat days per week, by randomised arm and overall, for the ITT population

Statistic	Randomised arm		Total (N = 453)
	Immediate tonsillectomy (N = 233)	Conservative management (N = 220)	
Total number of sore throat days in 24-month follow-up (raw data)			
n (%)	224 (96)	205 (93)	429 (95)
Median (IQR)	23 (11–46)	30 (14–65)	27 (12–52)
Mean (SD)	34.4 (39.0)	56.1 (85.6)	44.8 (66.4)
95% CI	29.3 to 39.6	44.3 to 67.9	38.5 to 51.1
Min, max	0, 376	0, 691	0, 691
Average number of sore throat days per week			
Median (IQR)	0.33 (0.19–0.66)	0.57 (0.27–1.35)	0.43 (0.21–1.00)
Mean (SD)	0.66 (0.97)	1.16 (1.50)	0.90 (1.28)
95% CI	0.54 to 0.79	0.96 to 1.37	0.78 to 1.02
Min, max	0, 6.3	0, 7	0, 7

Max, maximum; min, minimum.

Efficacy analysis: primary outcome

As described in [Data quality and completeness](#), a total of 32,052 STARs were within the follow-up period and, taking account of withdrawal points, were retained for the primary analysis. Of the 455 participants randomised into NATTINA, 429 (94%) provided some primary outcome data (at least one STAR) and, therefore, are included in the analyses outlined in this section. [Appendix 11](#) summarises how the participants with primary outcome data were assessed for inclusion.

Primary analysis

This primary analysis was undertaken on an ITT basis. The response variable is the total number of sore throat days over 105 weeks from randomisation to last follow-up. As the primary outcome is a count variable, all analyses use the NBR model. This model can allow for the variable number of STARs by including the proportion of STARs (out of a total of 105) as an exposure variable. All NBR analyses include the exposure variable.

TABLE 11 Unadjusted NBR showing IRR and 95% CIs, with exposure variable of proportion of weekly STAR responses for difference in total number of sore throat days between randomised arms: ITT population

Total sore throat days over 24 months	IRR	95% CI	p-Value
Arm (reference value: conservative management): immediate tonsillectomy	0.575	0.467 to 0.708	<0.001
Constant	116.200	100.038 to 134.973	<0.001
Ln (exposure)	1		
Ln (alpha)	0.107	-0.016 to 0.231	
Alpha (dispersion parameter)	1.113	0.984 to 1.260	

n = 429, log-likelihood = -2083.715

Negative binomial regression of the primary outcome

Table 11 shows the results of the unadjusted analysis (i.e. no covariates or stratification factors included in the model) of the primary outcome on the ITT population. The difference between arms indicates that participants randomised to receive tonsillectomy are estimated to have a total number of sore throat days 0.575 times smaller than participants randomised to receive conservative management (95% CI 0.47 to 0.71; $p < 0.001$).

Count data can be modelled using the Poisson model; however, NBR regression was judged to be more appropriate in this case because the coefficient of dispersion was high. For the raw data (both arms combined), the mean sore throat weekly rate reported during follow-up was 0.90, with variance 1.6384. Therefore, the coefficient of dispersion, variance/mean, is 1.82, implying that there is overdispersion of the weekly sore throat rate.

In the NBR model presented in **Table 11**, the dispersion parameter of 1.113 is significantly greater than zero; therefore, these results also confirm that data are over-dispersed and are better modelled using a negative binomial model than a Poisson model.

Several adjusted NBR analyses were conducted. The first of these was carried out to compare the NATTINA arms while adjusting for the stratification variables used at the point of randomisation in the trial: recruiting centre (as a random effect) and baseline severity (as a fixed effect). This is the primary analysis of the NATTINA trial and is referred to as model 1, and is presented in **Table 12**. The estimated difference between randomised arms, when adjusting for site and baseline severity, indicates that immediate tonsillectomy participants are estimated to have 0.528 times fewer total sore throat days in the 24-month follow-up than conservative management participants (95% CI 0.43 to 0.65; $p < 0.0001$). There is a significant reduction in sore throat days in those participants randomised to immediate tonsillectomy. In other words, for every 10 sore throat days reported by participants in the conservative management arm over the 24-month (105-week) follow-up, there are only 5.28 sore throat days reported in participants randomised to tonsillectomy. The median number of sore throat days in the conservative management arm is 30 days. Hence, the model would predict a reduction of sore throats in this case by 14 days.

Baseline severity was found to not be statistically significantly related to the total number of sore throat days reported in the 24-month follow-up, despite the fact that it was a stratification factor in the trial design. Participants with moderate or severe baseline severity in terms of TOI-14 score are estimated to have 1.213 and 1.031 times more sore throat days, respectively, than those with mild category scores while holding the other variables constant in the model. Including site as a random effect demonstrates

TABLE 12 Results of multilevel mixed-effect NBR (model 1): ITT population

	Type	IRR	95% CI of IRR		p-value
Arm (reference value: conservative management): immediate tonsillectomy	Fixed	0.528	0.428 to 0.650		<0.001
Baseline severity (ref is mild)					
Moderate	Fixed	1.213	0.923	1.594	0.166
Severe		1.031	0.778	1.367	0.831
Constant		122.478	87.521	171.399	<0.001
Ln (exposure)		1			
Ln (alpha)		0.025	-0.103	0.153	
Alpha (dispersion parameter)		1.025	0.902	1.165	

LR, likelihood ratio.

Note

The variance parameter (95% CI) for random variable 'site' is 0.188 (0.064 to 0.553).
 $n = 429$, log-likelihood = -2074.766, LR test vs. negative binomial model: $\bar{\chi}^2 = 17.12$. Prob> = $\bar{\chi}^2 < 0.0001$.

the limited impact of variation between sites on primary outcome measure, but is retained as a stratification variable in the trial analysis. [Table 13](#) shows the summary statistics of the raw total number of sore throats reported by arm and baseline severity. This confirms that the number of sore throat days reported does not seem to be dependent on baseline severity. The analysis that allowed interaction with baseline severity (see [Appendix 13](#)) showed that there was a weak tendency for those classed as mild at baseline to have more of an improvement with tonsillectomy than those classed as moderate or severe at baseline; however, the interaction terms were not statistically significant. Hence, overall, we see a strong benefit in those randomised to the tonsillectomy arm and have no evidence that there is major variation in the impact of the tonsillectomy pathway by baseline participant characteristics.

Given that alpha is, again, significantly greater than zero (1.025), the data are over dispersed and are estimated better using a NBR than using a Poisson model.

TABLE 13 Summary statistics of the raw total number of sore throats reported by arm and baseline severity

Severity	Randomised arm		Total (N = 453)
	Immediate tonsillectomy (N = 233)	Conservative management (N = 220)	
Mild			
n (%)	48 (21)	41 (19)	89 (20)
Median (IQR)	28.5 (10–52)	35 (16–85)	31 (10–54)
Mean (SD)	34.0 (30.2)	72.5 (121.0)	51.7 (86.7)
Min, max	1, 140	2, 691	1, 691
Moderate			
n (%)	92 (39)	88 (40)	180 (40)
Median (IQR)	23 (11.5–47)	29 (12.5–59)	27 (12–51)
Mean (SD)	36.9 (38.2)	46.1 (49.3)	41.4 (44.1)
Min, max	1, 224	0, 288	0, 288
Severe			
n (%)	84 (36)	76 (35)	160 (35)
Median (IQR)	22 (12.5–37.5)	31.5 (16–66)	25 (14–50)
Mean (SD)	32.0 (44.3)	58.9 (95.1)	44.8 (74.0)
Min, max	0, 376	1, 660	0, 660
No STAR data			
n (%)	9 (4)	15 (7)	24 (5)
Median (IQR)	N/A	N/A	N/A
Mean (SD)	N/A	N/A	N/A
Min, max	N/A	N/A	N/A
Total (with STAR)			
n (%)	224 (96)	205 (93)	429 (95)
Median (IQR)	23 (11–46)	30 (14–65)	27 (12–52)
Mean (SD)	34.4 (39.0)	56.1 (85.6)	44.8 (66.4)
Min, max	0, 376	0, 691	0, 691

N/A, not applicable.

Exploration of Sore Throat Alert Return profile by randomised arm over the 105 weeks

The primary analysis (model 1) shows a statistically significant reduction in sore throats in those randomised to the immediate tonsillectomy arm. It is important to understand where this strength of difference comes from because there is crossover between arms and tonsillectomy causes sore throat days post operatively.

The weekly average number of sore throats reported by the primary ITT analysis population by arm is shown in [Figure 6](#). The graphs also include embedded histograms showing the timing of the tonsillectomy intervention. The effect of the tonsillectomy on the number of sore throat days is seen clearly in the immediate tonsillectomy arm.

Participants who were originally randomised to the conservative management arm underwent tonsillectomy much later than the 8-week protocol target for the participants in the immediate tonsillectomy arm.

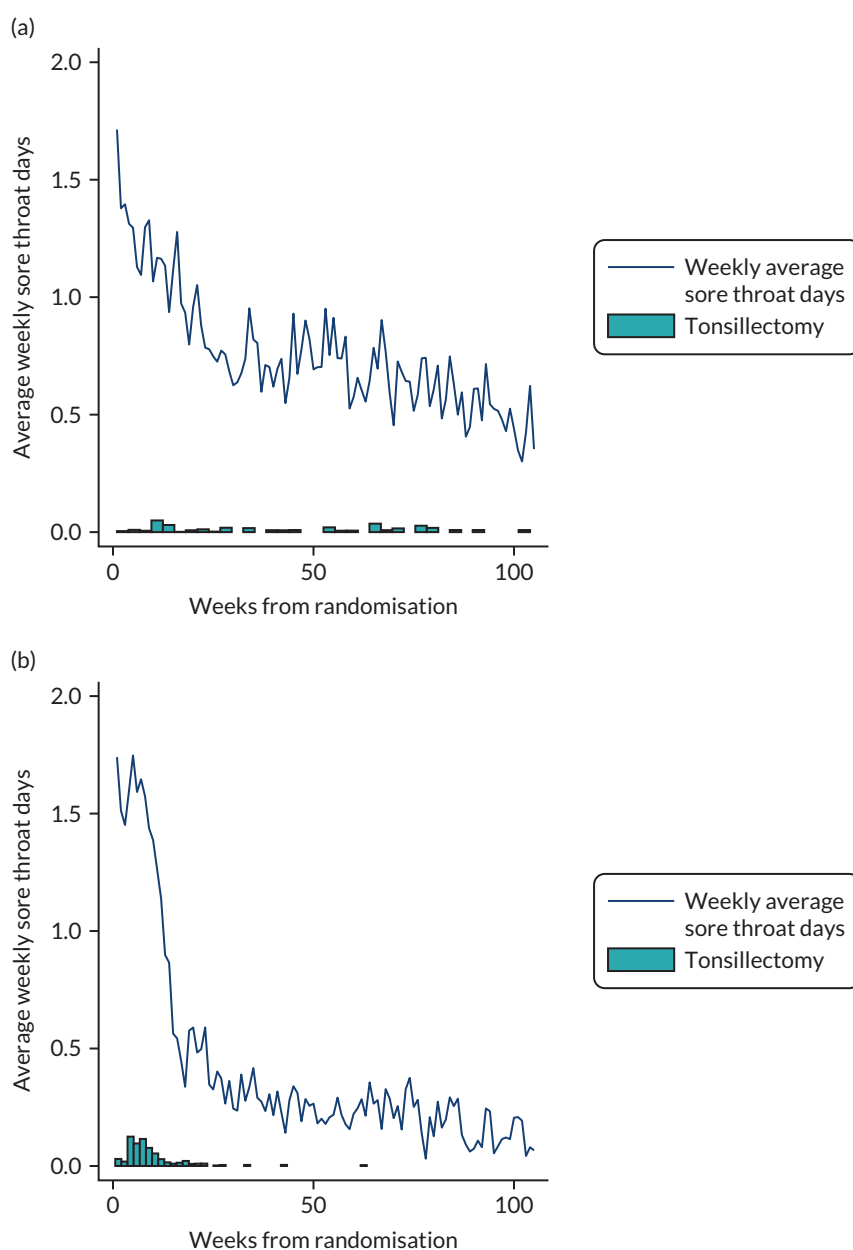


FIGURE 6 Weekly average sore throat days with histograms showing point of tonsillectomy by arm. (a) Conservative management ($n = 205$); and (b) immediate tonsillectomy ($n = 224$). Histogram of tonsillectomies: conservative management, $n = 59$; immediate tonsillectomy, $n = 172$.

The weekly average of number of sore throats also falls over time in this arm, but the decline is less obviously linked to the intervention (as in the immediate tonsillectomy arm) owing to there having been fewer tonsillectomies, which were more dispersed over follow-up. It can also be seen that at the end of the follow-up period, the average weekly rate in the conservative management arm was around 0.5 sore throat days per week, whereas in the tonsillectomy arm it was substantially lower (around 0.2 days per week).

The pattern of reporting of sore throat days in relation to the timing of any tonsil surgery is detailed in [Appendix 12, Table 44](#), and is summarised in [Figure 7](#). The patterns for sore throat days looked similar in both arms for the three periods for those who received surgery and for the one period for those who did not receive surgery. In both arms, there tended to be an increase in sore throats immediately following surgery, then thereafter post surgery the rates dropped below those seen for the no surgery group.

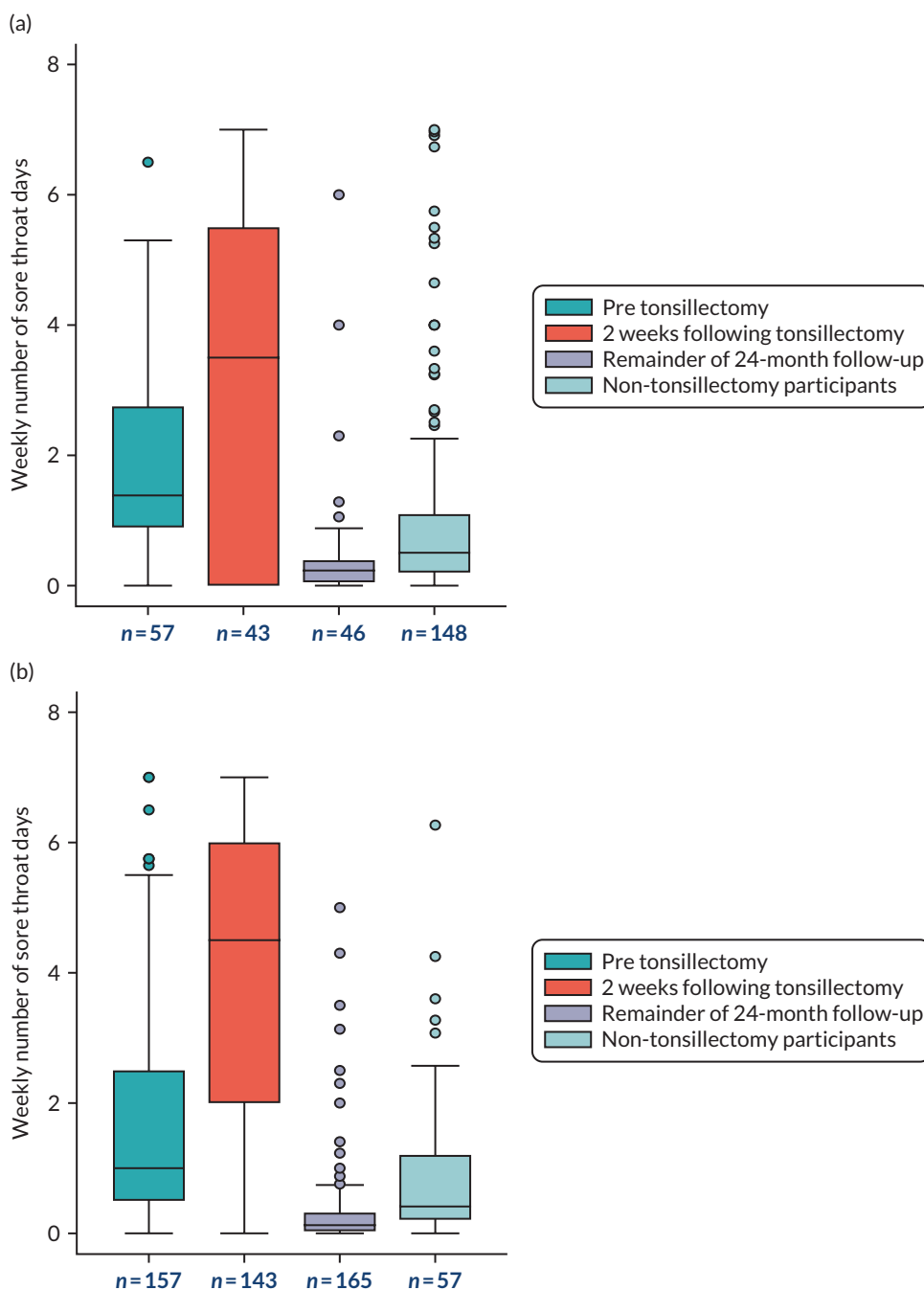


FIGURE 7 Average weekly sore throat rate for participants not receiving tonsillectomy and those who received tonsillectomy. (a) Conservative management; and (b) immediate tonsillectomy.

RESULTS

The box plots shown in [Figure 7](#) show the distribution of sore throats by time period. The box plots show that the pre-tonsillectomy rates look similar, that there is a sharp increase in sore throats reported immediately following tonsillectomy, and that rates are very low after the first 2 post-operative weeks. More than 2 weeks post tonsillectomy, the sore throat weekly rate is very much lower in those who received tonsillectomy than in those who did not receive tonsillectomy.

In [Figure 7](#), the average weekly sore throat rate in three time periods is shown: pre-tonsillectomy (mid-blue), 2 weeks following tonsillectomy (orange) and the remainder of the 24-month follow-up period post operation (light purple). The non-tonsillectomy participants are also displayed (light blue).

Addressing missing data

Although 72% of STARs were returned, we were interested to understand if there was differential missingness in the STAR rate by arm or subgroup of interest. [Table 14](#) shows the cumulative percentage of missing primary outcome data over the 24 months following randomisation. This is the length of time measuring the primary outcome. The numbers and percentages are presented by category, by arm and overall.

Eighty-one out of 455 (18%) randomised participants had complete data. A total of 227 (50%) participants provided >90% of their STAR responses and 263 (58%) returned >80%. A slightly higher proportion of the tonsillectomy arm completed at least 90% of STARs than the conservative management arm.

We explored this further by dividing the 24-month follow-up period into 6-monthly intervals. Weekly STAR responses can be categorised during each 6-month interval from the point of randomisation as:

- complete, in which participants completed all of the expected weekly STAR responses within that time frame
- partial, in which some but not all expected weekly STAR responses were received
- no returns, in which participants did not return any of the expected weekly STAR responses within that time frame.

[Figure 8](#) shows the STAR responses categorised as complete, partial or no returns in a bar chart at each of the 6-month follow-up intervals.

TABLE 14 Cumulative percentage of weekly STAR responses 24 months following randomisation

Cumulative percentage of STAR responses (%)	Randomised arm, n (%)		
	Immediate tonsillectomy	Conservative management	Total, n (%)
100	42 (18)	39 (18)	81 (18)
>90	128 (55)	99 (45)	227 (50)
>80	146 (62)	117 (53)	263 (58)
>70	156 (67)	124 (56)	280 (62)
>60	163 (70)	131 (59)	294 (65)
>50	171 (73)	139 (63)	310 (68)
>25	185 (79)	160 (72)	345 (76)
>0	224 (96)	205 (93)	429 (94)
No returns	10 (4)	16 (7)	26 (6)
Total	234 (100)	221 (100)	455 (100)

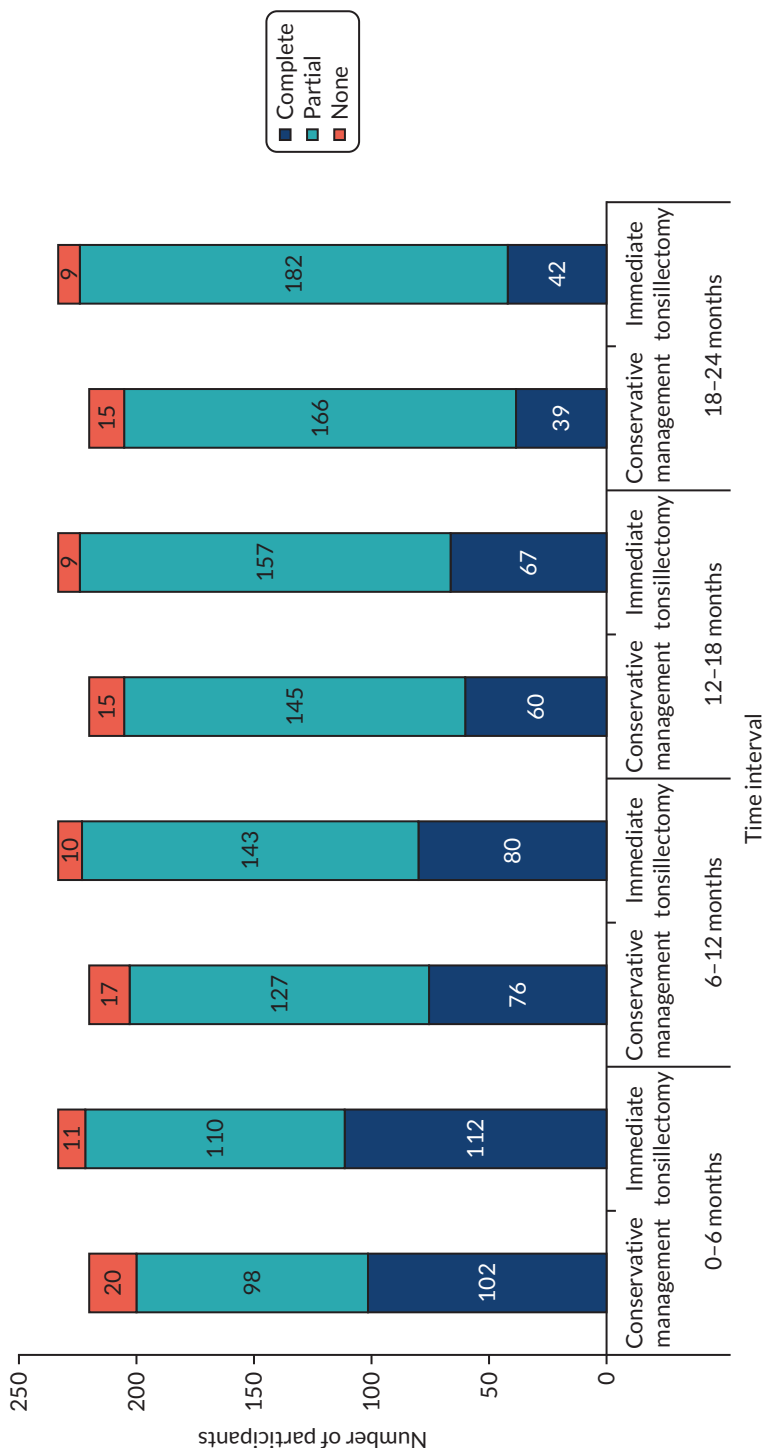


FIGURE 8 Bar chart showing observed number of STARs by categories (complete, partial and none) by arm at time points (6, 12, 18 and 24 months).

[Table 15](#) shows the expected number of responses, observed complete responses, observed partial responses and missing responses (no weekly STAR response returns) for 6, 12, 18 and 24 months when other trial questionnaires were due to be completed. [Tables 14](#) and [15](#) show that there are slight differential STAR rates by arm. For example, 146 (62%) participants in the immediate tonsillectomy arm completed more than 80% of their returns compared with just 117 (53%) participants in the conservative management arm.

We looked at the time of the last STAR for each participant. We define retention as the time in the trial, which is calculated as the time from randomisation to the last weekly STAR response. The retention times in the trial for the participants who sent at least one weekly STAR response are presented in box plots in [Figure 9](#). This allows an assessment of whether or not retention appeared to be balanced by randomised arm. Although 50% of participants in both arms completed almost all STARs, the box plots in [Figure 9](#) show that participants randomised to the immediate tonsillectomy arm were more likely than those in the conservative management arm to continue sending STAR responses further into the follow-up period. [Table 14](#) breaks down the distribution by completeness rather than by last return.

Sensitivity analyses

Two sensitivity analyses of the primary analysis were conducted. The first (NBR model 2) was a re-analysis, but replaced the baseline TOI-14 severity stratification ordinal variable with a continuous variable TOI-14. The second (NBR model 3) is a further-adjusted sensitivity analysis incorporating other potentially important baseline factors, sex, age, ethnicity, education level and employment status, in addition to the two variables of continuous baseline severity and recruitment site. Non-linear continuous covariates (age at randomisation and baseline severity TOI-14 score) were assessed by transforming them using simple first-degree polynomial transformations. Age at randomisation was found to not have a significant univariate relationship with the outcome measure at the primary end point ($p > 0.1$), so was omitted from model 3. There was no reduction in Akaike information criterion through simple transformations. To build the most parsimonious, clinically interpretable model, baseline TOI-14 was retained in models 2 and 3 as an untransformed continuous covariate, under the assumption of linearity with outcome.

As previously described, the exposure variable was incorporated into the NBR model developed to account for missing follow-up data. Further details of the sensitivity analyses are presented in [Appendix 7](#), but the key IRR and associated CIs are shown in the forest plots in [Figure 10](#). [Figure 10](#) shows the results for the primary adjusted analysis NBR model 1, along with models 2 and 3. It also shows an additional, unplanned, sensitivity analysis including only those with $\geq 80\%$ STARs. This was motivated by our concerns over differential missingness of STARs, as shown in [Figure 9](#).

All three sensitivity analyses showed very similar results to model 1. Therefore, we conclude that the primary analysis is robust (see [Figure 10](#)).

In response to reviewer feedback, an unplanned sensitivity analysis was conducted to further explore the impact of the missing STARs. The primary analysis adjusted for the proportion of STARs, which implicitly assumes that the STARs completed were a representative sample of the sore throat days throughout the 105 weeks. A more conservative assumption is to assume that a missing return means that there were no sore throats to report. We re-analysed the data using model 1 but without including the exposure variable. The overall IRR increased slightly to 0.614 (95% CI 0.507 to 0.775). This extreme conservative assumption weakens the strength of the signal, but the overall conclusions are unchanged.

TABLE 15 Proportion of participants responding to weekly STAR responses split by observed complete, partial and no response at 6-monthly intervals post randomisation

Time frame (months)	Randomised arm				Conservative management (N = 220)				Total (N = 453)			
	Immediate tonsillectomy (N = 233)											
	Complete, n (%)	Partial, n (%)	No returns, n (%)	Median (IQR)	Complete, n (%)	Partial, n (%)	No returns, n (%)	Median (IQR)	Complete, n (%)	Partial, n (%)	No returns, n (%)	Median (IQR)
0-6	112 (48)	110 (47)	11 (5)	25 (18-26)	102 (46)	98 (45)	20 (9)	25 (12.5-26)	214 (47)	208 (46)	31 (7)	25 (16-26)
6-12	80 (34)	143 (61)	10 (4)	50 (32-52)	76 (35)	127 (58)	17 (8)	49 (18.5-52)	156 (34)	270 (60)	27 (6)	50 (22-52)
12-18	67 (29)	157 (67)	9 (4)	74 (39-78)	60 (27)	145 (66)	15 (7)	70 (20.5-78)	127 (28)	302 (67)	24 (5)	72 (26-78)
18-24	42 (18)	182 (78)	9 (4)	98 (50-104)	39 (18)	166 (75)	15 (7)	89 (21.5-104)	81 (18)	348 (77)	24 (5)	95 (28-104)

Note that '%' is n/N x 100. Four participants returned no STARs from 0 to 6 months, but engaged thereafter. A further three participants returned no STARs from 0 to 12 months, but engaged thereafter.

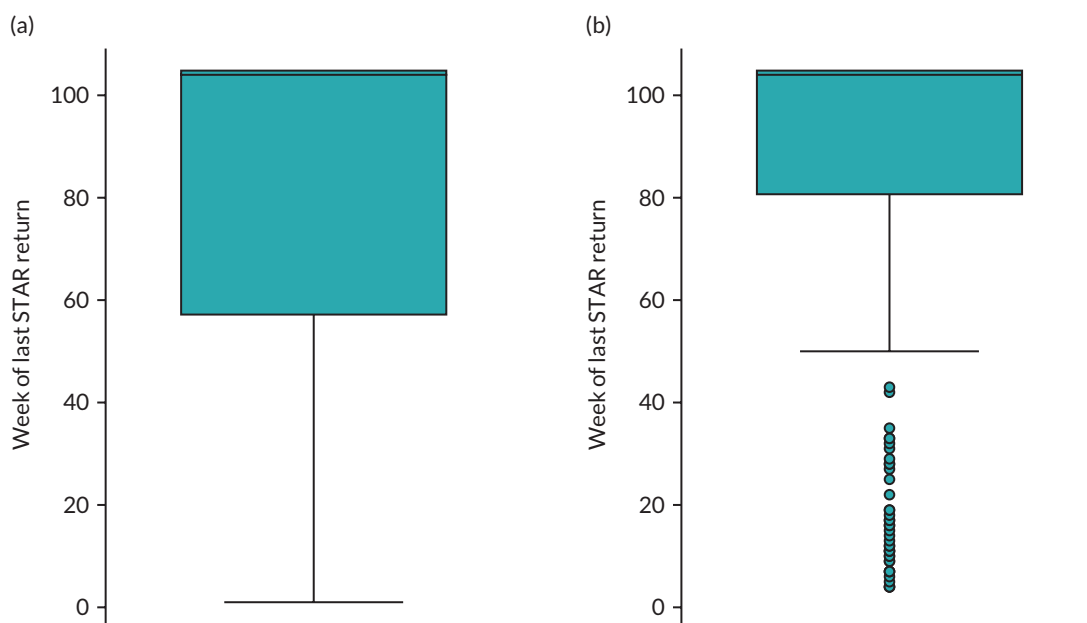


FIGURE 9 Box plots showing distribution of time from randomisation to last STAR response in the 24-month follow-up period by arm. Conservative management ($n = 205$); and (b) immediate tonsillectomy ($n = 224$).

Further sensitivity analyses outlined in the SAP were also undertaken to confirm the effect of tonsillectomy. Full results for these can also be found in [Appendix 7](#), but a brief description of what was analysed is included here. The results are shown in the forest plot in [Figure 11](#).

Per-protocol analysis

The per-protocol analysis was restricted to the 224 participants who complied with the protocol (see [Table 7](#)): 95 participants in the tonsillectomy arm and 139 participants in conservative management arm, excluding the 10 participants with no STARs. See [Appendix 7](#) for details of the analysis. Participants who withdrew were omitted from these analyses.

As might be expected, the per-protocol analyses showed a stronger effect of the intervention, with a reduction of 0.415 times the sore throat days. Therefore, the results of the primary analysis were found to be robust compared with these sensitivity analyses, as shown in the forest plot (see [Figure 11](#) and [Appendix 7](#)).

Addressing crossover (per-treatment analysis)

Participants were able to switch over from conservative management to tonsillectomy, and also from tonsillectomy to conservative management. A new group variable was created indicating whether or not tonsillectomy was received by the participant. This replaced 'arm randomised' to capture the effects of tonsillectomy for per-treatment analysis.

For this 'per-tonsillectomy' analysis, we included 224 participants who received tonsillectomy and 205 who did not receive tonsillectomy, making a total of 429 participants who had also completed at least one STAR (primary outcome measure for the primary analyses).

A second per-treatment analysis was carried out according to four categories:

1. randomised to receive immediate tonsillectomy and tonsillectomy completed (at any time during follow-up) ($n = 167$)
2. randomised to receive immediate tonsillectomy but tonsillectomy not completed ($n = 57$)
3. randomised to receive conservative management but tonsillectomy received ($n = 57$)
4. randomised to receive conservative management with no tonsillectomy completed ($n = 148$).

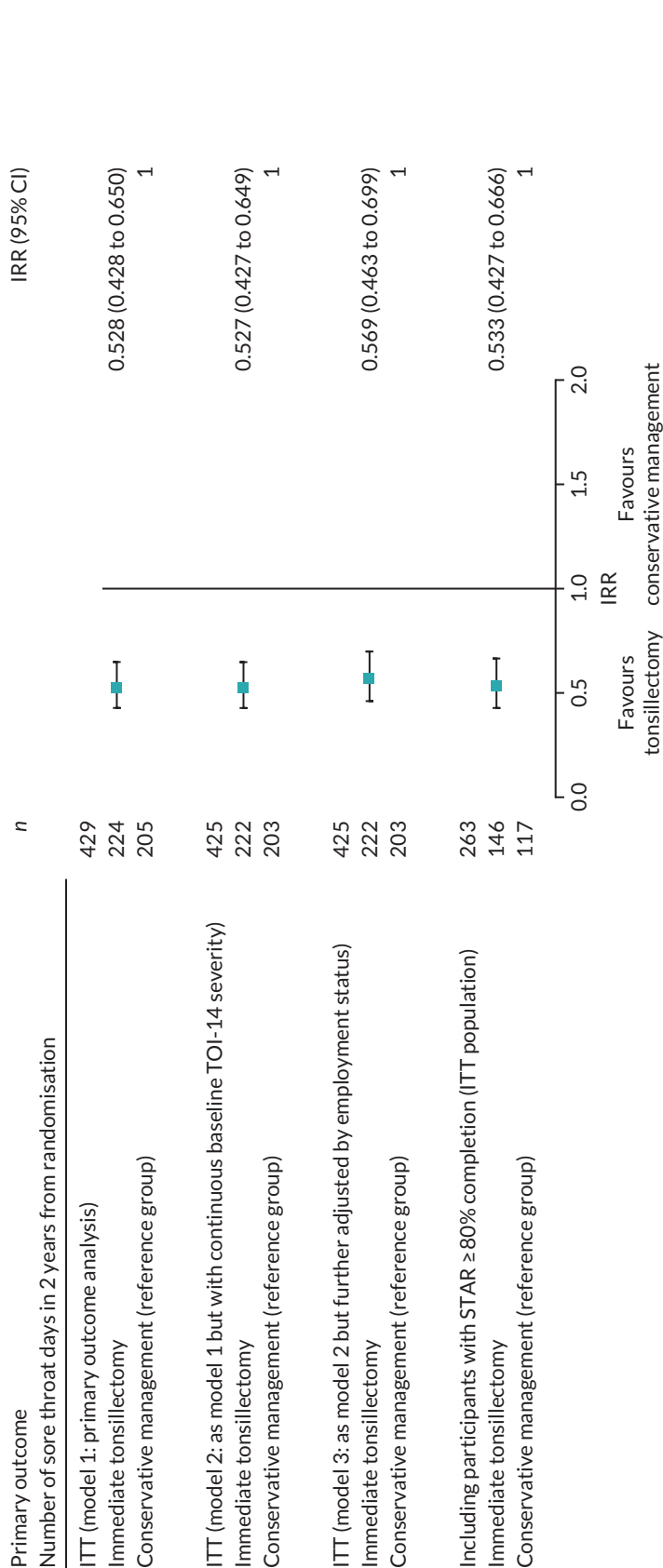


FIGURE 10 Forest plot showing IRRs and 95% CIs for primary ITT analysis and sensitivity analyses on ITT population.

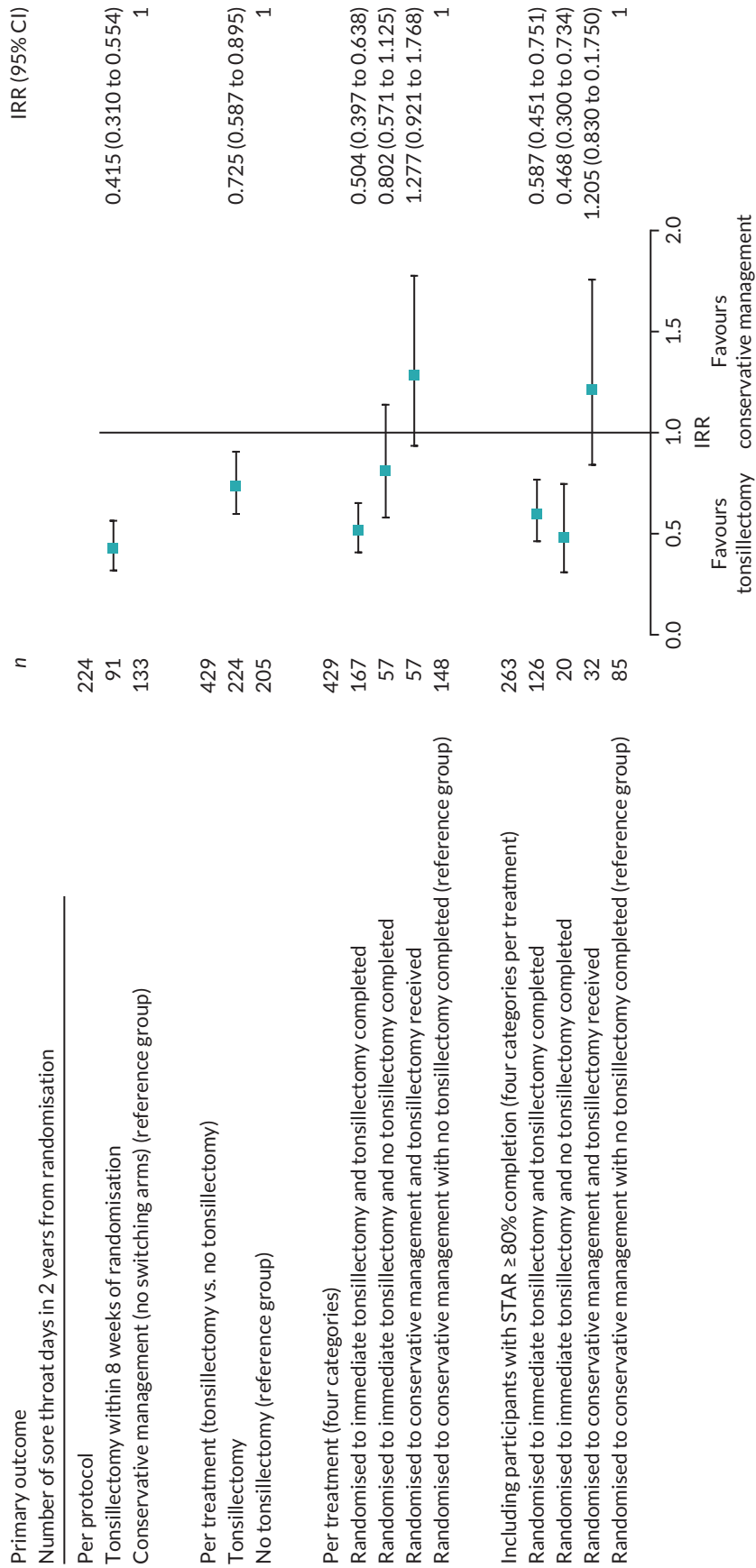


FIGURE 11 Forest plot with sensitivity analyses restricting the ITT population to per-protocol and per-treatment groups.

This reflects whether or not participants actively chose to swap from their randomised arm. This was carried out to assess if there were differences in severity and/or outcome among those who deliberately chose to swap from their randomised arm compared with those who were content to remain as randomised.

The instrumental variable sensitivity analyses (CACE and interactions of baseline severity with randomised arm allocation) were not included in the forest plot, and are presented in [Appendix 13](#), [Table 45](#).

We repeated the four category per-treatment analysis, limiting the population to those who returned at least 80% of STARs ($n = 263$). These sensitivity analyses were adjusted by stratification variables with reference categories as in the original analyses. The pattern is similar to the full per-treatment analysis (and some of the groups now have very small numbers), but those randomised to receive tonsillectomy who then chose not to have tonsillectomy had significantly fewer sore throats than those who were randomised to and received conservative management. This would be consistent with this group tending to switch away from surgery because of their lesser sore throat burden.

The forest plots in [Figures 10](#) and [11](#) demonstrate that the primary and sensitivity ITT analyses show a strong signal of differences between the two randomised arms in terms of IRRs and 95% CIs.

The per-protocol and two-group per-treatment analyses were consistent with the primary analysis NBR model 1 in that they too showed a statistically significant reduction in total sore throats when comparing the immediate tonsillectomy arm with the conservative management arm. The only slight signals at odds with this were those seen in the crossover groups in the four-group analysis. We therefore felt that it was important to explore the patterns of STARs and sore throat profiles over the 24 months by the various analysis populations. See [Appendix 11](#) for further details.

Those who received tonsillectomy having crossed over from conservative management [$(n = 57)$ in [Figure 11](#) per-treatment four categories group] saw not a fall but an (non-statistically significantly) increase in sore throat days (IRR 1.277, 95% CI 0.921 to 1.768; $p = 0.141$) (see [Appendix 12](#), [Table 44](#)). The anomaly here is that those having tonsillectomy after crossing over from conservative management reported more sore throats than those remaining in the conservative management arm. This could be because of two factors. First, these participants had higher rates of sore throat and, therefore, were more motivated to try tonsillectomy, whereas those who remained in the conservative management arm, as randomised, may have tended to be those with lower rates of sore throat. Second, the effect of surgery itself induced additional sore throats. The later that the surgery is received, the less time there is for benefit from the surgery to compensate for the (late) cost of additional, postoperative sore throat days. The four graphs in [Figure 12](#) of average weekly sore throats over time appear to support this interpretation. Those in the conservative management arm who switch to tonsillectomy seem to have higher initial rates of weekly sore throats than those in the tonsillectomy arm who switch to conservative management. Conversely, participants who were randomised to receive tonsillectomy, but who chose not to receive it, started off with fewer sore throat days than the other participants in that arm. This may explain why they opted not to have tonsillectomy (see [Figure 12](#)).

[Figure 13](#) shows the rate of weekly STARs (exposure) by the four per-treatment categories. These box plots show that those randomised to receive immediate tonsillectomy and who received surgery had the highest return rates. Those randomised to receive immediate tonsillectomy but who did not receive tonsillectomy had the lowest return rates.

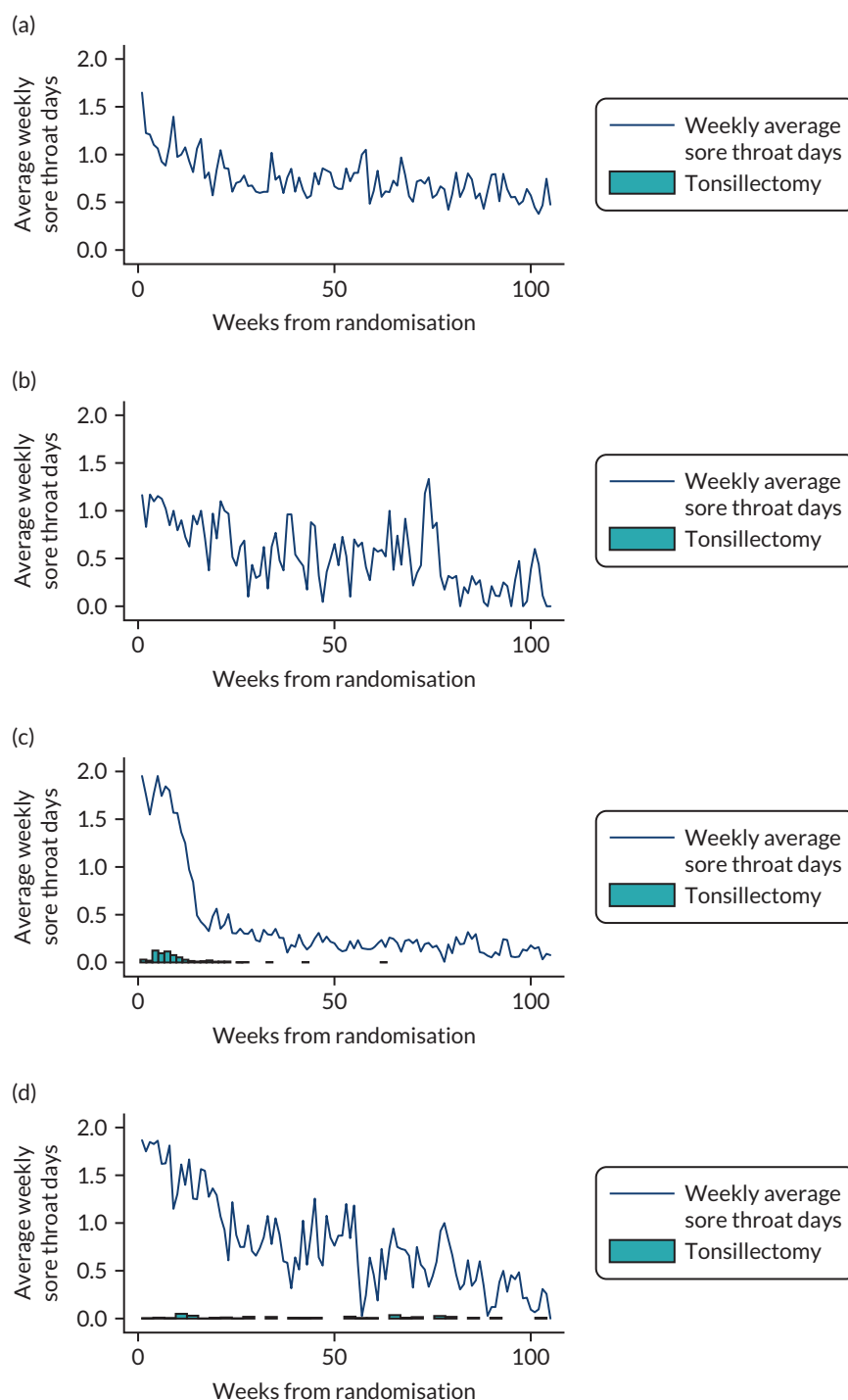


FIGURE 12 Weekly average sore throat days by randomised arm and occurrence of tonsillectomy. (a) No surgery as randomised ($n = 148$); (b) no surgery but randomised to surgery ($n = 57$); (c) surgery as randomised ($n = 167$); and (d) surgery crossed ($n = 57$). Graphs by tonsillectomy and randomised arm.

Secondary outcome measures

The majority of the secondary outcome measures are presented in [Chapter 4](#). The secondary outcomes, as described in the SAP, that are presented in this section are TOI-14 over time by arm and QoL SF-12 domains, and MCS and PCS, over time. Both questionnaires were completed at baseline, 6 months (postal), 12 months (visit), 18 months (postal) and 24 months (visit). [Table 16](#) shows the missing data.

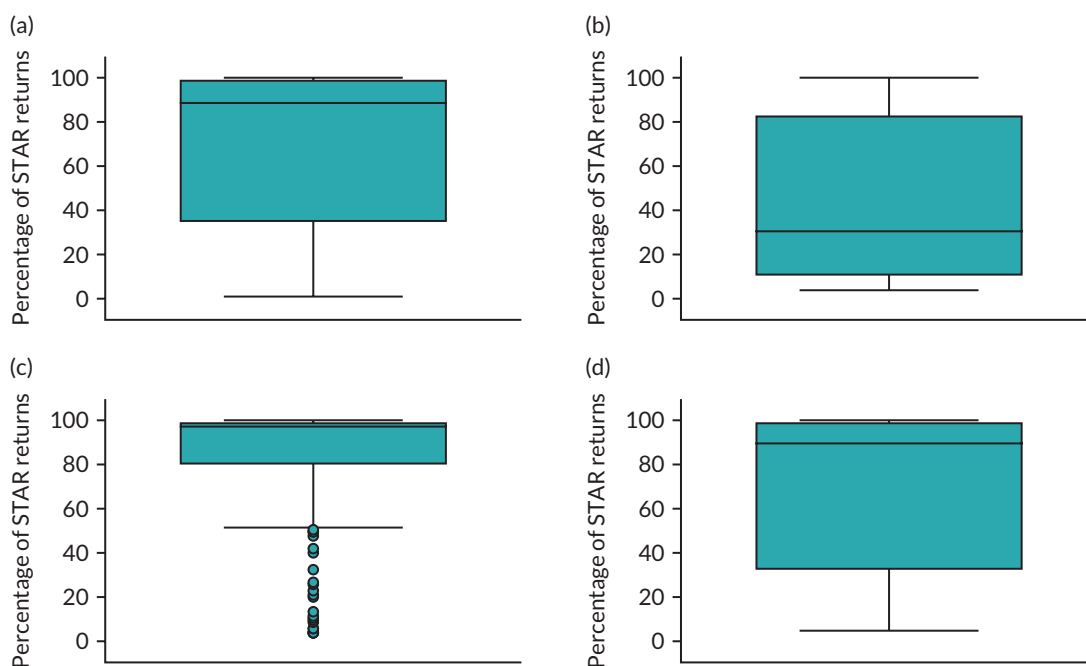


FIGURE 13 Box plots showing percentage of total STARs by four category per-treatment groups ($n = 429$). (a) No surgery as randomised ($n = 148$); (b) no surgery but randomised to surgery ($n = 57$); (c) surgery as randomised ($n = 167$); and (d) surgery crossed ($n = 57$). Graphs by surgery received and arm randomised to.

TABLE 16 The SF-12 and TOI-14 complete, partial or missing over time in the ITT population

Numbers missing, partial and complete	TOI-14 questionnaire, n (%)			SF-12 questionnaire, n (%)		
	Immediate tonsillectomy ($N = 233$)	Conservative management ($N = 220$)	Total ($N = 453$)	Immediate tonsillectomy ($N = 233$)	Conservative management ($N = 220$)	Total ($N = 453$)
Baseline						
Complete	230 (99)	216 (98)	446 (98)	228 (98)	215 (98)	443 (98)
Partial	1 (<1)	1 (<1)	2 (<1)	3 (1)	1 (<1)	4 (<1)
Missing completely	2 (<1)	3 (1)	5 (1)	2 (<1)	4 (2)	6 (1)
6 months						
Complete	104 (45)	81 (37)	185 (41)	108 (46)	84 (38)	192 (42)
Partial	1 (<1)	4 (2)	5 (1)	2 (>1)	1 (<1)	3 (<1)
Missing completely	128 (55)	135 (61)	263 (58)	123 (53)	135 (61)	258 (57)
12 months						
Complete	121 (52)	116 (53)	237 (52)	122 (52)	116 (53)	238 (53)
Partial	1 (<1)	1 (<1)	2 (<1)	0 (0)	1 (<1)	1 (<1)
Missing completely	111 (48)	103 (46)	214 (47)	111 (48)	103 (47)	214 (47)
18 months						
Complete	81 (35)	70 (32)	151 (33)	82 (35)	74 (34)	156 (34)
Partial	2 (<1)	2 (<1)	4 (<1)	5 (2)	0 (0)	5 (1)
Missing completely	150 (64)	148 (67)	298 (66)	146 (63)	146 (66)	292 (64)
24 months						
Complete	99 (42)	100 (45)	199 (44)	100 (43)	99 (45)	199 (44)
Partial	0 (0)	0 (0)	0 (0)	0 (0)	1 (<1)	1 (<1)
Missing completely	134 (58)	120 (55)	254 (56)	133 (57)	120 (55)	253 (56)

The analysis set is the ITT population. The analysis was repeated for per-protocol and per-treatment populations (see [Appendix 14](#)).

Missing secondary outcome data

[Appendix 15](#) includes details of the distribution of missing data among the baseline severity categories (mild, moderate and severe) according to STAR responses and TOI-14 questionnaires.

Tonsillectomy Outcome Inventory-14

The responses at the five time points to the TOI-14 are shown in [Figure 14](#) and are summarised using appropriate descriptive statistics along with 95% CIs in [Table 17](#). These were similar at baseline as expected, but there was a marked difference at follow-up visits (after surgery) in the overall TOI-14 scores (non-overlapping CIs). [Appendix 14](#) details the four TOI-14 subscales in a similar fashion – throat discomfort, general health, resource impact, and social and psychological:

- Questionnaires missing one or two items were completed by imputation of the item score calculated as the average of the completed items.
- Analysis and charts in the rest of this section include the recovered questionnaires with imputed missing items.

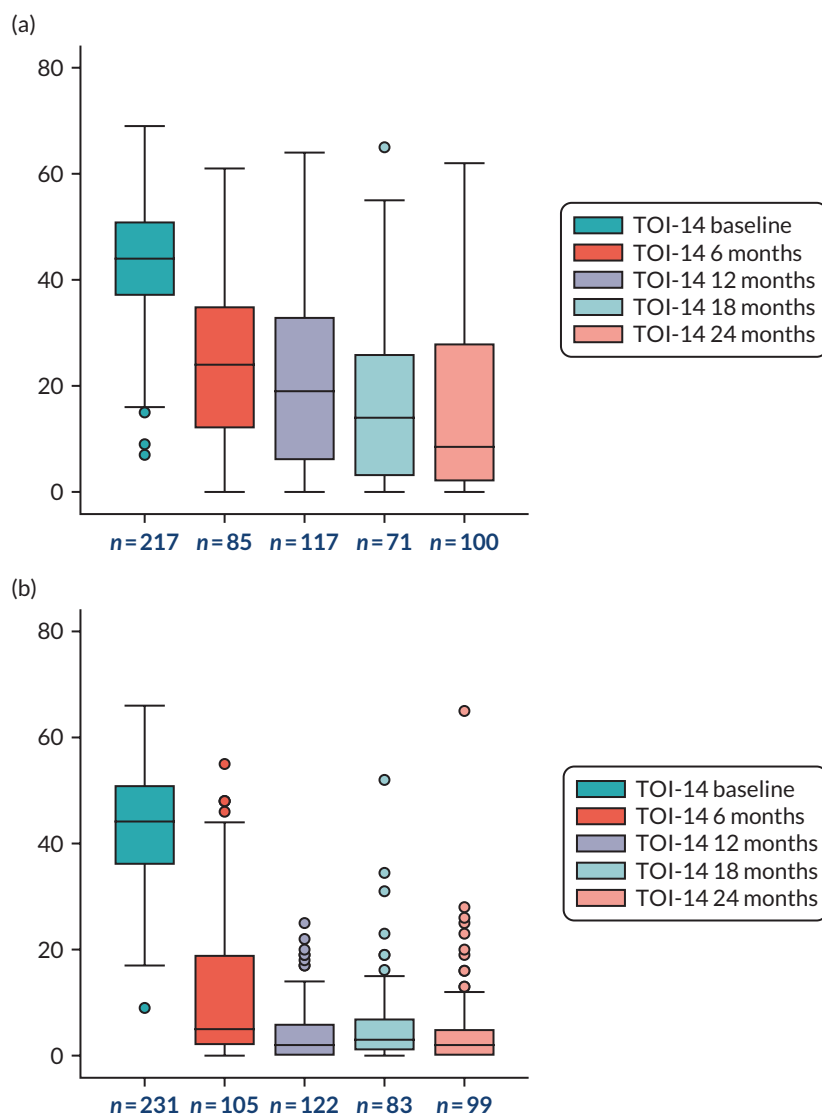


FIGURE 14 The TOI-14 at 6, 12, 18 and 24 months: ITT population. (a) Conservative management; and (b) immediate tonsillectomy.

TABLE 17 The TOI-14 overall scores at time points (% is the proportion completed out of eligible randomised)

TOI-14 overall score	Tonsillectomy (N = 233)	Conservative management (N = 220)	Total (N = 453)
Baseline			
n (%)	231 (99)	217 (99)	448 (99)
Median (IQR)	44.2 (36–51)	44 (37–51)	44 (36–51)
Mean (SD)	43.8 (11.2)	43.2 (11.6)	43.5 (11.4)
95% CI about mean	42.3 to 45.2	41.6 to 44.7	42.4 to 44.5
Min, max	9, 66	7, 69	7, 69
6 months (post)			
n (%)	105 (45)	85 (39)	190 (42)
Median (IQR)	5 (2–19)	24 (12–35)	14 (3–31)
Mean (SD)	12.8 (14.8)	24.7 (15.6)	18.1 (16.3)
95% CI about mean	9.9 to 15.7	21.4 to 28.1	15.8 to 20.5
Min, max	0, 55	0 to 61	0, 61
12 months (visit)			
n (%)	122 (55)	117 (50)	239 (53)
Median (IQR)	2 (0–6)	19 (6–33)	6 (1–19)
Mean (SD)	4.3 (5.3)	21.7 (17.6)	12.8 (15.5)
95% CI about mean	3.4 to 5.3	18.5 to 24.9	10.8 to 14.8
Min, max	0, 25	0, 64	0, 64
18 months (post)			
n (%)	83 (36)	71 (32)	154 (34)
Median (IQR)	3 (1–7)	14 (3–26)	5 (2–16)
Mean (SD)	5.7 (8.4)	16.4 (15.8)	10.6 (13.4)
95% CI about mean	3.9 to 7.5	12.6 to 20.1	8.5 to 12.7
Min, max	0, 52	0, 65	0, 65
24 months (visit)			
n (%)	99 (42)	100 (45)	199 (44)
Median (IQR)	2 (0–5)	8.5 (2–28)	3 (1–13)
Mean (SD)	4.7 (8.7)	15.4 (17.1)	10.1 (14.6)
95% CI about mean	2.9 to 6.4	12.0 to 18.8	8.0 to 12.1
Min, max	0, 65	0, 62	0, 65

Max, maximum; min, minimum.

Details of imputation are included in [Appendix 14, Table 51](#).

Reductions in the TOI-14 scores (symptom improvements) were seen in both randomised arms, but were greater in the immediate tonsillectomy participants, similar to the benefits demonstrated in [Figure 14](#). Summary statistics are presented in [Table 17](#), in which the non-overlapping CIs show the differences at follow-up visits between the randomised arms.

Box plots for the per-protocol and four category per-treatment population groups are presented in [Appendix 14, Figures 28 and 29](#).

The difference between the means, with CIs and the corresponding *p*-value at each time point, is presented in [Table 18](#). The difference is calculated as conservative management minus immediate tonsillectomy and higher values equate to worse symptoms.

There is no difference at baseline, which was expected owing to the randomised nature of the trial. Baseline TOI-14 scores were used as stratification factors for the trial. Significant differences can be seen at the follow-up time points, however, with conservative management having higher average TOI-14 scores than tonsillectomy. This shows how the participants who were randomised to immediate tonsillectomy report significantly better outcome at every follow-up point, with a larger difference apparent at the 12-month visit at which more participants provided TOI-14 questionnaires (*n* = 239). This supports the conclusions of the box plots in [Figure 14](#).

In addition, scores were statistically analysed using longitudinal repeated-measure maximum-likelihood models developed for longitudinal data. The dependent variable was the overall TOI-14 score for an individual participant at a particular time point. Both variation between participants and variation between responses nested within participants were modelled as random effects with a normal distribution. Differences between groups and changes over time were modelled as fixed effects. The analysis was adjusted for the treatment groups and stratification factor. There are statistically significant differences between randomised arms, with both randomised arms reporting improvements from baseline. The differences reported at follow-up time points beyond the initial 6-month time point are not significantly different, which shows that the majority of the improvements occur within the first 6 months after randomisation for participants randomised to either arm.

The mixed-model repeated-measure analysis carried out on the TOI-14 scores at the five time points showed that there is also a statistically significant difference in improvement between the randomised arms, with those randomised to receive immediate tonsillectomy reporting greater improvement than those randomised to receive conservative management. In summary, the mean scores tended to be lower in the tonsillectomy arm than in the conservative management arm by 13.17 (95% CI -17.41 to -8.92), which indicates a reduction in symptoms. See [Appendix 14, Table 47](#), for the full results table. The longitudinal repeat-measure maximum-likelihood model confirmed the conclusions drawn from the graphical interpretation for TOI-14 in [Figure 14](#). Summary statistics by arm and trial time points are shown in [Table 17](#).

Similar conclusions can be drawn for the TOI-14 subscales, as outlined in [Appendix 14, Tables 46 and 48–50](#), and [Figures 24–27](#).

Quality-of-life analysis: Short Form questionnaire-12 items

The analysis set is the ITT set. Quality-of-life scores based on the SF-12 were calculated using the Optum scoring software (Quality Metric Incorporated) at baseline and 6, 12, 18 and 24 months post randomisation. Scores are described with summary statistics and are graphically presented over time. The software takes

TABLE 18 Test difference in TOI-14 means by arm over time with 95% CIs

Time point	Number	Difference in means	Standard error	95% CI	<i>p</i> -value
Baseline	448	-0.634	1.077	-2.750 to 1.483	0.556
6 months	190	11.942	2.232	7.538 to 16.346	<0.0001
12 months	239	17.386	1.694	14.036 to 20.736	<0.0001
18 months	154	10.668	2.088	6.526 to 14.809	<0.0001
24 months	199	10.713	1.918	6.923 to 14.504	<0.0001

baseline levels into account and adjusts for follow-up time points accordingly. The algorithm contained within the software generates physical (SF-12 PCS) and mental (SF-12 MCS) health composite scores for comparison with the normative data for the UK population. In the normative data, the mean score is transformed to be 50 (SD 10), adjusted for sex and age, with scores greater than 50 indicating better physical or mental health than the mean and scores less than 50 indicating worse physical or mental health than the mean.

At baseline, the mean scores are just below 50, indicating worse mental and physical health scores than the UK population average. Mental health scores are further below 50 than physical health scores, which is an indication that the population suffers more with their mental health than their physical health when entering the trial.

For participants randomised to receive immediate tonsillectomy, improvements can be seen in both health measures, with physical health improving more, as reported at the follow-up collection time points. Smaller improvements are seen in the physical health score for participants randomised to receive conservative management than for those randomised to receive tonsillectomy. In addition, less improvement is seen in the mental health score for those randomised to the conservative management arm than for those randomised to the tonsillectomy arm, with average scores in the conservative management arm still falling below the population average and average scores in the tonsillectomy arm falling above the population average.

In addition, scores were statistically analysed using longitudinal repeated-measure maximum-likelihood models developed for longitudinal data. The dependent variable was the overall QoL score (SF-12 MCS and PCS) for an individual participant at a particular time point. Both variation between participants and variation between responses nested within participants were modelled as random effects with a normal distribution. Differences between arms and changes over time were modelled as fixed effects. The analysis was adjusted for the randomised arms and stratification factor.

Both randomised arms reported improvements from baseline. The differences reported at follow-up time points beyond the initial 6-month time point were not significantly different, showing that the majority of the improvements in both arms occurred within the first 6 months after randomisation.

There is also a statistically significant difference in improvement between the randomised arms. Both scores tended to be higher in the tonsillectomy arm than in the conservative management arm (SF-12 MCS 3.71 units higher, 95% CI 2.10 to 5.47; SF-12 PCS 2.77 units higher, 95% CI 0.30 to 5.23). See [Appendix 16, Tables 55 and 56](#), for the full results tables. The longitudinal repeat-measure maximum-likelihood models confirm the conclusions drawn from the graphical interpretation for SF-12 MCS and PCS in [Figures 15 and 16](#). Summary statistics by arm and trial time points are shown in [Table 19](#).

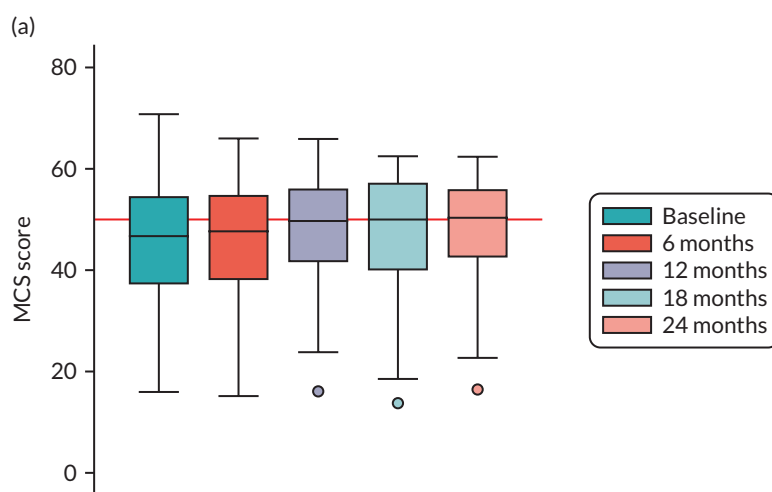


FIGURE 15 Box plots of normalised mental health scores (SF-12 MCS) at five time points by randomised arm. (a) Conservative management; and (b) immediate tonsillectomy. Red line is average normalised score for UK population. (*continued*)

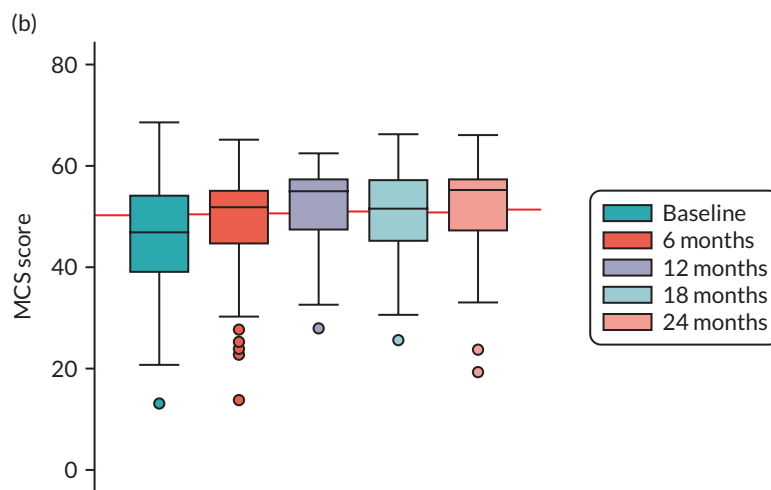


FIGURE 15 Box plots of normalised mental health scores (SF-12 MCS) at five time points by randomised arm. (a) Conservative management; and (b) immediate tonsillectomy. Red line is average normalised score for UK population.

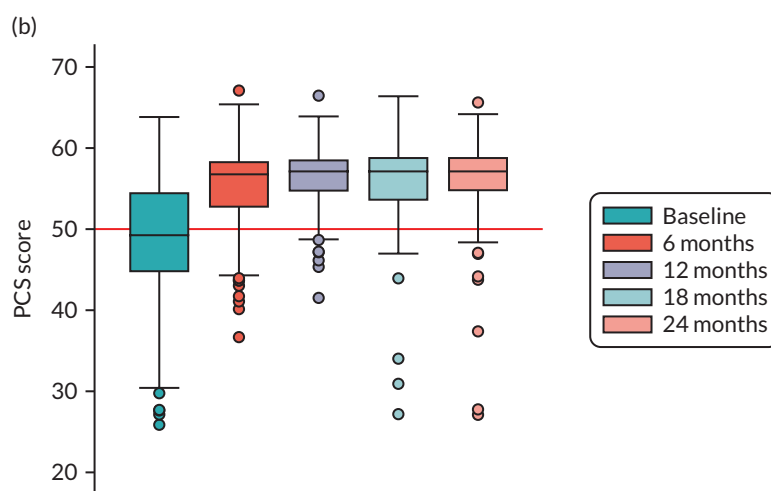
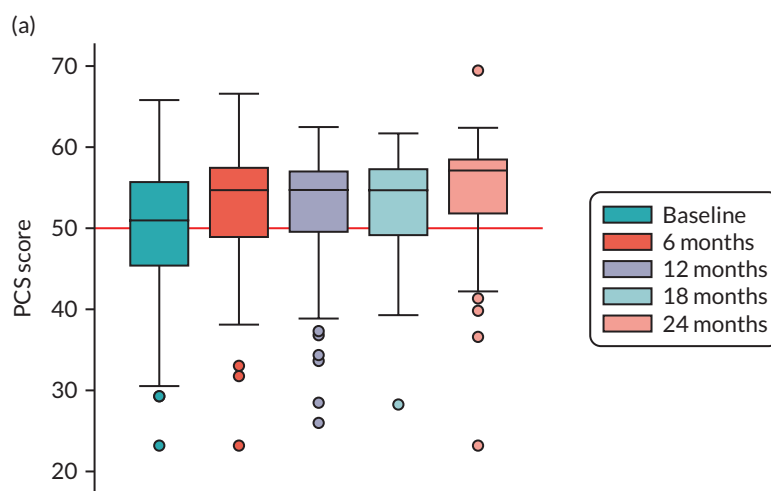


FIGURE 16 Box plots of normalised physical health scores (SF-12 PCS) at five time points by randomised arm. (a) Conservative management; and (b) immediate tonsillectomy. Red line is average normalised score for UK population.

TABLE 19 The SF-12 components (SF-12 MCS and PCS) summary statistics at time points, ITT population; lower scores = poorer QoL

Summary statistic	Immediate tonsillectomy (N = 233)		Conservative management (N = 220)	
	SF-12 MCS	SF-12 PCS	SF-12 MCS	SF-12 PCS
Baseline				
n (%)	231 (99)	231 (99)	216 (98)	216 (98)
Median (IQR)	46.9 (38.9–54.3)	49.3 (44.7–54.6)	46.7 (37.2–54.6)	51.0 (45.3–55.8)
Mean (SD)	45.9 (10.2)	48.9 (7.5)	45.6 (10.5)	50.0 (7.8)
Min, max	13.1, 68.6	25.9, 63.8	16.0, 70.8	23.2, 65.8
6 months				
n (%)	110 (47)	110 (47)	84 (38)	84 (38)
Median (IQR)	51.9 (44.5–55.3)	58.8 (52.7–58.4)	47.7 (38.1–54.8)	54.7 (48.8–57.6)
Mean (SD)	49.3 (9.4)	55.1 (5.9)	46.2 (10.8)	52.7 (7.5)
Min, max	13.8, 65.2	36.7, 67.1	15.2, 66.0	23.2, 66.6
12 months				
n (%)	122 (52)	122 (52)	117 (53)	117 (53)
Median (IQR)	55.0 (47.3–57.5)	57.1 (54.6–58.6)	49.7 (41.6–56.1)	54.7 (49.4–57.1)
Mean (SD)	52.3 (8.2)	56.3 (4.1)	48.5 (10.0)	52.8 (7.0)
Min, max	27.9, 62.5	41.5, 66.5	16.1, 65.9	26.0, 62.5
18 months				
n (%)	84 (36)	84 (36)	74 (34)	74 (34)
Median (IQR)	51.6 (45.0–57.4)	57.1 (53.5–58.9)	50.0 (40.0–57.3)	54.7 (49.0–57.4)
Mean (SD)	50.7 (8.2)	55.7 (6.5)	47.16 (11.56)	53.1 (6.1)
Min, max	25.6, 66.3	27.2, 66.4	13.8, 62.5	28.3, 61.7
24 months				
n (%)	100 (43)	100 (43)	100 (45)	100 (45)
Median (IQR)	55.2 (47.1–57.5)	57.1 (54.7–58.9)	50.3 (42.5–56.0)	57.1 (51.7–58.6)
Mean (SD)	52.3 (9.0)	56.0 (6.0)	48.6 (9.5)	54.7 (6.6)
Min, max	19.3, 66.1	27.1, 65.6	16.5, 62.4	23.2, 69.5

Max, maximum; min, minimum.

The box plots in [Figures 15](#) and [16](#) show the summary statistics graphically.

For the SF-12 MCSs in [Figure 15](#), the medians are higher than the benchmarked average for the UK population (red reference line) for the tonsillectomy arm in follow-up. This is better than for the conservative management arm, for which the medians are closer to the population mean.

For the SF-12 PCSs in [Figure 16](#), greater improvement compared with the UK population average can be seen in the tonsillectomy arm, for which the lower quartiles are well above the average. This implies that >75% of participants in the tonsillectomy arm experience sufficient improvement in physical symptoms to achieve physical QoL above the national average. Improvement is also seen in the conservative management arm, but not as widespread throughout all of the participants in that arm.

Summary

- The participants in this trial reported a median of 27 (IQR 12 to 52) sore throats over the full 24-month follow-up.
- When the primary outcome (total sore throats) is compared between the two randomised arms on an ITT basis, a 47% (95% CI 35% to 57%) reduction in sore throats is seen in the tonsillectomy arm. Sensitivity analyses on the ITT population analyses have little effect on the result.
- The analysis was repeated on the per-protocol population of 224 participants. This showed a stronger reduction on the number of sore throats of 58% (95% CI 45% to 69%).
- Around 25% of participants did not receive the treatment to which they were randomised, which meant that some opted not to receive a tonsillectomy and some opted to cross over to receive a tonsillectomy. There is evidence to suggest that those with larger numbers of sore throats following randomisation were more likely to either opt for or remain in the tonsillectomy arm. Conversely, those with slightly lower sore throat rates were more likely to remain in the conservative management arm or opt out of the tonsillectomy arm.
- Despite these crossovers, the ITT, per-protocol and per-treatment analyses all confirm that there is a significant reduction in total sore throats for those randomised to tonsillectomy.
- The benefits of tonsillectomy were also seen in the secondary outcome measures, which were collected at baseline and then every 6 months.
- TOI-14 scores improve in both arms, but show a greater improvement in the tonsillectomy arm than in the conservative management arm.
- Quality-of-life measures (SF-12 MCS and PCS) also show significant and beneficial differences in the tonsillectomy arm over time.
- There were 54 episodes of post-operative haemorrhage reported in 44 participants. This equates to 44 out of 231 participants undergoing tonsillectomy (19%). Of these participants, 37 were reported to have SAEs requiring re-admission: 9 reported as mild events, 22 as moderate events and 6 as severe event. No deaths were reported. Seventeen were recorded as AEs, for which participants did not attend hospital. All episodes of bleeding were managed conservatively with no returns to theatre.

Chapter 4 Economic evaluation

Introduction

This chapter reports the within-trial economic evaluation that was conducted to determine whether or not any clinical benefit associated with tonsillectomy was worthwhile for the NHS. The primary aim of the economic evaluation was to determine the relative efficiency of tonsillectomy compared with conservative management in adults suffering from recurring sore throats. Three different methodologies were used to estimate the relative efficiency of tonsillectomy: CEA, CUA and CBA. All analyses estimated costs using the same methodology but differed in how outcomes were measured. The CEA estimated the number of sore throat days, whereas the CUA estimated QALYs based on responses to the SF-12, and the CBA estimated benefits in monetary terms using the responses to a contingent valuation survey conducted as part of NATTINA. All economic analyses were based on an ITT principle (see [Chapter 2](#)).

The perspective of the trial was that of the NHS and PSS. The primary costs were healthcare resource use costs, that is the average total cost to the NHS and PSS to manage participants' recurrent sore throats during their 24-month follow-up. Sensitivity analyses took a broader perspective, in which the costs falling on participants were also considered. These costs included direct (e.g. out-of-pocket purchase of pain medication) and indirect (e.g. time off paid work/usual activities) costs.

The following outcomes were reported for the economic evaluation:

- healthcare costs of managing recurring sore throats to the NHS and PSS
- direct and indirect costs to participants
- total number of sore throats reported over 24 months
- QALYs estimated by the SF-6D derived from responses to the SF-12
- participants' WTP to avoid 1 sore throat day derived from response to the contingent valuation survey
- regression models to estimate the predictors of costs and effects, used to inform the calculation of incremental costs, effects and cost-effectiveness
- net monetary benefits (i.e. difference in mean benefit between the two arms minus the difference in mean cost between the two arms).

Methods

This analysis was designed and conducted to best practice, conforming to the Consolidated Health Economic Evaluation Reporting Standards.⁹⁶

Cost data collection

The costs collected as part of the economic evaluation were based on the cost of the interventions themselves and the use of health care and PSS over the 24-month follow-up period. All costs are reported in 2018/19 Great British pounds.

Healthcare service use

Tonsillectomy

The eCRF collected information on whether or not tonsillectomy was undertaken, the type of procedure that was undertaken, staff present, post-surgical medications and the length of admission. Staff present included the grade of the most senior operating surgeon and senior anaesthetist. The number, type and

grade of all other staff present [e.g. nurses and operating department practitioner (ODP)] were assumed based on clinical advice. The duration of procedure (minutes) was computed from the theatre record of times of entry into and exit from the theatre and recorded on the eCRF. The length of initial admission was estimated using admission and discharge dates recorded on the eCRF. Data on consumable (e.g. gloves, scrubs and gauze) and reusable (e.g. tonsil tray) resources were estimated based on the surgical technique of tonsillectomy and personal communication (Graham Stobbs, The Newcastle upon Tyne Hospitals NHS Foundation Trust, 22 September 2020).

Any discharge medications prescribed were costed based on the medication type recorded on the eCRF. Assumptions on duration and dose were based on clinical recommendations and the *British National Formulary*.⁹⁷

Information on AEs associated with surgical treatment was collected via telephone calls with the research nurse at 1 and 2 weeks post surgery. Treatments associated with AEs included drug therapies (prescribed or administered in hospital), hospital admissions, additional tests, consultations and further surgical treatment. The reported healthcare resources associated with an AE are presented as descriptive statistics (see [Appendix 17, Table 57](#)). Costs associated with AEs were not included in any analysis because it was assumed that these costs were captured in the participant questionnaires or the STAR/STARLET (see [Health service use over the trial follow-up](#) for an explanation of the STAR/STARLET).

Health service use over the trial follow-up

Follow-up use of NHS and PSS was collected via a self-completed participant questionnaire administered at 6-monthly intervals: baseline, at trial clinic visits at 12 and 24 months, and by post at 6 and 18 months post randomisation. The questionnaire captured information on use of primary care, secondary care and PSS. Primary care resource use referred to GP consultations, nurse consultations, speaking to a pharmacist and visits to walk-in clinics. The primary care consultations could take place at the general practice, the participant's home or over the telephone. Secondary care resource use referred to visits to an A&E department, outpatient clinic and hospital admissions either as a day case or overnight. Any PSS used were captured in an open-ended question asking participants to provide details on any additional care that they had received over the previous 6 months.

Information on primary and secondary care use was also collected in the STAR questionnaire, which was completed after the participant reported having a sore throat on the weekly STAR text-messaging service. The STAR questionnaire consisted of the following items:

1. severity of the sore throat
2. the TOI-14 throat-specific QoL tool
3. healthcare resource use (primary and secondary care)
4. SF-12
5. prescribed and bought medications
6. time away from work/usual activities associated with the sore throat episode.

Approximately 24 months into the trial, the STAR questionnaire was abbreviated to the STARLET questionnaire (see [Chapter 2](#) and [Appendix 1, Table 28](#) – Substantial amendment 7). STARLET omitted the TOI-14 and healthcare resource use items (i.e. retained items 1, 4, 5 and 6). Items 5 and 6 (medications and time away from work/usual activities) were condensed to binary (yes/no) questions in the STARLET. With the introduction of STARLET, the severity of the sore throat was now also captured by text message; therefore, costs and utility values could be assigned to participants based on the severity of their sore throat. Healthcare visits reported in the original STAR are included as descriptive statistics (see [Report Supplementary Material 1, Table S1](#)).

Information on prescribed medications reported in the STAR/STARLET questionnaires was included in the analysis because these data were not collected in the participant cost questionnaire. Over-the-counter medications were included only when the perspective widened to account for costs incurred by participants.

Sensitivity analysis considered healthcare resource use based on the information collected via GP linkage data. A comparison of resource use collected from participant questionnaires and the GP linkage allowed us to determine the feasibility of collecting GP linkage data and the reliability of the data collected in participant-reported questionnaires.

Participant costs

Direct and indirect participant costs were captured in the STAR/STARLET and the time and travel questionnaires. The STAR/STARLET questionnaire collected information on how much time participants spent away from work or usual activities as a result of their sore throat. The time and travel questionnaire, administered with the 18-month questionnaire, collected information on how participants travelled to each type of healthcare appointment (including out-of-pocket expenses), how much time they spent at each appointment, what they would have been doing if they did not attend that appointment and whether or not anybody attended their appointment with them. These data were collated to estimate an average unit time and travel cost for each healthcare appointment, which was combined with the reported use of each healthcare service.

Derivation of costs

Participant-level data on resource use were combined with unit costs and were presented in the following categories:

- tonsillectomy costs
- subsequent treatment costs
- participant costs.

Tonsillectomy costs

The cost of tonsillectomy was based on the national tariff for tonsillectomy (CA60A).⁹⁸ A sensitivity analysis estimated the cost of tonsillectomy on an individual-participant level (microcosting) (see [Tonsillectomy](#)). The unit costs associated with staffing, medications and hospital admissions were collected from routine sources (see [Report Supplementary Material 1, Table S2](#)).^{97,99–104} Unit costs for nursing staff, ODP and specialty registrars (e.g. ST5) were adjusted for employer contributions (national insurance and pension).^{105,106} Salary costs were converted into hourly rates and multiplied by the duration of the surgery (estimated from time in and out on the eCRF) to estimate the total staff costs per surgery. All participants were assumed to have a 30-minute pre-operative assessment with a band 5 nurse. The cost of consumable and reusable equipment was collected using personal communication from the Newcastle upon Tyne study centre (Graham Stobbs, 2020). Resources that were needed to perform the tonsillectomy included scrubs, gloves, silk ties, suction tubing, tonsil tray and anaesthetic consumables (see [Report Supplementary Material 1, Table S3](#)). Participants who had their tonsillectomy performed as a day case were assumed to stay on a ward for 6 hours, the cost of which was obtained from personal communication (Graham Stobbs, 2020). For any tonsillectomies that required hospital admission, the admission cost was obtained from routine sources¹⁰² and multiplied by the duration of the admission. It was assumed that the eCRFs would be well-completed, so mean imputation was used for participants with missing staff costs.

Subsequent treatment costs

Subsequent treatment costs were split into three categories: secondary care costs, primary care costs and medication costs.

Secondary care costs

Secondary care resource use refers to visits to an A&E department, outpatient visits and hospital admissions, either as a day case or overnight. Unit costs were collected from routine sources^{98,101} and multiplied by the number of contacts reported for each participant (see [Health service use over the trial follow-up](#)). The total cost of secondary care resource use (subsequent to the initial surgery) was estimated for every participant and summarised as the average total cost for each randomised arm.

Primary care costs

General practitioners or nurses could provide primary care consultations, and these were subcategorised depending on how the consultation was delivered: general practice, home visit, telephone or out-of-hours consultations. Unit costs were obtained from the unit costs of community care.¹⁰¹ The total cost of primary care resource use was estimated for every participant and summarised as the average total cost for each randomised arm.

Medication costs

If a participant experienced a sore throat episode, they could be treated with prescribed medications. If a participant reported receiving a prescription, they were asked to provide information on the name and dose of the prescribed medication(s) in the original STAR questionnaire. Unit costs associated with each prescribed medication were taken from the *British National Formulary*.⁹⁷ Daily prescription tariffs were estimated using these data and summarised using the self-reported severity of the sore throat episode to estimate an average daily prescription tariff for mild, moderate and severe sore throats (see [Report Supplementary Material 1, Table S2](#)). These prescription tariffs were applied to every sore throat day based on the self-reported severity of the sore throat episode and were multiplied by the duration of the sore throat episode to estimate the average total prescription cost per participant. It was originally anticipated that prescription costs would be applied to a sore throat day if participants only responded 'Yes' to receiving a prescribed medication question in the STAR/STARLET. However, as participants did not always complete the STAR/STARLET within 1 week of their reported sore throat episode, not all returned STAR/STARLET questionnaires could be matched to a sore throat episode; therefore, we could not clearly identify whether or not a participant was prescribed medication to manage their sore throat episode. This assumption reduced the heterogeneity in prescription costs between participants; however, this assumption allowed us to estimate variability in prescription costs based on clinical severity. The frequency of medications reported in the STAR/STARLET is provided in [Report Supplementary Material 1, Table S4](#).

Participant costs

Costs to participants and their main caregivers were considered in a sensitivity analysis. Participant costs were collected in the STAR/STARLET and time and travel questionnaires. Details of over-the-counter medications (i.e. name, dose and frequency) that were taken for a sore throat were provided in the STAR questionnaire. Using a methodology similar to that described for prescribed medications in the preceding subsection, these medications were costed using online sources¹⁰¹ and were combined to generate a daily medication tariff for each severity of sore throat reported (mild, moderate and severe) (see [Report Supplementary Material 1, Table S2](#)). For each sore throat reported, the unit cost derived from the STAR data was assigned based on the severity of their sore throat and the duration was assumed to be the same number of sore throat days reported by the participant. The frequency of medications reported in the STAR/STARLET is provided in [Report Supplementary Material 1, Table S4](#).

Travel costs, collected from the time and travel questionnaire, were summarised as unit costs and assigned to each healthcare contact (see [Report Supplementary Material 1, Table S2](#)). Time costs were estimated using wage rates and leisure time costs collected from the Department of Transport¹⁰⁷ and Office for National Statistics.¹⁰⁸ These unit costs were combined with reported time off paid work and/or usual activities owing to a sore throat day, which were reported in the STAR/STARLET (see [Appendix 17, Table 58](#)), and/or attendance at healthcare appointments, which was reported in the time and travel questionnaire. Time away from paid work and/or usual activities post tonsillectomy were also considered given that tonsillectomy is a painful procedure.³³ The average time someone would be

expected to be away from work after a tonsillectomy was 14 days, and it was assumed that these time costs would be captured in the STAR/STARLET.^{34,35} The total direct and indirect cost was estimated for each participant and summarised as the average total participant cost for each randomised arm.

General practitioner linkage

Healthcare resource use was also captured in the GP linkage case report form. Initial consent included access to participants' GP health records to assess healthcare use during each participant's 24-month follow-up and also 12 months prior to randomisation. The following data were collected on the GP linkage case report form:

- AEs
- attendance at general practice/walk-in clinic/A&E for sore throat or related event
- hospitalisations and emergency referrals
- prescriptions issued
- any additional interventions required.

The GP linkage data were used to estimate the total healthcare resource use for each participant for the duration of the follow-up as a secondary analysis. Unit costs were collected from routine sources, as described above. Prescription costs were not considered in this analysis because the data from the STAR/STARLET were used to inform prescription costs. The frequency of medications reported in the GP linkage case report form is provided in [Report Supplementary Material 1, Table S5](#).

Effectiveness measures

Three effectiveness measures were used in this economic evaluation, one for each of the analyses: CEA – number of sore throat days; CUA – QALYs; and CBA – WTP to avoid a sore throat day.

Estimation of health outcomes for the cost-effectiveness analysis

The effectiveness measure in the CEA, number of sore throat days, was equivalent to the primary outcome, derived from the STAR text message (see [Chapter 2](#)), giving the mean total number of sore throat days estimated for each arm on an ITT basis.

Estimation of quality-adjusted life-years for the cost-utility analysis

The SF-12 was completed at scheduled (baseline and at 6, 12, 18 and 24 months post randomisation) and unscheduled time points as part of the STAR/STARLET when participants reported a sore throat. SF-12 responses were mapped onto the SF-6D, a preference-based utility index, using a standard algorithm to produce a health state utility score.⁸⁶ The SF-6D is made up of six multilevel dimensions: physical functioning, role limitations, social functioning, bodily pain, mental health and vitality. The area-under-the-curve approach puts a time weight onto each utility score. The time-weighted average of the scores based on the responses to the SF-12 throughout the follow-up period allows us to generate QALY values for each participant.¹⁰⁹ One QALY is equivalent to 1 year in perfect health.

All information on QALYs is presented as the average utility value at each scheduled time point and the average total QALYs for each randomised arm. The primary analysis estimated QALYs based on the utility values estimated at each data collection time point ([Equation 1](#), QALY equation):

$$\text{QALY}^{\beta} = \left[\left(\frac{\text{SF-6D}_{\text{bl}} + \text{SF-6D}_{6\text{mths}}}{2} \right) \times \left(\frac{6}{12} \right) \right] + \left[\left(\frac{\text{SF-6D}_{6\text{mths}} + \text{SF-6D}_{12\text{mths}}}{2} \right) \times \left(\frac{6}{12} \right) \right] \\ + \left[\left(\frac{\text{SF-6D}_{12\text{mths}} + \text{SF-6D}_{18\text{mths}}}{2} \right) \times \left(\frac{6}{12} \right) \right] + \left[\left(\frac{\text{SF-6D}_{18\text{mths}} + \text{SF-6D}_{24\text{mths}}}{2} \right) \times \left(\frac{6}{12} \right) \right], \quad (1)$$

where QALY^β is quality-adjusted life year, bl is baseline and mths is months.

The sensitivity analysis incorporated the utility scores associated with sore throats into the QALY equation ([Equation 2](#), example QALY equation with one sore throat between baseline and 6 months). The utility scores associated with each sore throat were assigned based on the severity of the sore throat (see [Appendix 17](#), [Table 58](#)) and multiplied by the self-reported number of sore throat days, both collected via text message:

$$\text{QALY}^{\beta} = \left[\left(\frac{\text{SF-6D}_{\text{bl}} + \text{SF-6D}_{6\text{mths}}}{2} \right) \times \left(\frac{182.5 - \#ST\text{days}}{365} \right) + \left[\left(\text{SF-6D}_{\text{ST}} \right) \times \left(\frac{\#ST\text{days}}{365} \right) \right] \right] \\ + \left[\left(\frac{\text{SF-6D}_{6\text{mths}} + \text{SF-6D}_{12\text{mths}}}{2} \right) \times \left(\frac{6}{12} \right) \right] + \left[\left(\frac{\text{SF-6D}_{12\text{mths}} + \text{SF-6D}_{18\text{mths}}}{2} \right) \times \left(\frac{6}{12} \right) \right] \\ + \left[\left(\frac{\text{SF-6D}_{18\text{mths}} + \text{SF-6D}_{24\text{mths}}}{2} \right) \times \left(\frac{6}{12} \right) \right], \quad (2)$$

where QALY^β is quality-adjusted life-year, bl is baseline, mths is months, # is number of and ST is sore throat.

Estimation of willingness to pay for the cost-benefit analysis

As with all economic evaluations, there are limitations associated with the effectiveness measures chosen. In this instance, cost per number of sore throat days avoided may be difficult for policy-makers to interpret and QALYs may not fully capture individuals' preferences to avoid sore throats. An alternative technique used was the contingent valuation method, which allowed respondents to state their preferences, in monetary terms,¹¹⁰ to avoid a sore throat day at baseline. The contingent valuation questionnaire collected individuals' expression, for a given level of income, of their WTP to avoid a sore throat day, with higher monetary values indicating that they would derive greater benefit. The questionnaire took the form of a payment card.¹¹⁰⁻¹¹² Participants were asked 'Please select below the maximum amount you are sure you would be willing to pay to avoid 1 sore throat day'. The range of the payments was £0 to £30, with participants who were willing to pay £30 asked to provide their maximum WTP value. Further details on how WTP values were estimated are provided in [Report Supplementary Material 2](#).

The WTP results were analysed for the following reasons:

- Estimate the average WTP value associated with avoiding a sore throat day in each randomised arm. We estimated the maximum WTP value for each participant, totalled the maximum WTP values for both management strategies and divided them by the number of participants in each arm. We also explored how valuations might vary according to participant characteristics (e.g. family income, sex and age) using a regression analysis.
- Quantify in monetary terms the effect of a sore throat day on participants' health. We multiplied the above estimated maximum WTP values for each participant with their total number of reported sore throat days. These values were averaged for each randomised arm, giving an estimated average total reduction in participants' health in monetary terms per randomised arm.

Comparative incremental analyses of costs and effects

Both unadjusted and adjusted analyses were performed to estimate the cost-effectiveness of tonsillectomy compared with conservative management. All results were presented as unadjusted point estimates of mean total costs and effects, and adjusted point estimates of the mean incremental costs, effects and cost-effectiveness. Both costs and effects incurred during the second year of follow-up were discounted at the UK recommended rate of 3.5%.¹¹³ If an arm was found to be more costly and more effective, then it was considered to be the dominant strategy and, hence, cost-effective. If tonsillectomy was found to be, on average, more effective but more costly, then consideration had to be made as to whether or not the extra costs are worth the extra effectiveness. In this situation, decisions were based on the following:

- The incremental cost-effectiveness ratio (ICER). This ratio equates to the difference in average total costs divided by the difference in average total effects [e.g. £20 per sore throat day avoided could

equate to £20 (difference in average total costs)/1 day (difference in average total number of sore throat days) = £20 (ICER)]

- Society's WTP threshold for an additional unit of effect.

The same cost estimates were used in all three analyses, but the analyses differed in terms of outcome measures.

Cost-effectiveness analysis

The CEA was based on the incremental cost per sore throat day avoided. The average total cost and average total number of sore throat days were estimated for each arm, as point estimates of the mean incremental costs and sore throat days, and the incremental cost per sore throat day avoided.

Cost-utility analysis

The CUA was based on the incremental cost per QALY gained. The average total cost and average total QALYs were estimated for each arm and presented as point estimates of the mean incremental costs and QALYs and the incremental cost per QALY gained. The incremental cost per QALY has to be considered against what society is willing to pay for a QALY. The precise value for this is unknown but, in England, NICE typically conclude that when the incremental cost per QALY is <£20,000 an intervention is efficient.¹¹³

Cost-benefit analysis

The CBA estimated the average total cost and average total reduction in QoL (based on the number of reported sore throat days and WTP to avoid a sore throat day) per arm. Both costs and benefits were expressed in commensurate monetary units (Great British pounds), which enabled comparisons to be made between strategies.¹¹⁴ The net benefit of tonsillectomy was derived by estimating the difference in benefits minus the difference in costs. The decision rule for CBA is, therefore, relatively simple compared with the other two analyses. If the net benefit value is positive, this represents a gain in welfare and tonsillectomy is the preferred management strategy. If the net benefit value is negative, the potential benefits associated with tonsillectomy do not outweigh the additional costs and conservative management would be the preferred management strategy.¹¹⁴

Adjusted analysis: seemingly unrelated regression

An adjusted analysis was used for all analyses to estimate the point estimates of the mean incremental costs, effects and cost-effectiveness using seemingly unrelated regression (SUR).¹¹⁵ Costs and effects were estimated simultaneously using SUR, which accounts for possible correlations between costs and effects when using individual participant-level data.¹¹⁶ Additional covariates were also controlled for using SUR (e.g. age, randomisation arm, baseline severity/utility and income).

Sensitivity analysis

In the base-case analysis, missing cost and utility data were imputed using chained multiple imputation methods,¹¹⁷ which assumed data were missing at random. Differences in baseline utility and costs between those missing data and those with complete data were undertaken to test this assumption. Sensitivity analyses were conducted to assess the robustness of the results to realistic variations in the levels of underlying data.

The following sensitivity analyses were undertaken to determine what effect, if any, changing the assumptions in the base-case analysis had on the overall conclusions:

- Surgery costs were estimated using microcosting.
- Participant costs were considered.
- Healthcare costs were estimated using GP linkage data.

- Utility data associated with sore throat episodes collected in the STAR/STARLET data were incorporated into the QALY equation.
- Healthcare costs were estimated for participants with cost data at least at one time point post randomisation and estimated effects were based on the data available.

Deterministic sensitivity analyses were used to address any uncertainty in the assumptions used in our base-case analysis, including using the microcosting tonsillectomy cost as the intervention cost and the difference in healthcare costs when the GP linkage data were used.

Stochastic sensitivity analyses, using the bootstrapping technique,¹¹⁸ explored the impact of the statistical imprecision surrounding estimates of costs, effects and cost-effectiveness. In essence, the process entails extracting and estimating the difference in costs and effects from a paired sample (one participant from each arm) from the trial data set. Those are then returned to the data set and the process is repeated 1000 times. The bootstrapped results from all three analyses were presented on a cost-effectiveness plane. The cost-effectiveness plane was used to illustrate the distribution of incremental costs and incremental effects from which we can clearly identify the spread of uncertainty in our results.¹¹⁹

The bootstrapped results from the CEA and CUA were presented as cost-effectiveness acceptability curves (CEACs). CEACs determine the treatment option that maximises net benefits at each WTP value for an additional unit of health effect (i.e. a reduction in sore throat days and QALYs).¹²⁰ For the CEA, the threshold from which to interpret these results can be informed by the values obtained from the WTP exercise. A similar approach to the CEAC is not required for the CBA. For the CBA, the intervention with the greatest net monetary benefit is considered efficient. Therefore, for the stochastic analysis, the results were presented in terms of the probability that the intervention was associated with the greatest net monetary benefit.

Results

Data validity and completeness

[Table 20](#) presents the response rates for the participant-completed data collection tools used to inform the economic analysis. Predictably, the baseline questionnaires, which included the contingent valuation survey, were comprehensive, yielding primary outcome data for nearly 95% of participants. Attrition over time peaked for postal questionnaires. Only 20% of participants had complete data on health utilisation and the SF-12. The pattern of non-response was similar across both arms of the trial and at all time points. There was no difference in baseline costs ($p = 0.67$) or baseline utilities ($p = 0.50$) between those with data and those with no data available over the 24-month follow-up.

Resource use and costs

Over the 24-month follow-up period, participants reported a number of contacts with different healthcare professionals (see [Appendix 17, Table 59](#)). The most common visits were to GPs and pharmacists. On average, participants in the conservative management arm reported using more healthcare services than participants in the tonsillectomy arm, except at 6 months. The difference at 6 months appears to be a result of more secondary care visits (outpatient and inpatient), which could be participants reporting their surgery. The majority of tonsillectomy surgeries were performed as day procedures by either a consultant or a specialist registrar (see [Appendix 17, Table 60](#)).

The unit costs used in the analysis are provided in [Report Supplementary Material 1, Table S2](#). [Table 21](#) presents the average total costs per cost category and per randomised arm. Participants

TABLE 20 Response rates for data collection tools used to inform the economic analysis

Data response rates	Randomised arm, n (%)	
	Conservative management (N = 220)	Tonsillectomy (N = 233)
Participant questionnaire: health utilisation questions		
Baseline	215 (98)	231 (99)
6 months	85 (39)	110 (47)
12 months	117 (53)	122 (52)
18 months	74 (34)	87 (37)
24 months	100 (45)	99 (42)
Complete data at all time points	47 (21)	49 (21)
Primary outcome data		
Number of sore throat days	205 (93)	224 (96)
Participant questionnaire: SF-12		
Baseline	215 (98)	229 (98)
6 months	84 (38)	108 (46)
12 months	116 (53)	122 (52)
18 months	74 (34)	83 (36)
24 months	99 (45)	100 (43)
Complete data at all time points	44 (20)	46 (20)
Contingent valuation^a		
Completed	208 (95)	226 (97)
a Administered at baseline.		
Notes		
N = total number of participants in each randomised arm; n = number of participants contributing data; shaded rows represent the postal questionnaires.		

in both arms reported similar average baseline healthcare resource use. The average total cost of treatment (tonsillectomy and post-operative medications) was higher in the tonsillectomy arm than in the conservative management arm, which was expected given that 74% of participants in this arm received a tonsillectomy compared with 27% of participants in the conservative management arm. Total healthcare costs could be estimated for 308 participants as they responded to at least one health utilisation questionnaire post randomisation. On average, and as would be expected given the data reported in [Appendix 17, Table 59](#), those in the conservative management arm reported higher primary and secondary care costs than those in the tonsillectomy arm over the total 24-month follow-up period. After the total costs were combined to estimate the average total cost to the NHS, on average, tonsillectomy was more costly than conservative management. This result was consistent when missing cost data were imputed using the assumption that data were missing at random.

Data from the STAR/STARLET used to inform direct and indirect costs incurred by participants are summarised in [Appendix 17, Table 58](#), and [Report Supplementary Material 1, Tables S1–S4 and S6](#). These daily tariffs were combined with each reported sore throat episode based on the severity of

TABLE 21 Average total costs for each cost category by randomised arm (ITT)

Cost category	Randomised arm			
	Conservative management (N = 220)		Tonsillectomy (N = 233)	
	Costs (£), mean (SD)	n	Costs (£), mean (SD)	n
Healthcare resources: baseline	348 (325)	215	378 (340)	231
Tonsillectomy ^a	400 (662)	220	1101 (657)	233
Postoperative medications	2 (3)	220	4 (4)	233
Healthcare resources: 6 months	76 (184)	147	122 (260)	161
Healthcare resources: 12 months	96 (227)	147	11 (33)	161
Healthcare resources: 18 months	65 (168)	147	12 (43)	161
Healthcare resources: 24 months	45 (162)	147	21 (119)	161
Total healthcare resource costs (6 + 12 + 18 + 24 months)	282 (397)	147	166 (296)	161
Prescription medications owing to a sore throat	20 (32)	205	12 (20)	224
Total NHS costs	803 (880)	147	1473 (624)	161
Total NHS costs: base-case analysis using multiple imputation	879 (784)	217	1365 (714)	231
Participant costs				
Time and travel costs to healthcare appointments	71 (129)	220	41 (89)	233
Time away from usual activities owing to sore throat	1780 (2170)	205	1148 (1293)	224
Time away from paid work owing to sore throat	2277 (2948)	205	1459 (1855)	224
Over-the-counter medications	90 (140)	205	55 (63)	224
Total participant costs	3936 (5186)	220	2600 (3205)	233
Total NHS (using multiple imputation) and participant costs	4803 (5394)	220	3954 (3418)	233
a Number of participants who received a tonsillectomy (conservative management arm, n = 59; tonsillectomy arm, n = 172).				
Note n = number of participants contributing data.				

the episode and multiplied by the duration of the sore throat episode. When costs to participants were considered, conservative management was, on average, more costly than tonsillectomy {mean difference £889 [standard error (SE) £432], 95% CI £40 to £1738}. The difference in the average total costs was mainly driven by the indirect costs associated with sore throat episodes (see [Appendix 17, Table 58](#)).

Effectiveness outcomes

Number of sore throat days

On average, those in the tonsillectomy arm experienced fewer sore throat days than those in the conservative management arm [mean difference 21.27 (SE 6.24) days, 95% CI 9 to 34 days] ([Table 22](#)).

Severity of sore throat days

As previously mentioned, self-report severity of sore throat days was used to assign (1) medication costs, (2) participant time costs away from usual activities and (3) utility values. Participants provided

TABLE 22 Summaries of the outcome measures used in the CEA and CUA by randomised arm

Outcome measure	Randomised arm			
	Conservative management (N = 220)		Tonsillectomy (N = 233)	
	Mean (SD)	n	Mean (SD)	n
Number of sore throat days ^a	55.42 (84.20)	205	34.15 (38.65)	224
SF-12				
Baseline	0.687 (0.13)	215	0.680 (0.13)	229
6 months	0.730 (0.13)	84	0.779 (0.13)	108
12 months	0.751 (0.15)	116	0.830 (0.12)	122
18 months ^a	0.721 (0.14)	74	0.772 (0.12)	83
24 months ^a	0.741 (0.14)	99	0.807 (0.12)	100
QALYs ^a	1.517 (0.23)	44	1.577 (0.19)	46
QALYs: base-case analysis using multiple imputation ^a	1.444 (0.17)	215	1.551 (0.15)	229
QALYs: including sore throat utility decrements ^a	1.378 (0.18)	215	1.495 (0.18)	229

a Outcomes reported during the second year of follow-up are discounted at 3.5%.

Note
n = number of participants contributing data.

the severity of their sore throat days via text message for nearly 60% of all self-reported sore throat days ($n = 3150$: conservative management arm, $n = 1859$; tonsillectomy arm, $n = 1291$). Of the total number of sore throats reported by participants in the STAR and STARLET, those randomised to receive tonsillectomy reported, on average, a higher proportion of their sore throats to be mild than those randomised to receive conservative management (tonsillectomy: mild, 0.39; conservative management: mild, 0.32).

Data from the STAR/STARLET were used to inform missing severity data [$n = 3563$ (67%) sore throat days with severity information], after which severity was inferred based on the duration of the sore throat episode.

[Appendix 17, Table 58](#), provides further details on the number of sore throat days, the number of medications taken (over the counter and prescribed), time away from work and usual activities, and utility values associated with sore throat episodes by severity.

Quality-adjusted life-years

On average, participants in both arms of the trial reported their health status to be less than full health throughout the trial follow-up period (see [Table 22](#)). At baseline, on average, utility scores were similar between the two randomised arms; however, over time, those in the tonsillectomy arm were more likely than those in the conservative management arm to report higher utility scores. Similarly, when these utility values were incorporated into the QALY equation, on average, those in the tonsillectomy arm reported higher QALYs than those in the conservative management arm [mean difference 0.06 (SE 0.04); 95% CI -0.03 to 0.15]. This result was consistent when missing utility data were imputed [mean difference 0.11 (SE 0.02), 95% CI 0.08 to 0.14]. The inclusion of the utility decrements associated with each sore throat episode resulted in both randomised arms reporting lower average total QALYs; however, the difference in QALYs between the two arms remained similar [mean difference 0.12 (SE 0.02), 95% CI 0.08 to 0.15]. The difference in average total QALYs was statistically significant in the base-case analysis (multiple

imputation) and when utility decrements associated with sore throat days were considered (see [Appendix 17, Table 61](#)).

Willingness to pay

On average, participants were willing to pay £43 (95% CI £2 to £100) to avoid 1 sore throat day. There was a slight difference in the mean WTP values reported, with those in the tonsillectomy arm willing to pay more than those in the conservative management arm to avoid a sore throat day. However, the CI for the mean difference included £0, and was wide enough to include potentially important differences favouring either management [conservative management: mean £41 (SD £102)] or tonsillectomy [mean £45 (SD £122)] [mean difference £4 (SE £11), 95% CI -£17 to £25]. The minimum (£0) and maximum (£999.99) reported WTP values were the same for both arms.

Economic evaluation

Incremental cost-effectiveness analysis

[Table 23](#) presents the unadjusted average total cost and average total sore throat days per randomised arm and the adjusted incremental costs and incremental QALYs that were used to estimate the ICER. Tonsillectomy was, on average, more costly and more effective (in terms of number of sore throat days avoided) than conservative management (see [Table 23](#)). For the adjusted analysis, the ICER (incremental cost per sore throat day avoided) was £24 (see [Table 23](#)). This value is difficult for decision-makers to interpret given that there is no current threshold available with which it can be compared. It was for this reason that a formal valuation exercise using contingent valuation was conducted. The results of this are shown in this section.

The deterministic results presented below also do not reflect the imprecision in estimates of costs and sore throat days. For this we need the stochastic sensitivity results, which consider the impression surrounding costs, effects and cost-effectiveness ([Figures 17 and 18](#)). [Figure 17](#) illustrates the 1000 iterations of the CEA bootstrapped results for the adjusted analysis. In 100% of these iterations, tonsillectomy is more costly and more effective than conservative management. As [Figure 18](#) shows,

TABLE 23 Cost-effectiveness analysis results: outcome – sore throat days, results (multiple imputation)

Investigation strategy	Cost (£) (95% CI) ^a	Incremental cost (£) (95% CI) ^b	Effect (95% CI) ^a	Incremental effect (95% CI) ^b	ICER (£)	Probability that tonsillectomy is cost-effective for different threshold values for society's willingness to pay to avoid a sore throat day				
						£0	£20	£30	£50	£100
Conservative management (costs, n = 217; outcomes, n = 205)	879 (774 to 984)		55.42 (44 to 67)			1.00	0.72	0.29	0.06	0.01
Tonsillectomy (costs, n = 231; outcomes, n = 224)	1365 (1273 to 1458)	503 (362 to 664)	34.15 (29 to 39)	-21.01 (-33 to -9)	24	0.00	0.28	0.71	0.94	0.99

a Point estimates are based on the unadjusted analysis (costs n = 448, sore throat days n = 429).

b Incremental estimates are based on the adjusted analysis (n = 425).

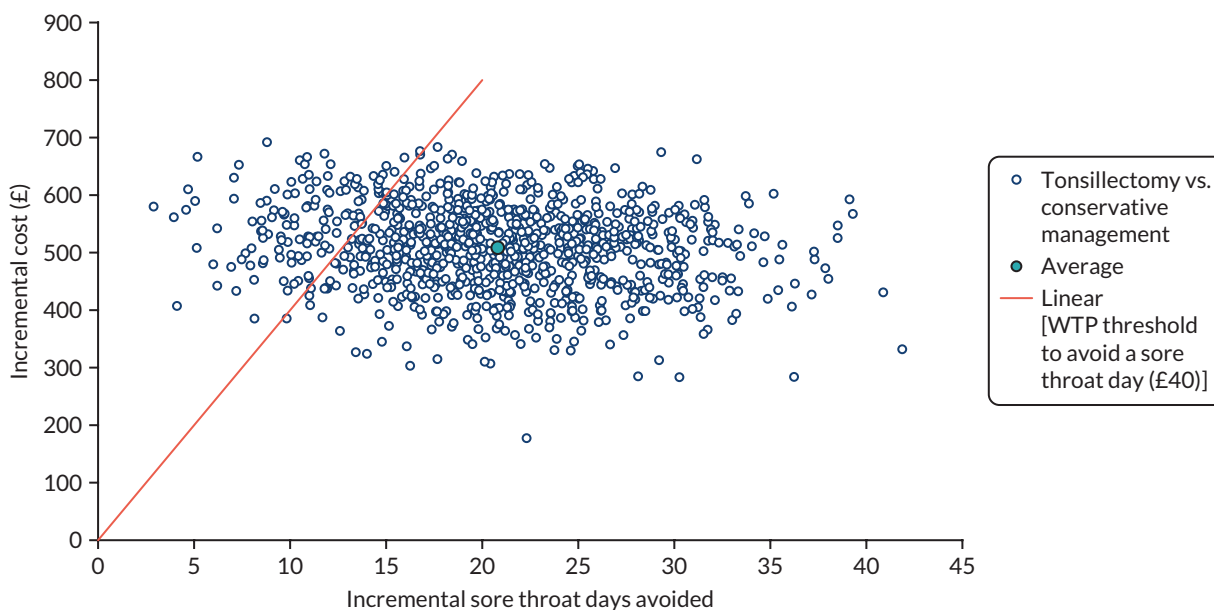


FIGURE 17 Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CEA multiple imputation results.

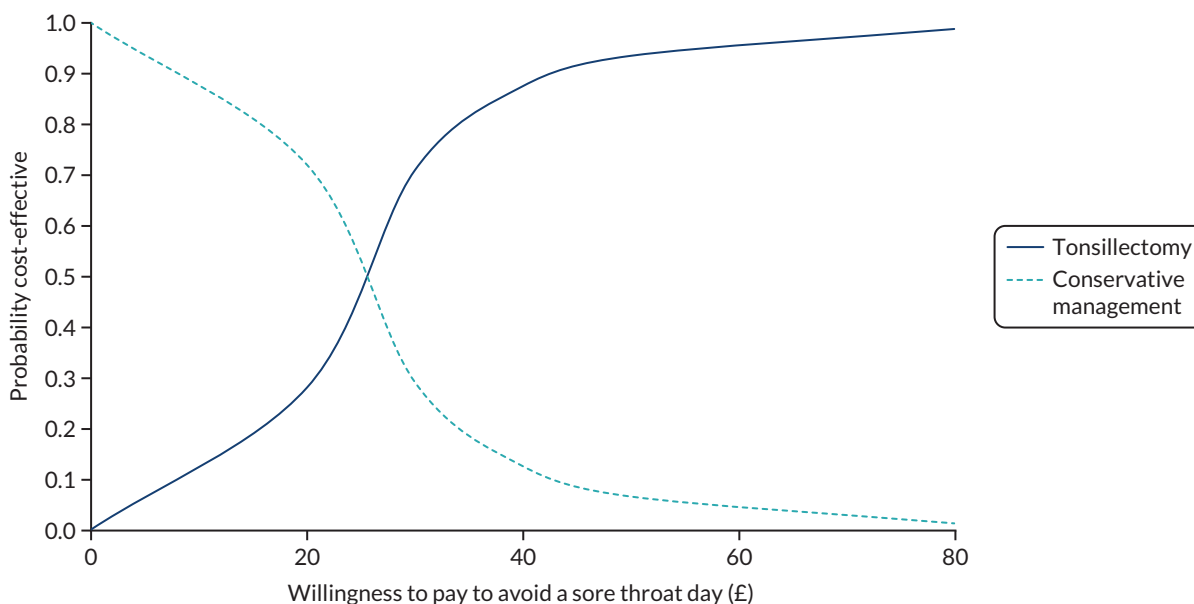


FIGURE 18 Cost-effectiveness acceptability curves for tonsillectomy vs. conservative management using the adjusted bootstrapped CEA multiple imputation results.

at a WTP threshold of £25 to avoid 1 sore throat day, tonsillectomy begins to have the higher probability of being considered cost-effective. This probability increases as the WTP threshold to avoid a sore throat day increases.

The results (see [Table 23](#), and [Figures 17](#) and [18](#)) also show that if decision-makers used the WTP value of £40 to avoid 1 sore throat day, which was derived from the contingent valuation question, tonsillectomy has 87% probability of being considered cost-effective.

TABLE 24 Cost-utility analysis results: outcome – QALYs, results (multiple imputation)

Investigation strategy	Cost (£) (95% CI) ^a	Incremental cost (£) (95% CI) ^b	Effect (95% CI) ^a	Incremental effect (95% CI) ^b	ICER (£)	Probability that tonsillectomy is cost-effective for different threshold values for society's willingness to pay for an additional QALY				
						£0	£10,000	£20,000	£30,000	£50,000
Conservative management (costs, n = 217; outcomes, n = 215)	879 (774 to 984)		1.444 (1.42 to 1.46)			1.00	0.00	0.00	0.00	0.00
Tonsillectomy (costs, n = 231; outcomes, n = 229)	1365 (1273 to 1458)	488 (349 to 626)	1.551 (1.53 to 1.57)	0.118 (0.09 to 0.14)	4136	0.00	1.00	1.00	1.00	1.00

a Point estimates are based on the unadjusted analysis (costs n = 448, QALYs n = 444).

b Incremental estimates are based on the adjusted analysis (n = 444).

Incremental cost-utility analysis

[Table 24](#) presents the unadjusted average total cost, the average total QALY per randomised arm and the adjusted incremental costs and incremental QALYs that were used to estimate the ICER. On average, tonsillectomy is more costly and more effective (in terms of QALYs gained) than conservative management. The incremental cost per QALY gained was over £4000. Tonsillectomy has a 100% probability of being considered cost-effective at a threshold value of £20,000 for an additional QALY (see [Table 24](#)).

The results of the stochastic sensitivity analysis, illustrating the uncertainty surrounding our estimates of cost, effects and cost-effectiveness, are presented in [Appendix 17, Figures 30 and 31](#). The majority of bootstrapped iterations show that tonsillectomy is, on average, more costly and more effective than conservative management.¹¹³

Incremental cost-benefit analysis

The CBA results are presented in [Table 25](#). In this analysis, the monetary measure of effect is negative because it quantifies the reduction in health experienced by participants in both randomised arms owing

TABLE 25 Cost-benefit analysis results: outcome – willingness to pay and number of sore throat days: results (multiple imputation)

Investigation strategy	Cost (£) (95% CI) ^a	Incremental cost (£) (95% CI) ^b	Effect (£) (95% CI) ^a	Incremental effect (£) (95% CI) ^b	Incremental net benefit (£)	Probability that tonsillectomy has a higher net benefit
Conservative management (costs, n = 217; outcomes, n = 196)	879 (774 to 984)		-1882 (-2474 to -1289)			
Tonsillectomy (costs, n = 231; outcomes, n = 217)	1365 (1273 to 1458)	516 (370 to 663)	-1218 (-1629 to -808)	697 (-24 to 1419)	181	0.69

a Point estimates are based on the unadjusted analysis (costs, n = 448; sore throat days and provided a WTP value in the contingent valuation study, n = 413).

b Incremental estimates are based on the adjusted analysis (n = 401).

to the number of sore throat days they reported. It is presented as a negative because sore throat days are undesirable (a 'dis-benefit' in economic parlance) and the value represents the total loss of benefits caused by them. As tonsillectomy reduces this number relative to conservative management, the incremental gain from tonsillectomy is a positive value. These monetary values of the (dis)benefits are the value that participants would be willing to pay to avoid a sore throat day [conservative management: mean £40.77 (SD £102); tonsillectomy: mean £44.68 (SD £122)] multiplied by the number of sore throat days reported by participants [conservative management: mean 55.19 days (SD 84 days); tonsillectomy: mean 34.14 days (39 days)].

Similar to the CEA and CUA results, tonsillectomy was, on average, more costly and more effective than conservative management (as there were fewer sore throat days). The incremental effect shown in [Table 25](#) shows that, on average, the tonsillectomy arm was better off than the conservative management arm in terms of avoiding 1 sore throat day. In the unadjusted analysis, tonsillectomy would be more efficient because it has a positive net benefit (i.e. the difference in benefits is greater than the difference in costs). In the adjusted analysis (see [Table 25](#); see [Appendix 17, Figure 32](#)), the probability of tonsillectomy having a higher net benefit was 69% (there is a 69% chance that the benefits of tonsillectomy would be worth the additional costs compared with conservative management).

Sensitivity analysis

Microcosting the intervention

The CUA results, which estimate the cost of the intervention (tonsillectomy) using microcosting, are presented in [Appendix 17, Table 62](#), and [Figures 33](#) and [34](#). The average cost of tonsillectomy (£1327) was slightly less than the NHS tariff for tonsillectomy (£1492) when surgery costs were estimated using microcosting. Our overall conclusions did not change with this analysis because tonsillectomy was still, on average, more costly and more effective, with an ICER of £3534. Similar to the base-case analysis, the probability of tonsillectomy being considered cost-effective increased as the value that we placed on an additional QALY increased. At a £20,000 threshold for an additional QALY, tonsillectomy had a 100% probability of being considered cost-effective.

Participant costs

The CUA results, which incorporate participant costs with healthcare costs (NHS and PSS), are presented in [Appendix 17, Table 63](#), and [Figures 35](#) and [36](#). In this analysis, tonsillectomy dominated conservative management because it was less costly and more effective in terms of QALYs gained. Tonsillectomy had a 98% probability of being considered cost-effective if society was not willing to pay for a QALY and this probability increased to 100% as the threshold value for a QALY increased.

General practitioner linkage

The CUA results estimating costs using GP linkage data are presented in [Appendix 17, Table 64](#), and [Figures 37](#) and [38](#). Data on healthcare resource use were provided by GPs for 197 participants. GPs were asked to provide data on healthcare resource use for the previous 3 years (1 year pre randomisation and 2 years post randomisation); there were data available for the full 3 years requested for 77% of participants ($n = 151$): 19% ($n = 37$) had pre-randomisation data only and 4% ($n = 9$) had post-randomisation data only. On average, the total costs estimated for both arms were lower than when healthcare costs were estimated using participant-reported data. On average, tonsillectomy was more costly and more effective than conservative management, with an ICER of £5300 for an additional QALY. As our threshold value for an additional QALY increased, so did the probability of tonsillectomy being considered cost-effective. At a threshold value of £20,000 for an additional QALY, tonsillectomy had a 100% probability of being considered cost-effective.

Quality-adjusted life-years estimated with utility decrements associated with sore throats

The CUA results estimating QALYs, which include the utility decrement associated with each sore throat episode (see *Equation 2*), are presented in [Appendix 17](#), [Table 61](#), and [Figures 39](#) and [40](#). On average, tonsillectomy was more costly and more effective than conservative management in terms of QALYs gained, with an ICER of £4000. Similar to the base-case analysis, if society was not willing to pay for an additional QALY, tonsillectomy would have a 0% probability of being considered cost-effective, but, as the threshold value increased, the probability of tonsillectomy being considered cost-effective increased to 100%.

Estimating average total costs using available data only

Sensitivity analyses estimating costs for those participants with at least one healthcare utilisation questionnaire completed and effects with only those with complete data are presented in [Report Supplementary Material 1](#), [Tables S7–S9](#) and [Figures S1–S5](#). In all analyses, our conclusions remain the same as the base-case analyses, although there is slightly more uncertainty in our results based on the spread of the iterations on the cost-effectiveness plane for all three analyses (CEA, CUA and CBA).

Chapter 5 Qualitative study

Introduction

NATTINA was concerned with establishing the clinical effectiveness and cost-effectiveness of tonsillectomy in adults, which is one of the more commonly performed routine surgical procedures in the UK.² Primary health care in the UK now restricts referrals for treatments that are deemed to be of limited clinical value, with tonsillectomy ranked top as a 'relatively ineffective' procedure.¹²¹ In the main NATTINA trial, participants were randomly allocated into immediate or conservative management (i.e. deferred surgery). Our experience of a randomised trial of tonsillectomy in children^{3,64} together with other published ENT surgical trials²² highlighted the problem of retaining participants in a non-surgical cohort. These findings, along with patient and public engagement, influenced our trial design and decision to use deferred surgery as the conservative management option rather than no surgery.

Although our feasibility study suggested that ENT staff and GPs were willing to randomise patients to NATTINA, not all ENT staff were in equipoise concerning the treatment pathways. In our recent qualitative work, recurrent sore throats were reported to severely impact participants' family, work and social life. ENT staff stated that patients were facing increasing barriers to accessing secondary care services. GPs also reported being under pressure to reduce 'limited clinical value' surgical procedures.⁸¹ In addition, participants reported reluctance to be randomised into the conservative management arm if they had already waited a substantial time before being referred.⁸²

Therefore, it was essential to include a process evaluation as part of the trial design. The emphasis of a process evaluation is on providing greater confidence in conclusions by assessing what was delivered and how it was delivered, and to assess the generalisability of findings by understanding the role of context.¹²² Therefore, the aim of the process evaluation of the pilot and main trial was to examine acceptability of the trial, treatments and unforeseen consequences from the perspective of participants and stakeholders, including ENT staff and GPs.

Methods

Design

An embedded, qualitative study using semistructured interviews.

Setting and sample

The sample for the qualitative process evaluation consisted of recruiting otolaryngology staff, GPs who had patients taking part in the NATTINA trial and ENT patients, including both those recruited to the NATTINA trial and those who declined to participate. Sampling was purposive, seeking maximum variation in terms of age, sex, phase of trial (pilot/main) and randomised arm (including trial participants who crossed over). Sample size was determined by reaching data saturation, whereby no new themes emerged in three consecutive interviews.¹²³

Recruitment

Recruiting staff (otolaryngologists, research nurses, nurse practitioners and clinic managers) and GPs who had patients taking part in the NATTINA trial were identified through trial records. NATTINA trial participants were asked to consent to be contacted for a qualitative process evaluation interview when consenting to either the pilot or the main trial. Patients who declined participation in either the pilot or the main trial were informed about the qualitative process evaluation interviews by the clinical team and invited to complete an expression of interest form. Otolaryngology staff, GPs and a sample of NATTINA trial participants and declining patients were contacted by telephone, e-mail or letter inviting them to

participate in a telephone or face-to-face interview with either Dr Lorraine McSweeney or Dr Lyndsay Lindley (both trained and experienced qualitative researchers) at a time and location convenient for the participants. Invitations were accompanied by a PIS. Verbal consent was given at the time of the telephone interviews with signed written consent returned post interview. Signed written consent was given at the time of face-to-face interviews.

Interviews

Semistructured interviews were based on flexible topic guides (see [Report Supplementary Material 3](#)) derived from the literature and issues raised by the NATTINA PPI group, as well as in conjunction with the trial otolaryngologists and GPs. Interviews with trial participants covered expectations and motivations for participating, views and experiences of the randomised arms, and views about sore throat. Interviews with declining participants explored expectations, reasons for not participating and views about sore throat. Recruiting staff and GP interviews explored the practicality and suitability of the treatments, research tasks and randomisation, and any barriers to or enablers of treatment delivery. Interviews for the qualitative process evaluation were carried out during the internal pilot and towards the end of the main trial. Trial participants and declining participants were reimbursed for their involvement in the qualitative process evaluation via a £15 shopping voucher.

Data management and analyses

Interviews, lasting up to 1 hour, were digitally recorded. We then transcribed them accurately word for word. The Framework Analysis, which we used to generate a matrix output [rows (cases), columns (codes) and 'cells' of data],¹²⁴ is regarded as the optimum approach by which the aims of qualitative health research are linked to quantitative investigation.¹²⁵ This method allowed coding to be transparent and shared.¹²⁵ NVivo software (QSR International, Warrington, UK) facilitated coding. Data were repeatedly read and coded by Dr Lorraine McSweeney/Dr Lyndsay Lindley, with reference to anticipated, participant-reported or data-generated issues. Themes were discussed with the qualitative lead (Professor Catherine Houghton) and the trial team to reduce bias.

Results

In total, 47 interviews were conducted between July 2015 and June 2019. Interviewees comprised a sample of NATTINA trial participants ($n = 19$), including five crossover participants (two tonsillectomy to conservative management and three conservative management to tonsillectomy) and 10 tonsillectomy participants and four conservative management participants, plus a sample of patients who had declined to participate in NATTINA ($n = 4$). Ages ranged from 17 to 42 years, and there were 6 males and 17 females ([Table 26](#)). Interviews also took place with a sample of GPs who had a patient participating in NATTINA ($n = 9$); a sample of ENT/research staff from recruiting sites, including consultants, registrars, nurses and research nurses ($n = 10$); and a sample of ENT/research staff situated at sites that had failed to recruit any participants to the trial ($n = 5$) ([Table 27](#)).

Findings

General practitioners

We contacted 181 GPs, of whom nine (5%) consented to participate. A further 17 GPs responded that they were unable to participate owing to time constraints, five had left the practice/retired, two NATTINA participants had left the practice and four practice managers acted as gatekeepers denying access. There were 142 non-responders (78%). All nine GP participants undertook telephone interviews.

Six main themes emerged from the GP interview data:

1. GPs adhere to usual surgery practice.
2. GPs demonstrate knowledge of SIGN and local guidelines.

TABLE 26 Interviewees (NATTINA trial participants/decliners)

Site number	Date	Randomised	Age (years)	Sex
01	July 2015	Decline	19	Female
01	July 2015	Decline	40	Male
02	April 2016	Decline	17	Female
01	April 2016	Tonsillectomy	40	Female
05	May 2016	Crossover conservative management to tonsillectomy	21	Female
05	May 2016	Decline	22	Female
05	May 2016	Tonsillectomy	26	Female
09	May 2016	Crossover conservative management to tonsillectomy	29	Female
09	May 2016	Tonsillectomy	32	Male
09	May 2016	Tonsillectomy	29	Female
10	June 2016	Crossover conservative management to tonsillectomy	19	Female
10	June 2016	Crossover tonsillectomy to conservative management	29	Female
19	May 2017	Conservative management	32	Male
07	August 2017	Tonsillectomy	18	Female
22	March 2019	Tonsillectomy	19	Female
03	March 2019	Conservative management	19	Female
12	March 2019	Conservative management	22	Female
06	March 2019	Tonsillectomy	25	Male
03	March 2019	Tonsillectomy	31	Male
12	May 2019	Crossover tonsillectomy to conservative management	30	Female
27	May 2019	Conservative management	37	Male
27	June 2019	Tonsillectomy	34	Female
08	June 2019	Tonsillectomy	42	Female

TABLE 27 Interviewees (staff)

Staff type	Site number	Date	Recruited
GP	06	May 2017	Yes
GP	06	March 2017	Yes
GP	06	November 2016	Yes
GP	05	February 2017	Yes
GP	03	January 2017	Yes
GP	03	February 2017	Yes
GP	03	March 2017	Yes
GP	02	January 2016	Yes
GP	09	May 2017	Yes

continued

TABLE 27 Interviewees (staff) (continued)

Staff type	Site number	Date	Recruited
ENT/research staff	06	August 2016	Yes
ENT/research staff	03	Missing	Yes
ENT/research staff	02	September 2016	Yes
ENT/research staff	07	August 2016	Yes
ENT/research staff	07	September 2016	Yes
ENT/research staff	08	October 2016	Yes
ENT/research staff	09	August 2016	Yes
ENT/research staff	10	July 2016	Yes
ENT/research staff	29	January 2017	Yes
ENT/research staff	29	January 2017	Yes
ENT/research staff	N/A	July 2017	No
ENT/research staff	N/A	July 2017	No
ENT/research staff	N/A	July 2017	No
ENT/research staff	N/A	August 2017	No
ENT/research staff	N/A	August 2017	No

N/A, not applicable.

3. GPs refer only on patient request, if the patient’s work is affected and if they meet the SIGN/NICE criteria.
4. GPs have negative views of tonsillectomy.
5. GPs refer to ENT for a consultation only, not necessarily expecting the patient to be offered surgery.
6. GPs consider NATTINA appropriate and necessary.

General practitioners adhere to usual surgery practice

The use of swab tests to determine whether sore throat episodes were viral or bacterial to determine the necessity for antibiotic prescribing was becoming increasingly common: ‘[w]e do swabs, especially if people are mentioning that they want referral for tonsillectomy, that’s what I certainly would do, swab to prove or disprove this is a bacterial one’ (interview 1, GP).

There was a strong belief that GP and patient expectations of conservative treatment were slowly changing. Patients were reported to be more accepting of self-management strategies and to have a lower expectation of antibiotic prescribing: ‘I think that’s happening gradually, people are aware that we don’t give antibiotics anymore unless there are specific indications for it’ (interview 2, GP).

General practitioners demonstrate knowledge of SIGN and local guidelines

Concurrently with the increasing use of swab tests was the use of guideline criteria for antibiotic prescribing or for ENT referral if this was considered essential:

I’m also aware ... that basically the way that you get your tonsils taken out is if you have seven episodes of purulent tonsillitis within the space of 12 months or five episodes every year for 2 years or three episodes every year for 3 years.

Interview 6, GP

General practitioners refer only on patient request, if the patient's work is affected and if they meet the Scottish Intercollegiate Guidelines Network/National Institute for Health and Care Excellence criteria

Patient presentation with recurrent sore throat symptoms was reported to be a common occurrence, but referrals to ENT were reported to be rare: '[n]o, no, I never refer. I do refer folks' kids with sleep apnoea, just in case, but for adults normally, unless it's making their life a complete and awful misery ...' (interview 6, GP).

Referrals to ENT were generally patient-led conversations and would be considered if the patient met the guidelines for referral and was having to take considerable time off work or miss education:

Yes, usually what happens is often the push for tonsillectomy is generated by the patient rather than by us. They are having recurrent bouts of tonsillitis which is adversely affecting their life and compromising either their academic studies or their prospects of continuing with a job.

Interview 5, GP

General practitioners have negative views of tonsillectomy

Referrals to ENT were rarely considered for several reasons: GPs felt that adult tonsillectomies were no longer routinely performed, surgery was not always considered to be effective, the negatives of surgery outweighed the benefits, or procedures such as tonsillectomies were being vetted and restricted:

Certainly, commissioning groups are more likely to question procedures such as tonsillectomies, and whether they're necessary. We're moving towards a situation where referrals and/or surgical procedures are vetted, really, before they're approved and funded. So, we may reach a stage where anybody that needs a procedure like a tonsillectomy has to have an individual funding request approved before they can have their procedure.

Interview 8, GP

General practitioners refer to ear, nose and throat only for a consultation not necessarily expecting the patient to be offered surgery

There was emphasis that the patient was being referred for an expert opinion with no guarantee that surgery would be offered:

I tell them it is for a consultation, it is with the ENT to discuss their options for ENT to decide whether or not they think it is appropriate and if they do, they would get a later appointment for an operation. But I do tell them that they are not going in and having their tonsils out there and then.

Interview 1, GP

General practitioners consider NATTINA to be appropriate and necessary

With respect to their patient being involved in the NATTINA trial, several GPs had no knowledge of this prior to being contacted to participate in an interview. However, some conceded that there was information about their patient's participation in the patient's notes:

I don't think I was necessarily told ... let me just check ... maybe in a letter back from ENT. Even the letter from ENT didn't come to me, it went to one of my colleagues so I actually don't know anything about this person, I'm afraid.

Interview 1, GP

The GPs could foresee no negatives for patients of participating in NATTINA, apart from patients' possible changing expectations of care if randomised into the conservative management group (i.e. prescribing of antibiotics):

There might be expectations about what usual care is, which they might feel always should be antibiotics, or that because they're part of a study, they have an expectation of a right to an antibiotic that otherwise they wouldn't have had.

Interview 8, GP

It was agreed that the trial was relevant as more evidence was needed to determine the clinical effectiveness of tonsillectomies and to provide patients with an informed choice. The trial tasks required of the participants were believed to not be onerous and to be a good use of technology:

I think that is a really good way of using technology [text/e-mail] and also getting the information that you need, because if you sent a paper questionnaire you wouldn't get it back. I think that's a good and innovative way of doing it, yes.

Interview 5, GP

Ear, nose and throat/research staff (recruiters)

The staff interviews with recruiting sites comprised ENT consultants, ENT nurses and research nurses. Twenty-three staff members involved with NATTINA were contacted by e-mail or telephone to arrange an interview. Of these, 10 (43%) consented and participated in a telephone interview.

Six main themes emerged from the staff interview data:

1. ENT staff feel that referral processes are changing.
2. ENT staff highlight several barriers to recruitment.
3. The recruitment process itself poses some difficulties for ENT staff and varies depending on capacity of, and support for, ENT staff.
4. Maintaining patient communication is problematic, particularly with younger patients.
5. ENT staff report trial processes to be acceptable.
6. Many ENT staff believe that tonsillectomy is the most effective treatment for recurrent tonsillitis.

Ear, nose and throat staff feel that referral processes are changing

Staff reported that GPs were compliant with SIGN guidelines. It was felt that fewer patients were being referred to ENT or that GPs were giving patients the expectation that they would receive surgery and that they were not merely attending the appointment for a consultation: '[b]y the time they come ... to see an ENT doctor here, they've had it for a few times and they're almost convinced from the GP that they're going to have their tonsils out' (interview 16, ENT staff).

It was suggested that GPs should have been given advance warning of the trial and that perhaps they could be provided with 'consultation guidance' to enhance recruitment:

Obviously they've been made aware that the trial is up and running but my reason for saying that is if the GPs were more aware that they were sending a patient up to a specific tonsillitis/tonsillectomy clinic they wouldn't send them up with such clear referrals such as 'This patient wants to have their tonsils removed can you please see and arrange'.

Interview 10, ENT staff

Ear, nose and throat staff highlight several barriers to recruitment

Several potential barriers to recruitment were highlighted, the main one being the patients' expectation of surgery:

I think the main reason was that our patients were waiting for a long time, being made to jump through so many hoops by GPs and CCGs – the guidelines and protocols – and were just desperate to ask for the surgery. We don't have patients who would be willing, if you like, to sit on the fence and say, 'Maybe we'll give it another few months and see what happens'. That is the main difficulty that my team has.

Interview 18, ENT staff

Furthermore, staff reported that many patients were unwilling to take the risk of being randomised into the conservative management arm and having to continue with the same treatment they had experienced for, in some cases, many years:

So one of the handicaps I suppose is that one of the arms of the study is conservative treatment which they've already had in their mind for the last however long. They have experience of one limb, if you like, of the study already and now they want something else.

Interview 10, ENT staff

There was some discussion of participants in the conservative management arm being provided with more support and communication than originally planned:

I do think that the conservative arm; I personally think they should have more appointments. I think we should be seeing them face to face more often than we do. Just so that we can keep them on board really with completing the STAR forms and things and to give them a bit of encouragement.

Interview 13, ENT staff

In addition, it was felt that conservative management participants should be provided with guaranteed access to GP appointments and antibiotics when required:

So if you could give the patients the option of saying if you do go into the conservative arm, as well as the letter your GP is going to be getting to notify them of that, if there could be a little amendment in there just to say 'could this patient somehow be ...' Not prioritised, but, 'Is there any chance of having a more speedy appointment?'. I don't know if that's at all possible?

Interview 12, ENT staff

The crossover option was reported to provide patients with an added incentive to participate and provided an additional option:

The fact that they can switch in between the groups if they didn't feel like – if they felt actually I want to have my tonsils out or I've changed my mind I don't want to have the tonsils out anymore. That also helps some of them decide.

Interview 16, ENT staff

Patients who were unsure about having surgery or who were considered to be more 'health aware' were thought to be more likely to participate in the trial: '[t]he ones who are interested have only been the ones who know about clinical trials. They know a little bit about, I don't know, they're health conscious, so they know it's important to take part in things like that' (interview 16, ENT staff).

The recruitment process itself poses some difficulties for ear, nose and throat staff and vary depending on capacity of and support for ear, nose and throat staff

It was apparent that the recruitment process itself posed some difficulties; some sites did not have specific tonsil clinics or were unable to allocate extra time to the patient during consultation:

I think the difficulty is that we run a general clinic; we don't have a sore throat clinic. So the potential may be for a longer consultation, trying to convince the patient to take up the study. I think that would potentially pay off. But you know what it's like in the NHS: I don't have any special funding which comes with the studies, therefore I'm unable to offer the patients more than the allocated time ... I can't offer a potential recruit half an hour for a consultation in order to convince them.

Interview 18, ENT staff

Depending on the capacity and support available for each site, screening, contacting and recruiting of patients was varied. Those who had full-time access to a research nurse appeared to manage the process more efficiently:

Well we set up a specific clinic, which was generally held once a fortnight, with our research nurse, and they usually have in attendance the senior registrar ... I think having a specific clinic, has been the only way to do this ... I wouldn't have the 20 minutes, half an hour that we give the patients, in a routine clinic.

Interview 14, ENT staff

Some research nurses were allocated to several research projects simultaneously and were not always available when required: '[w]e are big team, but equally our research nurse is not allocated for ENT research, so potentially she could be doing other research' (interview 18, ENT staff).

Maintaining patient communication is problematic, particularly with younger patients

The majority of the ENT consultants took the lead in explaining the trial to the patients, but the patient contact/communication was carried out by the nurse/research nurse. In some cases, maintaining patient contact/communication was problematic; this was especially the case with the younger patients. There was also some concern of patient availability throughout the 24-month trial period:

I think sometimes people are reluctant to be followed up for 2 years. It's a long time especially when the majority of the patients that we're dealing with are young people who don't think about what they're doing in the next 6 months, never mind in 2 years.

Interview 13, ENT staff

Ear, nose and throat staff report trial processes to be acceptable

The staff interviewees reported the trial process to be acceptable, with only a few sites experiencing barriers to treatment. They alluded to having to negotiate surgery slots and compromising with participants over preferred surgery dates owing to holidays, for example:

So I had a few patients reach the 8-week time covered, mostly because there's no set of slot or they had to rearrange them or the patients, because they're young, they always going away or they think this is no good. They want to push it back to almost a time that they feel it's better for them.

Interview 16, ENT staff

The randomisation process was reported to be effective and straightforward:

I mean the randomisation has been quite easy, obviously when we've explained it to them they've understood it completely and they know that the computer will decide for them, in a sense. They've either got tonsillectomy or they've got conservative management.

Interview 17, ENT staff

The research tasks were considered to be satisfactory; however, the completion of questionnaires in the clinic was often time-consuming:

The questionnaires that we fill out in the clinic, they are a bit time-consuming, bearing in mind there are nine pages to them. That is a bit time-consuming. One of the questions that has come up for the patients we have randomised, there's an option where it asks them to summarise their financial loss over a given period. Like I say, there are only four but all four of those people have essentially just made that number up because as they've tried to quantify it, they can't.

Interview 12, ENT staff

The weekly alerts that the participants received were considered to be effective in maintaining frequent contact and acted as a reminder to the participant of their participation in the trial. In addition, it was felt that feedback about participants' compliance with the tasks would be useful:

It would be nice to have some feedback, if there's any sort of meaningful data, as to how compliant patients are, and how do we, I was going to say rank, but how do our patients fit in with compliance, so from their point of view, or do we need to chase it up.

Interview 14, ENT staff

Many ear, nose and throat staff believe that tonsillectomy is the most effective treatment for recurrent tonsillitis

To establish equipoise for the trial, staff were asked to discuss their own views and experiences of sore throat treatment. There was a belief that tonsillectomy was an effective treatment for recurrent tonsillitis; however, this view was reported not to be shared with patients:

You could not be comparing two more different treatments and trying to persuade people that in some ways we don't know when in fact it's really difficult under the umbrella of a trial not to ... It's really difficult to maintain an unbiased feel because I think most ENT surgeons feel that tonsillectomy is by far and away the superior treatment ... the treatment for chronic recurrent tonsillitis I think should be tonsillectomy but that's why we're doing the study of course and I must not be biased.

Interview 10, ENT staff

Staff reported a significant increase in the number of patients being admitted to A&E with tonsil infections: '[t]here is a significant increase in people coming in, with tonsil infections, into hospital, possibly because they either can't get access to primary care, or they're coming in to out-of-hours casualty' (interview 14, ENT staff).

Procedures dictated by the GP referral criteria, such as the use of throat swabs and documented usage of antibiotics, were described as 'impractical':

The only thing that occasionally puts us in a difficult position is that we're expected to demonstrate that the patient had, for example, positive swabs – streptococcal swab. But it's totally impractical because, if you ring your GP and say 'I have a sore throat', they'd say, 'wait for a couple of days'. Because then they'll come and get some antibiotics and they probably don't need to see you. So, I question the value of swabs.

Interview 18, ENT staff

Non-recruiting site interviews

Five ENT staff were interviewed across four non-recruiting sites. Four main themes emerged from the staff interview data at non-recruiting sites:

1. Non-recruiting sites show concern at their lack of recruitment for NATTINA.
2. Non-recruiting sites express reservations about the NATTINA trial design.

3. Non-recruiting sites struggle with equipoise and GPs' interpretation of SIGN guidance.
4. Non-recruiting sites experience site-specific barriers.

Non-recruiting sites show concern at their lack of recruitment for NATTINA

All four non-recruiting sites (covering five interviews) were typically research active and displayed concern at their lack of recruitment for NATTINA – this was considered to be an unusual situation for them:

We have a reasonable track record for recruiting to studies. We do a number of ENT trials and I think we've got about six or seven things going at the moment. For this one not to work for us I think probably seems like a reflection on the difficulties with the referral barriers that we have here locally, I would hope.

Interview 20, ENT staff

Non-recruiting sites express reservations about the NATTINA trial design

Interviewees expressed reservations about the NATTINA trial design. It was felt that recruitment should begin earlier in the care pathway, as interviewees reported that patients were desperate for intervention by the time that they attended the ENT appointment:

Most people who turn up clinically hoping for a tonsillectomy in the north west here, have struggled to get past the GP and will not countenance the idea of being randomised into waiting for another 2 years to suffer all the symptoms as they first presented with, on the off chance that somebody in the future might decide that it's a good or bad idea ... if you want to answer those questions you would be better off doing it as a GP study in GP land, where you can actually see how any people have documented tonsillitis.

Interview 22, ENT staff

Non-recruiting sites struggle with equipoise and general practitioners' interpretation of SIGN guidance

It is possible that some sites were struggling with equipoise given that they placed great emphasis on surgery being the best/only option for the patients they were seeing:

I don't believe that one treatment's not better than the other. I think that's maybe one of the difficulties you've got. I'm in no way convinced that not operating on them is a good idea. In fact, I kind of think it's a bad idea, personally, if you ask me. I try not to convey that to the patient.

Interview 22, ENT staff

In addition, personal views of the SIGN guidelines elicited some negative reactions. It was felt that GPs were not using SIGN appropriately, as guidelines only, and were being too stringent with the referral criteria, which meant that patients were in quite a severe condition by the time they were seen:

They've all, as I have said, struggled to get past their GP, because they all have SIGN guidelines, which are a load of rubbish; plus, they have to work really hard to get past the initial SIGN guideline before they're referred. By this stage, they're pretty miserable and pretty unhappy with the medical care they've received.

Interview 22, ENT staff

Non-recruiting sites experience site-specific barriers

One site reported that their main referral population came from the local Ministry of Defence (MOD) camp and they felt it unreasonable to expect the soldiers to take more time away from their duties:

We are perhaps in a slightly unique situation that we take a lot of referrals from the MODs. We have a lot of army camps in the area and so, these are young fellows who are in the armed forces, who have had recurrent tonsillitis and can't afford to have additional time off of their duties in the armed forces. So, they were not willing to be selected for the non-interventional arm of the trial, for that reason.

Interview 20, ENT staff

Another consultant did not believe that research staff, without clinical knowledge, should be approaching patients to ask them to consider not going through with surgery and to participate in NATTINA. Other barriers identified included the time needed for screening patients.

Interviews with NATTINA trial participants and declining patients

Interviews were carried out with 23 patients, including 19 NATTINA trial participants: five crossovers (two tonsillectomy to conservative management participants and three conservative management to tonsillectomy participants), 10 tonsillectomy participants and four conservative management participants. In total, 144 expression of interest forms were received from declining patients across 15 sites: four declining patients were interviewed.

Seven main themes emerged from the patient interview data:

1. Patients report requesting an ENT referral.
2. There are variations in the ways that patients are informed about NATTINA.
3. Variations exist in patient consultations and recruitment processes.
4. Understandings of, and motivations for, participation also vary.
5. Patients decline to participate because they do not want to risk being randomised into conservative management.
6. Patients randomised into conservative management are disappointed.
7. Patients prefer electronic questionnaires.

Patients report requesting an ear, nose and throat referral

There was heightened awareness among patients of their general practice's procedures for sore throat treatment. Many patients spoke of the futility of requesting antibiotics because they either knew that they were unlikely to be prescribed or felt that they were no longer an effective treatment:

Yes. Well, I've always suffered with it. I know when it's coming on because I've had it so many times. A lot of the time, because I know that they wouldn't give me antibiotics, they say 'ride it out', I just end up riding it out ... If I'm ill – the thing is, you don't want to be wasting a GP's time when you know that in time your body will deal with it anyway. You know what the GP is going to say. It's 'There's no point in me giving you any antibiotics'. What's the point then in going to the GP? If that makes sense.

Interview 13, NATTINA conservative management, male, aged 32 years

Patients spoke about not wanting to waste their GP's time by visiting each time they were suffering from a sore throat. However, patients were mindful of having to record sore throat episodes for consideration of a referral to ENT:

About 5 years ago, I got tonsillitis for the first time. I didn't know it was tonsillitis, and I was ill for about a week, I think, 7 days before I went to the doctor. I was really sick. I think I lost loads of weight, but I was physically sick. Then the doctor told me I had tonsillitis, I had antibiotics. That year, I had tonsillitis six times. I got referred to ENT ... they said they wouldn't take my tonsils out because I hadn't had it seven times, which was their criteria.

Interview 9, NATTINA tonsillectomy, male, aged 32 years

Owing to the difficulties of having to record episodes, there was often a need for patients to request an ENT referral despite not formally being considered eligible: '[s]o I went to the doctors with regular tonsillitis and they kept just telling me to go and have antibiotics and I was pretty forceful and just said that 'I want to get referred, because this is ridiculous' (interview 8, NATTINA crossover conservative management to tonsillectomy, female, aged 29 years).

There are variations in the ways that patients are informed about NATTINA

There was variation from site to site in the ways in which patients were informed about NATTINA. Some patients received information pre appointment in the post or by telephone, whereas others were informed about the trial in person on the day of their appointment. In addition, some patients had no knowledge of, or were not shown, the video clip. However, overall, patients felt that they had received enough information about the trial to make an informed choice:

I've got to say, the team have a pretty well-polished chat. So, although I work in [medical research] myself, I think that they didn't take for granted that I knew anything. They explained the whole process really very well, actually, and as I said, for me, it was great anyhow but I felt very much like, you know, it was explained to me that at any point, you know, I could come out of that if I was getting sick all the time.

Interview 21, NATTINA conservative management, male, aged 37 years

Variations exist in patient consultations and recruitment processes

In the same manner, the consultation and recruitment process differed between sites. This was largely dependent on whether or not the site had access to a research nurse. In some sites, the explanation about the trial was given by the consultant, whereas in others the research nurse would see the patient before they met with the consultant. In a small minority of cases, there was some confusion about the randomisation process: 'I think how I interpreted it, was that they would give you a group. I didn't understand that I'd be randomised into a group until I was randomised into a group, if that makes sense' (interview 12, NATTINA crossover tonsillectomy to conservative management, female, aged 29 years).

Understandings of and motivations for participation also vary

There was a belief that participation depended on the severity of symptoms or that agreeing to participate was the only way to be offered surgery:

I think they did a random survey of who to choose to have them out within a period of time, and I actually got chosen for that because I was having such bad problems.

Interview 14, NATTINA tonsillectomy, female, aged 18 years

I don't think I would have been offered a tonsillectomy anyway, so it felt like at least I was doing something rather than nothing ... I think just because they said it was a chance. You didn't know whether you would or whether ... I don't know if it was on merit or if it was ... I thought that it was maybe, you might get one or you might not get one. I didn't feel like I would be able to get one just normally.

Interview 7, NATTINA tonsillectomy, female, aged 26 years

Overall, however, patients had a good understanding of why they were being asked to take part in NATTINA; they understood the need to find out whether or not surgical intervention was better than normal care:

The study is about whether having a tonsillectomy will then reduce people coming back with sore throats. I'm imagining a tonsillectomy is quite an expensive procedure and they're trying to balance up whether a tonsillectomy is a better idea than people coming back for repeat prescriptions and ongoing problems with their throat.

Interview 7, NATTINA tonsillectomy, female, aged 26 years

Patients' other motivations for participation included wanting to help others, furthering research, believing that there was little commitment in taking part and knowing that they could drop out of the trial if they felt the need to:

Because you are helping other people who might be sick, who might be in pain, and they don't know what to do. So when you are taking part in this kind of research, you give a feedback, and it might help others, so that was my main reason.

Interview 23, NATTINA tonsillectomy, female, aged 42 years

Patients decline to participate because they do not want to risk being randomised to conservative management

Patients who declined to participate did so mainly because they had been suffering from tonsillitis episodes for many years and did not want to risk being randomised into the conservative management arm and having to wait longer for surgery:

I had already made my mind up before ... from reading the literature which was sent out ahead of my consultation appointment I had already determined that it was not for me. Had this been some time ago, then I would have been quite happy to participate, but the nature of how I have suffered with tonsillitis over a number of years and it is getting worse now and it doesn't fit in with ... it's impacting on family life, I had decided, 'Right, I am getting them out'.

Interview 2, declining patient, male, aged 40 years

Although there was some discussion of guilt in declining, patients were not made to feel uncomfortable or awkward for doing so:

... so I felt a little bit guilty, actually; but actually, the research nurse was just like, 'You've got three kids, you've had to give up work for a month, think of yourself'. So that was really helpful. There was absolutely no emotional pressure put on.

Interview 6, declining patient, female, aged 22 years

Patients randomised to conservative management are disappointed

There was an overwhelming sense that participants who had been randomised to the conservative management arm were disappointed with this outcome: 'I was a little bit disappointed I think because I was expecting to have them removed ...' (interview 13, NATTINA, conservative management, male, aged 32 years).

Moreover, those who had been randomised to the tonsillectomy arm reported that they would have 'struggled' with a conservative management decision:

... disappointment, that I would have to go at least another year. I know there's the halfway point of the study in which you would be able to have a tonsillectomy. I think I would have reached that and would have gone, 'Just do it, please'.

Interview 9, NATTINA tonsillectomy, male, aged 32 years

In addition, there was a general feeling that the conservative management arm offered no real benefit to the participant:

It seemed to me as if the people that were on the half of the trial that didn't get the op were just ... there was the assurance that at any point you could say, 'I want to have my tonsils out'. But you didn't get any special treatment, in a way. You didn't get a fast track to antibiotics or to medical help, it was just like, right, back you go, out into the world and cope with it ... you would just be thrown back out to the wolves.

Interview 6, declining patient, female, aged 22 years

Crossing over from one arm of the trial to the other was reported to be relatively straightforward. Participants who had crossed over to the tonsillectomy arm did so because they experienced a further severe sore throat episode:

I was a bit frustrated. I think by the time, once you've come to the hospital stage you're already expecting to have your tonsils out really, so I think that was a little bit of an issue with it ... so yes, by that time I was a bit frustrated, because I was kind of hoping to already have them out. And then I think I had another bout of tonsillitis within 2 weeks of being randomised and was like, 'OK, this is ...', I just said I'm not really happy.

Interview 8, NATTINA crossover CM to TX, female, aged 29 years

Participants who had crossed over to the conservative management arm did so because they felt that the severity of their symptoms did not justify the risks associated with surgery:

Yes. I do feel like having a tonsillectomy could be an option. But I just don't feel like I'm ready for it yet. And with all the bleeding and the time off work, I don't know whether the positives – If we're still going to get ... Yes, because I didn't feel ... I just don't feel like I'm – Even though I do have really bad spells and that's why I was referred – I've had the quinsy and everything else. Because it's been cleared up since Christmas time, and with all the negatives, I felt like I just didn't want to make that decision just yet. I wanted to think about it for a while.

Interview 12, NATTINA crossover TX to CM, female, aged 29 years

Despite the challenging recovery period experienced after surgery and the complications experienced for some participants, those who had received surgery reported being happy to have done so: '[i]t's been really, really positive to have the tonsillectomy. It's been amazing. Life changing, almost. I know it sounds quite dramatic, but ...' (interview 7, NATTINA tonsillectomy, female, aged 26 years).

For some (in both the tonsillectomy and the conservative management arm), the randomisation process relieved them of having to make the difficult decision of whether or not to undergo surgery:

So I just thought, you know, if I am in two minds about it, I would be up for being randomised, so ...

Interview 18, NATTINA tonsillectomy, male, aged 25 years

In fact, the good thing for me was that I was actually, although I was sick of having tonsillitis, I wasn't that keen on the idea of surgery, necessarily. So, when I got, you know, to be part of a randomised trial was excellent because it took the need to make a decision out of my hands.

Interview 21, NATTINA conservative management, male, aged 37 years

Participants prefer electronic questionnaires

The STAR alerts were well accepted and deemed to be 'quick' and 'easy':

It's really good, actually. Compared to other trials I've been involved with, and on the side lines of, with the use of technology it's much less intrusive because it comes to you and you're not having to remind yourself to do something, it's really handy. The fact you can reply immediately and have it done with and then can forget about it again means that I'm reminded for all of about a minute a week that I'm on a trial. Which is actually quite a nice way, I think, to participate in a trial. It's not using up any of my resources or time, which is good.

Interview 21, NATTINA conservative management, male, aged 37 years

However, both completing and returning the STAR form questionnaires were not considered to be straightforward. Participants commented that electronic and online versions would be more practical. This, they felt, would improve response rates because the weekly questionnaires were often not posted or a few would pile up before they returned them:

It just comes through by text. I reply, zero, and it's done for a week. It doesn't take up any time ... if there was a STAR form online, that might be easier ... just because I sometimes keep it in my bag for a week before I remember to post it ... I think if it was online, because I'm on the computer all day, it would be maybe a bit quicker to just fill it in online.

Interview 7, NATTINA tonsillectomy, female, aged 26 years

A number of participants had changed their address or other contact details but did not know how to pass this on to the trial teams, so were not able to receive questionnaires in the post or appointments for follow-up 12-/24-month visits:

Some of the paperwork ... I'm a very clueless person and since having my tonsils out I've had to move and I've had to move college, so I haven't been able to keep up to date with them as much as I'd want to, and I hope that hasn't affected the research.

Interview 14, NATTINA tonsillectomy, female, aged 18 years

Most participants continued to receive the texts. In one case, the participant's mobile telephone number and address had changed and she did not attend her tonsillectomy: '[a]t the moment I feel, because the contact seems to have stopped, for whatever reason, it feels a bit like I've signed up for this thing and it's not really going to go anywhere, or be of any use' (interview 4, NATTINA tonsillectomy, female, aged 40 years).

Regarding GDPR issues, one participant felt that it had not been clear that the texts came from an outside agent (it had been over 2 years since he entered the trial and read the PIS). However, he felt that it was acceptable providing that the university oversaw the approvals for storage and use of the data: 'I suppose I would have preferred it if it was just done by, yes, Newcastle University. Or maybe even it was just a bit more explicit' (interview 18, NATTINA tonsillectomy, male, aged 25 years).

All other participants who were asked about data protection were happy with the arrangements. Responses ranged from accepting that giving out contact details is universal ['everyone and their aunt' has your e-mail address (interview 18, NATTINA tonsillectomy, male, aged 25 years)] to expressing trust in the handling of data by the NHS/university: 'I'd hate to even wonder at the amount of different people that have my telephone number or e-mail address' (interview 19, NATTINA tonsillectomy, male, aged 31 years).

Discussion

The trial processes were generally deemed to be acceptable, with only a few sites experiencing barriers to treatment. The use of technology to collect data was particularly well received and this is reflected in our respective response rates to texts compared with paper questionnaires. In fact, based on early feedback on non-compliance with the STAR form questionnaires, the STARLET questionnaires were introduced. Unfortunately, a number of participants had changed their address or other contact details and did not know how to pass their updated details on to the trial teams; therefore, they were not able to receive questionnaires in the post or appointments for follow-up 12-/24-month visits. Simple measures to capture changes in contact details might have further improved response rates in such a young, transient trial population.

However, there were some challenges with recruitment. Despite the TMG strongly advising having dedicated clinics for recruitment to the trial, some centres chose not to run these. For this reason, it was these centres that reported finding recruitment difficult. Although there were dedicated and paid research nurses at each centre, the trial still suffered from a lack of dedicated personnel. Clinical research network funding supports nurses for portfolio studies; however, a high turnover of staff exists and prioritisation is a constant battle, particularly for a relatively low headcount recruiting trial, such as a surgical trial.

Another challenge for surgical trials in general is the need for participants to pause their normal life processes for at least 1 week, if not more, which is not the case for medical treatments. ENT staff alluded to having to negotiate surgery slots and compromising with participants over preferred surgery dates, which, at times, meant that they had to deviate from trial protocols.

It was not surprising to find that those staff who lacked equipoise were also those who struggled to recruit to the trial. This was despite the fact that some participants reported that they themselves were in a position of equipoise and were pleased that 'the computer' chose their treatment choice for them. Non-recruiting staff also suggested that recruitment should begin earlier in the care pathway, as they reported that patients were desperate for intervention by the time they attended the ENT appointment. Although this is a seemingly appealing trial design because such a tiny proportion of sore throat consultations are eligible for tonsil surgery, hundreds of GPs would need to be aware of, and participate in, such a trial design, making it impractical. All non-recruiting sites were typically research active and displayed concern at their lack of recruitment for NATTINA; therefore, regional variation in referral processes and access to tonsil surgery should also be acknowledged.

There were also reported problems with communication between primary and secondary care, with GPs restricting or delaying referring patients up to ENT, and ENT not able to inform the correct GP when their patient was involved in the trial. GPs felt that it was rarely necessary to refer patients. They were aware of guidelines and would refer only if requested by a patient who fulfilled the guidelines criteria and/or who were missing considerable amounts of work.⁸⁰ Patients corroborated this point, reporting having to request an ENT referral despite not formally being considered eligible. Several ENT staff felt that GPs were not using SIGN appropriately as guidelines only and were being too stringent with the referral criteria, meaning that patients were in quite a severe condition by the time they were seen.² These ENT staff had previously suggested that patients with recurrent acute tonsillitis may be experiencing undue delay for treatment as patients in their trial experienced two to three times more tonsillitis episodes than the minimum required by current SIGN guidelines.² In secondary care, some of the document creation systems did not allow a change to the designated recipient of a letter, therefore not allowing ENT staff to inform the relevant GP of their patient's involvement in the trial or to respond to the referring clinician. In addition, patients would often access other providers, such as walk-in centres, because they knew that they would not get antibiotics from their GP; this is reflected in the increase in acute admissions for severe tonsillitis.¹¹

Patients who were unsure about having surgery or who were considered to be more 'health aware' were thought to be more likely to participate in the trial. Patients who declined to participate did so mainly because they had been suffering from tonsillitis episodes for many years and did not want to risk being randomised into the conservative management arm and having to wait longer for surgery. This aligns with our quantitative data that reveal that participation in the trial depended on disease severity at baseline. Given the findings from our feasibility study, it was not surprising to find that patients declined to participate in the trial because they did not want to risk being randomised to the conservative management arm and were disappointed with the conservative management arm.⁸²

It was concerning to hear that some patients did not fully understand the process of randomisation, and that some believed that participation in the trial depended on the severity of their symptoms. It was even more worrying to hear that some patients report that they thought that agreeing to participate was

the only way to be offered surgery. Not all patients were offered or chose to watch the trial recruitment DVD, and this may have contributed to patients' lack of understanding.

It was interesting to hear that GPs and ENT staff both suggested that patients randomised to conservative management might have increased expectations of access to GP consultations and antibiotic prescribing. In our original trial design, we had suggested an early maximum medical therapy arm, in which we had proposed that patients should be able to access a package of steroids and antibiotics when necessary to make the conservative management arm more attractive.

The weekly alerts that the patients received were considered to be effective in maintaining frequent contact and acted as a reminder to the patient of their participation in the trial. However, in hindsight, feedback about patients' compliance with the tasks could have also been provided to trial sites.

Patients who had received surgery were unanimous in reporting to be happy to have done so despite the challenging recovery period experienced after surgery and complications for some patients. This aligns well with the high level of patient satisfaction reported in the Scottish tonsillectomy audit, with 98% of patients who returned the questionnaire being glad that the operation had been performed.¹²⁶

Chapter 6 Discussion

Statement and interpretation of main results

The output of NATTINA is a landmark contribution to the evidence base for the efficacy, clinical effectiveness, safety and cost-effectiveness of adult tonsil removal for recurrent sore throats. In terms of the primary outcome – the number of sore throat days in 24 months after randomisation – the ITT, the per-protocol and the per-treatment analyses all demonstrated a significantly greater reduction in sore throats from surgical as opposed to non-surgical intervention.

Economically, on average, tonsillectomy was more costly and more effective than conservative management in terms of sore throat days avoided and QALYs gained in those with recurrent acute tonsillitis. Tonsillectomy has an 87% probability of being considered cost-effective if we are willing to pay £40 to avoid 1 sore throat day and 100% if we are willing to pay at least £6500 for 1 QALY. Tonsillectomy had a 69% probability of having a higher net benefit (i.e. benefits in monetary units exceed costs) than conservative management. Tonsillectomy, despite the impact of post-operative pain and a post-operative bleeding rate in excess of 20%, had a higher probability of being considered cost-effective than conservative management if we are willing to pay for the extra benefit associated with tonsillectomy.

As this was a pragmatic trial designed to reflect current NHS patient pathways, all patients were required to fulfil UK guidelines for NHS tonsillectomy for recurrent sore throats in adults. These guidelines were, however, based on consensus rather than level 1 evidence.¹⁸ NATTINA was a superiority trial that was designed to meet a commissioning brief set out by the NIHR Health Technology Assessment (HTA) programme (HTA 12/146). The brief stipulated that there should be multidimensional outcomes to assess QoL, the cost-utility of the randomised comparisons, the need for further interventions, and the number of days of sore throat and adverse effects, including primary and secondary haemorrhage.

The proposed initial sample of 510 participants had a 90% power to detect an effect size of 0.33 in a quantitative outcome, assuming a type 1 error rate of 5%. This effect size was based on a number of considerations, including our prior NESSTAC study.^{3,64,89} It allowed for a 25% attrition; that is, to detect the target difference at 90% power, 382 participants needed to provide data for analysis. Recruitment to trials in which interventions are very different, such as between surgical and medical arms, can be difficult.¹²⁷ We applied our prior experience of using a pre-recorded question-and-answer video-recording (covering topics key to patient decision-making prior to joining a surgical trial, and incorporating learning from our feasibility study and PPI group meetings) to support recruitment. This recording was uploaded onto the trial website, used by sites during recruitment conversations with potential participants, and could be accessed by patients prior to clinic visit one.⁷⁷ Recruitment remained slower than projected despite the opening of 27 sites and our efforts to present tonsillectomy as a procedure for which evidence was limited and for which its merits (and harms) compared with conservative management were unclear. In March 2018, a substantial amendment was approved to revise our target number of participants from 510 to 444 (to achieve 85% power to detect the 0.33 effect size planned for), with a decision that recruitment would continue beyond 444, until 30 April 2018, if eligible patients were identified. By the close of recruitment, 455 participants had been recruited and were randomised (two participants were randomised in error, i.e. 453 authentic participants).

Our attrition rate was much lower than the 25% that we had anticipated in the original sample size calculations, in which we assumed that our final analysis would include 382 participants. The observed attrition rate was only 5%, as 95% of those randomised had usable data (i.e. had completed

at least one STAR; see [Appendix 11](#)), resulting in the ITT analysis sample size of 429 participants. Overall, the individual weekly STAR response rate was high, with 72% of all possible weekly STARs returned (see [Table 5](#)), a testament to the value of text message prompts in medical settings generally.¹²⁸⁻¹³⁰

The primary outcome measure was the number of sore throat days during the 24 months following randomisation, collected weekly using the participant's preferred method of communication (e-mail, text message or IVR via telephone).

The baseline characteristics of the NATTINA recruits were balanced across randomised arms (see [Table 4](#)). It is well recognised that tonsillitis in older children and adults has a clear female preponderance,^{131,132} and indeed genetic effects may account for >60% of the variance in susceptibility.¹⁶ Nonetheless, we were struck by the fact that approximately 80% of participants were female. The peak age in early adulthood was to be expected.

The NATTINA participants reported a median of 27 (IQR 12–52) total sore throat days over the full 24-month follow-up. When this primary outcome is compared between the two randomised arms on an ITT basis, a 47% (95% CI 35% to 57%) reduction in sore throats is seen in the tonsillectomy arm. The per-protocol analysis of 224 participants showed a stronger reduction in the number of sore throats of 58% (95% CI 45% to 69%). The adjusted ITT NBR analysis – model 1 – indicated that participants randomised to receive tonsillectomy were estimated to have 0.528 times the total number of sore throat days of the conservative management participants (see [Table 12](#); $p < 0.001$). There was no evidence that baseline severity affected the number of sore throat days (see [Table 15](#)). Likewise, using a continuous rather than a categorical baseline TOI-14 variable in the NBR model did not materially change the results, providing reassurance that the primary analysis was robust.

We observed no clear impact of baseline severity. A weak association ($r = 0.39-0.55$) was previously found between self-report benefit from tonsillectomy on the Glasgow Benefit Inventory and pre-operative severity.¹³³ The best pre-operative predictor was the number of days with throat pain and with fever during the preceding few months but with low precision. The per-protocol and per-treatment analyses were consistent in showing a reduction in sore throats following tonsillectomy: the signal strength of the different analyses are summarised in the forest plots (see [Figures 10 and 11](#)).

The benefits of tonsillectomy are also seen in the secondary outcome measures: TOI-14 scores and QoL measures (SF-12 MCS and PCS).

There were 54 episodes of post-operative haemorrhage reported in 44 participants. This equates to 44 out of 231 participants undergoing tonsillectomy (19%). Of these participants, 37 were reported as SAEs requiring re-admission: nine were reported as mild events, 22 were reported as moderate events and six were reported as severe events. No deaths were reported. Seventeen were recorded as AEs, for which participants did not attend hospital. All episodes of bleeding were managed conservatively with no returns to theatre.

Prior to NATTINA, the level of evidence for tonsillectomy in adults was of low quality,²² a likely contributing factor to variation in tonsillectomy rates.^{24,134,135} There has been a continuing question mark over the value of this procedure, with tonsillectomy for recurrent tonsillitis being placed on a list of Procedures of Limited Clinical Effectiveness (PoLCE), with clear guidance to GPs that referral should take place only when certain stringent criteria are met. Thus, the proportion of adults undergoing tonsillectomy has declined. These changes in practice have occurred despite there being no evidence that the natural history of tonsillitis in the population at large has changed; indeed, *pari passu* the fall in the number of tonsillectomies has been accompanied by an increase in emergency admissions with uncontrolled tonsillitis.⁷⁴

The falling numbers of tonsil operations in England may have been responsible for our observation that NATTINA recruitment suffered from stringent secondary care access criteria applied in varying degrees by participating centres in England. This affected our recruitment projections and led to the abandonment of the formal pilot phase, after discussion with the HTA programme's trial monitoring group. Conversely, the trial benefited from having four active sites in Scotland, for which recruitment proceeded more closely along to that predicted as part of the trial design work. We are indebted, however, to the generous co-operation from the other 23 sites that opened for recruitment.

Crossover

In a surgical trial, it is particularly important to assess the impact of crossover (see [Table 6](#)) owing to the long-acknowledged risk of accepting a null hypothesis if the non-surgical arm ultimately contains too many individuals who have undergone surgery, resulting in an underestimation of the treatment effect size.¹³⁶ In other words, although random allocation eliminates one source of bias it needs to be taken into account as an explanatory factor in the 'real-world' results that we observed.

Around 25% of participants did not receive the treatment that they were randomised to, which meant that some opted to not receive a tonsillectomy and some opted to crossover to tonsillectomy. Despite these crossovers, the ITT, per-protocol and per-treatment analysis all confirm that there was a significant reduction in total sore throats for those randomised to tonsillectomy.

Of the NATTINA participants randomised to the immediate tonsillectomy arm, 26% did not receive the intervention, although only 9% actively requested to crossover to conservative management. Of the 172 participants who underwent tonsillectomy, 95 did so within the specified 8 weeks following randomisation. Of the participants allocated to receive conservative management, 33% requested to crossover, with almost 27% undergoing surgery (see [Table 7](#)). Interestingly, although the numbers requesting crossover were very different between the two arms, the proportions of those who received 'other arm' management were roughly similar.

The impact of the crossover between treatments is to reduce the difference that would be expected to occur between groups. However, our ITT analysis provided evidence that the tonsillectomy arm provided additional benefits compared with the conservative management arm. Further analysis by treatment received explored the extent of the benefit further.

There is evidence to suggest that those with higher rates of sore throats following randomisation were more likely to either crossover to, or remain in, the tonsillectomy arm. Conversely, those with slightly lower sore throat rates were more likely to remain in the conservative management arm or opt out of the tonsillectomy arm. This is intuitively sensible given that those with less severe symptoms are less likely to be willing to undergo an invasive procedure.

Clinical effectiveness

The primary analysis was ITT based on a sample of 429 participants comparing the total number of sore throat days over 24 months from randomisation to the date of last follow-up, while adjusting for the stratification variables (recruiting centre, as a random effect, and baseline severity, as a fixed effect). The data were transformed by dividing each participant's total number of sore throat days by the proportion of the 105 weeks for which they returned data.

Secondary outcome measures

Tonsillectomy Outcome Inventory-14

The TOI-14 questionnaire⁶⁶ allowed us to stratify for baseline severity in the randomisation process. As with the SF-12,⁸³ it was recorded at baseline and at 6, 12, 18 and 24 months post randomisation. The economic evaluation included the Health Service Utilisation questionnaire, which was also recorded at the same time points. The value to trial participants of avoiding a sore throat day was elicited using

a Contingent Valuation Questionnaire, which was administered at baseline. The Time and Travel Questionnaire, which measured participant costs of accessing health care, was completed at 18 months. We also recorded AEs and gathered the views and experiences of patients and clinicians regarding tonsillectomy and conservative management.⁸²

Compliance with questionnaire follow-up was conspicuously less successful (see [Table 6](#)) than the text message weekly reporting of sore throat days. The 18-month postal assessments were completed by the lowest proportion (39%) of participants. We had factored two follow-up face-to-face visits into the study design on the premise that this might encourage retention and compliance particularly within the conservative management arm, so that people in that arm did not feel that they had been ‘forgotten’. However, the inconvenience and financial implications of having to attend a hospital appointment were clearly disincentives. Issues with postal data collection are well known, and over the years that NATTINA ran the likelihood of achieving high return rates with postal questionnaires appeared to us to be decreasing, as the population at large turned more and more to using electronic media in day-to-day life.

NATTINA secondary outcome data collection remained disappointing despite prompts and participant gift vouchers. A Cochrane review of randomised evidence found that monetary incentives and offers of monetary incentives increased postal and electronic questionnaire responses.¹³⁷ A review of non-randomised evidence found low-grade evidence to support the use of telephone or online follow-up compared with postal questionnaire administration.¹³⁸ This same review found similarly limited evidence in support of shortened versions of questionnaires compared with longer versions of questionnaires, electronically transferred monetary incentives compared with cash incentives, cash compared with no incentive; and telephone or text message reminders to non-responders.¹³⁸ However, telephone reminders are not always clearly successful.^{137,139}

Tonsillectomy Outcome Inventory-14 scores were similar for both arms at baseline (see [Table 2](#)). Almost 80% of participants had a baseline TOI-14 total score that placed them in the moderate or severe categories (i.e. TOI-14 score of ≥ 36). The overall mean score was 43.5 (see [Table 17](#)). The tool is a useful metric to assess clinical comparability of tonsillar cohorts, especially when confronting the established geographic variations in the rates of surgical intervention for sore throats. One small comparable surgical cohort from Finland reported a mean TOI-14 pre-operative score of 33 (95% CI 27 to 39).⁸⁴

Tonsillectomy Outcome Inventory-14 scores improved over the 24 months in both randomised arms, with the improvement being more pronounced and earlier in the tonsillectomy arm than in the conservative management arm (see [Table 18](#)), with the difference between arms greatest at 12 months.

Tonsillectomy Outcome Inventory-14 response rates varied among different cohorts (see [Appendix 14](#)). Participants who had tonsillectomy as randomised had the highest response rate (>40% returning four or five TOI-14 questionnaires).

Short Form questionnaire-12 items

One potential issue identified in previous tonsil disease research was that generic health-related QoL questionnaires may not be sensitive enough to capture changes in QoL associated with a sore throat day.⁵ This would not be a limitation of the tool per se, but rather a difficulty of capturing the transitory health effect of a sore throat with a tool that asks only about health over a specific time period. This period is ‘today’ for the EuroQoL-5 Dimensions, ‘the previous week’ for the SF-12 acute version and for ‘the previous four weeks’ for the SF-12 standard version. Within a trial, it is common to administer tools such as the SF-12 (or EuroQoL-5 Dimensions) at set points during the follow-up periods. If the intervals between administration are quite long it is quite possible that an episode of sore throat could occur and fully resolve between these scheduled administrations. It is for this reason that we administered the standard version of the SF-12 (4-week recall) 6-monthly during the trial follow-up and as part of the STAR/STARLET.

Both the SF-12 MCSs and PCSs tended to be higher (indicating better health) in the tonsillectomy arm than in the conservative management arm (SF-12 MCS 3.71 units higher, 95% CI 2.09 to 5.47; SF-12 PCS 2.77 units higher, 95% CI 0.30 to 5.23). Most of the improvement occurred within the first 6 months across both arms.

The scoring and analysis system for the SF-12 questionnaire, despite its shortened length, is relatively cumbersome. Its license is also costly for many settings. Witsell *et al.*¹⁴⁰ used the SF-12 questionnaire in 71 tonsillectomy patients but had only about 40 responders at 6 and 12 months. In this US population, the mean baseline score for SF-12 PCS (47.5) and MCS (42.8) showed both to be slightly more impaired than the NATTINA arms (c. 49 and 46, respectively; see [Table 19](#)). At 6 months, the Witsell *et al.*¹⁴⁰ mean SF-12 MCS was improved by 4 points and the PCS by 7 points. These health-related QoL gains are similar to those in the NATTINA tonsillectomy arm in the ITT analysis (3 and 6, respectively), bearing in mind that 27% of the 233 randomised to tonsillectomy did not undergo the operation, which would tend to reduce the gains.

In a study of knee surgery, the baseline PCS-12 score was around 36.¹⁴¹ In other words, as one might expect, there is a higher impact on life quality when mobility is impaired than in the tonsillectomy young adult population where the mean was close to 50. Conversely, in the NATTINA participants, the mean baseline SF-12 MCS was only 45, whereas the MCS of the orthopaedic surgery group pre-operatively was clearly higher: 54. Some of these differences might also reflect the episodic nature of tonsillitis (see [Economic evaluation](#)) and the intervening generally good health of young adults. However, many coming forward for surgery articulated concerns about the implication of recurrent absences on academic and employment records.⁸² Patients reported that they would not consider taking part in NATTINA because they did not want the risk of being randomised to deferred surgery:

It [the tonsillitis] had too much of an impact. It was happening at least twice a month as well, so it was really interfering with my attendance and stuff, and work, and money.

McSweeney *et al.*⁸²

Safety

There were 244 AEs reported by 119 participants, and 51 SAEs, which were defined as potentially life-threatening, requiring prolonged in-hospital stay or requiring re-admission for treatment. There were 47 re-admissions to hospital in the 231 participants undergoing tonsillectomy (20%). This is at the upper end of the re-admission rate in a Scottish national survey.⁸ Four participants had a prolonged in-hospital stay following surgery. There were 54 episodes of post-operative haemorrhage reported in 44 participants. This equates to 44 out of 231 participants undergoing tonsillectomy (19%). Of these participants, 37 (15% of those receiving tonsillectomy) required re-admission. This compares, for example, with 0.5% in a 2009 Norwegian survey.¹⁴² Thus, overall, the rates appear high, but the use of telephone call check-ups at 1 and 2 postoperative weeks served to maximise ascertainment of the true rate of operative complications. Supplementing mailed questionnaires with telephone interviews may increase the validity of surgical outcome studies.³² The haemorrhage rate observed in NATTINA is important supporting information to inform the content of doctor-patient shared decision-making going forward, given that most published rates, which are reliant on patient self-report, show lower haemorrhage rates.

In a regional study in New Zealand,¹⁴³ there was a time trend of increasing haemorrhage rates, which the authors postulate may relate to a transition to greater use of diathermy technique and a routine use of post-operative non-steroidal anti-inflammatory drugs and steroids.¹⁴³ However, a meta-analysis failed to link peri-operative steroids with bleeding risk.¹⁴⁴ The Cochrane review²² of the use of steroids in controlling post-operative symptoms, such as pain, nausea and vomiting, included over 6000 participants in 64 studies. Intravenous steroids produced a clinically significant reduction in post-operative nausea and vomiting (adult odds ratio 0.32, 95% CI 0.16 to 0.67). Steroids administered intravenously or locally reduced immediate post-operative pain, but there are very few local steroid studies in adults.

Haemorrhage rates were unaffected.³⁹ A systematic review and meta-analysis on treatments for post-tonsillectomy pain in adults and adolescents identified 29 RCTs with over 1800 subjects. The authors of this review concluded that use of several analgesics produced greater pain relief than single agents.⁴⁰

Economic evaluation

Across the analyses, there was evidence of a difference in costs, number of sore throat days and QALYs between the randomised arms. The tonsillectomy arm was, on average, more costly than the conservative management arm in terms of costs to the NHS. The mean (SD) cost for all NHS health care was higher in the tonsillectomy arm [£1358 (£720)] than in the conservative management arm [£858 (£814)]. When we considered participant costs, tonsillectomy was less costly, on average, than conservative management. This aligns with the results of a prior small German study of 97 participants.¹⁴⁵

On average, those in the tonsillectomy arm experienced fewer sore throat days than those in the conservative management arm [mean difference -21.27 (SE 6.24), 95% CI 9 to 34] (see [Table 22](#)).

On average, those in the tonsillectomy arm reported more QALYs than those in the conservative management arm [mean difference 0.06 (SE 0.04), 95% CI -0.03 to 0.15] (see [Table 22](#)).

If we did not value avoiding a sore throat day or gaining a QALY, from the perspective of the NHS and PSS tonsillectomy would have a 0% probability of being considered cost-effective. However, as the importance of avoiding a sore throat day or gaining a QALY increases (i.e. the threshold value for an additional unit of a sore throat day avoided or a QALY gained increased), so does the probability of tonsillectomy being considered cost-effective. For example, should society be willing to pay £40 to avoid a sore throat day, tonsillectomy would have an 87% probability of being considered cost-effective. Similarly, in the CUA, should society be willing to pay £5000 for an additional QALY, tonsillectomy had an 87% probability of being considered cost-effective, and this probability increased to 100% as the threshold value for an additional QALY increased to ≥£6500 and above.

For the CBA, the results of the Contingent Valuation Questionnaire were combined with information on cost. This analysis found that, on average, tonsillectomy was associated with a higher net benefit than conservative management; however, when taking into account the variability in cost and WTP estimates among trial participants, tonsillectomy had a 69% probability of having a higher net benefit. This is lower than that seen in the CUA and is caused by the comparatively wide CI surrounding the estimate of WTP.

As with any economic analysis, sensitivity analysis is essential. The statistical imprecision surrounding estimates was explored using a stochastic sensitivity analysis, which used bootstrapping techniques. This analysis was used to estimate the probability that tonsillectomy would be considered cost-effective compared with conservative management. Other forms of uncertainty were explored using deterministic sensitivity analyses. In all of these sensitivity analyses our conclusions remained unchanged, except when costs falling on participants were considered. When we considered the direct and indirect costs associated with sore throat days to participants, tonsillectomy dominated conservative management as it was, on average, less costly and more effective. In this analysis, even if we were not willing to pay for an additional QALY, tonsillectomy had a 98% probability of being considered cost-effective. In the base-case analysis, which did not consider utility decrements associated with a sore throat episode, there were higher average QALYs reported in the tonsillectomy arm than in the conservative management arm [mean difference 0.11 (SE 0.02), 95% CI 0.08 to 0.14] as expected given that more sore throat days were reported in the conservative management arm. This difference in QALYs was maintained when utility decrements associated with a sore throat episode were included [mean difference 0.12 (SE 0.02), 95% CI 0.08 to 0.15]. Our results suggest that generic QoL questionnaires, particularly the SF-12, are suitable to capture changes in QoL in individuals suffering with recurring sore throats.

There is a substantial economic cost associated with managing recurring sore throats incurred not only by the healthcare provider but also by the individuals suffering from sore throats.^{5,35,49,71} Costs incurred by the individual include direct (e.g. medication) and indirect costs (e.g. time away from usual activities owing to sore throat episode or tonsillectomy). In our analysis, when costs incurred by participants were considered, tonsillectomy was less costly than conservative management (98% probability, which is analogous to a p -value of 0.02 for a one-sided test). This finding is similar to that of Burns *et al.*,⁵ whose intervention (dexamethasone) was, on average, more costly from a healthcare provider perspective, but less costly after incorporating participant costs. This highlights the importance of considering costs incurred by participants in evaluating treatments to manage sore throats.

Qualitative study

Patients reported having to request an ENT referral despite not formally being recognised by primary care services as eligible because their episodes of tonsillitis had not been officially documented in primary care records. Several ENT staff felt that GP commissioning groups were not implementing SIGN guidelines appropriately, that is as guidelines only, and were instead using the criteria as *sine qua non* qualifiers for access to tonsillectomy within the NHS.² Some secondary care systems did not accurately direct GP correspondence to the correct individual, which reduced the recorded 'episode count' within the patient's record.

Patients who were unsure about having surgery or who were considered to be more 'health aware' were thought to be more likely to participate in the trial. Most patients who declined to participate did so mainly because they had suffered a lot of tonsillitis and did not want to risk being randomised into the conservative management arm. Conversely, some patients reported that they thought that agreeing to participate was the only way to be offered surgery.

General practitioners and ENT staff both suggested that patients randomised to conservative management might have increased expectations of access to GP consultations and antibiotic prescribing. Some patients who declined randomisation cited the lack of any 'special treatment' in the conservative management arm: '[b]ut you didn't get any special treatment, in a way. You didn't get a fast track to antibiotics or to medical help, it was just like, right, back you go, out into the world and cope with it' (example quotation from a non-randomised participant). Our original study design split the conservative management arm into a care-as-usual subgroup and an early maximum medical therapy subgroup with ready access to a package of steroids and antibiotics. This was rejected by the funder. In retrospect, its inclusion might have been counter-productive because of the greater overall sample size required adequately to power NATTINA with this additional subgroup. However, the strong evidence of benefit from tonsillectomy in patients with recurrent tonsillitis and the compelling accounts of patients about the impact of their tonsillitis raises the question as to whether or not this group of patients would benefit from earlier identification and about their optimal management in primary care. Patients who had received surgery were unanimously happy to have done so, which is in line with prior descriptive satisfaction scores with tonsil surgery.²⁰

Strengths and limitations

Strengths

Trial processes were generally deemed to be acceptable and few sites experienced barriers to treatment. The popularity of our use of technology to collect data was reflected in our respective response rates to texts compared with paper questionnaires. Drop-out rates and missing primary outcome data were higher in the conservative management arm than in the tonsillectomy arm, and the NATTINA statistical

analysis accounted for this discrepancy. The additional analyses undertaken supported the primary ITT analysis conclusion. Both the per-protocol and per-treatment analyses are in concordance with the ITT results, demonstrating the reduction in sore throat days in the tonsillectomy arm. The counting of sore throat days, as specified by the funder and as recommended by Cochrane,²² proved a valuable pragmatic metric for a large clinical trial such as NATTINA.

To our knowledge, this is the first economic evaluation investigating the cost-effectiveness of tonsillectomy compared with conservative management in the management of recurring sore throats in adults, a major NATTINA strength. Although the response rates to the data collection tools used in the economic evaluation were lower than anticipated, this was not surprising given the average age of the participants. A similar trend was seen by Bhattacharyya and Kepnes.⁷¹

Our embedded economic evaluation considered costs incurred by the NHS, Personal Social Services and participants, and a wide range of benefits (clinical effectiveness, QoL and patient benefits measured in monetary terms). All three strands were concordant and were further supported by the results of our robust sensitivity analyses that actively sought to explore the impact of changes that could change conclusions. Another strength of the economic evaluation was how the impact of sore throat days on participants' well-being was incorporated into the analyses. First, participants completed the SF-12 questionnaire when they reported a sore throat episode, and a utility decrement associated with each sore throat episode was integrated into the QALY equation. In this analysis, we identified the negative effect that sore throat days have on participants' QoL, as the average total QALY per randomised arm was lower than the average total QALY per randomised arm reported in the base-case analysis [base-case analysis mean QALYs: conservative management 1.44 (SD 0.17), tonsillectomy 1.55 (SD 0.15); sensitivity analysis incorporating sore throat days mean QALYs: conservative management 1.38 (SD 0.18), tonsillectomy 1.50 (SD 0.18)]. Second, participants completed a contingent valuation questionnaire at baseline, which assigned a value on the effect that a sore throat day has on their overall well-being. The results of this questionnaire helped us to quantify the burden of sore throat days not only on participants' health-related QoL but also on all aspects of their life. It also puts the results of the CEA into context for decision-making and was used in the CBA to quantify the reduction in patient benefit associated with sore throat days.

One of the strengths of NATTINA is the number of TOI-14 data that we generated, with well over 900 questionnaires available, roughly equally distributed between trial participants and those declining participation. Those declining the trial had a higher proportion of severe tonsillitis scores, although severity categories were represented within NATTINA. [Appendix 3, Figure 19](#), shows box plots of TOI-14 scores for patients declining and accepting to enter NATTINA.

Limitations

There appeared some evidence that those with the most severe extent of disease were reluctant to enter the trial (see [Appendix 3, Table 31](#), for reasons for declining). Screening log data showed that 1002 out of 1403 (71%) of those giving a reason for declining to participate did so because they wanted to receive a tonsillectomy. We must accept that some symptoms may not be due to tonsillitis, although in a paediatric study, parents appeared clearly to be able to distinguish between the two.¹⁴⁶

Despite the TMG strongly advising to have dedicated clinics for recruitment to the trial, some centres chose not to run these and tended to have not as good recruitment as a result. Staff who lacked equipoise struggled to recruit to the trial, even when patients themselves were in a position of equipoise. All non-recruiting sites were typically research active and displayed concern at their lack of recruitment for NATTINA; therefore, regional variation in referral processes and access to tonsil surgery should also be acknowledged. Clinical research network funding supports nurses for portfolio studies; however, there is a high turnover of staff and prioritisation is a constant battle, particularly for a relatively low headcount recruiting study, such as a surgical trial.

Another challenge for surgical trials is the need for the operative cohort participants to pause their normal life processes for at least 1 week, if not more, which is not the case for medical treatment trials. ENT staff alluded to having to negotiate surgery slots and compromising with patients over preferred surgery dates, which at times meant that they had to deviate from trial protocols.

Not all patients were offered, or chose, to watch the trial recruitment video, which may have contributed to patients' lack of understanding.

The ITT analysis is likely to offer a conservative underestimate of the true impact of tonsillectomy in reducing sore throat days, as a result of patients crossing over to receive tonsillectomy.

Based on early feedback on non-compliance with the STAR form questionnaires, the STARLET questionnaires were introduced. Unfortunately, a number of participants had changed address or other contact details but did not know how to pass this on to the trial teams, so were not able to receive questionnaires in the post or appointments for follow-up 12-/24-month visits. Simple measures to capture changes in contact details might have further improved response rates in such a young, transient study population. Overall, the design required several amendments to the methodology as lessons were learnt, notably from the response patterns of young adults to information-gathering. Most were incorporated with minimal impact on data robustness, but each was time-consuming for the NCTU staff.

The qualitative study uncovered problems of communication at the primary–secondary care interface, including issues in informing the correct GP that their patient was involved in the trial.

One of the main challenges of the economic evaluation was the progressive loss of data over the 2-year follow-up period. Utilisation data responses were near complete at baseline: 98% for conservative management and 99% for tonsillectomy. These figures reduced to 53% and 52%, respectively, at 12 months, and to 45% and 42%, respectively, at 24 months. SF-12 data responses were similar. The poor response rate to the STAR observed during the trial required its refinement into the STARLET and assumptions were made to maximise the data available, which included creating cost and utility tariffs that were assigned to each sore throat day depending on the self-reported severity of the sore throat episode. In doing so, we estimated the variability in costs and utilities based on clinical severity.

Existing research into medical management of sore throat has primarily focused on one-off episodes of sore throat and, therefore, is of less value in understanding how best to support a recurrent tonsillitis population. The strong evidence of benefit from tonsillectomy in patients with recurrent tonsillitis and the compelling accounts of patients about the impact of their tonsillitis raises questions as to whether or not this group of patients would benefit from earlier identification and about their optimal management in primary care. Optimal strategies for non-surgical management of recurrent tonsillitis should be explored in future research.

Conclusions

NATTINA has shown that tonsillectomy in adults is a clinically effective and cost-effective intervention compared with conservative management in patients over 16 years with recurrent sore throat. Participants with recurrent tonsillitis who met current NHS guidelines to undergo tonsillectomy experienced significantly fewer sore throat days over 24 months if they underwent tonsillectomy than similar participants who were treated conservatively. NATTINA participants in the tonsillectomy arm reported, on average, fewer healthcare contacts, fewer sore throat days and higher QALYs, which resulted in them having less time away from work and/or usual activities during follow-up. Bhattacharyya and Kepnes⁷¹ came to a similar conclusion and stated that the clinical decision-making process for tonsillectomy needs to consider not only health and healthcare improvements but also economic improvements.

Implications for practitioners and health services

The SIGN adult guidance states 'A Cochrane review found limited evidence of benefit of tonsillectomy in adults. In adults with proven recurrent group A streptococcal pharyngitis GAHSP, a small well conducted RCT demonstrated benefit for tonsillectomy in adults. Tonsillectomy reduced the incidence of GAHSP in the 90-day postoperative period with a number needed to treat (NNT) of 5', but concluded that, although 'tonsillectomy is recommended for recurrent severe sore throat in adults', both study numbers and follow-up duration were deficient' [© SIGN 2010. This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license, which permits others to copy and redistribute this work, for non-commercial use, provided the original work is properly cited. See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>].¹⁸ It is helpful to reflect on the NATTINA results in the light of the findings of the 2014 Cochrane review,²² which concluded that there was insufficient information available on the effectiveness of adenotonsillectomy/tonsillectomy compared with non-surgical treatment in adults to draw a firm conclusion. The pooled data from two studies in 156 adult participants found a mean of 10.6 days fewer in those receiving surgery. However, there was significant statistical heterogeneity in the analyses and only short-term data available, with poor detail on complications. Given the numbers of NHS tonsillectomy operations still carried out, despite restricted access, the decision by NIHR to fund a trial on this procedure was clearly an important one and the NATTINA results justify the research investment, which represents only a small fraction of the current annual costs of tonsil disease nationwide.

Despite the fact that there appeared to be some evidence that those with the most severe extent of disease were reluctant to enter the trial through lack of equipoise in favour of surgery, there is compelling evidence, especially when patient costs are included, that adult tonsillectomy for recurrent tonsillitis is a worthwhile procedure.

Issues of research implementation in tonsillectomy have been highlighted in a recent review.¹⁴⁷ Within the UK, tonsillectomy is listed as a 'procedure of limited clinical value'; NATTINA quantifies the clinical effectiveness and cost-effectiveness of tonsillectomy in adults with recurrent tonsillitis. Pre NATTINA, UK guidelines were a translation of level 1 evidence in children, applied to adults. Access to tonsillectomy in the UK was governed by application of national guidance, which is predicated on a qualifying number of episodes of tonsillitis.¹⁸ However, the frequency of episodes was an arbitrary number established by Paradise *et al.*¹⁴⁸ who used seven episodes in 1 year, five in 2 years and three in 3 years as an entry criterion in a trial in 1971. Ever since, there has effectively been reiteration of tonsil evidence gathering on this ad hoc basis ever since.

It is worth noting the socioeconomic gradient of tonsillitis, as the economic analysis indicates that increased use of surgery would narrow health and socioeconomic inequalities. Contemporary data are lacking, but an early study suggested that tonsillitis patients who were more deprived were more likely to consult but less likely to be referred to hospital.¹⁴⁹ We have identified communication issues at the primary-secondary care interface in our qualitative work. Barriers to accessing secondary care services appeared in one study to translate into adult patients having on average 27 episodes of tonsillitis over 7 years prior to referral to secondary care.² This goes far beyond the three episodes per annum for 3 years articulated in the SIGN advice.¹⁸ 'Guidelines' are being operationalised as stringent provider-driven thresholds, rather than as the focus of patient-centred shared decision-making. Thus, although NATTINA is the first definitive trial to demonstrate that tonsillectomy performed in accordance with the current UK national guidelines is cost-effective for patients and the NHS, the question remains as to whether or not the threshold is set too high and whether or not people with recurrent tonsillitis below the threshold would also benefit. The NATTINA results indicate that consideration should be given to an urgent review of the guidance to incorporate the evidence from NATTINA.

The implementation of guidance predicated on counting of episodes can also be problematic. Patients may not have exactly the same number of episodes on each successive year, but one or two of those episodes might be extremely severe or indeed have led to them being hospitalised. Therefore, many ENT surgeons would favour integration of a severity metric along with the frequency metric and the integration of both of these parameters with individual circumstances in an updated decision-making process. There are multiple difficulties in identifying episodes of sore throat. Asking people to reflect on the number of sore throat episodes that happened up to 5 years ago is an unrealistic demand. Documentation of sore throats by healthcare services is affected by patients' awareness of the variable access to antibiotic prescription. Some patients will conclude that there is no point in attending primary care if they know that the recommendation will be for over-the-counter therapy. Conversely, certain patients will actively seek appointments if they become aware that episodes have to be documented to qualify for an ENT appointment. Out-of-hours healthcare service provision occurs at multiple locations. All such factors distort our understanding of the natural history of recurrent sore throats in adults. In the NESSTAC trial,⁶⁴ it was also found to be difficult when looking at diary reports of sore throats experienced to define when one episode stopped and another episode started. Many episodes have a remission and then flare up again: should that be counted as two episodes or as one?

Implications for patients and the public

Therefore, there is a need to convert the findings of NATTINA into a practical decision support tool for patients and surgeons.¹⁵⁰ This is a well-recognised model for implementing medical evidence. At present, the opportunity for shared or clinical decision-making is compromised by the rigid application of NHS tonsillectomy guidelines as a rationing threshold.¹⁸ As detailed above, we question the predication of UK guidelines on 'episodes' of sore throat. Individual funding requests have to be submitted on the basis that the beneficiary is uniquely disadvantaged by the nature of the disease; however, this is a hard argument to make given the high prevalence of tonsillitis in the adult population. The NHS individual funding request guide for patients states the following:

In the application, your clinician will need to explain why your clinical circumstances are different and show all available clinical evidence for why they believe you would benefit more from the treatment than other patients with the same condition.

NHS England¹⁵¹ (Contains public sector information licensed under the Open Government Licence v3.0)

Despite the fact that so many thousands of tonsil operations are still carried out in the NHS, there is also a widespread belief that 'they don't do that anymore' among both clinicians and patients; is it time to rehabilitate the reputation of tonsillectomy despite its obvious drawbacks? At the very least, a more balanced approach to decision-making is now possible given that we have clinical and economic evidence on which to predicate such decision processes.

NATTINA has justified the current guidelines as identifying a level at which tonsillectomy is cost-effective, but this is not necessarily the minimum level at which the procedure can be justified on health and economic grounds. The threshold of symptoms that merit offering an adult patient tonsillectomy is, therefore, likely to be lower than those currently defined.

Implications for research

The top research priority to emerge from NATTINA is to determine the optimum timing of tonsillectomy in adults with recurrent acute tonsillitis.

Such work would have several strands, including how to incorporate the experiences of those with the most severe level of disease who, although represented within the trial cohort, were disproportionately observed in the declining cohort. The magnitude of the clinical effect observed supports this additional research to refine the level of disease burden that merits tonsillectomy. Work is also required to optimise metric for disease burden severity, and to exploit the novel real-time data collection methods elaborated in the NATTINA design. The counting of sore throat days was a useful metric in NATTINA. Although our design used weekly text prompts to collect such details, a study application (app) might be a less labour-intensive portal for daily sore throat data collection.

Guideline revision, in particular how guidelines translate into healthcare commissioning, could follow directly from NATTINA after some qualitative follow-up with relevant stakeholders.

Those with expertise in shared decision-making could begin to form a decision support tool taking account of the novel NATTINA level 1 evidence base until threshold optimisation can be delivered by a future trial.

There is also a need to better understand optimum treatment strategies for tonsillitis in primary care. Even among those patients meeting SIGN guidance, some patients were reluctant to undergo a surgical procedure; there will always be some patients with severe recurrent sore throats who will require conservative management. There is an established reluctance, for well-rehearsed reasons, to overuse antibiotics.⁵¹ However, there is recognition that a delayed antibiotic prescription does have a clinical effect.⁶¹ In our qualitative study, some patients voiced a desire to be able to access fast-track antibiotics when their lived experience of recurrent sore throat indicated that this was necessary; previous research has indicated that patients' own self-assessment of sore throat severity was a good predictor of a bacterial sore throat.⁷² The place of oral steroids, if any, remains to be fully understood. At 48 hours, a single dose of oral steroid appeared of some clinical,⁶² but not economic,⁵ benefit. The relevant Cochrane review concluded that oral or intramuscular corticosteroids, in addition to antibiotics, moderately increased the likelihood of both resolution and improvement of pain in participants with sore throat, but that further research into the harms and benefits of short courses of steroids was needed to permit informed decision-making owing to their limited benefit.⁶³

NATTINA makes an important contribution to the health-related QoL tonsillectomy literature, the amount of which remains relatively low given the number of surgical interventions. More work is required to expand on the conclusions of a review that patients with coexisting chronic conditions are likely to benefit less, and that younger patients and those with more severe infective tonsillitis are likely to benefit more.¹⁵ Further research has also been called for into the potential value of community point-of-care testing.⁴⁸

Acknowledgements

The NATTINA trial was possible only because of the generous support and enthusiasm of a large number of individuals.

The investigators would also like to acknowledge the NIHR Clinical Research Networks that provided NHS research support to the study.

Clinicians and research staff from across the UK contributed to the study at individual sites, and we express our gratitude to all of the staff at sites in the UK for their support of study recruitment, intervention delivery and data collection. We would like to thank the following centres: Aneurin Bevan University Health Board; Bradford Teaching Hospitals NHS Foundation Trust; City Hospital Sunderland NHS Foundation Trust (now 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 & Arran; NHS Grampian; NHS Greater Glasgow and Clyde; NHS Tayside; North Cumbria University Hospitals NHS Trust – Cumberland Infirmary (now 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.

We would also like to thank all of the site staff who have taken part in the qualitative interviews.

We would also like to thank the NCTU staff, including previous team members; trial managers Amy Cranston, Isabel Rubie, Nicola Goudie, Faye Wolstenhulme, Rebecca Forbes and Georgiana Browne; clinical trial administrator Julia Phillipson; secretaries Janet Jobling and Amy Collins; data manager Laura Simms; database officer Andy Cutts; and senior trial managers Lesley Hall, Chris Speed and Jared Thornton.

Furthermore, we would like to thank the health economists Peter McMeekin (study co-applicant) and William King, and the statisticians Sean Hiu and Nick Steen (study co-applicant).

We would like to thank Paul Sergeant for his continued support with management of the trial primary outcome.

We would like to thank all of the members of the PPI group who have provided helpful advice and guidance when reviewing the trial and patient-facing documents.

We would like to thank the members of the NATTINA TSC: Professor John Birchall (chairperson); Miss Victoria Ward (independent clinician); Professor Catherine Hewitt (independent statistician); and Ms Sabena Monk, Mr Jim Ainslie and Mr Ian Armstrong (independent lay members).

We thank the members of the NATTINA DMC for all their valuable guidance: Mr Andrew Swift (chairperson), Professor Tim Woolford (independent clinician) and Professor Robert West (independent statistician).

Finally, but most importantly, we would like to thank every participant who took part in the NATTINA trial.

Contributions of authors

Janet A Wilson (<https://orcid.org/0000-0002-6416-5870>) (Professor of Otolaryngology) was the trial chief investigator, provided substantial input to the design and execution of the trial, authored the introductory and discussion chapters of the final report and had substantial input into review of all other sections of the final report.

Tony Fouweather (<https://orcid.org/0000-0002-2292-0495>) (Trial Statistician) co-wrote the statistical analysis plan, carried out the main trial analysis, wrote the results chapter of the final report, and reviewed and contributed to the final report.

Deborah D Stocken (<https://orcid.org/0000-0001-8031-1738>) (Professor of Clinical Trials Research, Statistical Co-applicant) contributed to the trial design, protocol and statistical analysis plan; was a co-applicant for funding; interpreted trial data; and reviewed and contributed to the final report.

Tara Homer (<https://orcid.org/0000-0002-6664-0671>) (Health Economist and Senior Research Associate) wrote the health economics analysis plan, conducted the health economics analysis, wrote the economic evaluation chapter and reviewed and contributed to the final report.

Catherine Haighton (<https://orcid.org/0000-0002-8061-0428>) (Professor of Public Health) led the qualitative aspects of the trial, contributed to the design of the trial, was a co-applicant for funding, interpreted qualitative data, wrote the qualitative chapter of the final report, and reviewed and contributed to the final report.

Nikki Rousseau (<https://orcid.org/0000-0001-8826-3515>) (University Academic Fellow in Healthcare Technology Evaluation, Co-Applicant) contributed to the design of the protocol, was a co-applicant for funding, contributed to the management and conduct of the study, and co-wrote the final report.

James O'Hara (<https://orcid.org/0000-0002-4096-3296>) (Consultant Otolaryngologist, Head and Neck surgeon, Co-Applicant) was a co-applicant for funding, contributed to the trial design and costings, contributed to trial recruitment, aided in the analysis of the results, and reviewed and contributed to the final report.

Luke Vale (<https://orcid.org/0000-0001-8574-8429>) (Professor of Health Economics, Co-Applicant) contributed to the design of the protocol, was a co-applicant for funding, contributed to the design of the health economics analysis plan, and reviewed and contributed to the final report.

Rebecca Wilson (<https://orcid.org/0000-0002-7695-4375>) (Trial Manager) was trial manager for the study, provided day-to-day management of trial conduct, performed monitoring to ensure that the trial was conducted to good clinical practice requirements, and edited and co-wrote the final report.

Sonya Carnell (<https://orcid.org/0000-0002-7648-4297>) (Senior Trial Manager) led the NCTU team's involvement in the trial and reviewed and edited the final report.

Scott Wilkes (<https://orcid.org/0000-0003-2949-7711>) (Professor of General Practice and Primary Care and Head of Sunderland School of Medicine, Co-applicant) contributed to the design of the protocol, was a co-applicant for funding, and reviewed and contributed to the final report.

Jill Morrison (<https://orcid.org/0000-0003-3965-1516>) (Professor of General Practice, Clerk of Senate and Vice Principal, Co-Applicant) provided advice about liaison and communication with primary care during the trial, was a co-applicant for funding, and reviewed and contributed to the final report.

Kim Ah-See (<https://orcid.org/0000-0002-7085-7486>) (Consultant ENT & Head and Neck Surgeon, Co-Applicant, Site PI) was a co-applicant for funding, contributed to trial recruitment, and reviewed and contributed to the final report.

Sean Carrie (<https://orcid.org/0000-0003-1722-1814>) (Consultant Otolaryngologist and Head and Neck Surgeon, Co-Applicant, Site PI) contributed to the design of the protocol, was a co-applicant for funding, contributed to trial recruitment and reviewed the final report.

Claire Hopkins (<https://orcid.org/0000-0003-3993-1569>) (Consultant ENT Surgeon and Professor of Rhinology, Site PI) contributed to the design of the protocol, contributed to trial recruitment and reviewed the final report.

Nicola Howe (<https://orcid.org/0000-0001-9446-8314>) (Database Manager) was the database manager for the study, designing the database and monitoring data entered by sites for the duration of the study, as well as cleaning and preparing data in preparation for analysis, and for summarisation and presentation to the TMG as and when required, and she contributed to the final report.

Musheer Hussain (<https://orcid.org/0000-0002-3169-3087>) (Consultant ENT Surgeon, Co-Applicant, Site PI) was a co-applicant for funding, contributed to trial recruitment, and reviewed and contributed to the final report.

Lyndsay Lindley (<https://orcid.org/0000-0002-5982-9679>) (Qualitative Research Associate) conducted the follow-up qualitative interviews and analysed the interview data for the qualitative trial.

Kenneth MacKenzie (<https://orcid.org/0000-0002-2952-7717>) (Consultant Otolaryngologist and Head and Neck Surgeon, Site PI) contributed to trial recruitment, and reviewed and contributed to the final report.

Lorraine McSweeney (<https://orcid.org/0000-0003-1044-6201>) (Qualitative Research Associate) conducted the qualitative interviews, analysed the interview data for the feasibility and pilot studies, and reviewed and contributed to the final report.

Hisham Mehanna (<https://orcid.org/0000-0002-5544-6224>) (Chairperson of Head and Neck Surgery, Co-Applicant, Site PI) provided input into the origination of the trial (grant writing and planning of the trial), was a co-applicant for funding, contributed to trial recruitment, and reviewed and contributed to the final report.

Christopher Raine (<https://orcid.org/0000-0003-0839-0878>) (Consultant ENT Surgeon, Co-Applicant, Site PI) was a co-applicant for funding, contributed to trial recruitment, and reviewed and contributed to the final report.

Ruby Smith Whelan (<https://orcid.org/0000-0003-4389-8685>) (Clinical Trials Administrator) provided high-level administrative support for the trial, aided management of the trial, and reviewed and contributed to the final report.

Frank Sullivan (<https://orcid.org/0000-0002-6623-4964>) (Professor of Primary Care Medicine, Co-Applicant) contributed to the design of the protocol, was a co-applicant for funding, interpreted trial data, and reviewed and contributed to the final report.

Alexander von Wilamowitz-Moellendorff (<https://orcid.org/0000-0002-8735-1879>) (Senior Clinical Research Associate, former Trial Manager) was trial manager from April 2015 to February 2018 for the trial, provided day-to-day management of trial conduct, performed monitoring to ensure that the trial was conducted to good clinical practice requirements, and reviewed and contributed to the final report.

Dawn Teare (<https://orcid.org/0000-0003-3994-0051>) (Trial Senior Statistician) edited the statistical analysis plan, led the final study analysis, and co-wrote and reviewed the final report.

Publications

Rubie I, Haighton C, O'Hara J, Rousseau N, Steen N, Stocken D, *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).

Haighton CA, Wilson J. 2015. Tonsillectomy or adenotonsillectomy reduces the number of sore throats in children however insufficient information is available on the effectiveness in adults. *BMJ Evid Based Med* 2015;**20**(2).

McSweeney L, Wilson J, Rousseau N, Wilkes S, Haighton C. Stakeholders' views of recurrent sore throat, tonsillitis and their management: a qualitative interview study (NATTINA Part 1). *Clin Otolaryngol* 2016;**42**:301–6.

McSweeney L, O'Hara J, Rousseau N, Stocken D, Steen N, Sullivan F, *et al.* 'Thinking that somebody's going to delay [a tonsillectomy] for one to two years is quite horrifying really': a qualitative feasibility study for the NAtional Trial of Tonsillectomy IN Adults (NATTINA Part 2). *Clin Otolaryngol* 2016;**42**:578–83.

McSweeney LA, Wilson JA, Wilkes S, Haighton CA. 2018. Is SIGN guidance for GP management of tonsillitis suitable? A qualitative study. *Fam Pract* 2018;**35**:633–7.

Data-sharing statement

Anonymised data from this trial may be available to the scientific community subject to regulatory and ethics approval. Requests for data should be directed to the corresponding author.

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it is important that there are safeguards to make sure that they are stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: <https://understandingpatientdata.org.uk/data-citation>.

References

1. Bhattacharyya N, Kepnes LJ, Shapiro J. Efficacy and quality-of-life impact of adult tonsillectomy. *Arch Otolaryngol Head Neck Surg* 2001;**127**:1347–50. <https://doi.org/10.1001/archotol.127.11.1347>
2. 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. <https://doi.org/10.1111/coa.12850>
3. Lock C, Wilson J, Steen N, Eccles M, Brittain K, Carrie S, et al. Childhood tonsillectomy: who is referred and what treatment choices are made? Baseline findings from the North of England and Scotland Study of Tonsillectomy and Adenotonsillectomy in Children (NESSTAC). *Arch Dis Child* 2010;**95**:203–8. <https://doi.org/10.1136/adc.2009.165530>
4. Mooney H. Cut useless medical treatments, says Audit Commission. *BMJ* 2011;**342**:d2438. <https://doi.org/10.1136/bmj.d2438>
5. Burns RM, Wolstenholme J, Jawad S, Williams N, Thompson M, Perera R, et al. Economic analysis of oral dexamethasone for symptom relief of sore throat: the UK TOAST study. *BMJ Open* 2018;**8**:e019184. <https://doi.org/10.1136/bmjopen-2017-019184>
6. Mantzourani E, Evans A, Cannings-John R, Ahmed H, Hood K, Reid N, et al. Impact of a pilot NHS-funded sore throat test and treat service in community pharmacies on provision and quality of patient care. *BMJ Open Qual* 2020;**9**:e000833. <https://doi.org/10.1136/bmjoq-2019-000833>
7. Curtis L, Burns A. *Unit Costs of Health and Social Care 2020*. Canterbury: PSSRU, University of Kent; 2020.
8. Milner TD, Hilmi O, Marshall J, MacKenzie K. Pan-Scotland tonsillectomy outcomes: a national cross-sectional study. *Clin Otolaryngol* 2021;**46**:138–45. <https://doi.org/10.1111/coa.13608>
9. Al-Hussaini A, Owens D, Tomkinson A. Health costs and consequences: have UK national guidelines had any effect on tonsillectomy rates and hospital admissions for tonsillitis? *Eur Arch Oto-Rhino-L* 2013;**270**:1959–65. <https://doi.org/10.1007/s00405-013-2345-z>
10. NHS Digital. *Hospital Admitted Patient Activity 2018–19*. 2019. URL: <https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/2018-19#> (accessed September 2021).
11. NHS Digital. *Hospital Admitted Patient Care Activity 2019–20*. 2020. URL: <https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/2019-20#> (accessed 17 November 2020).
12. 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. *Clinical Otolaryngology* 2019;**44**:7–13. <https://doi.org/10.1111/coa.13233>
13. NHS Digital. *Hospital Episode Statistics, Admitted Patient Care – England, 2009–10*. 2010. URL: <https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/hospital-episode-statistics-admitted-patient-care-england-2009-10#> (accessed 1 December 2020).
14. Windfuhr JP, Chen Y-S. Do changing trends in tonsil surgery affect hemorrhage rates? A longitudinal study covering 1,452,637 procedures. *Eur Arch Oto-Rhino-L* 2019;**276**:2585–93. <https://doi.org/10.1007/s00405-019-05532-3>

15. Andreou N, Hadjisymeou S, Panesar J. Does tonsillectomy improve quality of life in adults? A systematic literature review. *J Laryngol Otol* 2013;**127**:332–8. <https://doi.org/10.1017/S0022215113000273>
16. Kvestad E, Kvaerner KJ, Roysamb E, Tambs K, Harris JR, Magnus P. Heritability of recurrent tonsillitis. *Arch Otolaryngol Head Neck Surg* 2005;**131**:383–7. <https://doi.org/10.1001/archotol.131.5.383>
17. Dan JM, Havenar-Daughton C, Kendric K, Al-Kolla R, Kaushik K, Rosales SL, et al. Recurrent group A *Streptococcus* tonsillitis is an immunosusceptibility disease involving antibody deficiency and aberrant T(FH) cells. *Sci Transl Med* 2019;**11**. <https://doi.org/10.1126/scitranslmed.aau3776>
18. Scottish Intercollegiate Guidelines Network (SIGN). *Management of Sore Throat and Indications for Tonsillectomy: A National Clinical Guideline*. Edinburgh: SIGN; 2010.
19. National Institute for Health and Care Excellence (NICE). *Sore Throat (Acute): Antimicrobial Prescribing NICE Guideline [NG84]*. London: NICE; 2018.
20. Blair RL, McKerrow WS, Carter NW, Fenton A. The Scottish tonsillectomy audit. Audit Sub-Committee of the Scottish Otolaryngological Society. *J Laryngol Otol* 1996;**110**:1–25. <https://doi.org/10.1017/S0022215100136175>
21. Stalfors J, Ericsson E, Hemlin C, Hultcrantz E, Månsson I, Roos K, Hessén Söderman AC. Tonsil surgery efficiently relieves symptoms: analysis of 54 696 patients in the National Tonsil Surgery Register in Sweden. *Acta Otolaryngol* 2012;**132**:533–9. <https://doi.org/10.3109/00016489.2011.644252>
22. 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;**11**:CD001802. <https://doi.org/10.1002/14651858.CD001802.pub3>
23. Blakley BW, Magit AE. The role of tonsillectomy in reducing recurrent pharyngitis: a systematic review. *Otolaryngol Head Neck Surg* 2009;**140**:291–7. <https://doi.org/10.1016/j.otohns.2008.12.013>
24. Suleman M, Clark MPA, Goldacre M, Burton M. Exploring the variation in paediatric tonsillectomy rates between English regions: a 5-year NHS and independent sector data analysis. *Clin Otolaryngol* 2010;**35**:111–7. <https://doi.org/10.1111/j.1749-4486.2010.02086.x>
25. Yeo JC, Ah-See KW, Mackenzie K. Variation in practice: an analysis of Scottish Surgical Profiles ENT data. *Scott Med J* 2013;**58**:22–5. <https://doi.org/10.1177/0036933012474589>
26. Alho OP, Koivunen P, Penna T, Teppo H, Koskela M, Luotonen J. Tonsillectomy versus watchful waiting in recurrent streptococcal pharyngitis in adults: randomised controlled trial. *BMJ* 2007;**334**:939. <https://doi.org/10.1136/bmj.39140.632604.55>
27. 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* 2013;**185**:E331–6. <https://doi.org/10.1503/cmaj.121852>
28. Mathew R, McCombe A. Rationing based on 'procedures of limited clinical effectiveness': what are the risks? *Br J Hosp Med* 2012;**73**:184–5. <https://doi.org/10.12968/hmed.2012.73.4.184>
29. Heiser C, Landis BN, Giger R, Cao Van H, Guinand N, Hörmann K, Stuck BA. Taste disorders after tonsillectomy: a long-term follow-up. *Laryngoscope* 2012;**122**:1265–6. <https://doi.org/10.1002/lary.23270>
30. Smithard A, Cullen C, Thirlwall AS, Aldren C. Tonsillectomy may cause altered tongue sensation in adult patients. *J Laryngol Otol* 2009;**123**:545–9. <https://doi.org/10.1017/S0022215108003277>

31. Ovesen T, Kamarauskas G, Dahl M, Mainz J. Pain and bleeding are the main determinants of unscheduled contacts after outpatient tonsillectomy. *Dan Med J* 2012;**59**:A4382.
32. Haye R, Døsen LK, Gay C, TarAngen M, Shiryaeva O. Self-reported complications after tonsillectomy: comparison of responders and nonresponders to a mailed questionnaire. *Int J Otolaryngol* 2020;2020:4561858. <https://doi.org/10.1155/2020/4561858>
33. Mathiesen O, Jørgensen DG, Hilsted KL, Trolle W, Stjernholm P, Christiansen H, *et al.* Pregabalin and dexamethasone improves post-operative pain treatment after tonsillectomy. *Acta Anaesthesiol Scand* 2011;**55**:297–305. <https://doi.org/10.1111/j.1399-6576.2010.02389.x>
34. Hsu AP, Tan KL, Tan YB, Han HJ, Lu PK. Benefits and efficacy of tonsillectomy for recurrent tonsillitis in adults. *Acta Otolaryngol* 2007;**127**:62–4. <https://doi.org/10.1080/00016480500540501>
35. 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**:104–8. <https://doi.org/10.1046/j.1365-2273.2001.00433.x>
36. Pynnonen M, Brinkmeier JV, Thorne MC, Chong LY, Burton MJ. Coblation versus other surgical techniques for tonsillectomy. *Cochrane Database Syst Rev* 2017;**8**:CD004619. <https://doi.org/10.1002/14651858.CD004619.pub3>
37. Cardozo AAJ, Hallikeri C, Lawrence H, Sankar V, Hargreaves S. Teenage and adult tonsillectomy: dose-response relationship between diathermy energy used and morbidity. *Clin Otolaryngol* 2007;**32**:366–71. <https://doi.org/10.1111/j.1749-4486.2007.01529.x>
38. Diakos EA, Gallos ID, El-Shunnar S, Clarke M, Kazi R, Mehanna H. Dexamethasone reduces pain, vomiting and overall complications following tonsillectomy in adults: a systematic review and meta-analysis of randomised controlled trials. *Clin Otolaryngol* 2011;**36**:531–42. <https://doi.org/10.1111/j.1749-4486.2011.02373.x>
39. Titirungruang C, Seresirikachorn K, Kasemsuwan P, Hirunwiwatkul P. The use of steroids to reduce complications after tonsillectomy: a systematic review and meta-analysis of randomized controlled studies. *Eur Arch Oto-Rhino-L* 2019;**276**:585–604. <https://doi.org/10.1007/s00405-018-5202-2>
40. Tolska HK, Hamunen K, Takala A, Kontinen VK. Systematic review of analgesics and dexamethasone for post-tonsillectomy pain in adults. *Br J Anaesth* 2019;**123**:e397–411. <https://doi.org/10.1016/j.bja.2019.04.063>
41. Hessén Söderman AC, Ericsson E, Hemlin C, Hultcrantz E, Månsson I, Roos K, Stalfors J. Reduced risk of primary postoperative hemorrhage after tonsil surgery in Sweden: results from the National Tonsil Surgery Register in Sweden covering more than 10 years and 54,696 operations. *Laryngoscope* 2011;**121**:2322–6. <https://doi.org/10.1002/lary.22179>
42. Evans AS, Khan AM, Young D, Adamson R. Assessment of secondary haemorrhage rates following adult tonsillectomy – a telephone survey and literature review. *Clin Otolaryngol Allied Sci* 2003;**28**:489–91. <https://doi.org/10.1046/j.1365-2273.2003.00763.x>
43. Heiser C, Landis BN, Giger R, Cao Van H, Guinand N, Hörmann K, Stuck BA. Taste disturbance following tonsillectomy – a prospective study. *Laryngoscope* 2010;**120**:2119–24. <https://doi.org/10.1002/lary.20971>
44. Avent ML, Cosgrove SE, Price-Haywood EG, van Driel ML. Antimicrobial stewardship in the primary care setting: from dream to reality? *BMC Family Practice* 2020;**21**. <https://doi.org/10.1186/s12875-020-01191-0>

45. Centor RM. Adolescent and adult pharyngitis: more than 'strep throat': comment on 'Large-scale validation of the Centor and Mclsaac Scores to predict group A streptococcal pharyngitis'. *Arch Intern Med* 2012;**172**:852–3. <https://doi.org/10.1001/archinternmed.2012.1741>
46. Cooper RJ, Hoffman JR, Bartlett JG, Besser RE, Gonzales R, Hickner JM, Sande MA, Centers for Disease Control and Prevention. Principles of appropriate antibiotic use for acute pharyngitis in adults: background. *Ann Emerg Med* 2001;**37**:711–9. [https://doi.org/10.1067/S0196-0644\(01\)70090-X](https://doi.org/10.1067/S0196-0644(01)70090-X)
47. Little P, Gould C, Williamson I, Warner G, Gantley M, Kinmonth AL. Clinical and psychosocial predictors of illness duration from randomised controlled trial of prescribing strategies for sore throat. *BMJ* 1999;**319**:736–7. <https://doi.org/10.1136/bmj.319.7212.736>
48. Fraser H, Gallacher D, Achana F, Court R, Taylor-Phillips S, Nduka C, *et al*. Rapid antigen detection and molecular tests for group A streptococcal infections for acute sore throat: systematic reviews and economic evaluation. *Health Technol Assess* 2020;**24**(31). <https://doi.org/10.3310/hta24310>
49. Salkind AR, Wright JM. Economic burden of adult pharyngitis: the payer's perspective. *Value Health* 2008;**11**:621–7. <https://doi.org/10.1111/j.1524-4733.2007.00286.x>
50. van der Velden AW, Pijpers EJ, Kuyvenhoven MM, Tonkin-Crine SK, Little P, Verheij TJ. Effectiveness of physician-targeted interventions to improve antibiotic use for respiratory tract infections. *Br J Gen Pract* 2012;**62**:e801–7. <https://doi.org/10.3399/bjgp12X659268>
51. Little P, Gould C, Williamson I, Warner G, Gantley M, Kinmonth AL. Reattendance and complications in a randomised trial of prescribing strategies for sore throat: the medicalising effect of prescribing antibiotics. *BMJ* 1997;**315**:350–2. <https://doi.org/10.1136/bmj.315.7104.350>
52. Mortazhejri S, Hong PJ, Yu AM, Hong BY, Stacey D, Bhatia RS, Grimshaw JM. Systematic review of patient-oriented interventions to reduce unnecessary use of antibiotics for upper respiratory tract infections. *Syst Rev* 2020;**9**:106. <https://doi.org/10.1186/s13643-020-01359-w>
53. Zwart S, Sachs AP, Ruijs GJ, Gubbels JW, Hoes AW, de Melker RA. Penicillin for acute sore throat: randomised double blind trial of seven days versus three days treatment or placebo in adults. *BMJ* 2000;**320**:150–4. <https://doi.org/10.1136/bmj.320.7228.150>
54. Little P, Williamson I, Warner G, Gould C, Gantley M, Kinmonth AL. Open randomised trial of prescribing strategies in managing sore throat. *BMJ* 1997;**314**:722–7. <https://doi.org/10.1136/bmj.314.7082.722>
55. Urbánek K, Kolár M, Cekanová L. Utilisation of macrolides and the development of *Streptococcus pyogenes* resistance to erythromycin. *Pharm World Sci* 2005;**27**:104–7. <https://doi.org/10.1007/s11096-004-6607-0>
56. Melo-Cristino J, Santos L, Ramirez M, Grupo de Estudo Português de Bactérias Patogénicas Respiratórias. The Viriato Study: update of antimicrobial susceptibility data of bacterial pathogens from community-acquired respiratory tract infections in Portugal in 2003 and 2004. *Rev Port Pneumol* 2006;**12**:9–30. [https://doi.org/10.1016/S2173-5115\(06\)70385-2](https://doi.org/10.1016/S2173-5115(06)70385-2)
57. Casey JR, Pichichero ME. Metaanalysis of short course antibiotic treatment for group A Streptococcal tonsillopharyngitis. *Pediatr Infect Dis J* 2005;**24**:909–17. <https://doi.org/10.1097/01.inf.0000180573.21718.36>
58. Carbon C, Chatelin A, Bingen E, Zuck P, Rio Y, Guetat F, Orvain J. A double-blind randomized trial comparing the efficacy and safety of a 5-day course of cefotiam hexetil with that of a 10-day course of penicillin V in adult patients with pharyngitis caused by group A beta-haemolytic streptococci. *J Antimicrob Chemother* 1995;**35**:843–54. <https://doi.org/10.1093/jac/35.6.843>

59. Mui S, Rasgon BM, Hilsinger RL. Efficacy of tonsillectomy for recurrent throat infection in adults. *Laryngoscope* 1998;**108**:1325–8. <https://doi.org/10.1097/00005537-199809000-00012>
60. Dalalah D, Magableh S. A remote fuzzy multicriteria diagnosis of sore throat. *Telemed J E Health* 2008;**14**:656–65. <https://doi.org/10.1089/tmj.2007.0120>
61. Moore M, Stuart B, Hobbs FR, Butler CC, Hay AD, Campbell J, *et al.* Symptom response to antibiotic prescribing strategies in acute sore throat in adults: the DESCARTE prospective cohort study in UK general practice. *Br J Gen Pract* 2017;**67**:e634–42. <https://doi.org/10.3399/bjgp17X692321>
62. Hayward GN, Hay AD, Moore MV, Jawad S, Williams N, Voysey M, *et al.* Effect of oral dexamethasone without immediate antibiotics vs placebo on acute sore throat in adults: a randomized clinical trial. *JAMA* 2017;**317**:1535–43. <https://doi.org/10.1001/jama.2017.3417>
63. de Cassan S, Thompson MJ, Perera R, Glasziou PP, Del Mar CB, Heneghan CJ, Hayward G. Corticosteroids as standalone or add-on treatment for sore throat. *Cochrane Database Syst Rev* 2020;**5**:CD008268. <https://doi.org/10.1002/14651858.CD008268.pub3>
64. 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 Technol Assess* 2010;**14**(13). <https://doi.org/10.3310/hta14130>
65. Koskenkorva T, Koivunen P, Penna T, Teppo H, Alho OP. Factors affecting quality-of-life impact of adult tonsillectomy. *J Laryngol Otol* 2009;**123**:1010–4. <https://doi.org/10.1017/S0022215109005271>
66. Skevas T, Klingmann C, Plinkert PK, Baumann I. [Development and validation of the Tonsillectomy Outcome Inventory 14.] *HNO* 2012;**60**:801–6. <https://doi.org/10.1007/s00106-012-2545-7>
67. Hopkins C, Fairley J, Yung M, Hore I, Balasubramaniam S, Haggard M. The 14-item Paediatric Throat Disorders Outcome Test: a valid, sensitive, reliable, parent-reported outcome measure for paediatric throat disorders. *J Laryngol Otol* 2010;**124**:306–14. <https://doi.org/10.1017/S0022215109992386>
68. Erskine S, Kwong FNK, Robinson G, Powell E, Moor J, JA W. Use of the Tonsil Outcome Inventory in adults undergoing elective tonsillectomy – a prospective study (abstract). *Clin Otolaryngol* 2012;**37**:43.
69. Powell J, Powell EL, Conroy K, Hopkins C, Moor JW, Wilson JA. Throat-related quality of life in peritonsillar abscess sufferers: application of the adult tonsil outcome inventory. *J Laryngol Otol* 2013;**127**:1190–3. <https://doi.org/10.1017/S0022215113003071>
70. Powell J, Powell E, Conroy K, Hopkins C, Moor J, Wilson J. Throat symptoms: the importance in peritonsillar abscess (abstract). *Otolaryngol Head Neck Surg* 2012;**147**:51. <https://doi.org/10.1177/0194599812451438a49>
71. Bhattacharyya N, Kepnes LJ. Economic benefit of tonsillectomy in adults with chronic tonsillitis. *Ann Otol Rhinol Laryngol* 2002;**111**:983–8. <https://doi.org/10.1177/000348940211101106>
72. Little P, Hobbs FD, Mant D, McNulty CA, Mullee M, PRISM investigators. Incidence and clinical variables associated with streptococcal throat infections: a prospective diagnostic cohort study. *Br J Gen Pract* 2012;**62**:e787–94. <https://doi.org/10.3399/bjgp12X658322>
73. NHS Digital. *Hospital Episode Statistics, Admitted Patient Care – England, 2011–12*. 2012. URL: <https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/hospital-episode-statistics-admitted-patient-care-england-2011-12#> (accessed September 2021).

74. Lau AS, Upile NS, Wilkie MD, Leong SC, Swift AC. The rising rate of admissions for tonsillitis and neck space abscesses in England, 1991–2011. *Ann R Coll Surg Engl* 2014;**96**:307–10. <https://doi.org/10.1308/003588414X13946184900363>
75. Vaikjarv R, Mandar R, Kasenomm P. Peritonsillar abscess is frequently accompanied by sepsis symptoms. *Eur Arch Oto-Rhino-L* 2019;**276**:1721–5. <https://doi.org/10.1007/s00405-019-05424-6>
76. Windfuhr JP, Alizoti P, Hendricks C. Regional variability of hemorrhage following tonsil surgery in 1,520,234 cases. *Eur Arch Oto-Rhino-L* 2020;**277**:3169–77. <https://doi.org/10.1007/s00405-020-06080-x>
77. Trial website NATTINA. URL: <http://research.ncl.ac.uk/nattina> (accessed April 2021).
78. Nunes EV, Pavlicova M, Hu MC, Campbell AN, Miele G, Hien D, Klein DF. Baseline matters: the importance of covariation for baseline severity in the analysis of clinical trials. *Am J Drug Alcohol Abuse* 2011;**37**:446–52. <https://doi.org/10.3109/00952990.2011.596980>
79. Rubie I, Haighton C, O'Hara J, Rousseau N, Steen N, Stocken DD, *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. <https://doi.org/10.1186/s13063-015-0768-0>
80. McSweeney LA, Wilson JA, Wilkes S, Haighton CA. Is Scottish Intercollegiate Guidelines Network guidance for GP management of tonsillitis suitable? A qualitative study. *Fam Pract* 2018;**35**:633–7. <https://doi.org/10.1093/fampra/cmy017>
81. 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**:301–6. <https://doi.org/10.1111/coa.12720>
82. McSweeney LA, O'Hara JT, Rousseau NS, Stocken DD, Sullivan F, Vale L, *et al.* 'Thinking that somebody's going to delay [a tonsillectomy] for one to two years is quite horrifying really': a qualitative feasibility study for the National Trial of Tonsillectomy IN Adults (NATTINA Part 2). *Clin Otolaryngol* 2017;**42**:578–83. <https://doi.org/10.1111/coa.12781>
83. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;**34**:220–33. <https://doi.org/10.1097/00005650-199603000-00003>
84. Laajala A, Autio TJ, Ohtonen P, Alho OP, Koskenkorva TJ. Interpretation of Tonsillectomy Outcome Inventory-14 scores: a prospective matched cohort study. *Eur Arch Otorhinolaryngol* 2020;**277**:1499–505. <https://doi.org/10.1007/s00405-020-05832-z>
85. Plath M, Sand M, Federspil PA, Plinkert PK, Baumann I, Zaoui K. Normative tonsillectomy outcome inventory 14 values as a decision-making tool for tonsillectomy. *Eur Arch Otorhinolaryngol* 2021;**278**:1645–51. <https://doi.org/10.1007/s00405-020-06374-0>
86. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care* 2004;**42**:851–9. <https://doi.org/10.1097/01.mlr.0000135827.18610.0d>
87. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, *et al.* Guidelines for the content of statistical analysis plans in clinical trials. *JAMA* 2017;**318**:2337–43. <https://doi.org/10.1001/jama.2017.18556>
88. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, *et al.* CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;**340**:c869. <https://doi.org/10.1136/bmj.c869>

89. 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. *Otolaryngol Head Neck Surg* 2012;**146**:122–8. <https://doi.org/10.1177/0194599811422011>
90. Dumville JC, Torgerson DJ, Hewitt CE. Reporting attrition in randomised controlled trials. *BMJ* 2006;**332**:969–71. <https://doi.org/10.1136/bmj.332.7547.969>
91. Ye C, Beyene J, Browne G, Thabane L. Estimating treatment effects in randomised controlled trials with non-compliance: a simulation study. *BMJ Open* 2014;**4**:6. <https://doi.org/10.1136/bmjopen-2014-005362>
92. Kjellsson G, Clarke P, Gerdtham UG. Forgetting to remember or remembering to forget: a study of the recall period length in health care survey questions. *J Health Econ* 2014;**35**:34–46. <https://doi.org/10.1016/j.jhealeco.2014.01.007>
93. Elsevier. *MACRO. Advanced Data Collection for Clinical Research Website*. URL: www.elsevier.com/en-gb/solutions/veridata/macro (accessed April 2021).
94. Great Britain. *Data Protection Act 1998*. London: Her Majesty's Stationery Office; 1998.
95. Great Britain. *Data Protection Act 2018*. London: Her Majesty's Stationery Office; 2018.
96. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, *et al.* Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMJ* 2013;**346**:f1049. <https://doi.org/10.1136/bmj.f1049>
97. Joint Formulary Committee. *British National Formulary* (online). London: BMJ Group and Pharmaceutical Press. URL: www.medicinescomplete.com (accessed October 2020).
98. Department of Health and Social Care. *NHS Reference Costs 2017–18*. URL: <https://improvement.nhs.uk/resources/reference-costs/> (accessed October 2019).
99. Oron Y, Marom T, Russo E, Ezri T, Roth Y. Don't overlook the complications of tonsillectomy. *Fam Pract* 2010;**59**:E4–9.
100. Department of Health and Social Care. *NHS Reference Costs 2016–17*. URL: <https://improvement.nhs.uk/resources/reference-costs/> (accessed October 2019).
101. Curtis L, Burns A. *Unit Costs of Health and Social Care 2019*. Canterbury: PSSRU, University of Kent; 2019.
102. Department of Health and Social Care. *NHS Reference Costs 2018–19*. URL: www.england.nhs.uk/national-cost-collection/ (accessed October 2020).
103. British Medical Association. *Pay Scales for Junior Doctors in England*. URL: www.bma.org.uk/pay-and-contracts/pay/junior-doctors-pay-scales/pay-scales-for-junior-doctors-in-england (accessed April 2021).
104. NHS Health Education England. *Agenda for Change: Pay Rates*. 2018. URL: www.healthcareers.nhs.uk/working-health/working-nhs/nhs-pay-and-benefits/agenda-change-pay-rates (accessed 12 February 2021).
105. GOV.UK. *National Insurance Rates and Categories*. URL: www.gov.uk/national-insurance-rates-letters/contribution-rates (accessed April 2021).
106. NHS Employers. *Pension Contributions and Tax Relief*. URL: www.nhsemployers.org/your-workforce/pay-and-reward/pensions/pension-contribution-tax-relief (accessed April 2021).
107. Department of Transport. *Transport Analysis Guidance: WebTAG*. 2013. URL: www.dft.gov.uk/webtag/documents/expert/unit3.5.6.php (accessed 19 December 2014).

108. Office for National Statistics. *Annual Survey of Hours and Earnings, 2014 Provisional Results*. 2014. URL: www.ons.gov.uk/ons/rel/ashe/annual-survey-of-hours-and-earnings/2014-provisional-results/stb-ashe-statistical-bulletin-2014.html: Office of National Statistics 2014 (accessed July 2020).
109. Matthews JN, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. *BMJ* 1990;**300**:230–5.
110. Donaldson C, Mason H, Shackley P. Contingent valuation in health care. In Jones AM, editor. *The Elgar Companion to Health Economics*. Second edition. Cheltenham: Edward Elgar Publishing Limited; 2012.
111. Bishop R, Mitchell R, Carson R. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, DC: Resources for the Future; 1990. <https://doi.org/10.2307/1243175>
112. Mitchell R, Carson R. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, DC: John Hopkins University Press; 1989.
113. National Institute for Health and Care Excellence (NICE). *Guide to the Methods of Technology Appraisal 2013*. London: NICE; 2013.
114. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the Economic Evaluation of Health Care Programmes*. Third edition. New York, NY: Oxford University Press; 2005.
115. Fiebig DG. Seemingly unrelated regression. In Baltagi BH, editor. *A Companion to Theoretical Econometrics*. Malden, MA: Blackwell Publishing Ltd; 2003. pp. 101–21.
116. Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health Econ* 2004;**13**:461–75. <https://doi.org/10.1002/hec.843>
117. 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**:40–9. <https://doi.org/10.1002/mpr.329>
118. Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. *Stat Med* 2000;**19**:3219–36. [https://doi.org/10.1002/1097-0258\(20001215\)19:23<3219::AID-SIM623>3.0.CO;2-P](https://doi.org/10.1002/1097-0258(20001215)19:23<3219::AID-SIM623>3.0.CO;2-P)
119. Black WC. The CE plane: a graphic representation of cost-effectiveness. *Med Decis Making* 1990;**10**:212–4. <https://doi.org/10.1177/0272989X9001000308>
120. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ* 2001;**10**:779–87. <https://doi.org/10.1002/hec.635>
121. Gainsbury S. NHS spending: McKinsey exposes hard choices to save 20bn pounds sterling. *Health Serv J* 2009;**119**:12–3.
122. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 2015;**350**:h1258. <https://doi.org/10.1136/bmj.h1258>
123. Francis JJ, Johnston M, Robertson C, Glidewell L, Entwistle V, Eccles MP, Grimshaw JM. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychol Health* 2010;**25**:1229–45. <https://doi.org/10.1080/08870440903194015>
124. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol* 2013;**13**:117. <https://doi.org/10.1186/1471-2288-13-117>

125. Pope C, Ziebland S, Mays N. Analysing qualitative data. *BMJ* 2000;**230**:114. <https://doi.org/10.1136/bmj.320.7227.114>
126. Blair R, McKerrow W, Carter N, Fenton A. The Scottish tonsillectomy audit. *J Laryngol Otol* 1996;**110**:1–25. <https://doi.org/10.1017/S0022215100136175>
127. Cook JA, Ramsay CR, Norrie J. Recruitment to publicly funded trials – are surgical trials really different? *Contemp Clin Trials* 2008;**29**:631–4. <https://doi.org/10.1016/j.cct.2008.02.005>
128. Khonsari S, Subramanian P, Chinna K, Latif LA, Ling LW, Gholami O. Effect of a reminder system using an automated short message service on medication adherence following acute coronary syndrome. *Eur J Cardiovasc Nurs* 2015;**14**:170–9. <https://doi.org/10.1177/1474515114521910>
129. Montes JM, Medina E, Gomez-Beneyto M, Maurino J. A short message service (SMS)-based strategy for enhancing adherence to antipsychotic medication in schizophrenia. *Psychiatry Res* 2012;**200**:89–95. <https://doi.org/10.1016/j.psychres.2012.07.034>
130. Vahlberg B, Lundström E, Eriksson S, Holmbäck U, Cederholm T. Effects on walking performance and lower body strength by short message service guided training after stroke or transient ischemic attack (The STROKEWALK Study): a randomized controlled trial. *Clin Rehabil* 2021;**35**:276–87. <https://doi.org/10.1177/0269215520954346>
131. Powell HR, Mehta N, Daly N, Watters GW. Improved quality of life in adults undergoing tonsillectomy for recurrent tonsillitis. Is adult tonsillectomy really a low priority treatment? *Eur Arch Otorhinolaryngol* 2012;**269**:2581–4. <https://doi.org/10.1007/s00405-012-2095-3>
132. Windfuhr JP. Specified data for tonsil surgery in Germany. *GMS Curr Top Otorhinolaryngol Head Neck Surg* 2016;**15**:Doc08. <https://doi.org/10.3205/cto000135>
133. Koskenkorva T, Koivunen P, Läärä E, Alho OP. Predictive factors for quality of life after tonsillectomy among adults with recurrent pharyngitis: a prospective cohort study. *Clin Otolaryngol* 2014;**39**:216–23. <https://doi.org/10.1111/coa.12263>
134. 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**:40–4. <https://doi.org/10.1016/j.ijporl.2017.01.014>
135. Van Den Akker EH, Hoes AW, Burton MJ, Schilder AG. Large international differences in (adeno)tonsillectomy rates. *Clin Otolaryngol Allied Sci* 2004;**29**:161–4. <https://doi.org/10.1111/j.0307-7772.2004.00771.x>
136. Weinstein GS, Levin B. Effect of crossover on the statistical power of randomized studies. *Ann Thorac Surg* 1989;**48**:490–5. [https://doi.org/10.1016/S0003-4975\(10\)66846-4](https://doi.org/10.1016/S0003-4975(10)66846-4)
137. Brueton VC, Tierney J, Stenning S, Harding S, Meredith S, Nazareth I, Rait G. Strategies to improve retention in randomised trials. *Cochrane Database Syst Rev* 2013;**12**:MR000032. <https://doi.org/10.1002/14651858.MR000032.pub2>
138. Elfeky A, Gillies K, Gardner H, Fraser C, Ishaku T, Treweek S. Non-randomised evaluations of strategies to increase participant retention in randomised controlled trials: a systematic review. *Syst Rev* 2020;**9**:224. <https://doi.org/10.1186/s13643-020-01471-x>
139. Akl EA, Briel M, You JJ, Sun X, Johnston BC, Busse JW, et al. Potential impact on estimated treatment effects of information lost to follow-up in randomised controlled trials (LOST-IT): systematic review. *BMJ* 2012;**344**:e2809. <https://doi.org/10.1136/bmj.e2809>
140. Witsell DL, Orvidas LJ, Stewart MG, Hannley MT, Weaver EM, Yueh B, et al. Quality of life after tonsillectomy in adults with recurrent or chronic tonsillitis. *Otolaryngol Head Neck Surg* 2008;**138**(Suppl. 1):1–8. <https://doi.org/10.1016/j.otohns.2007.08.015>

REFERENCES

141. Bade MJ, Struessel T, Dayton M, Foran J, Kim RH, Miner T, *et al.* Early high-intensity versus low-intensity rehabilitation after total knee arthroplasty: a randomized controlled trial. *Arthritis Care Res* 2017;**69**:1360–8. <https://doi.org/10.1002/acr.23139>
142. Kvaerner KJ. Benchmarking surgery: secondary post-tonsillectomy hemorrhage 1999–2005. *Acta Otolaryngol* 2009;**129**:195–8. <https://doi.org/10.1080/00016480802078101>
143. Macassey EA, Baguley C, Dawes P, Gray A. 15-year audit of post-tonsillectomy haemorrhage at Dunedin Hospital. *ANZ J Surg* 2007;**77**:579–82. <https://doi.org/10.1111/j.1445-2197.2007.04154.x>
144. Geva A, Brigger MT. Dexamethasone and tonsillectomy bleeding: a meta-analysis. *Otolaryngol Head Neck Surg* 2011;**144**:838–43. <https://doi.org/10.1177/0194599811399538>
145. Senska G, Ellermann S, Ernst S, Lax H, Dost P. Recurrent tonsillitis in adults: quality of life after tonsillectomy. *Dtsch Arztebl Int* 2010;**107**:622–8. <https://doi.org/10.3238/arztebl.2010.0622>
146. Capper R, Canter RJ. How well do parents recognize the difference between tonsillitis and other sore throats? *Clin Otolaryngol Allied Sci* 2001;**26**:458–64. <https://doi.org/10.1046/j.1365-2273.2001.00500.x>
147. Mandavia R, Schilder AGM, Dimitriadis PA, Mossialos E. Addressing the challenges in tonsillectomy research to inform health care policy: a review. *JAMA Otolaryngol Head Neck Surg* 2017;**143**:943–7. <https://doi.org/10.1001/jamaoto.2017.0964>
148. 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. Results of parallel randomized and nonrandomized clinical trials. *N Engl J Med* 1984;**310**:674–83. <https://doi.org/10.1056/nejm198403153101102>
149. Chaturvedi N, Ben-Shlomo Y. From the surgery to the surgeon: does deprivation influence consultation and operation rates? *Br J Gen Pract* 1995;**45**:127–31.
150. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, *et al.* Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev* 2017;**4**:CD001431. <https://doi.org/10.1002/14651858.CD001431.pub5>
151. NHS England. *Individual Funding Requests for Specialised Services a Guide for Patients*. London: NHS England; 2017.

Appendix 1 Substantial amendments

TABLE 28 Changes to the protocol requiring regulatory approval

Amendment number	Protocol version and date	Description
Substantial amendment 1	2.0; 11 November 2014	<ul style="list-style-type: none"> Change to the process of STAR data collection (see Primary outcome data collection) Wording of question on STAR form amended from 'severe' to 'bad' Now collecting patients' e-mail addresses and telephone numbers on the expression of interest form The secondary objective wording was split into two separate objectives, as the previous protocol version did not correctly explain the use of the SF-12 questionnaire in determining QoL The timing of the qualitative process evaluation interviews was changed for them to take place during the internal pilot, instead of at the end of the pilot Comparison data form was updated to state that we would like to collect a reason for decline Addition of qualitative interview process evaluation topic guide, NATTINA DVD and website Several further protocol clarifications and administrative changes to trial documents were made
Substantial amendment 2	3.0; 23 February 2015	<ul style="list-style-type: none"> Participant time and travel questionnaire would no longer be administered at the baseline visit and instead would be collected at the 18-month follow-up visit. Additional questions were added to the questionnaire in relation to participant costs. Separate questionnaire implemented for the 18-month follow-up visit, including the participant time and travel questions Bradford replaced York as one of the internal pilot sites and the PI was replaced at the Sunderland site Data Monitoring and Ethics Committee independent statistician added Secondary objectives and outcomes clarified in relation to the analysis of the SF-12 Clarification added in relation to surgery purposefully delayed over 8 weeks owing to severe tonsillitis being accepted in the protocol Economic analysis section corrected to state that the contingent valuation questionnaire was for completion at baseline rather than the end of the trial Several further clarifications and administrative changes to trial documents were made
Substantial amendment 3	3.0; 23 February 2015	<ul style="list-style-type: none"> A letter of invitation was introduced to send to participants along with the standard clinic appointment letter and PIS

continued

TABLE 28 Changes to the protocol requiring regulatory approval (*continued*)

Amendment number	Protocol version and date	Description
Substantial amendment 4	4.0; 18 August 2016	<ul style="list-style-type: none"> • Exclusion criteria were updated to provide clarification that tonsilloliths should be excluded only if this was the primary reason for referral. Additional clarity provided that bleeding diathesis included haemophilia, sickle cell disease and platelet dysfunction • Identification processes updated to allow patients initially assessed in other centres to be approached to participate in the NATTINA trial • Sites to restart collecting expression of interest forms for qualitative interviews for comparison data to be obtained • Sample size reduced from 600 to 510 owing to changes in trial design in relation to not including an initially planned subgroup randomisation of the conservative management arm • Trial pathway updated to make it optional for participants to watch the recruitment video • Definition of end of study was updated • Several further clarifications and administrative changes to trial documents were made
Substantial amendment 5	4.0; 18 August 2016	<ul style="list-style-type: none"> • Introduction of a second weekly STAR contact, as described in the primary outcome section • Implementation of the use of the STAR system to send reminders to participants regarding 12- and 24-month due appointments, and 6- and 18-month questionnaires. STAR reminders reduced from up to seven per week to three to prevent increased number of total contacts via the system • Letter to participants explaining these changes and updated STAR instructions introduced
Substantial amendment 6	5.0; 15 May 2017	<ul style="list-style-type: none"> • PIS summary sheet introduced to send with main PIS, in response to site and participant feedback • Recruitment period extended by 12 months, from 18 months to 30 months in total • Further administrative changes to the protocol were made
Substantial amendment 7	5.0; 15 May 2017	<ul style="list-style-type: none"> • Further information added to the patient information sheet to provide a balanced view of treatment options available in response to crossover rates • STAR system provider changed company name from In-teleme to CI Data • Simplification of the existing electronic and paper STAR questionnaire, as described in the primary outcome section • Further administrative changes to the protocol were made
Substantial amendment 8	6.0; 26 March 2018	<ul style="list-style-type: none"> • Recruitment end date extended by 4 months to 20 April 2018, totalling 34-month recruitment. Participants followed up for 24 months meaning all clinical interventions would be completed by 30 April 2020 • Analysis window extended by 3 months. Target number of participants revised from 510 to 444 to achieve 85% power to detect the 0.33 effect size planned for; however, recruitment would continue beyond 444 if eligible patients were identified, although recruitment would not exceed the original 510 participants, or beyond 30 April 2018
Substantial amendment 9	5.0; 15 May 2017	<ul style="list-style-type: none"> • Time and travel questionnaire no longer collected at 18 months to help to alleviate participant burden and increase response rates. Enough of these data had already been collected to inform this part of the analysis, as it was implemented approximately 24 months following the start of recruitment • Letter introduced to remind participants when the 24-month visit was due • Further clarifications and administrative changes to trial documents were made

TABLE 28 Changes to the protocol requiring regulatory approval (*continued*)

Amendment number	Protocol version and date	Description
Substantial amendment 10	5.0; 15 May 2017	<ul style="list-style-type: none"> A GDPR supplementary information sheet was introduced to clarify the different roles of the NCTU and sponsor, to be provided to participants who had not yet completed the trial when they attended for their next follow-up visit
Substantial amendment 11	7.0; 11 September 2019	<ul style="list-style-type: none"> Approval was gained to utilise the CI Data system to enable participants to complete the TOI-14 form electronically via a link in a text message or e-mail, instead of completing the paper form if this would be their preference. This was to attempt to increase TOI-14 completion rates and, therefore, secondary objective data. The new electronic methods of TOI-14 completion did not become available to the participants owing to issues with the company due to provide this service

Appendix 2 Reasons for requesting to switch randomised arms

The reasons given for participants requesting to crossover are summarised in [Table 29](#) (for those randomised to the conservative management arm). Further details are also given for those choosing 'other' in the [Table 29](#) footnotes (for those randomised to the conservative management arm).

TABLE 29 Summary of the 134 reasons that the 73 participants randomised to the conservative management arm gave for crossing over to receive tonsillectomy

Reasons given for crossing over to tonsillectomy	Total (n)
Concerns about time off work/study owing to tonsillitis	35
Too many sore throat days	62
Unhappy with usual care/conservative management	17
Advice from friend/family	9
Hospital admission with tonsil-related illness	4
Other reason	7 ^a
Total	134

a The seven 'other' reasons are:

1. Disturbed school.
2. Wanted to stay in the study, but cross over.
3. Three reasons given: concerns about time off work/study due to tonsillitis, too many sore throat days and unhappy with usual care/conservative management.
4. Off sick from college.
5. Why won't the form let me leave some of these unticked?
6. Details not provided by participant.
7. Patient had tonsillitis after consent, affecting work so crossed over for surgery.

Note that one patient specified all reasons owing to MACRO problem.

[Table 30](#) shows the 24 reasons that the 21 participants randomised to tonsillectomy gave for crossing over to conservative management. Further details are also given for those choosing 'other' in the [Table 30](#) footnotes.

TABLE 30 Reasons for patients randomised to tonsillectomy for crossing over to conservative management

Reasons given for crossing over to conservative management	Frequency
Fear of surgery	9
No time	1
Advice from friend/family	1
Other reason	13 ^a
Total	24

a The 13 'other' reasons are:

1. Felt his sore throats had improved enough to no longer need the surgery.
2. Needle phobic.
3. Sore throats have resolved.
4. Reduction in frequency of symptoms.
5. No longer feels tonsillitis severe enough to want surgery.
6. Will have surgery if she has any more episodes of tonsillitis.
7. Episodes of tonsillitis reduced.
8. Possibility of still having the sore throat symptoms even after surgery.
9. Did some further reading about the surgery and changed her mind. Will review decision in a year.
10. Personal problems.
11. Other medical issues.
12. Decided against having the surgery, no specific reason given.
13. No reason given.

Appendix 3 Additional information for eligible participants who declined to be randomised

TABLE 31 Reasons given at screening stage for declining to participate

Reasons for declining	Total (n)
Wanted tonsillectomy off study	1002
Wanted usual care off study	60
Did not think study was fair	13
Not comfortable with randomisation	15
Uncertain about being a research participant	4
Concerns about time involved	14
Concerns about confidentiality	2
Too much to consider	14
Patient- specific concerns about undergoing the operation of tonsillectomy	5
Not enough time to consider patient info	2
No reason given	53
Other reason (not stated)	1
Other reason (stated)	37
Multiple reasons	13
Cannot be contacted/did not respond/DNA	13
Ineligible	2
Missing	147
Not interested	6
Total	1403
DNA, did not attend.	

In total, 609 out of the 1403 declining participants provided baseline information, including TOI-14. [Table 32](#) gives the numbers of complete and partially completed TOI-14 questionnaires for those declining to take part. TOI-14 scores with one or two missing items were imputed to give 607 (99.7%) scores for a comparative analysis (accepters vs. decliners).

TABLE 32 Completeness of TOI-14 for declining patients

Completeness (number of TOI-14 items) provided	Total, n (%)
All 14 items: complete questionnaire (0 missing)	551 (91)
13 items (one missing)	48 (8)
12 items (two missing)	8 (1)
8 items (six missing)	1 (<1)
0 items (all 14 missing)	1 (<1)
Total	609 (100)

Figure 19 shows the distribution of TOI-14 scores in the 609 participants who declined who provided additional data, compared with the distribution for those who agreed to be randomised.

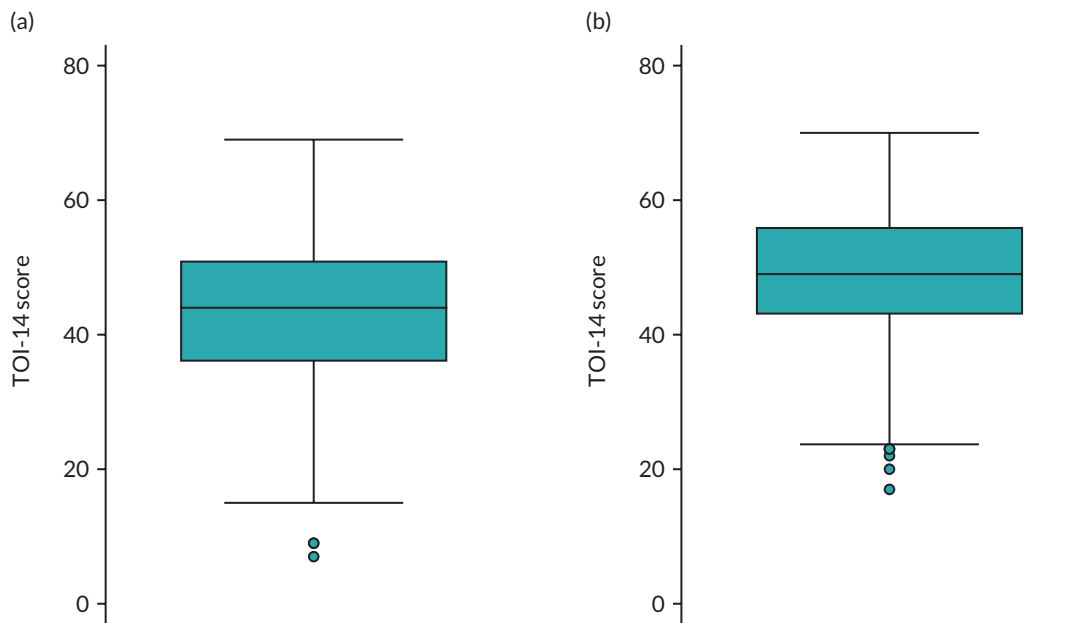


FIGURE 19 Box plots of baseline TOI-14 scores for participants accepting and declining to be randomised into NATTINA. (a) Accepters; and (b) decliners. Graphs by group: accepters, $n = 448$; decliners, $n = 607$.

Appendix 4 Stratification factors

Table 33 shows the distribution of participants by recruiting centre.

TABLE 33 Distribution of participants by recruiting centre (n=27)

Site	Randomised arm, n (%)		Total, n (%)
	Immediate tonsillectomy	Conservative management	
Glasgow	39 (17)	39 (18)	78 (17)
Aberdeen	29 (12)	28 (13)	57 (13)
Bradford	28 (12)	27 (12)	55 (12)
Guy's	24 (10)	23 (10)	47 (10)
Newcastle	23 (10)	22 (10)	45 (10)
Dundee	18 (8)	19 (9)	37 (8)
Wigan	12 (5)	11 (5)	23 (5)
Oxford	11 (5)	10 (5)	21 (5)
Nottingham	9 (4)	9 (4)	18 (4)
Birmingham	9 (4)	6 (3)	15 (3)
West Suffolk	5 (2)	6 (3)	11 (2)
Sunderland	4 (2)	6 (3)	10 (2)
Ayrshire & Arran	7 (3)	4 (2)	11 (2)
Blackburn	2 (<1)	4 (2)	6 (1)
Frimley	2 (<1)	3 (1)	5 (1)
Newport	3 (1)	1 (<1)	4 (<1)
Sheffield	2 (<1)	1 (<1)	3 (<1)
Cumberland	2 (<1)	1 (<1)	3 (<1)
Worcester	2 (<1)	1 (<1)	3 (<1)
Dorchester	1 (<1)	0 (0)	1 (<1)
Boston	1 (<1)	0 (0)	1 (<1)
Middlesbrough	1 (<1)	0 (0)	1 (<1)
Salisbury	0 (0)	0 (0)	0 (0)
Plymouth	0 (0)	0 (0)	0 (0)
Lancashire	0 (0)	0 (0)	0 (0)
Great Yarmouth	0 (0)	0 (0)	0 (0)
East and North Hertfordshire NHS Trust	0 (0)	0 (0)	0 (0)
Total	234 (100)	221 (100)	455 (100)

Table 34 shows the mis-stratification variable TOI-14 baseline severity distribution by number of participants in each arm for the ITT population. Seven, including the two ineligible participants and the participant from the unresponsive site, did not have baseline TOI-14 data collected.

TABLE 34 Number of participants mis-stratified based on TOI-14 severity categories

Mis-stratified as	Baseline TOI-14		
	Mild	Moderate	Severe
Mild	0	5	0
Moderate	8	0	8 ^a
Severe	3	12	0

a One with imputed TOI-14 score indicating severe was originally stratified as moderate.

Appendix 5 Recoding of employment status and education levels

Employment status

Some participants specified their employment status as 'other' and also provided a free-text description. The TMG agreed to recode these to more appropriate categories. These are listed below:

- Employment 'other' category ($n = 11$).
- Recoded as –
 - Housewife/mother: recode as 'Looking after family or home'.
 - Freelance contractor: recode as 'Self-employed'.
 - Employment supplementary allowance: recode as 'Unemployment (actively seeking work)'.
 - College student: recode as 'Full-time student/at school'.
 - X2 SVQ (Scottish Vocational Qualification) Level 3: recode as 'Full-time student/at school'.
 - Housewife: recode as 'Looking after family or home'.
 - In paid employment and full-time student/at school: recode as 'Full time student/at school' (as the employment is a working student job).
 - Part-time work and part-time college: recode as 'Paid employment (full or part time)'.
 - UNOB (unobtained): recode as 'MISSING'.
 - About to quit my full time job to become self-employed: recode as 'Self-employed'.
 - Paid work temporary contract: recode as 'Paid employment (full or part time)'.

Education level

Some participants specified their education level as 'other' and also provided a free-text description. The TMG agreed to recode these to more appropriate categories. These are listed below:

- Education 'other' category ($n = 9$).
- Recoded as –
 - There are four National Vocational Qualification (NVQ) level 2: recode as 'O-level/O grade/General Certificate of Secondary Education (GCSE)/Standard Grade/vocational [e.g. HNC (Higher National Certificate)]'. Internet search says equivalent to five GCSE grade A–E.
 - One is NVQ level 4: recode as 'Degree/Professional/Vocational (e.g. NVQ level 4)'.
 - One IBSC (Intercalated Bachelor of Science) Global health: recode as 'Degree/Professional/Vocational (e.g. NVQ level 4)'.
 - One AS (advanced subsidiary) level: recode as 'Higher/A-level/National grade/vocational [e.g. Higher National Diploma (HND)]'.
 - One just says Diploma: recode as 'Higher/A-level/National grade/vocational (e.g. HND)'.
 - And a final one says 'declined to answer': recode as missing?

Appendix 6 Tables and histograms showing 6- and 18-month postal return and 12- and 24-month visit compliance

Some visits/form returns took place well after the compliance window when the participants were contacted after non-attendance (see the maximum time from randomisation in [Table 35](#)).

TABLE 35 Summary statistics of time from randomisation to visits/postal returns of secondary outcome questionnaires

Time point	Randomised arm		Total
	Immediate tonsillectomy	Conservative management	
6 months			
<i>n</i> (%)	109 (48)	85 (41)	194 (45)
Median	27 weeks + 4 days	27 weeks + 2 days	27 weeks + 3 days
Lower quartile	26 weeks + 4 days	26 weeks + 3 days	26 weeks + 3 days
Upper quartile	30 weeks + 4 days	30 weeks + 0 days	30 weeks + 3 days
Mean (SD)	28.7 (3.3) weeks	28.9 (4.5) weeks	28.8 (3.8) weeks
Minimum	24 weeks + 4 days	23 weeks + 0 days	23 weeks + 0 days
Maximum	42 weeks + 2 days	53 weeks + 0 days	53 weeks + 0 days
12 months			
<i>n</i> (%)	121 (55)	117 (59)	238 (57)
Median	54 weeks + 4 days	56 weeks + 4 days	56 weeks + 0 days
Lower quartile	52 weeks + 4 days	52 weeks + 2 days	52 weeks + 2 days
Upper quartile	60 weeks + 3 days	62 weeks + 0 days	61 weeks + 0 days
Mean (SD)	57.9 (10.6) weeks	60.0 (14.6) weeks	59.0 (12.7) weeks
Minimum	38 weeks + 0 days	45 weeks + 6 days	38 weeks + 0 days
Maximum	111 weeks + 4 days	169 weeks + 2 days	169 weeks + 2 days
18 months			
<i>n</i> (%)	87 (40)	75 (38)	162 (39)
Median	80 weeks + 2 days	80 weeks + 4 days	80 weeks + 2.5 days
Lower quartile	78 weeks + 5 days	79 weeks + 1 day	78 weeks + 6 days
Upper quartile	84 weeks + 5 days	82 weeks + 5 days	83 weeks + 6 days
Mean (SD)	82.2 (5.5) weeks	81.7 (4.3) weeks	81.9 (5.0) weeks
Minimum	67 weeks + 0 days	77 weeks + 2 days	67 weeks + 0 days
Maximum	109 weeks + 1 day	105 weeks + 0 days	109 weeks + 1 day

continued

TABLE 35 Summary statistics of time from randomisation to visits/postal returns of secondary outcome questionnaires (continued)

Time point	Randomised arm		Total
	Immediate tonsillectomy	Conservative management	
24 months			
<i>n</i> (%)	99 (45)	100 (52)	199 (48)
Median	106 weeks + 0 days	107 weeks + 0 days	106 weeks + 1 day
Lower quartile	103 weeks + 3 days	104 weeks + 0 days	103 weeks + 6 days
Upper quartile	110 weeks + 0 days	111 weeks + 2 days	110 weeks + 4 days
Mean (SD)	108.3 (9.5) weeks	108.8 (11.4) weeks	108.6 (10.5) weeks
Minimum	95 weeks + 3 days	52 weeks + 4 days	52 weeks + 4 days
Maximum	147 weeks + 6 days	169 weeks + 2 days	169 weeks + 2 days

% are number of forms completed out of number expected at each point allowing for withdrawals.

Charts showing the time in days between randomisation and each follow-up visit with reference lines added to show the compliance window are provided (Figures 20–23). These are in the form of histograms, with compliance reference lines added at ± 6 weeks.

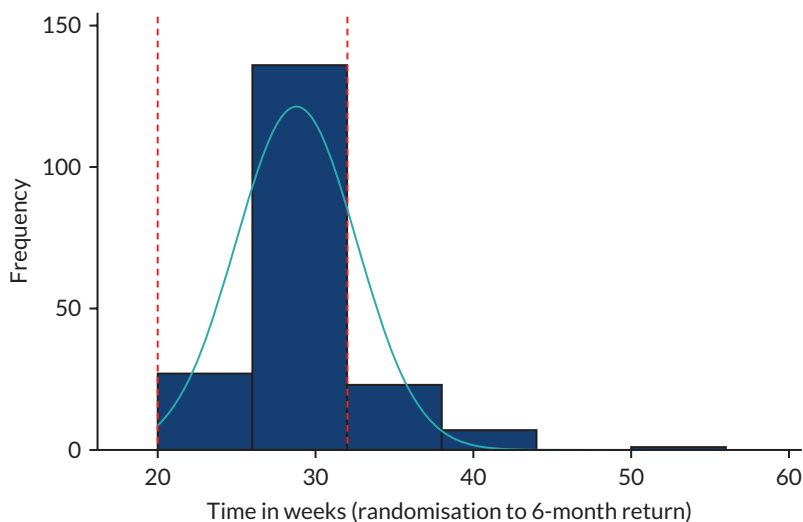


FIGURE 20 The 6-month postal returns. Red dashed lines show compliance window for postal return (20–32 weeks post randomisation) ($n = 194$). There were 195 completed TOI-14/SF-12 questionnaires. In total, 194 of these have the date completed recorded in MACRO and one has the date missing so is omitted from the graph.

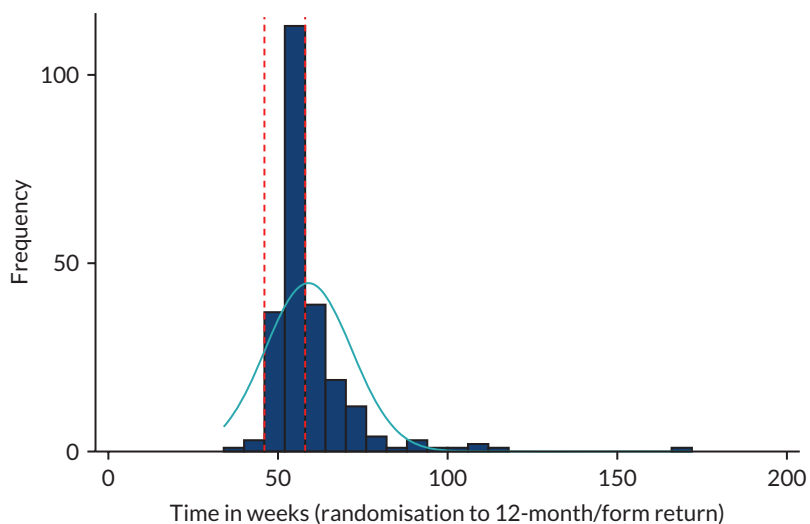


FIGURE 21 The 12-month visit. Red dashed lines show compliance window for 12-month visit/form return (46–58 weeks) ($n = 238$). There are 238 completed TOI-14/SF-12 completed. In total, 237 of these have the date completed recorded in MACRO and one has the date missing so is omitted from the graph.

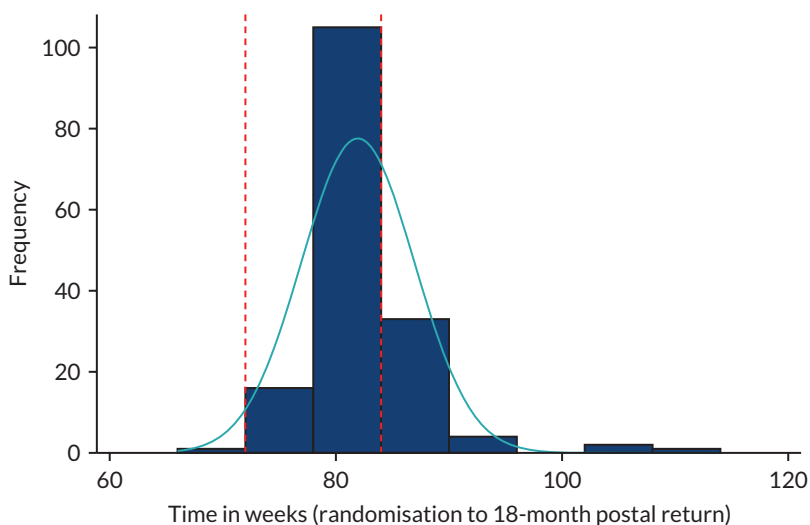


FIGURE 22 The 18-month postal returns. Red dashed lines show compliance window for 18-month postal return (72–84 weeks) ($n = 162$). There are 161 completed TOI-14/SF-12 completed. All of these have the date completed recorded in MACRO and one has the date missing. A further date is given for one participant who did not complete TOI-14 and provided only 9 of the 12 SF-12 items. This participant is included in the histogram.

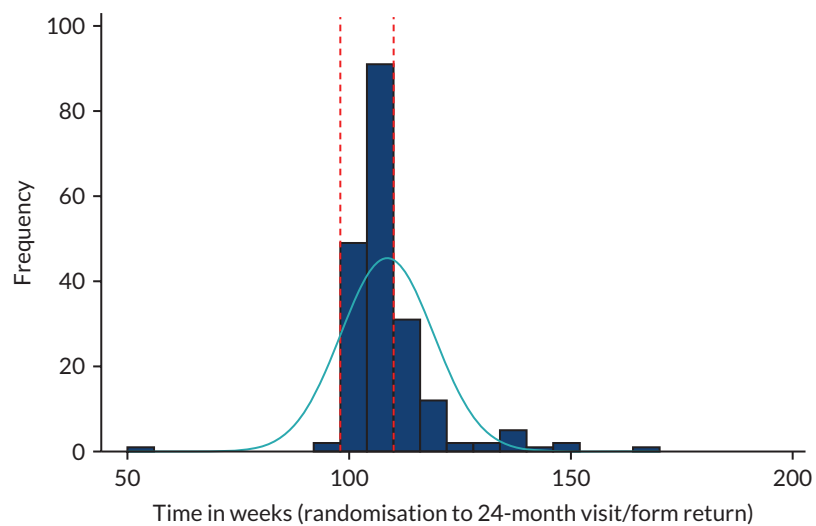


FIGURE 23 The 24-month visit. Red dashed lines show compliance window for 24-month visit/form return (98–110 weeks) ($n = 199$). There are 200 completed TOI-14/SF-12 completed. In total, 199 of these have the date completed recorded in MACRO and one has the date missing so is omitted from the graph.

Appendix 7 Further details of sensitivity analyses

Model 2

Model 2 was adjusted for the stratification factor recruiting centre (as a random effect) at randomisation and baseline severity in terms of TOI-14 as a continuous measure (as a fixed effect).

The estimated difference between the randomised arms, when accounting for site and baseline severity (as a continuous variable), indicates that patients randomised to receive immediate tonsillectomy were estimated to have total number of sore throat days reported after 24-month follow-up 0.527 times lower than those randomised to the conservative management arm (95% CI 0.43 to 0.65; $p < 0.0001$). This is very similar to the results from model 1.

Model 3

Model 3 was adjusted for baseline severity (TOI-14 as a continuous measure) and other important clinical and demographic baseline factors, specifically age, sex, ethnicity, employment status and education attained.

Univariate analyses found that sex ($p = 0.076$) and employment status (looking after family/home, $p = 0.061$; full-time education/student, $p = 0.001$; long-term sick/disabled, $p = 0.004$) could potentially be included in model 3 ($p < 0.1$). The non-linear continuous covariates, age at randomisation and continuous baseline severity were explored for suitable first-order fractional polynomial transformations but none was appropriate ($p < 0.1$). Transformation using natural log Ln (TOI-14) and complex transformation did not improve the fit significantly, so model 3 used the continuous (untransformed) TOI-14 to represent baseline severity. Employment status was the only baseline demographic data that remained significant ($p = 0.0003$) when included in model 3, also adjusted for site (random effect). The model also included the exposure variable. The estimated difference between randomised arms, when accounting for these covariates, indicated that patients randomised to receive immediate tonsillectomy were estimated to have a total number of sore throat days reported after 24-month follow-up 0.569 times lower than those randomised to conservative management (95% CI 0.46 to 0.70; $p < 0.0001$). These results are very similar to those found in model 1.

Addressing crossover (per-treatment analysis)

The estimated difference between 'per-tonsillectomy' arms, when accounting for the same covariates as model 1, found that patients receiving tonsillectomy were estimated to have a total number of sore throat days reported after 24-month follow-up 0.725 times lower than those randomised to receive conservative management (95% CI 0.587 to 0.895; $p = 0.003$).

Next, a further new indicator variable for whether or not tonsillectomy was carried out was included. This variable had four categories: randomised to tonsillectomy and tonsillectomy completed; randomised to tonsillectomy and no tonsillectomy completed; randomised to conservative management and tonsillectomy received; and randomised to conservative management with no tonsillectomy completed.

The 'as randomised by tonsillectomy received' analysis compared three patient categories with the reference category of 'randomised to conservative management with no tonsillectomy completed'.

The results can be viewed in the forest plot in [Figure 7](#). The group 'randomised to tonsillectomy and tonsillectomy received' shows a strong reduction in sore throat days (IRR 0.504, 95% CI 0.4 to 0.6; $p < 0.001$). The two remaining categories are the patients who crossed over from their randomised treatment and are relatively small groups; therefore, the 95% CIs are wide. The patients who appear to have an increase in the number of sore throats compared with the reference are those who were randomised to conservative management and received tonsillectomy. The patients who were randomised to receive tonsillectomy and had no tonsillectomy completed tended to have fewer sore throats than those randomised to conservative management and received no tonsillectomy.

Per-protocol analysis

Per-protocol analysis showed an IRR in favour of the participants receiving tonsillectomy as randomised. The IRR was reduced further to 0.415 (95% CI 0.310 to 0.554; $p < 0.001$) when adjusted by stratification variables.

It was also of interest to examine the sore throat days reported pre and post tonsillectomy. The period close to tonsillectomy was of interest because it was expected that the number of sore throats would be increased immediately following tonsillectomy. [Table 13](#) and [Figure 6](#) summarise the data. The patterns for sore throat days looked similar in both arms for the three periods for those who received tonsillectomy and for the one period for those who did not receive tonsillectomy. In both arms, there tended to be an increase in sore throats immediately following tonsillectomy, then post tonsillectomy the rates dropped to below the rates seen for the group who did not receive tonsillectomy.

Intention-to-treat population restricted to those with 80% STAR data

Although sample sizes were now much smaller ($n = 263$), for the primary analysis the IRR was similar to when all 429 participants were included with any amount of STAR data (IRR 0.533, 95% CI 0.427 to 0.666; $p < 0.001$). Similarly for the four category per-treatment, no tonsillectomy randomised to tonsillectomy IRR (CI) was now slightly lower (IRR 0.468, 95% CI 0.300 to 0.734; $p = 0.001$), tonsillectomy as randomised was slightly higher (IRR 0.584, 95% CI 0.451 to 0.757; $p < 0.001$), whereas the tonsillectomy crossed over from conservative management was similar (IRR 1.205, 95% CI 0.830 to 1.750; $p = 0.326$).

Appendix 8 Withdrawals

TABLE 36 Summary statistics for time from randomisation to withdrawal in weeks by arm

Time from randomisation to withdrawal	Randomised arm		Overall
	Immediate tonsillectomy	Conservative management	
<i>n</i>	16	27	43
Median	31 weeks + 4 days	39 weeks + 1 day	33 weeks + 0 days
Lower quartile	16 weeks + 6 days	1 week + 6 days	13 weeks + 5 days
Upper quartile	52 weeks + 1 day	53 weeks + 4 days	53 weeks + 4 days
Minimum	0 weeks + 2 days	0 weeks + 0 days	0 weeks + 0 days
Maximum	140 weeks + 1 day ^a	103 weeks + 5 days	140 weeks + 1 day ^a

a One participant withdrew beyond the 24-month follow-up period (105 weeks).

TABLE 37 Participants who withdrew from the trial (n = 43)

Arm	Randomisation date	Withdrawal visit	Time from randomisation to withdrawal (weeks)	Withdrawal date	Withdrawal reasons	Further information if withdrew reason was 'other'	Other information
Conservative management	8 December 2015	Baseline	0	8 December 2015	Other	He changed his mind after randomisation, and asked to have the surgery off the study. He was randomised to the conservative management arm. Withdrew consent	
Conservative management	29 February 2016	Baseline	0	2 March 2016	Participant withdrew consent	N/A	Patient felt a sore throat coming on during the night, and decided that she could not wait any longer, and would like to have the surgery not as part of the study. She was randomised to the conservative management arm
Conservative management	24 July 2015	Other	39	19 April 2016	Other	Unwilling to complete STAR forms	
Immediate tonsillectomy	23 October 2015	Other	31	27 May 2016	Participant withdrew consent	N/A	No longer wishes to have a tonsillectomy. Offered move to conservative management arm but is finding weekly alerts annoying
Conservative management	8 June 2016	Baseline	0	8 June 2016	Other	Participant had not fully understood the nature of the randomisation process	
Conservative management	22 June 2016	Baseline	0	22 June 2016	Participant withdrew consent	N/A	
Immediate tonsillectomy	17 November 2015	Baseline	33	5 July 2016	Other	Protocol deviation	Patient failed to have tonsillectomy within the time frame and could not contact her

Arm	Randomisation date	Withdrawal visit	Time from randomisation to withdrawal (weeks)	Withdrawal date	Withdrawal reasons	Further information if withdrawal reason was 'other'	Other information
Immediate tonsillectomy	27 May 2016	Baseline	14	31 August 2016	Other	PI decision to withdraw owing to non-attendance with appointments	
Conservative management	15 December 2015	12-month follow-up	40	20 September 2016	Participant withdrew consent	N/A	
Immediate tonsillectomy	18 August 2015	12-month follow-up	57	20 September 2016	Other	Site lost contact with patient, GP returned a form stating patient has declined the request as she opted out of the trial	
Immediate tonsillectomy	6 June 2016	Baseline	17	6 October 2016	Other	Participant was pregnant	
Conservative management	4 August 2016	Other	11	18 October 2016	Other	Patient found the STAR reminders and questionnaires too much to complete on a weekly basis	
Immediate tonsillectomy	28 June 2016	Baseline	17	24 October 2016	Other	Lost to follow-up	Patient has not attended several appointments and surgery has not gone ahead. I can confirm the patient has withdrawn
Conservative management	21 March 2016	2-week call	31	26 October 2016	Participant withdrew consent	N/A	
Conservative management	14 June 2016		20	3 November 2016	Participant withdrew consent	N/A	No other reason given, did not want to continue with the study
Conservative management	24 November 2016	Baseline	0	24 November 2016	Participant withdrew consent	N/A	Participant wanted tonsillectomy as soon as possible but had been randomised to conservative treatment. Did not want to cross over as no guarantee that surgery would be in the next week or so. Was going to go private

continued

TABLE 37 Participants who withdrew from the trial (n = 43) (continued)

Arm	Randomisation date	Withdrawal visit	Time from randomisation to withdrawal (weeks)	Withdrawal date	Withdrawal reasons	Further information if withdrew reason was 'other'	Other information
Conservative management	23 October 2015	Other	65	19 January 2017	Other	Not completing/responding to STARS	
Immediate tonsillectomy	23 September 2016	Other	17	19 January 2017	Other	Patient has moved to Glasgow and cannot proceed to tonsillectomy in Aberdeen	
Conservative management	26 April 2016	12-month follow-up	44	28 February 2017	Participant withdrew consent	N/A	
Immediate tonsillectomy	21 November 2016	Other	17	21 March 2017	Other	Patient decided to withdraw from the study because he does not want certain parts of his medical history to be known	
Conservative management	2 February 2017	Other	17	1 June 2017	Other	Patient got fed up of weekly questionnaires as her husband is being diagnosed with cancer	
Immediate tonsillectomy	17 July 2017	Baseline	0	19 July 2017	Participant withdrew consent	N/A	Patient wanted tonsillectomy off study. Didn't want to be burned with e-mails and questionnaires
Immediate tonsillectomy	20 December 2016		35	25 August 2017	Other	Patient requested to be taken off the waiting list for surgery. No other correspondence documented since this date	
Immediate tonsillectomy	3 October 2016	Baseline	47	30 August 2017	Other	Patient did not want to receive the weekly alerts anymore	
Conservative management	5 December 2016	12-month follow-up	39	5 September 2017	Participant withdrew consent	N/A	Patient continued to experience episodes of tonsillitis

Arm	Randomisation date	Withdrawal visit	Time from randomisation to withdrawal (weeks)	Withdrawal date	Withdrawal reasons	Further information if withdrew reason was 'other'	Other information
Conservative management	29 August 2017	Other	2	11 September 2017	Other	Asked to be listed for tonsillectomy without study involvement. Developed tonsillitis and felt that completing STAR forms when ill and attending the GP for antibiotics was too much	
Conservative management	8 September 2016	Baseline	54	18 September 2017	Other	Moving house, did not want to remain as postal/telephone participant, did not come to 12-month visit	
Conservative management	26 October 2016	12-month follow-up	52	23 October 2017	Participant withdrew consent	N/A	Trial is time-consuming and sending questionnaire back was not working for me
Conservative management	25 October 2017	Baseline	0	25 October 2017	Participant withdrew consent	N/A	After randomised to usual care decided she really wanted tonsillectomy
Immediate tonsillectomy	21 March 2017	Baseline	32	1 November 2017	Participant withdrew consent	N/A	Patient randomised to tonsillectomy arm but cancelled his surgery dates on numerous occasions. Appointment made to be seen back in clinic in September to re-list him for surgery but this was cancelled. Patient e-mailed to say that he no longer wished to continue in the trial
Conservative management	14 July 2017	1-week call	23	20 December 2017	Participant withdrew consent	N/A	Patient decided she was no longer willing to complete weekly STAR forms and did not wish to be contacted any further by NCTU or local research team

continued

TABLE 37 Participants who withdrew from the trial (n = 43) (continued)

Arm	Randomisation date	Withdrawal visit	Time from randomisation to withdrawal (weeks)	Withdrawal date	Withdrawal reasons	Further information if withdrew reason was 'other'	Other information
Immediate tonsillectomy	7 March 2016		96	8 January 2018	Participant withdrew consent	N/A	No reason given by patient. Patient informed me that he had withdrawn consent when I contacted him on 8 January 2018
Conservative management	2 March 2016	Other	104	26 February 2018	Participant withdrew consent	N/A	Patient is now working abroad
Conservative management	12 January 2017	Other	59	2 March 2018	Other	Participant's condition no longer requires treatment	Participant stated she did not feel she was a suitable candidate for the study as she has not been having any sore throats and advised she wished not to take part anymore
Conservative management	29 November 2016	12-month follow-up	67	14 March 2018	Other	Patient has no time owing to university	
Immediate tonsillectomy	26 May 2016	Tonsillectomy	103	15 May 2018	Other	Lost to follow-up	I have tried to contact this patient several times. She did not attend 12- or 24-month follow-up or return questionnaires posted out. Has not responded to e-mails. I can confirm this patient has withdrawn
Conservative management	22 May 2017	12-month follow-up	53	25 May 2018	Participant withdrew consent	N/A	Unhappy that she received text alerts after requesting e-mail alerts
Immediate tonsillectomy	11 April 2018	Other	7	30 May 2018	Other	AE(s): complete AE form	
Conservative management	17 October 2016		97	28 August 2018	Participant withdrew consent	N/A	Too busy

Arm	Randomisation date	Withdrawal visit	Time from randomisation to withdrawal (weeks)	Withdrawal date	Withdrawal reasons	Further information if withdrawal reason was 'other'	Other information
Conservative management	8 December 2017		59	22 January 2019	Participant withdrew consent	N/A	Participant was under the impression that the study follow-up was only for 12 months. She does not wish to continue receiving/completing STAR alerts
Conservative management	21 March 2018		52	21 March 2019	Participant withdrew consent	N/A	Participant was contacted over the telephone for her 12-month follow-up. She informed me she would like to withdraw from the study as she has not had a sore throat for several months now and that she had not completed her online questionnaires
Conservative management	26 April 2018		49	4 April 2019	Other	Patient has other problems in life and does not want to carry on with study	Patient feels they have too much going on
Immediate tonsillectomy	20 September 2016		140 ^a	29 May 2019	Participant withdrew consent	N/A	No reason given
N/A, not applicable. a Withdrew after 24-month follow-up had ended.							

Appendix 9 Time from randomisation to tonsillectomy

TABLE 38 Summary statistics of time to tonsillectomy by arm and overall

Time from randomisation to tonsillectomy	Randomised arm		Total (n = 231)
	Immediate tonsillectomy (n = 172)	Conservative management (n = 59) ^a	
Lower quartile	5 weeks + 1 day	14 weeks + 2 days	5 weeks + 6 days
Median	7 weeks + 2 days	33 weeks + 6 days	9 weeks + 1 day
Upper quartile	11 weeks + 0 days	65 weeks + 6 days	17 weeks + 6 days
Mean (SD)	9.3 (7.3)	40.1 (28.0) weeks	17.2 (20.5) weeks
Minimum	0 weeks + 4 days	1 week + 0 days	0 weeks + 4 days
Maximum	63 weeks + 2 days	104 weeks + 0 days	104 weeks + 0 days

a One tonsillectomy was carried out off trial (so not crossover) after 12 weeks + 1 day post randomisation.

Tonsillectomy methods were at the discretion of the participating centres and are shown in [Table 39](#).

TABLE 39 Type of tonsillectomy delivered in NATTINA

Type of tonsillectomy	Randomised arm, n (%)		Total (N = 231), n (%)
	Immediate tonsillectomy (N = 172)	Conservative management (N = 59)	
Cold dissection	100 (58)	21 (36)	121 (52)
Bipolar diathermy	60 (35)	32 (55)	92 (40)
Coblation	1 (<1)	0 (0)	1 (<1)
Laser	1 (<1)	0 (0)	1 (<1)
Mixed ^a	7 (4)	1 (2)	8 (3)
Off-trial tonsillectomy	0 (0)	1 (2)	1 (<1)
Not specified	3 (2)	4 (7)	7 (3)

a Combination of cold and bipolar dissection.

TABLE 40 Further details of complications with tonsillectomy reported (n = 10)

Tonsillectomy type	Tonsillectomy date	Randomised to tonsillectomy	Severity of AE (complication)	Compilation details (AE text)
Bipolar diathermy	30 July 2015	Yes	Severe	Throat pain
Cold dissection	21 August 2015	Yes	Severe	Throat pain
Bipolar diathermy	14 October 2016	Yes	Moderate	Vomiting
Cold dissection	16 November 2016	Yes	Mild	Sinus tachycardia post surgery: kept overnight following surgery. Discharged day after surgery
Cold dissection	5 December 2016	Yes	Severe	Hospitalisation (post day case surgery)
Cold dissection	18 July 2017	Yes	Moderate	Difficulty swallowing
Bipolar diathermy	15 February 2018	No	Moderate	Pain
Bipolar diathermy	19 February 2018	Yes	Mild	Tonsils removed with difficulty. Rim of left tonsil left in situ
Bipolar diathermy	16 March 2018	Yes	Mild	Pain
Missing info	26 July 2018	Yes	Mild	Small primary post-op bleed

Appendix 10 Additional safety information

TABLE 41 One-hundred and ninety-one reported AEs for 90 participants undergoing tonsillectomy (SAEs removed)

Severity	Randomised to immediate tonsillectomy arm (69 participants over all categories), n (%)		Crossed over from conservative management arm (21 participants over all categories), n (%)		Total (90 participants over all categories), n (%)	
	AEs reported	Participants	AEs reported	Participants	AEs reported	Participants
Mild	65 (45)	48	24 (52)	16	89 (47)	64
Moderate	60 (41)	33	17 (37)	11	77 (40)	44
Severe	20 (14)	17	5 (11)	5	25 (13)	22
Total	145 (100)	69	46 (100)	21	191 (100)	90

TABLE 42 Details of SAEs deemed to be caused by the tonsillectomy ranked by severity: severe, moderate and mild (n = 51)

SAE number	Severity	Tonsillectomy date	Start date	Recovery date	SAE description	Case description
SAE006	Severe	25 January 2016	29 January 2016	30 January 2016	Post-tonsillectomy bleed	Pain, difficulty swallowing, coughing up blood clots
SAE011	Severe	13 April 2016	15 April 2016	25 April 2016	Throat infection	Re-admitted to hospital with throat infection post tonsillectomy 1 day. Unable to swallow and take in fluids. IV co-amoxiclav 1–2 g three times per day given for 2 days. IV dexamethasone 8 mg three times per daily given once. Paracetamol 1 g four times daily. Difflam mouthwash 10 ml four times per day. Ibuprofen 400 mg three times daily. Ibuprofen 400 mg three times daily. Tramadol 50 mg three times per day. Codeine phosphate 60 mg four times per day
SAE025	Severe	5 December 2016	10 December 2016	19 December 2016	Infection (post tonsillectomy)	10 December 2016: onset of cough and sharp substance in throat. Cough = painful. Coughed up stitch on 11 December 2016. Telephoned 111, told to go to A&E came in then to ward. Two doses of IV antibiotics given (to follow). Discharged home on 12 December 2016 with oral antibiotics. 15 December 2016 1 × 1.2 IV co-amoxiclav 06.00 given in hospital

continued

TABLE 42 Details of SAEs deemed to be caused by the tonsillectomy ranked by severity: severe, moderate and mild (n=51) (continued)

SAE number	Severity	Tonsillectomy date	Start date	Recovery date	SAE description	Case description
SAE026	Severe	5 December 2016	5 December 2016	6 December 2016	Hospitalisation post day case surgery	Participant reports she was kept in overnight as she lost blood during surgery. Unable to locate notes so far (15 December 2016)
SAE070	Severe	16 January 2017	19 January 2017	20 January 2017	Post-tonsillectomy bleed	
SAE033	Severe	16 February 2017	21 February 2017	24 February 2017	Post-tonsillectomy bleed	
SAE040	Severe	31 July 2017	5 August 2017	10 August 2017	Post-operation bleeding	Post-operation bleeding 31 July 2017
SAE051	Severe	24 January 2018	29 January 2018	8 February 2018	Post-tonsillectomy bleed	7 days post tonsillectomy, noticed some blood in saliva over preceding 3 days seen GP who saw superior pole clot. Further bleeding in hospital. Conservatively managed with antibiotics and hydrogen peroxide gargles. Remained in hospital for observation
SAE001	Moderate	14 January 2016	16 January 2016	16 January 2016	Bleeding from tonsils	Bleeding from tonsils 16 January 2016 (2 days post surgery) stopped after 20 minutes. Had been shouting
SAE005	Moderate	14 January 2016	23 January 2016	25 January 2016	Post-tonsillectomy haemorrhage	Spontaneous bleeding from tonsillar bed
SAE007	Moderate	25 January 2016	6 February 2016	7 February 2016	Post-tonsillectomy bleed	Coughing up blood clots
SAE008	Moderate	19 February 2016	20 February 2016	20 February 2016	Post-tonsillectomy bleed	The evening following tonsillectomy she had bleeding from tonsils and was hypotensive. Admitted to hospital and treated with IV co-amoxiclav, antibiotics and analgesia. Discharged 22 February 2016
SAE014	Moderate	27 June 2016	2 July 2016	4 July 2016	Post-tonsillectomy bleed	Presented at hospital 5 days post op with two small bleeds. Admitted for observation. Further bleed overnight, treated with oral rinse. Observed for further 24 hours, no further bleeds
SAE017	Moderate	31 August 2016	5 September 2016	6 September 2016	Post-tonsillectomy bleed	Admitted to hospital on 5 September 2016 with bleeding from tonsils
SAE019	Moderate	9 May 2016	18 May 2016	20 May 2016	Post-tonsillectomy secondary bleed	Bleeding from right tonsillar fossa 10 days after tonsillectomy, hydrogen peroxide gargles and analgesia
SAE020	Moderate	15 September 2016	16 September 2016	29 September 2016	Post-tonsillectomy pain	Sore throat and dysphagia 6 days post tonsillectomy

TABLE 42 Details of SAEs deemed to be caused by the tonsillectomy ranked by severity: severe, moderate and mild (n=51) (continued)

SAE number	Severity	Tonsillectomy date	Start date	Recovery date	SAE description	Case description
SAE030	Moderate	22 December 2016	27 December 2016	28 December 2016	Post-tonsillectomy bleed	Admitted to hospital on 27 December 2016 following bleeding from surgery side
SAE031	Moderate	23 November 2016	4 December 2016	4 December 2016	Post-tonsillectomy bleed	Tonsillectomy performed 23 November 2016 began to bleed from tonsil bed on 4 December 2016 admitted to hospital and observed. Bleeding spontaneously resolved. Discharged on 5 December 2016 with antibiotics gargle and mouth wash. Now recovered
SAE032	Moderate	6 February 2017	14 February 2017	16 February 2017	Post-operative haemorrhage	Bleeding
SAE034	Moderate	15 December 2016	15 December 2016	17 December 2016	Post-tonsillectomy pain	Post-op pain following surgery
SAE036	Moderate	30 March 2017	4 April 2017	19 April 2017	Post-operative bleeding	Bleeding post tonsillectomy, participant had a second episode of bleeding and a third (same site)
SAE037	Moderate	4 May 2017	11 May 2017	18 May 2017	Haemorrhage	Admitted to inverness A&E (staying with family post op) with bleeding on 11 May 2017 stayed for 3 nights, discharged 14 May 2017. Received IV clarithromycin and IV fluid
SAE042	Moderate	3 August 2017	10 August 2017	11 August 2017	Post-tonsillectomy haemorrhage	Self-limiting post-tonsillectomy haemorrhage
SAE044	Moderate	1 November 2017	11 November 2017	13 November 2017	Post-tonsillectomy bleed	Four to five episodes of coughing up clots and fresh blood. Has not eaten much solid food post op for odynophagia. Clot noticed in right tonsillar fossa
SAE045	Moderate	3 November 2017	8 November 2017	8 November 2017	Post-tonsillectomy bleed	5 days post tonsillectomy was admitted to hospital after spitting up blood
SAE047	Moderate	8 December 2017	10 December 2017	22 December 2017	Post-tonsillectomy pain	Attended A&E with deep abscess in throat, painful left ear, difficulty eating and drinking 10 December 2017. Admitted to ENT ward 11 December 2017 antibiotics and pain relief administered. Discharged 12 December 2017
SAE053	Moderate	14 February 2018	21 February 2018	1 March 2018	Minor haemorrhage tonsillar bed	Participant admitted to Ninewells Hospital in Dundee on 21 February 2018 for overnight stay owing to minor haemorrhage in throat tonsillar bed observed overnight, discharged 22 February 2018. Attended A&E clinic on 1 March 2018 after attending GP for persistent symptoms

continued

TABLE 42 Details of SAEs deemed to be caused by the tonsillectomy ranked by severity: severe, moderate and mild (n = 51) (continued)

SAE number	Severity	Tonsillectomy date	Start date	Recovery date	SAE description	Case description
SAE055	Moderate	26 February 2018	3 March 2018	10 March 2018	Poorly controlled pain and dehydration post tonsillectomy	Not eaten or drunk much since discharge, also had small flecks of blood in saliva but no fresh blood
SAE056	Moderate	26 February 2018	2 March 2018	3 March 2018	Post-tonsillectomy bleeding and nausea	Patient felt unwell, vomited then expectorating fresh blood twice, approximately two spoonfuls. Remained nauseated
SAE061	Moderate	2 May 2018	7 May 2018	8 May 2018	Post-tonsillectomy bleed	Admitted to hospital 6 days post tonsillectomy following bleeding. Discharged the next day with oral co-amoxiclav
SAE062	Moderate	9 May 2018	11 May 2018	11 May 2018	Post-operative vomiting and bleeding	Presented to A&E on 11 May 2018 with frequent vomiting (every 15 minutes) since 4 a.m. Patient had vomited fresh blood. Admitted to ENT unit, IV fluids and IV paracetamol administered. Symptom resolved. Patient was discharged with advice on 11 May 2018
SAE063	Moderate	16 May 2018	18 May 2018	19 May 2018	Post-tonsillectomy infection	Presented at hospital with rigors and vomiting: 2 days post tonsillectomy
SAE067	Moderate	30 July 2015	8 August 2015	8 August 2015	Haemorrhage	Haemorrhage 9 days post tonsillectomy
SAE041	Moderate	5 June 2017	11 June 2017	13 June 2017	Post-tonsillectomy bleed and infection	Pain and clots of blood at back of throat 1 week after tonsillectomy. Struggling to eat and drink
SAE010	Moderate	9 March 2016	10 March 2016	17 October 2016	Throat infection	High temperature post tonsillectomy on 9 March 2016, started day after op, pain in throat, cervical light lymphadenopathy, exudate from tonsil bed, inflammation of tonsil bed, unable to eat or drink, dehydrated, difficulty opening mouth, admitted for IV antibiotics
SAE002	Mild	14 January 2016	15 January 2016	17 January 2016	Uvulitis	Admitted to ENT ward owing to difficulty with breathing and swallowing. Haemodynamically stable, no signs of bleeding
SAE003	Mild	14 January 2016	22 January 2016	23 January 2016	Post-tonsillectomy bleed	Post-tonsillectomy bleed 8 days post op woke up with a mixture of clots and fresh blood in mouth no active bleeding on admission slough on both tonsils
SAE012	Mild	15 June 2016	20 June 2016	22 June 2016	Post-operative bleeding	Admitted on 21 June 2016 with post-tonsillectomy bleed. This settled by itself. Discharged home on 22 June 2016 with oral co-amoxiclav antibiotics

TABLE 42 Details of SAEs deemed to be caused by the tonsillectomy ranked by severity: severe, moderate and mild (n=51) (continued)

SAE number	Severity	Tonsillectomy date	Start date	Recovery date	SAE description	Case description
SAE013	Mild	15 June 2016	23 June 2016	27 June 2016	Tonsillar haemorrhage	Admitted after 2-day history of post-tonsillectomy bleeding. Settled with conservative measures
SAE016	Mild	24 August 2016	28 August 2016	29 August 2016	Post-tonsillectomy pain	Admitted on 29 August 2016 following pain from tonsillectomy and unable to tolerate oral fluids. Also has rash brought on by osomorph
SAE022	Mild	24 August 2016	25 August 2016	31 August 2016	Pain and infection post tonsillectomy	In pain and unable to swallow, admitted to ward via causality for IV fluids, pain relief and antibiotics, stayed in hospital overnight discharged 26 August 2016
SAE023	Mild	21 September 2016	25 September 2016	2 October 2016	Throat infection post tonsillectomy	Unable to swallow post tonsillectomy, rising temperature, rigors and dehydration
SAE027	Mild	6 December 2016	14 December 2016	16 December 2016	Post-operative bleeding	Two episodes of blood-stained spit lasting 20 minutes each
SAE035	Mild	3 March 2017	7 March 2017	9 March 2017	Post-tonsillectomy bleeding	Tonsillectomy on 3 March 17, bleeding started evening of 7 March 17
SAE050	Mild	25 May 2017	30 May 2017	2 June 2017	Post-operative bleeding following tonsillectomy	Coughed up small amount of blood and finally coughed up large clot from back of her throat. Came to A&E. Has had bleeding on and off since tonsillectomy. No fever but general unwell feeling
SAE052	Mild	8 February 2018	13 February 2018	16 February 2018	Post-tonsillectomy bleed	Patient experienced bleeding from tonsillectomy site
SAE057	Mild	13 March 2018	19 March 2018	20 March 2018	Post-operative bleeding	Participant began bleeding in the early hours of 19 March 2018. Attended A&E and was admitted for overnight observation. He was discharged next morning
SAE065	Mild	25 May 2018	1 June 2018	4 June 2018	Painful throat, difficulty swallowing, fever 37.2°C	Admitted to ENT ward 1 June 2018 post tonsillectomy 25 May 2018 with 2 days, history of increasingly painful throat and difficulty swallowing
SAE066	Mild	30 July 2015	30 July 2015	31 July 2015	Intraoperative bleeding (excessive)	Surgeon reported more bleeding than usual during tonsillectomy (500 ml): bipolar technique
SAE068	Mild	16 November 2016	16 November 2016	17 November 2016	Drowsiness? Vasovagal post op	
SAE069	Mild	16 November 2016	16 November 2016	17 November 2016	Sinus tachycardia post op	

IV, intravenous.

TABLE 43 Severity and causality of the SAEs owing to the intervention reported

SAE severity	Randomised to immediate tonsillectomy arm, <i>n</i> (%)		Crossed over from conservative management arm, <i>n</i> (%)		Total, <i>n</i> (%)	
	SAE	Participants	SAE	Participants	SAE	Participants
Severe	7 (17)	6	1 (11)	1	8 (16)	7
Moderate	20 (48)	19	7 (78)	7	27 (53)	26
Mild	15 (36)	13	1 (11)	1	16 (31)	14
Overall	42 (100)	36	9 (100)	9	51 (100)	45

Appendix 11 Measuring Sore Throat Alert Return response

The patient population is the ITT group, but some participants did not have any primary outcome data so the population is reduced (see below):

- ITT – all ineligible and protocol violator participants will be included in the analysis on an ITT basis, with participants kept in their randomised treatment group.
- Total number of participants randomised, $n = 455$.
- Two ineligible participants would have been included in the ITT analysis if data were collected. However, as they were randomised into NATTINA by mistake after consenting to be involved in other trials they withdrew immediately when the error was realised, so no data were collected. Therefore, omitted from the ITT analysis.
- One participant from unresponsive site did not have baseline data collected. However, they did have stratification data from screening and the participant did make a few STAR responses. Therefore, this participant was included in the ITT analysis.
- There were a total of 43 participants who withdrew during the course of the trial. Some of these participants made some STAR responses so these are included in the analysis retaining all STAR responses received up until the point of withdrawal.
- A total of 26 participants did not make any STAR responses during the course of the trial (two were ineligible as detailed above and nine of these officially withdrew). Thus, they had no primary outcome data so could not be included in the primary ITT analysis.
- In addition, four other participants had some missing baseline data. However, these did have stratification data collected and the participants did make STAR responses. Therefore, these participants were included in the ITT analysis.
- Total included in primary ITT analysis was, therefore, 429 participants.

Appendix 12 Summary statistics for crossing or remaining in randomised arm and tonsillectomy versus no tonsillectomy, with rates

TABLE 44 Summary statistics for crossing or remaining in randomised arm and tonsillectomy vs. no tonsillectomy, with rates [average sore throat days per return (week)]: per-treatment

Statistic	Time randomisation to tonsillectomy	Sore throat rate			
		Pre tonsillectomy	2 weeks following tonsillectomy	Post 2 weeks tonsillectomy	No tonsillectomy
Tonsillectomy as randomised					
<i>n</i> (% of randomised)	167 (72)	157	143	165	N/A
Median (IQR)	7.6 (5.1–11.0)	1.00 (0.50–2.50)	4.50 (2.00–6.00)	0.13 (0.03–0.32)	
Mean (SD)	9.4 (7.4)	1.71 (1.62)	3.87 (2.36)	0.33 (0.69)	
Min, max	0.57, 63.3	0, 7	0, 7	0, 5	
Tonsillectomy crossed					
<i>n</i> (% of randomised)	57 (26)	56	43	46	N/A
Median (IQR)	33.9 (14.4–65.9)	1.37 (0.84–2.80)	3.50 (0.00–5.50)	0.23 (0.05–0.39)	
Mean (SD)	40.5 (28.1)	1.85 (1.50)	3.41 (2.54)	0.51 (1.07)	
Min, max	1, 104	0, 6.5	0, 7	0, 6	
No tonsillectomy as randomised					
<i>n</i> (% of randomised)	N/A	N/A	N/A	N/A	147
Median (IQR)					0.50 (0.20–1.10)
Mean (SD)					1.10 (1.54)
Min, max					0, 7
No tonsillectomy randomised to tonsillectomy					
<i>n</i> (% of randomised)	N/A	N/A	N/A	N/A	57
Median (IQR)					0.41 (0.21–1.20)
Mean (SD)					0.90 (1.20)
Range (LQ, UQ)					0, 6.27

LQ, lower quartile; max, maximum, min, minimum; NA, not applicable; UQ, upper quartile.

- Average time to tonsillectomy for those randomised to tonsillectomy was 9.4 weeks compared to 40.5 weeks for those crossing over from conservative management.
- On average patients randomised to immediate tonsillectomy who went on to receive the operation had a pre-tonsillectomy rate of 1.71, 2 weeks after tonsillectomy rate of 3.87 and for remaining follow-up period 0.33 sore throats per week.
- Higher sore throats rate pre tonsillectomy for those requesting to cross over to receive tonsillectomy.
- Similar sore throats rates for the 2-week period after tonsillectomy. But this would be expected owing to the procedure causing sore throats soon after tonsillectomy.
- Those crossing over to have tonsillectomy continued with higher sore throat rate for rest of follow-up period.
- Rates slightly higher on average for those patients randomised to conservative management who did not cross over.
- Those randomised to tonsillectomy who did not receive the intervention had a lower average rate than the pre-tonsillectomy phase of those receiving the tonsillectomy as randomised.

Appendix 13 Instrumental variables

Complier-average causal effect

The CACE approach assesses the effect of treatment scaled up to represent full participant inclusion accounting for non-compliance. This was analysed twice in each direction, with each arm set as the reference value. Results are shown in [Table 45](#).

Using the CACE approach for unadjusted NBR with conservative management set as the reference value and adjusted for proportion complying with tonsillectomy within 8 weeks, we observe an improved effect of tonsillectomy: for each 1 sore throat day reported by patients in the conservative management arm, those receiving tonsillectomy as per protocol can expect 0.256 (95% CI 0.143 to 0.412) sore throat days with no adjustment for stratification factors. When adjusted, the IRR was 0.195 (95% CI 0.114 to 0.332).

When the same procedure was applied with immediate tonsillectomy set as the reference category, unadjusted NBR complying with conservative management showed that those in the conservative management arm had 2.496 (95% CI 1.773 to 3.522) more sore throat days than those randomised to receive tonsillectomy. This increased to 2.878 (95% CI 2.040 to 4.067) more sore throat days when adjusted for stratification factors.

TABLE 45 Results of the CACE analyses (unadjusted and adjusted) with each arm set as the reference category

Reference	Proportion complied ^a	IRR (coefficient)	SE	CI (CI coefficient)	Adjusted IRR (CI)
Conservative management (unadjusted)	91/233 = 0.3906	0.575 (-0.553)	0.061 (0.106)	0.467 to 0.708 (-0.761 to -0.346)	Adjusted coefficient: -1.3636 Adjusted IRR is $\exp(-1.3636) = 0.256$ (0.143 to 0.412)
Conservative management (adjusted by stratification)	91/233 = 0.3906	0.528 (-0.639)	0.056 (0.106)	0.428 to 0.650 (-0.848 to -0.431)	Adjusted coefficient: -1.6359 Adjusted IRR is $\exp(-1.6359) = 0.195$ (0.114 to 0.332)
Immediate tonsillectomy (unadjusted)	133/220 = 0.6045	1.739 (0.553)	0.184 (0.106)	1.413 to 2.139 (0.346 to 0.761)	Adjusted coefficient is 0.9148 Adjusted IRR is $\exp(0.9148) = 2.496$ (1.773 to 3.522)
Immediate tonsillectomy (adjusted by stratification)	133/220 = 0.6045	1.895 (0.639)	0.202 (0.106)	1.539 to 2.334 (0.431 to 0.848)	Adjusted coefficient is 1.0571 Adjusted IRR is $\exp(1.0571) = 2.878$ (2.040 to 4.067)

^a Protocol compliance (tonsillectomy within 8 weeks or received conservative management).

Interaction model with interaction between baseline severity and randomised arm

- Adjusted immediate tonsillectomy, $n = 224$.
- Conservative management, $n = 205$.
- IRR 0.411 (95% CI 0.264 to 0.639). The reference group is conservative management.
- Moderate IRR 1.060 (95% CI 0.717 to 1.573), severe IRR 0.851 (95% CI 0.569 to 1.272). Reference is mild.
- Interaction terms: tonsillectomy \times moderate IRR 1.307 (95% CI 0.759 to 2.251) and tonsillectomy \times severe IRR 1.467 (95% CI 0.837 to 2.570).

The interaction with baseline severity showed that there was a tendency for those classed as mild at baseline to have more of an improvement with tonsillectomy than those classed as moderate or severe at baseline; however, the interaction terms were not statistically significant.

Appendix 14 Secondary analyses: Tonsillectomy Outcome Inventory-14

Tables 46–50 show the summary statistics for TOI-14 scores and the subscales. These are similar at baseline as expected, but there are marked differences at the follow-up visits (after surgery) in the overall TOI-14 scores (non-overlapping CIs). The box plots (Figures 24–27) for the subscales follow. The box plots and tables include the imputed item scores as described in Table 51.

TABLE 46 Tonsillectomy Outcome Inventory-14 throat discomfort score completion rate and scores at time points

	Randomised arm		Total (N = 453)
	Tonsillectomy (N = 233)	Conservative management (N = 220)	
Baseline			
n (%)	231 (99)	217 (99)	448 (99)
Median (IQR)	14 (12–16)	14 (12–16)	14 (12–16)
Mean (SD)	13.5 (3.3)	13.5 (3.4)	13.5 (3.4)
95% CI about mean	13.1 to 13.9	13.1 to 14.0	13.2 to 13.8
Min, max	2, 20	2, 20	2, 20
6 months (post)			
n (%)	105 (45)	85 (39)	190 (42)
Median (IQR)	4 (1–9)	10 (6–13)	6 (2–11)
Mean (SD)	5.2 (4.6)	9.2 (4.8)	7.0 (5.1)
95% CI about mean	4.3, 6.1	8.2 to 10.3	6.3 to 7.7
Min, max	0, 19	0, 19	0, 19
12 months (visit)			
n (%)	122 (55)	117 (50)	239 (53)
Median (IQR)	2 (0–4)	8 (4–12)	4 (1–8)
Mean (SD)	2.5 (2.5)	7.9 (5.0)	5.1 (4.7)
95% CI about mean	2.0 to 2.9	6.9 to 8.8	4.5 to 5.7
Min, max	0, 10	0, 19	0, 19
18 months (post)			
n (%)	83 (36)	71 (32)	154 (34)
Median (IQR)	2 (1–4)	6 (2–11)	3 (1–7)
Mean (SD)	2.8 (3.1)	6.5 (5.2)	4.5 (4.6)
95% CI about mean	2.1 to 3.5	5.3 to 7.8	3.8 to 5.3
Min, max	0, 18	0, 20	0, 20

continued

TABLE 46 Tonsillectomy Outcome Inventory-14 throat discomfort score completion rate and scores at time points (continued)

	Randomised arm		Total (N = 453)
	Tonsillectomy (N = 233)	Conservative management (N = 220)	
24 months (visit)			
n (%)	99 (42)	100 (45)	199 (44)
Median (IQR)	1 (0–4)	4.5 (1–9.5)	2 (0–6)
Mean (SD)	2.3 (3.3)	5.6 (5.0)	3.9 (4.5)
95% CI about mean	1.7–3.0	4.6–6.6	3.3–4.6
Min, max	0, 20	0, 19	0, 20

Max, maximum; min, minimum.

Note

% is the proportion completed out of eligible randomised.

TABLE 47 Mixed-model repeated measures for TOI-14

	Coefficient	SE	Test statistic	p-value	95% CI of coefficient
Time point (reference value: 6 months)					
12 months	-4.071	4.511	-0.90	0.367	-12.913 to 4.772
18 months	-6.687	4.877	-1.37	0.170	-16.246 to 2.872
24 months	-3.724	5.329	-0.70	0.485	-14.169 to 6.720
Baseline TOI-14	0.184	0.097	1.90	0.057	-0.006 to 0.374
Interaction time and baseline TOI-14					
12 months × baseline TOI-14	-0.003	0.099	-0.03	0.975	-0.198 to 0.192
18 months × baseline TOI-14	-0.037	0.107	-0.34	0.732	-0.248 to 0.174
24 months × baseline TOI-14	-0.165	0.117	-1.41	0.157	-0.395 to 0.064
Arm (reference value: conservative management)					
Immediate tonsillectomy	-13.167	2.167	-6.08	<0.001	-17.414 to -8.920
Interaction time and arm					
12 months × immediate tonsillectomy	-4.162	2.226	-1.87	0.062	-8.534 to 0.201
18 months × immediate tonsillectomy	1.164	2.470	0.47	0.637	-3.676 to 6.005
24 months × immediate tonsillectomy	2.475	2.640	0.94	0.348	-2.699 to 7.650
Constant	18.225	4.433	4.11	<0.001	9.536 to 26.913
Site: Var(_cons) Estimate	2.979	3.249			0.349 to 25.433

Sites: n = 20. Average observations per site: n = 39.1. Study ID: n = 308. Average observations per ID: n = 2.5. Log likelihood = -3080.210, LR test vs. negative binomial model: $\bar{\chi}^2 = 107.16$. Prob > = $\bar{\chi}^2 = 0.000$. × represents an interaction between variables.

Table 48 shows that scores were similar at baseline as expected, but that there were marked difference at follow-up visits (after surgery) in overall TOI-14 scores (non-overlapping CIs).

Table 49 shows that scores were similar at baseline as expected, but that there were marked difference at follow-up visits (after surgery) in overall TOI-14 scores (non-overlapping CIs).

Table 50 shows that scores were similar at baseline as expected, but that there were marked differences at follow-up visits (after surgery) in TOI-14 general health scores (non-overlapping CIs).

In **Table 51**, questionnaires missing one or two items were completed by imputation of item score calculated as the average of the completed items. The analysis and charts in rest of this section include the recovered questionnaires with imputed missing items.

TABLE 48 Tonsillectomy Outcome Inventory-14 general health score completion rate and scores at time points

	Randomised arm		Total (N = 453)
	Tonsillectomy (N = 233)	Conservative management (N = 220)	
Baseline			
n (%)	231 (99)	217 (99)	448 (99)
Median (IQR)	8 (6–9)	7 (6–9)	8 (6–9)
Mean (SD)	7.2 (2.3)	7.1 (2.3)	7.1 (2.3)
95% CI about mean	6.9 to 7.5	6.8 to 7.4	6.9 to 7.4
Min, max	0, 10	0, 10	0, 10
6 months (post)			
n (%)	105 (45)	85 (39)	190 (42)
Median (IQR)	1 (0–4)	4 (2–7)	2 (0–5)
Mean (SD)	2.2 (2.7)	4.2 (2.9)	3.1 (3.0)
95% CI about mean	1.7 to 2.7	3.6 to 4.8	2.7 to 3.5
Min, max	0, 10	0, 10	0, 10
12 months (visit)			
n (%)	122 (55)	117 (50)	239 (53)
Median (IQR)	0 (0–1)	3 (1–6)	1 (0–3)
Mean (SD)	0.8 (1.3)	3.5 (2.9)	2.1 (2.6)
95% CI about mean	0.6 to 1.1	2.9 to 4.0	1.8 to 2.5
Min, max	0, 7	0, 10	0, 10
18 months (post)			
n (%)	83 (36)	71 (32)	154 (34)
Median (IQR)	0 (0–1)	2 (0–4)	1 (0–3)
Mean (SD)	1.0 (1.7)	2.7 (2.8)	1.8 (2.4)
95% CI about mean	0.6 to 1.3	2.0 to 3.4	1.4 to 2.2
Min, max	0, 8	0, 10	0, 10

continued

TABLE 48 Tonsillectomy Outcome Inventory-14 general health score completion rate and scores at time points (continued)

	Randomised arm		Total (N = 453)
	Tonsillectomy (N = 233)	Conservative management (N = 220)	
24 months (visit)			
n (%)	99 (42)	100 (45)	199 (44)
Median (IQR)	0 (0–1)	1 (0–4)	0 (0–3)
Mean (SD)	0.9 (1.8)	2.6 (3.1)	1.8 (2.7)
95% CI about mean	0.5 to 1.3	2.0 to 3.2	1.4 to 2.1
Min, max	0, 10	0, 10	0, 10
Max, maximum; min, minimum.			
Note			
% is the proportion completed out of eligible randomised.			

TABLE 49 Tonsillectomy Outcome Inventory-14 resource impact score completion rate and scores at time points

	Randomised arm		Total (N = 453)
	Tonsillectomy (N = 233)	Conservative management (N = 220)	
Baseline			
n (%)	231 (99)	217 (99)	448 (99)
Median (IQR)	11 (8–14)	11 (9–14)	11 (8–14)
Mean (SD)	11.2 (4.0)	11.1 (4.3)	11.1 (4.1)
95% CI about mean	10.7 to 11.7	10.5 to 11.6	10.8 to 11.5
Min, max	1, 20	0, 20	0, 20
6 months (post)			
n (%)	105 (45)	85 (39)	190 (42)
Median (IQR)	0 (0–4)	4 (1–9)	2 (0–7)
Mean (SD)	2.5 (4.1)	5.2 (4.7)	3.7 (4.5)
95% CI about mean	1.8 to 3.3	4.2 to 6.2	3.1 to 4.4
Min, max	0, 15	0, 17	0, 17
12 months (visit)			
n (%)	122 (55)	117 (50)	239 (53)
Median (IQR)	0 (0–0)	4 (0–8)	0 (0–4)
Mean (SD)	0.6 (1.3)	5.0 (5.2)	2.8 (4.3)
95% CI about mean	0.4 to 0.8	4.1 to 6.0	2.2 to 3.3
Min, max	0, 6	0, 20	0, 20

TABLE 49 Tonsillectomy Outcome Inventory-14 resource impact score completion rate and scores at time points (continued)

	Randomised arm		Total (N = 453)
	Tonsillectomy (N = 233)	Conservative management (N = 220)	
18 months (post)			
n (%)	83 (36)	71 (32)	154 (34)
Median (IQR)	0 (0-1)	2 (0-6)	0 (0-3)
Mean (SD)	1.2 (2.5)	3.5 (4.3)	2.2 (3.6)
95% CI about mean	0.6 to 1.7	2.5 to 4.5	1.7 to 2.8
Min, max	0, 12	0, 17	0, 17
24 months (visit)			
n (%)	99 (42)	100 (45)	199 (44)
Median (IQR)	0 (0-0)	1 (0-6)	0 (0-2)
Mean (SD)	0.7 (2.0)	3.5 (4.8)	2.1 (4.0)
95% CI about mean	0.3 to 1.1	2.5 to 4.5	1.6 to 2.7
Min, max	0, 15	0, 19	0, 19
Max, maximum; min, minimum.			
Note			
% is the proportion completed out of eligible randomised.			

TABLE 50 Tonsillectomy Outcome Inventory-14 social/psychological score completion rate and scores at time points

	Randomised arm		Total (N = 453)
	Tonsillectomy (N = 233)	Conservative management (N = 220)	
Baseline			
n (%)	231 (99)	217 (99)	448 (99)
Median (IQR)	21 (9-15)	12 (9-15)	12 (9-15)
Mean (SD)	11.9 (4.3)	11.5 (4.4)	11.7 (4.4)
95% CI about mean	11.3 to 12.4	10.9 to 12.1	11.3 to 12.1
Min, max	0, 20	0, 20	0, 20
6 months (post)			
n (%)	105 (45)	85 (39)	190 (42)
Median (IQR)	0 (0-4)	5 (1-10)	2 (0-8)
Mean (SD)	2.8 (4.6)	6.1 (5.4)	4.3 (5.2)
95% CI about mean	1.9 to 3.7	4.9 to 7.2	3.5 to 5.0
Min, max	0, 18	0, 19	0, 19
continued			

TABLE 50 Tonsillectomy Outcome Inventory-14 social/psychological score completion rate and scores at time points (continued)

	Randomised arm		Total (N = 453)
	Tonsillectomy (N = 233)	Conservative management (N = 220)	
12 months (visit)			
n (%)	122 (55)	117 (50)	239 (53)
Median (IQR)	0 (0–0)	3 (0–9)	0 (0–4)
Mean (SD)	0.4 (1.3)	5.3 (5.9)	2.8 (4.9)
95% CI about mean	0.2 to 0.6	4.2 to 6.4	2.2 to 3.4
Min, max	0, 8	0, 20	0, 20
18 months (post)			
n (%)	83 (36)	71 (32)	154 (34)
Median (IQR)	0 (0–0)	1 (0–6)	0 (0, 3)
Mean (SD)	0.7 (2.1)	3.6 (5.0)	2.1 (4.0)
95% CI about mean	0.3 to 1.2	2.4 to 4.8	1.4 to 2.7
Min, max	0, 15	0, 20	0, 20
24 months (visit)			
n (%)	99 (42)	100 (45)	199 (44)
Median (IQR)	0 (0–0)	0 (0–8)	0 (0–3)
Mean (SD)	0.7 (2.5)	3.7 (5.1)	2.2 (4.3)
95% CI about mean	0.2 to 1.2	2.7 to 4.7	1.6 to 2.8
Min, max	0, 20	0, 20	0, 20

Max, maximum; min, minimum.

Note

% is the proportion completed out of eligible randomised.

TABLE 51 Details of partially missing TOI-14 questionnaires at trial visits and postal returns

Time point	Missing (n)	Details
Baseline	2	Missing one item: 1 participant (Q13), 1 participant (Q9)
6 months	5	Missing one item: 1 participant (Q9), 1 participant (Q2), 1 participant (Q11), 1 participant (Q5) Missing two items: 1 participant (Q1 and Q2)
12 months	2	Missing one item: 1 participant (Q5) Missing two items: 1 participant (Q5 and Q9)
18 months	4	Missing one item: 1 participant (Q13), 1 participant (Q8), 1 participant (Q13) Missing six items: 1 participant (Q2, Q6, Q7, Q8, Q11 and Q12)
24 months	0	None partially missing

Q, question.

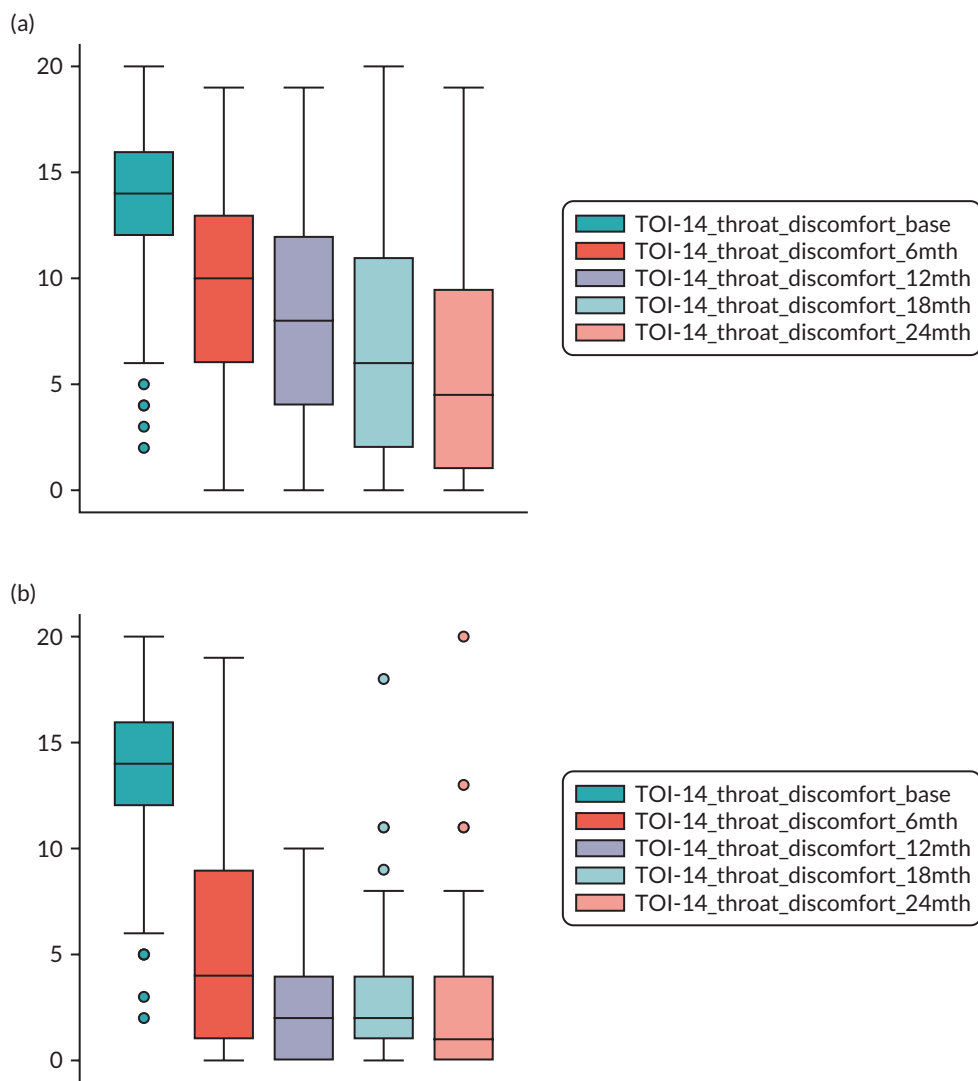


FIGURE 24 The TOI-14 subscale throat discomfort. (a) Conservative management; (b) immediate tonsillectomy.

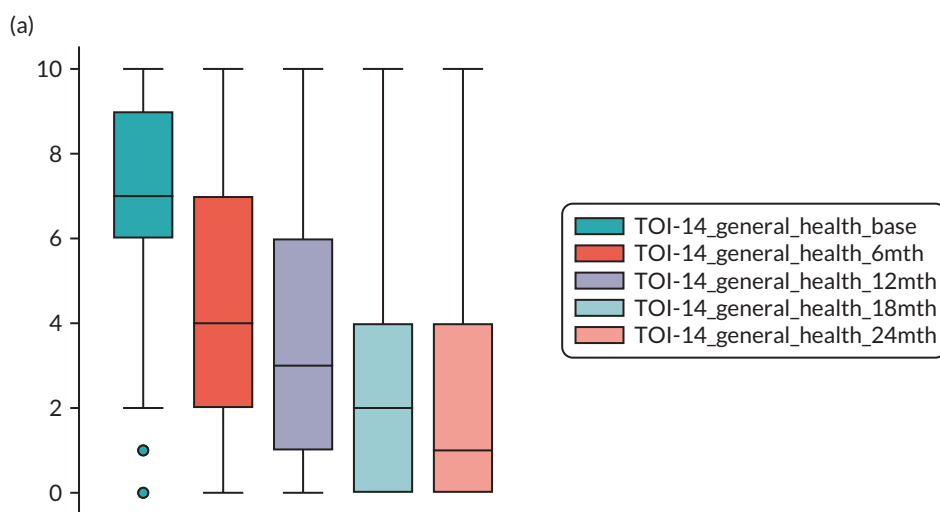


FIGURE 25 The TOI-14 subscale general health. (a) Conservative management; (b) immediate tonsillectomy. (continued)

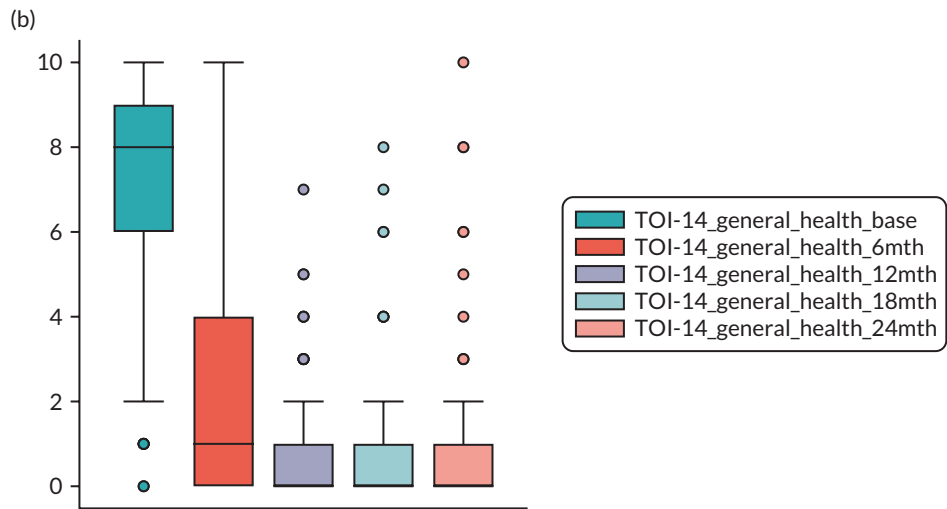


FIGURE 25 The TOI-14 subscale general health. (a) Conservative management; (b) immediate tonsillectomy.

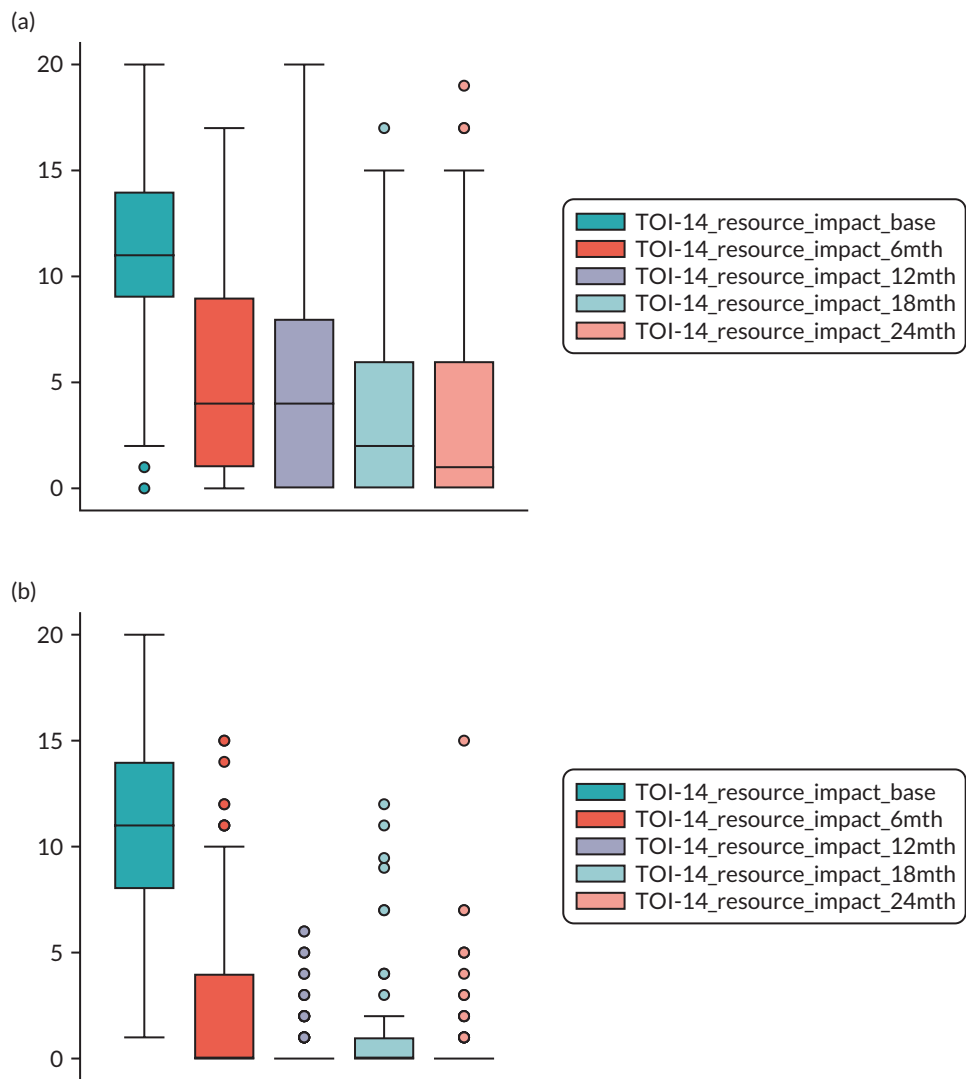


FIGURE 26 The TOI-14 subscale resource impact. (a) Conservative management; (b) immediate tonsillectomy.

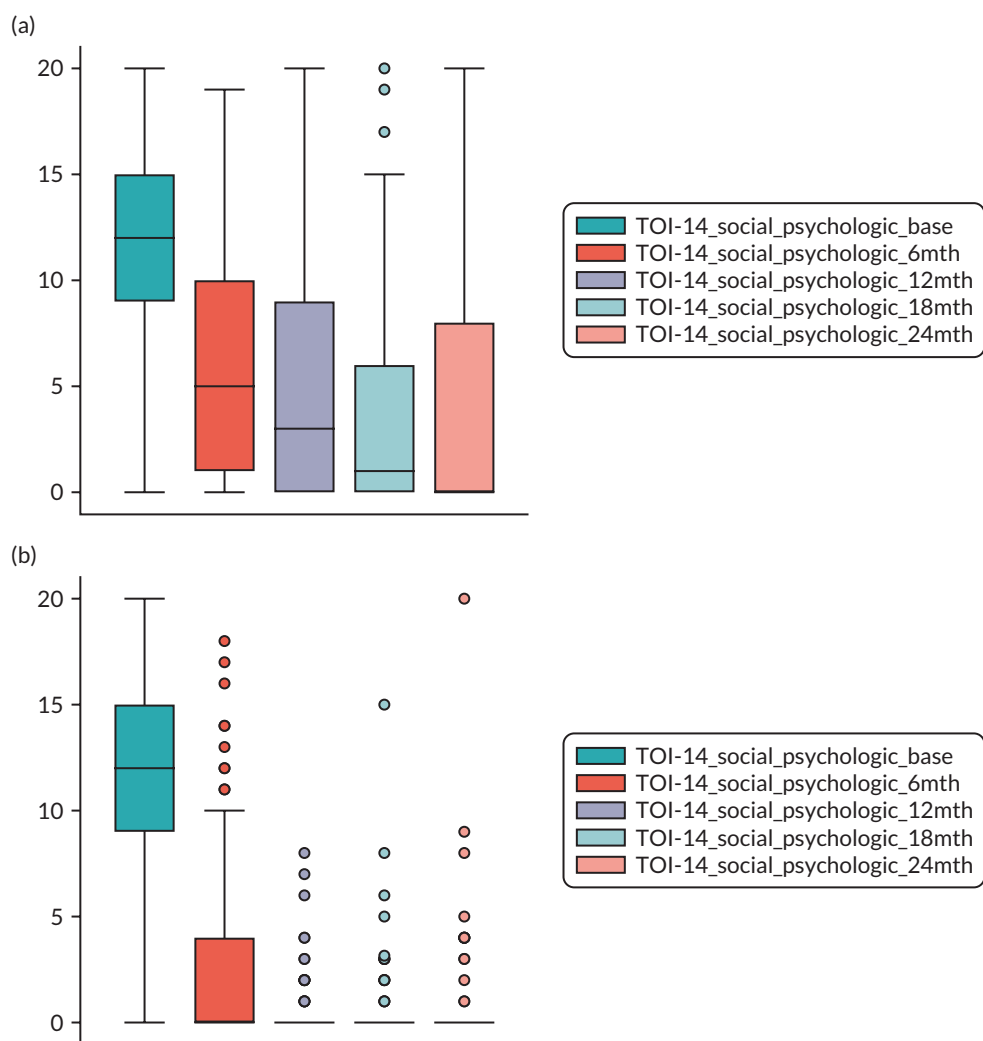


FIGURE 27 The TOI-14 subscale social psychological. (a) Conservative management; (b) immediate tonsillectomy.

Other total Tonsillectomy Outcome Inventory-14 box plots for different populations

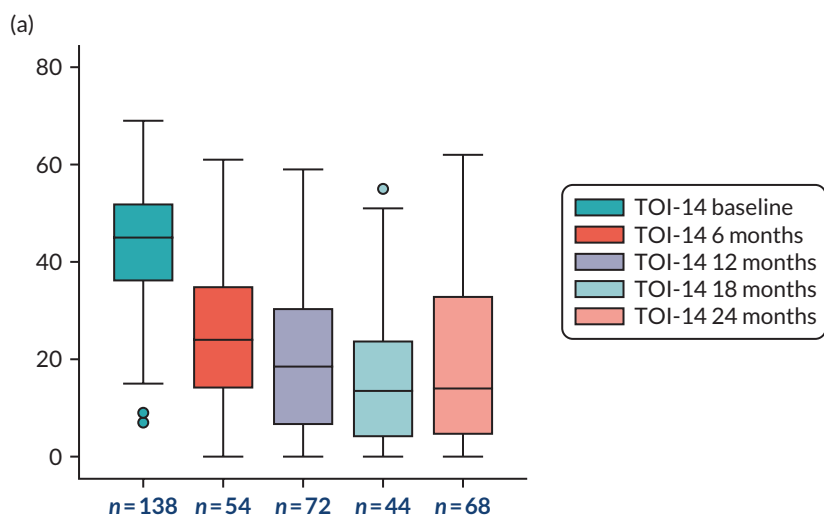


FIGURE 28 The TOI-14 at five time points in the per-protocol population: baseline and 6, 12, 18 and 24 months. (a) Conservative management ($n = 133$); and (b) immediate tonsillectomy ($n = 91$). Graphs by protocol compliance. (continued)

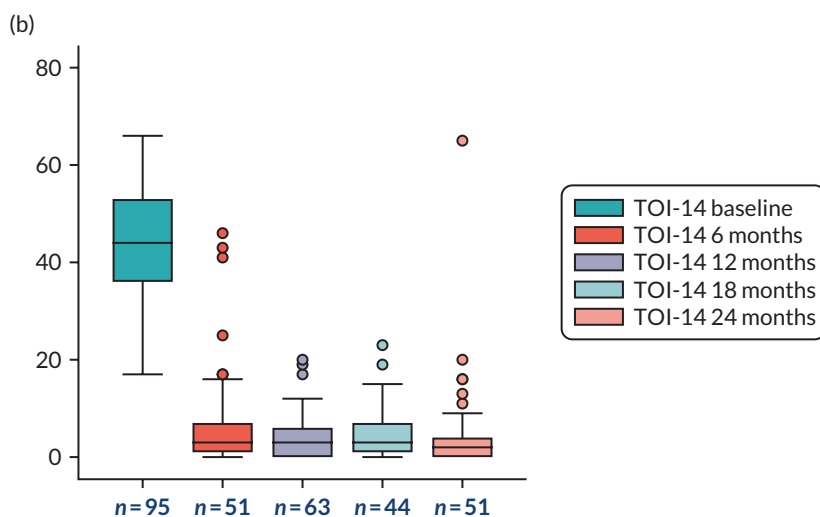


FIGURE 28 The TOI-14 at five time points in the per-protocol population: baseline and 6, 12, 18 and 24 months. (a) Conservative management (n = 133); and (b) immediate tonsillectomy (n = 91). Graphs by protocol compliance.

For the per-protocol population (*Figure 28*) although improvements were seen in both randomised arms, clearly larger improvements in TOI-14 scores throughout follow-up can be seen for the group receiving tonsillectomy within 8 weeks of randomisation compared with those receiving conservative management as randomised.

For the per-treatment population with four categories covering tonsillectomy or not in both arms is shown in *Figure 29*. Participants receiving tonsillectomy as randomised show clear reduction in TOI-14 scores over time. This could be due to them receiving tonsillectomy quicker than those originally randomised to conservative management who waited longer to receive tonsillectomy after having crossed over. Those not having tonsillectomy showed improvements from baseline which could account for their decision to not receive the tonsillectomy.

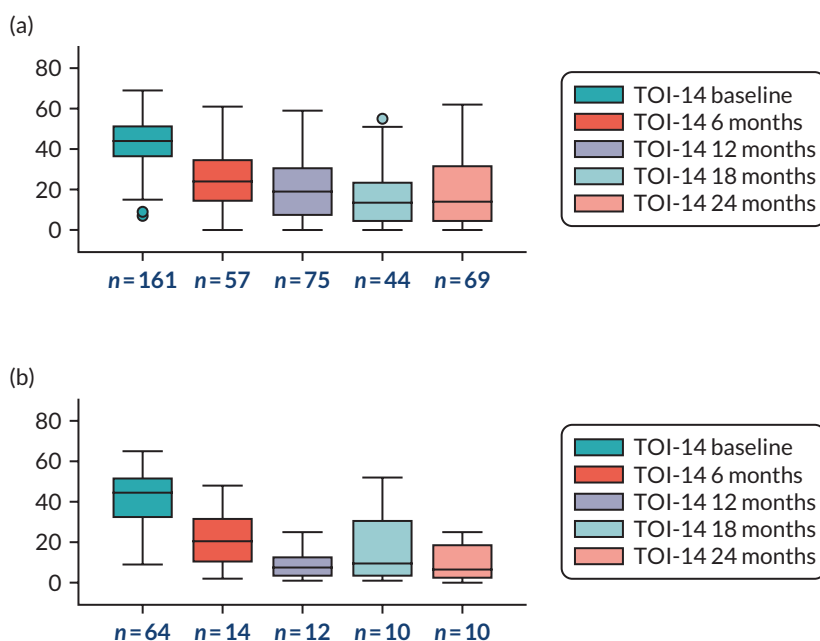


FIGURE 29 TOI-14 at five time points in the per-treatment population: baseline and 6, 12, 18 and 24 months. (a) No tonsillectomy as randomised; (b) no tonsillectomy randomised to tonsillectomy; (c) tonsillectomy as randomised; and (d) tonsillectomy crossed. Graphs by tonsillectomy received and arm. (*continued*)

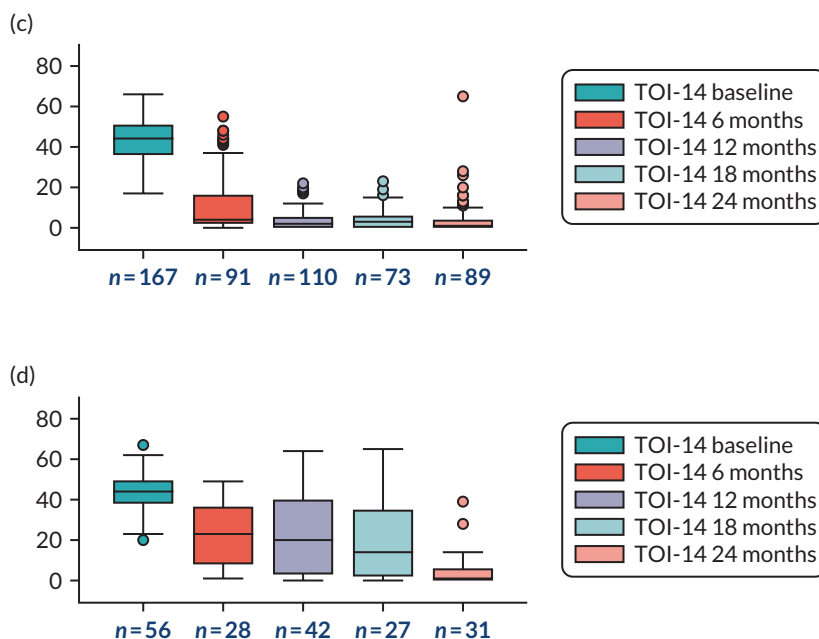


FIGURE 29 TOI-14 at five time points in the per-treatment population: baseline and 6, 12, 18 and 24 months. (a) No tonsillectomy as randomised; (b) no tonsillectomy randomised to tonsillectomy; (c) tonsillectomy as randomised; and (d) tonsillectomy crossed. Graphs by tonsillectomy received and arm.

The mixed-model repeated measures specify no patient-level random effects but instead model the correlation within the repeated measures over time by specifying that the residual errors are correlated. To reduce the chances of model mis-specification, the residual errors are assumed to be from a multivariate normal distribution with an unstructured covariance matrix. This imposes no restriction on the form of the correlation matrix of the repeated measures.

For the fixed effects, the effects of time are specified as categorical variables, randomised treatment group and the interaction between them. This implies a saturated model for the mean, so there is a separate mean parameter for each time point in each treatment group. The baseline covariates to be adjusted for are baseline severity and site. These are adjusted for as simple main effects. The results table for TOI-14 is shown in [Table 47](#).

Appendix 15 Addressing missing data

In addition to assessing missing primary outcome data (STAR responses), we also assessed the exposure variable (i.e. a measure of missing data) in relation to the four per-treatment categories in terms of how this relates to the baseline severity (mild, moderate and severe). We also carried out a similar check regarding the secondary outcome measure of TOI-14 in a similar manner.

TABLE 52 Missing data examination: summarise exposure variable by baseline severity

Severity	No surgery as randomised	No surgery randomised to surgery	Surgery as randomised	Surgery crossed
Mild				
<i>n</i>	33	14	34	8
Median (IQR)	0.95 (0.58–0.99)	0.66 (0.13–1.00)	0.95 (0.89–0.99)	0.64 (0.28–0.90)
Mean (SD)	0.76 (0.34)	0.59 (0.42)	0.83 (0.28)	0.59 (0.37)
Min, max	0.03, 1.00	0.04, 1.0	0.04, 1.00	0.05, 1.00
Moderate				
<i>n</i>	56	19	73	32
Median (IQR)	0.69 (0.10–0.99)	0.11 (0.08–0.65)	0.98 (0.89–1.00)	0.93 (0.30–0.99)
Mean (SD)	0.58 (0.42)	0.30 (0.33)	0.85 (0.27)	0.70 (0.37)
Min, max	0.01, 1.00	0.06, 0.95	0.04, 1.00	0.08, 1.00
Severe				
<i>n</i>	59	24	60	17
Median (IQR)	0.89 (0.49–0.98)	0.61 (0.17–0.81)	0.96 (0.73–0.99)	0.82 (0.49–0.97)
Mean (SD)	0.73 (0.30)	0.53 (0.36)	0.81 (0.28)	0.68 (0.36)
Min, max	0.05, 1.00	0.04, 1.00	0.04, 1.00	0.05, 1.00
Total				
<i>n</i>	148	57	167	57
Median (IQR)	0.89 (0.35–0.99)	0.30 (0.10–0.83)	0.97 (0.80–0.99)	0.90 (0.32–0.99)
Mean (SD)	0.68 (0.36)	0.47 (0.38)	0.83 (0.27)	0.68 (0.36)
Min, max	0.01, 1.00	0.04, 1.00	0.04, 1.00	0.05, 1.00

TABLE 53 Summary statistics for overall return rate split by baseline severity to see whether or not patterns are different for return of the TOI-14 (secondary outcome)

Severity	Randomised arm, n/N (%)		Surgery status (randomised arm and surgery received), n/N (%)				Overall, n/N (%)
	Conservative management	Tonsillectomy	No surgery as randomised	No surgery randomised to surgery	Surgery as randomised	Surgery crossed	
TOI-14 returns by baseline severity							
Mild (N = 95)							
0 (no returns)	0/44 (0)	0/51 (0)	0/36 (0)	0/15 (0)	0/36 (0)	0/8 (0)	0/95 (0)
<80%	29/44 (66)	32/51 (63)	23/36 (64)	12/15 (80)	20/36 (56)	6/8 (75)	61/95 (64)
≥80%	15/44 (34)	19/51 (37)	13/36 (36)	3/15 (20)	16/36 (44)	2/8 (25)	34/95 (36)
Moderate (N = 190)							
0 (no returns)	3/96 (3)	2/94 (2)	2/63 (3)	2/20 (10)	0/74 (0)	1/33 (3)	5/190 (3)
<80%	57/96 (59)	58/94 (62)	43/63 (68)	17/20 (85)	41/74 (55)	14/33 (42)	115/190 (61)
≥80%	36/96 (38)	34/94 (36)	18/63 (29)	1/20 (5)	33/74 (45)	18/33 (55)	70/190 (37)
Severe (N = 168)							
0 (no returns)	0/80 (0)	0/88 (0)	0/62 (0)	0/26 (0)	0/62 (0)	0/18 (0)	0/168 (0)
<80%	58/80 (73)	60/88 (68)	46/62 (74)	24/26 (92)	36/62 (58)	12/18 (67)	118/168 (70)
≥80%	22/80 (28)	28/88 (32)	16/62 (26)	2/26 (8)	26/62 (42)	6/18 (33)	50/168 (30)
Total	220	233	161	61	172	59	453 (100)
TOI-14 24-month returns by baseline severity							
Mild (N = 95)							
Not returned	24/44 (55)	27/51 (53)	18/36 (50)	12/15 (80)	15/36 (42)	6/8 (75)	51/95 (54)
Returned	20/44 (45)	24/51 (47)	18/36 (50)	3/15 (20)	21/36 (58)	2/8 (25)	44/95 (46)
Moderate (N = 190)							
Not returned	48/96 (50)	53/94 (56)	37/63 (59)	18/20 (90)	35/74 (47)	11/33 (33)	101/190 (53)
Returned	48/96 (50)	41/94 (44)	26/63 (41)	2/20 (10)	39/74 (53)	22/33 (67)	89/190 (47)
Severe (N = 168)							
Not returned	48/80 (60)	54/88 (61)	38/62 (61)	22/26 (85)	32/62 (52)	10/18 (56)	102/168 (61)
Returned	32/80 (40)	34/88 (39)	24/62 (39)	4/26 (15)	30/62 (48)	8/18 (44)	66/168 (39)
Total	220	233	161	61	172	59	453

TABLE 54 Crosstab by those who complete at least 80% STAR data by 80% TOI-14

	Completion, n (%)			Total
	No STAR	<80%	≥80%	
No TOI-14	1 (4)	3 (2)	1 (<1)	5 (1)
<80%	22 (92)	147 (89)	125 (48)	294 (65)
≥80%	1 (4)	16 (10)	137 (52)	154 (34)
Total	24 (100)	166 (100)	263 (100)	453 (100)

Appendix 16 Secondary analyses: quality of life (Short Form questionnaire-12 items)

The mixed-model repeated measures specify no patient-level random effects but instead model the correlation within the repeated measures over time by specifying that the residual errors are correlated. To reduce the chances of model mis-specification, the residual errors are assumed to be from a multivariate normal distribution with an unstructured covariance matrix. This imposes no restriction on the form of the correlation matrix of the repeated measures.

For the fixed effects, the effects of time are specified as categorical variables, randomised treatment group and the interaction between them. This implies a saturated model for the mean, so there is a separate mean parameter for each time point in each randomised arm. The baseline covariates to be adjusted for are baseline severity and site. These are adjusted for as simple main effects. The results tables for SF-12 MCS and PCS are shown in [Tables 55](#) and [56](#).

TABLE 55 Mixed-model repeated measures for SF-12 PCS

SF-12 PCS	Coefficient	SE	Test statistic	p-value	95% CI of coefficient
Time point (reference value: 6 months)					
12 months	5.255	3.327	1.58	0.114	-1.267 to 11.776
18 months	0.269	3.799	0.07	0.944	-7.178 to 7.715
24 months	11.595	3.705	3.13	0.002	4.333 to 18.856
Baseline PCS-12	0.366	0.059	6.15	<0.001	0.250 to 0.483
Interaction time and baseline PCS					
12 months × baseline PCS	-0.091	0.063	-1.44	0.151	-0.215 to 0.033
18 months × baseline PCS	0.003	0.072	0.04	0.969	-0.138 to 0.144
24 months × baseline PCS	-0.174	0.071	-2.46	0.014	-0.313 to -0.035
Arm (reference value: conservative management)					
Immediate tonsillectomy	3.781	0.860	4.40	<0.001	2.095 to 5.467
Interaction time and arm					
12 months × immediate tonsillectomy	0.461	0.925	0.50	0.619	-1.353 to 2.275
18 months × immediate tonsillectomy	0.304	1.053	0.29	0.773	-1.760 to 2.367
24 months × immediate tonsillectomy	-1.875	1.022	-1.83	0.067	-3.879 to 0.129
Baseline severity (reference: mild)					
Moderate	-0.113	0.680	-0.17	0.868	-1.446 to 1.220
Severe	-0.003	0.722	-0.00	0.997	-1.419 to 1.413
Constant	33.477	3.213	10.42	<0.001	27.178 to 39.775
Site: Var(_cons) Estimate	0.551	0.506			0.091 to 3.327

Site: $n = 19$ average observations per site = 41.6. Study ID: $n = 307$ average observations per ID = 2.6. Log likelihood = -2461.928, LR test vs. negative binomial model: $\bar{\chi}^2 = 98.02$. $Prob > = \bar{\chi}^2 = 0.000$. × represents an interaction between variables.

TABLE 56 Mixed-model repeated measures for SF-12 MCS

SF-12 MCS	Coefficient	SE	Test statistic	p-value	95% CI of coefficient
Time point (reference value: 6 months)					
12 months	4.020	3.235	1.24	0.214	-2.320 to 10.360
18 months	0.211	3.399	0.06	0.951	-6.451 to 6.872
24 months	5.994	3.426	1.75	0.080	-0.721 to 12.709
Baseline MCS	0.414	0.061	6.79	<0.001	0.295 to 0.534
Interaction time and baseline MCS					
12 months × baseline MCS	-0.055	0.068	-0.81	0.421	-0.189 to 0.079
18 months × baseline MCS	-0.016	0.071	-0.22	0.825	-0.154 to 0.123
24 months × baseline MCS	-0.101	0.072	-1.40	0.162	-0.241 to 0.040
Arm (reference value: conservative management)					
Immediate tonsillectomy	2.767	1.259	2.20	0.028	0.301 to 5.234
Interaction time and arm					
12 months × immediate tonsillectomy	0.951	1.368	0.70	0.487	-1.729 to 3.632
18 months × immediate tonsillectomy	1.185	1.374	0.86	0.389	-1.509 to 3.879
24 months × immediate tonsillectomy	0.389	1.483	0.26	0.793	-2.517 to 3.295
Baseline severity (reference: mild)					
Moderate	0.385	1.129	0.34	0.733	-1.828 to 2.598
Severe	0.699	1.189	0.59	0.557	-1.632 to 3.030
Constant	27.282	3.098	8.81	<0.001	21.210 to 33.355
Site: Var(_cons)	1.68e-21	-			-

Site: $n = 19$ average observations per site = 41.6. Study ID: $n = 307$ average observations per ID = 2.6. Log likelihood = -2750.859, LR test vs. negative binomial model: $\bar{\chi}^2 = 164.47$. Prob > = $\bar{\chi}^2 = <0.001$. × represents an interaction between variables.

Appendix 17 Economic evaluation

TABLE 57 Summary of post-surgical AE treatments, as recorded in the eCRF

	Conservative management (N = 60 who received tonsillectomies)	Tonsillectomy (N = 172 received tonsillectomies)
	Mean (SD)	Mean (SD)
Number of AEs reported	0.45 (0.72)	0.65 (0.80)
No action required	0.10 (0.30)	0.10 (0.31)
Total outpatient drug medications	0.23 (0.43)	0.28 (0.45)
Prescribed outpatient drug medications	0.20 (0.40)	0.27 (0.44)
Required hospitalisation	0.15 (0.36)	0.17 (0.38)
Number of inpatient nights	0.22 (0.58)	0.27 (0.76)
Admitted to ICU/HDU	0.00 (0.00)	0.00 (0.00)
Number of blood tests	0.07 (0.25)	0.10 (0.30)
Number of blood transfusions	0.00 (0.00)	0.01 (0.08)
Number of drips	0.05 (0.22)	0.09 (0.29)
Number of consultations	0.02 (0.13)	0.08 (0.27)
Number of observations	0.07 (0.25)	0.12 (0.33)
Number of drug therapies	0.12 (0.32)	0.15 (0.36)
Number of surgeries	0.02 (0.13)	0.02 (0.13)
Inpatient nights post surgery	0.00 (0.00)	0.03 (0.32)

TABLE 58 Combined STAR/STARLET summaries by severity

	Mild		Moderate		Severe	
	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n
Number of sore throat days ^a	2.55 (1.71)	1251	3.80 (2.01)	1463	5.15 (1.99)	849
	n = 508 ^b		n = 518 ^b		n = 360 ^b	
Utility score (SF-6D)	0.75 (0.23)	439	0.70 (0.13)	453	0.59 (0.12)	311
Time off work (days)	0.07 (0.43)	508	0.36 (1.04)	517	1.77 (2.31)	357
Time off usual activities (days)	0.28 (0.77)	508	0.91 (1.45)	517	2.81 (2.44)	357
Prescribed medications	0.48 (0.53)	508	0.65 (0.64)	518	1.01 (0.72)	360
Over-the-counter medications	0.86 (0.63)	508	1.15 (0.65)	518	1.13 (0.67)	360

^a Of the 5309 weekly text messages reporting a sore throat episode, 3563 (66%) had information on severity available from the text messages or STAR/STARLET.

^b n = number of STAR/STARLET questionnaires available (n = 1395), of which nine did not have information on severity of the sore throat episode.

TABLE 59 Summary of healthcare resource use, as reported in the health utilisation questionnaire

	Conservative management (N = 220)		Tonsillectomy (N = 233)	
	Mean (SD)	n	Mean (SD)	n
Baseline				
GP consultation: practice	0.98 (0.15)	215	0.97 (0.16)	231
Number of GP consultations: practice	3.74 (2.41)	211	3.80 (2.69)	227
GP consultation: home	0.03 (0.17)	215	0.03 (0.16)	231
Number of GP consultations: home	0.03 (0.20)	215	0.03 (0.24)	230
GP consultation: telephone	0.42 (0.49)	215	0.46 (0.50)	231
Number of GP consultations: telephone	0.98 (1.96)	209	0.94 (1.43)	229
Nurse consultation: practice	0.35 (0.48)	215	0.39 (0.49)	231
Number of nurse consultations: practice	0.64 (1.31)	211	0.82 (1.48)	227
Nurse consultation: home	<0.01 (0.07)	215	0.13 (0.11)	231
Number of nurse consultations: home	0.02 (0.27)	215	0.03 (0.31)	231
Nurse consultation: telephone	0.31 (0.46)	215	0.29 (0.45)	231
Number of nurse visits: telephone	0.65 (1.30)	213	0.55 (1.15)	228
NHS 111 consultation	0.27 (0.45)	215	0.36 (0.48)	231
Number of NHS 111 consultations	0.50 (1.02)	212	0.63 (1.27)	229
Out-of-hours clinic consultation	0.34 (0.48)	215	0.38 (0.49)	231
Number of out-of-hours clinic consultations	0.61 (1.24)	213	0.61 (1.02)	228
Walk-in clinic consultation	0.23 (0.42)	215	0.25 (0.43)	231
Number of walk-in clinic consultations	0.45 (1.28)	213	0.43 (1.10)	227
Pharmacist consultation	0.57 (0.50)	215	0.63 (0.48)	231
Number of pharmacist consultations	1.76 (2.40)	210	2.06 (2.76)	226
A&E visit	0.12 (0.32)	215	0.21 (0.41)	231
Number of A&E visits	0.19 (0.64)	214	0.25 (0.53)	231
Outpatient visit	0.47 (0.50)	214	0.47 (0.50)	230
Number of outpatient visits	0.63 (0.91)	211	0.55 (0.71)	225
Hospital admission	0.12 (0.33)	215	0.16 (0.36)	230
Number of days admitted	0.14 (0.42)	214	0.17 (0.41)	229
Emergency ambulance	0.03 (0.18)	215	0.04 (0.20)	231
Number of emergency ambulance trips	0.03 (0.18)	215	0.05 (0.26)	231
6 months				
GP consultation: practice	0.53 (0.50)	85	0.40 (0.49)	110
Number of GP consultations: practice	1.40 (1.98)	81	1.10 (1.98)	110
GP consultation: home	0.00 (0.00)	85	0.00 (0.00)	110
Number of GP consultations: home	0.00 (0.00)	85	0.00 (0.00)	110
GP consultation: telephone	0.19 (0.39)	85	0.18 (0.39)	110
Number of GP consultations: telephone	0.31 (0.77)	85	0.36 (1.05)	110
Nurse consultation: practice	0.08 (0.28)	85	0.12 (0.32)	110

TABLE 59 Summary of healthcare resource use, as reported in the health utilisation questionnaire (continued)

	Conservative management (N = 220)		Tonsillectomy (N = 233)	
	Mean (SD)	n	Mean (SD)	n
Number of nurse consultations: practice	0.09 (0.33)	85	0.16 (0.58)	109
Nurse consultation: home	0.00 (0.00)	85	0.00 (0.00)	110
Number of nurse consultations: home	0.00 (0.00)	85	0.00 (0.00)	110
Nurse consultation: telephone	0.13 (0.34)	85	0.13 (0.33)	110
Number of nurse visits: telephone	0.24 (0.72)	85	0.24 (0.83)	110
NHS 111 consultation	0.04 (0.19)	84	0.13 (0.33)	110
Number of NHS 111 consultations	0.10 (0.67)	84	0.14 (0.37)	110
Out-of-hours clinic consultation	0.07 (0.26)	84	0.15 (0.35)	110
Number of out-of-hours clinic consultations	0.13 (0.69)	84	0.16 (0.41)	109
Walk-in clinic consultation	0.09 (0.29)	85	0.06 (0.25)	110
Number of walk-in clinic consultations	0.12 (0.42)	85	0.10 (0.43)	109
Pharmacist consultation	0.40 (0.49)	85	0.17 (0.38)	110
Number of pharmacist consultations	1.10 (1.61)	84	0.32 (0.87)	109
A&E visit	0.06 (0.24)	85	0.08 (0.28)	110
Number of A&E visits	0.09 (0.43)	85	0.10 (0.38)	109
Outpatient visit	0.14 (0.35)	85	0.22 (0.41)	110
Number of outpatient visits	0.19 (0.52)	85	0.34 (0.76)	110
Hospital admission	0.08 (0.28)	85	0.22 (0.41)	110
Number of days admitted	0.09 (0.33)	85	0.25 (0.51)	110
Emergency ambulance	0.01 (0.11)	85	0.01 (0.10)	110
Number of emergency ambulance trips	0.05 (0.43)	85	0.02 (0.19)	110
12 months				
GP consultation: practice	0.50 (0.50)	117	0.16 (0.37)	122
Number of GP consultations: practice	1.32 (1.81)	116	0.28 (0.72)	122
GP consultation: home	0.01 (0.09)	117	0.00 (0.00)	122
Number of GP consultations: home	0.02 (0.18)	117	0.00 (0.00)	122
GP consultation: telephone	0.21 (0.41)	117	0.03 (0.18)	122
Number of GP consultations: telephone	0.33 (0.77)	116	0.02 (0.16)	121
Nurse consultation: practice	0.15 (0.36)	117	0.02 (0.13)	122
Number of nurse consultations: practice	0.33 (1.00)	117	0.02 (0.20)	122
Nurse consultation: home	0.00 (0.00)	117	0.00 (0.00)	122
Number of nurse consultations: home	0.00 (0.00)	117	0.00 (0.00)	122
Nurse consultation: telephone	0.15 (0.35)	117	0.01 (0.09)	122
Number of nurse visits: telephone	0.24 (0.67)	116	0.01 (0.09)	122
NHS 111 consultation	0.06 (0.24)	117	0.01 (0.09)	122
Number of NHS 111 consultations	0.12 (0.63)	117	0.01 (0.09)	122

continued

TABLE 59 Summary of healthcare resource use, as reported in the health utilisation questionnaire (*continued*)

	Conservative management (N = 220)		Tonsillectomy (N = 233)	
	Mean (SD)	n	Mean (SD)	n
Out-of-hours clinic consultation	0.09 (0.29)	117	0.02 (0.13)	122
Number of out-of-hours clinic consultations	0.16 (0.58)	116	0.02 (0.13)	122
Walk-in clinic consultation	0.09 (0.29)	117	0.02 (0.13)	122
Number of walk-in clinic consultations	0.19 (0.75)	117	0.02 (0.13)	122
Pharmacist consultation	0.32 (0.47)	117	0.07 (0.26)	122
Number of pharmacist consultations	1.05 (2.08)	115	0.15 (0.70)	121
A&E visit	0.04 (0.20)	117	0.00 (0.00)	122
Number of A&E visits	0.09 (0.52)	117	0.00 (0.00)	122
Outpatient visit	0.07 (0.25)	117	0.02 (0.13)	122
Number of outpatient visits	0.12 (0.62)	116	0.02 (0.13)	122
Hospital admission	0.06 (0.24)	117	0.00 (0.00)	122
Number of days admitted	0.08 (0.35)	117	0.00 (0.00)	122
Emergency ambulance	0.01 (0.09)	117	0.00 (0.00)	122
Number of emergency ambulance trips	0.01 (0.09)	117	0.00 (0.00)	122
18 months				
GP consultation: practice	0.34 (0.48)	74	0.15 (0.36)	87
Number of GP consultations: practice	0.81 (1.54)	73	0.33 (0.93)	86
GP consultation: home	0.00 (0.00)	74	0.00 (0.00)	87
Number of GP consultations: home	0.00 (0.00)	74	0.00 (0.00)	87
GP consultation: telephone	0.12 (0.33)	73	0.08 (0.27)	87
Number of GP consultations: telephone	0.21 (0.69)	72	0.09 (0.36)	86
Nurse consultation: practice	0.14 (0.34)	74	0.05 (0.21)	87
Number of nurse consultations: practice	0.19 (0.54)	74	0.05 (0.26)	86
Nurse consultation: home	0.00 (0.00)	74	0.00 (0.00)	87
Number of nurse consultations: home	0.00 (0.00)	74	0.00 (0.00)	87
Nurse consultation: telephone	0.12 (0.33)	74	0.07 (0.25)	87
Number of nurse visits: telephone	0.22 (0.75)	74	0.07 (0.30)	86
NHS 111 consultation	0.11 (0.31)	74	0.05 (0.21)	87
Number of NHS 111 consultations	0.11 (0.31)	74	0.03 (0.18)	86
Out-of-hours clinic consultation	0.08 (0.27)	74	0.06 (0.23)	87
Number of out-of-hours clinic consultations	0.11 (0.42)	74	0.04 (0.19)	85
Walk-in clinic consultation	0.08 (0.27)	74	0.02 (0.15)	87
Number of walk-in clinic consultations	0.09 (0.34)	74	0.02 (0.15)	87
Pharmacist consultation	0.22 (0.41)	74	0.11 (0.32)	87
Number of pharmacist consultations	0.73 (2.17)	74	0.20 (0.63)	86
A&E visit	0.07 (0.25)	74	0.01 (0.11)	87
Number of A&E visits	0.07 (0.25)	74	0.00 (0.00)	86

TABLE 59 Summary of healthcare resource use, as reported in the health utilisation questionnaire (*continued*)

	Conservative management (N = 220)		Tonsillectomy (N = 233)	
	Mean (SD)	n	Mean (SD)	n
Outpatient visit	0.20 (0.40)	74	0.05 (0.21)	87
Number of outpatient visits	0.26 (0.62)	73	0.07 (0.40)	86
Hospital admission	0.15 (0.36)	74	0.00 (0.00)	87
Number of days admitted	0.14 (0.35)	73	0.00 (0.00)	87
Emergency ambulance	0.01 (0.12)	74	0.00 (0.00)	87
Number of emergency ambulance trips	0.01 (0.12)	74	0.00 (0.00)	87
24 months				
GP consultation: practice	0.29 (0.46)	100	0.15 (0.36)	99
Number of GP consultations: practice	0.70 (1.51)	99	0.45 (1.58)	99
GP consultation: home	0.01 (0.10)	100	0.00 (0.00)	99
Number of GP consultations: home	0.07 (0.70)	100	0.00 (0.00)	99
GP consultation: telephone	0.10 (0.30)	100	0.03 (0.17)	99
Number of GP consultations: telephone	0.22 (0.86)	99	0.09 (0.57)	99
Nurse consultation: practice	0.06 (0.24)	100	0.06 (0.24)	99
Number of nurse consultations: practice	0.11 (0.47)	100	0.16 (1.04)	99
Nurse consultation: home	0.00 (0.00)	100	0.00 (0.00)	99
Number of nurse consultations: home	0.00 (0.00)	100	0.00 (0.00)	99
Nurse consultation: telephone	0.11 (0.31)	100	0.03 (0.17)	99
Number of nurse visits: telephone	0.19 (0.70)	99	0.06 (0.42)	99
NHS 111 consultation	0.04 (0.20)	100	0.01 (0.10)	99
Number of NHS 111 consultations	0.08 (0.62)	99	0.01 (0.10)	99
Out-of-hours clinic consultation	0.07 (0.26)	100	0.03 (0.17)	99
Number of out-of-hours clinic consultations	0.15 (0.69)	99	0.03 (0.17)	99
Walk-in clinic consultation	0.07 (0.26)	100	0.01 (0.10)	99
Number of walk-in clinic consultations	0.13 (0.55)	99	0.01 (0.10)	99
Pharmacist consultation	0.21 (0.41)	100	0.09 (0.29)	99
Number of pharmacist consultations	0.55 (1.49)	97	0.15 (0.54)	99
A&E visit	0.02 (0.14)	100	0.02 (0.14)	99
Number of A&E visits	0.02 (0.20)	99	0.03 (0.22)	99
Outpatient visit	0.04 (0.20)	100	0.03 (0.17)	99
Number of outpatient visits	0.05 (0.33)	99	0.12 (0.86)	99
Hospital admission	0.03 (0.17)	100	0.01 (0.10)	99
Number of days admitted	0.04 (0.28)	99	0.02 (0.20)	99
Emergency ambulance	0.00 (0.00)	100	0.01 (0.10)	99
Number of emergency ambulance trips	0.00 (0.00)	100	0.01 (0.10)	99

TABLE 60 Summary of tonsillectomy data, as recorded in the eCRF

Resource use	Conservative management (N = 220) (n = 59 receiving tonsillectomies)	Tonsillectomy (N = 233) (n = 172 receiving tonsillectomies)
Surgical treatment, mean (SD); n		
Number of tonsillectomies	0.27 (0.44); 220	0.74 (0.44); 233
Length of admission	0.22 (0.49); 59	0.25 (0.47); 171
Time in surgery	43.44 (19.14); 52	47.48 (18.54); 153
Paracetamol	0.69 (0.46); 59	0.70 (0.46); 172
Ibuprofen	0.68 (0.47); 59	0.58 (0.49); 172
Codeine phosphate	0.25 (0.44); 59	0.35 (0.48); 172
Co-codamol	0.15 (0.36); 59	0.17 (0.38); 172
Voltarol	0.03 (0.18); 59	0.03 (0.17); 172
Dihydrocodeine	0.32 (0.47); 59	0.19 (0.39); 172
Tramadol	0.02 (0.13); 59	0.05 (0.21); 172
Diffiam spray	0.22 (0.42); 59	0.15 (0.35); 172
Amoxicillin	0.07 (0.25); 59	0.08 (0.27); 172
Inhaler	0.02 (0.13); 59	0.00 (0); 172
Ondansetron	0.00 (0); 59	0.00 (0); 172
Morphine	0.03 (0.18); 59	0.05 (0.21); 172
Diclofenac	0.02 (0.13); 59	0.03 (0.18); 172
Naproxen	0.00 (0); 59	0.02 (0.13); 172
Clarithromycin	0.00 (0); 59	0.01 (0.08); 172
Benzylamine mouthwash	0.00 (0); 59	0.02 (0.15); 172
Senna	0.00 (0); 59	0.01 (0.08); 172
Cyclizine	0.00 (0); 59	0.01 (0.11); 172
Nefopam	0.00 (0); 59	0.02 (0.13); 172
Surgery type, n (%)		
	n = 59	n = 172
Bipolar diathermy	33 (56)	64 (37)
Coblation	0 (0)	1 (1)
Cold dissection	21 (36)	93 (54)
Other	1 (1)	11 (6)
Other surgery type details, n (%)		
	n = 1	n = 11
Tonsillectomy	1 (100)	0 (0)
Bilateral	0 (0)	6 (55)
L tonsil; bipolar diathermy, R tonsil; cold dissection	0 (0)	1 (9)
Excision of tonsil bilateral laser tonsillectomy	0 (0)	1 (9)
Excision of both tonsils tie on lower pole, haemostasis with bipolar cautery	0 (0)	1 (9)

TABLE 60 Summary of tonsillectomy data, as recorded in the eCRF (continued)

Resource use	Conservative management (N = 220) (n = 59 receiving tonsillectomies)	Tonsillectomy (N = 233) (n = 172 receiving tonsillectomies)
Excision of tonsil, bilateral dissection tonsillectomy	0 (0)	1 (9)
Bipolar dissection, cold steel dissection and ties to lower poles	0 (0)	1 (9)
Surgeon grade, n (%)	n = 59	n = 172
Consultant	30 (51)	91 (53)
Specialist registrar	24 (41)	52 (30)
Associate specialist	0 (0)	8 (5)
Staff grade	1 (2)	7 (4)
Core trainee	1 (2)	1 (1)
Other	1 (2)	9 (5)
Other surgeon grades, n (%)	n = 1	n = 9
CST	0 (0)	1 (12)
Fellow	0 (0)	2 (25)
ST3	0 (0)	1 (12)
ST4	0 (0)	4 (50)
ST7	1 (100)	0 (0)
Anaesthetist grade, n (%)	n = 59	n = 172
Consultant	47 (80)	136 (79)
Specialist registrar	3 (5)	11 (6)
Associate specialist	1 (2)	2 (1)
Other	3 (5)	10 (6)
Other anaesthetist grade, n (%)	n = 3	n = 10
Anaesthetic practitioner	0 (0)	1 (10)
Fellow	0 (0)	1 (10)
Novice consultant	0 (0)	1 (10)
ST5	2 (67)	1 (10)
ST7	1 (33)	6 (10)
Anaesthetic practitioner	0 (0)	1 (10)

CST, core surgical trainee; ST, specialist trainee.

TABLE 61 Cost-utility analysis sensitivity analysis: utility decrements associated with a sore throat episode are included in the QALY equation

Investigation strategy	Cost (£) (95% CI) ^a	Incremental cost (£) (95% CI) ^b	Effect (95% CI) ^a Effect (95% CI) ^b	Incremental effect (95% CI) ^b	ICER (£)	Probability that tonsillectomy is cost-effective for different threshold values for society's WTP for an additional QALY				
						£0	£10,000	£20,000	£30,000	£50,000
Conservative management (costs, n = 217; outcomes, n = 215)	879 (774 to 984)		1.38 (1.35 to 1.40)			1.00	0.00	0.00	0.00	0.00
Tonsillectomy (costs, n = 231; outcomes, n = 229)	1365 (1273 to 1458)	488 (341 to 620)	1.49 (1.47 to 1.52)	0.122 (0.09 to 0.15)	4000	0.00	1.00	1.00	1.00	1.00

a Point estimates are based on the unadjusted analysis (costs, n = 448; QALYs, n = 444).

b Incremental estimates are based on the adjusted analysis (n = 444).

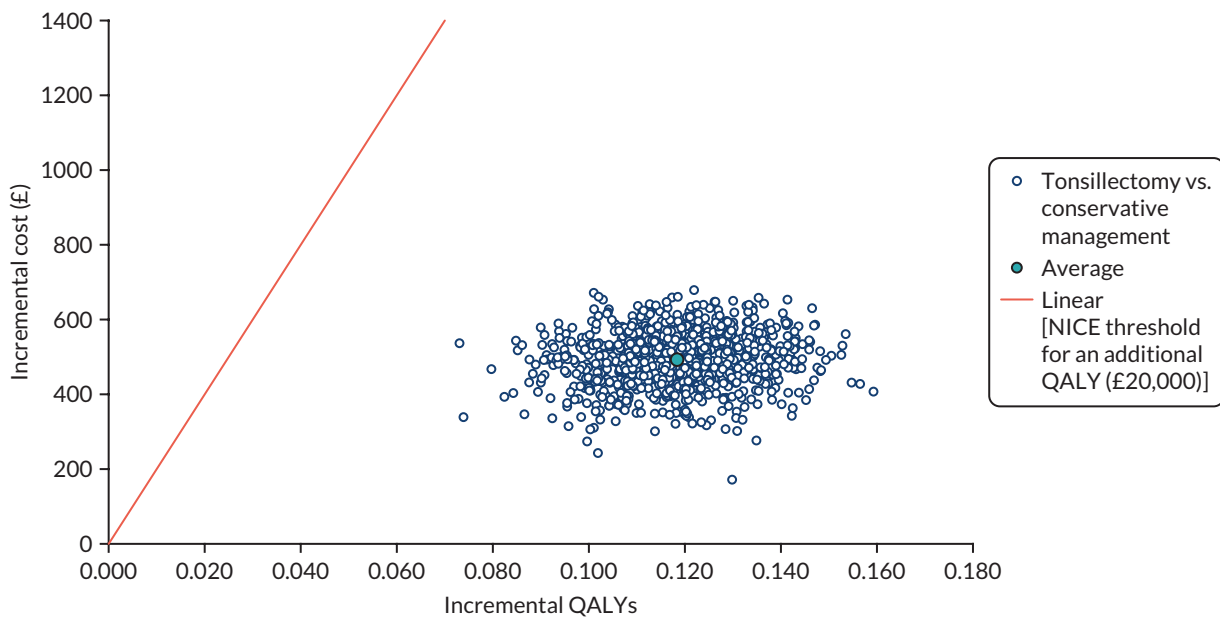


FIGURE 30 Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA multiple imputation results.

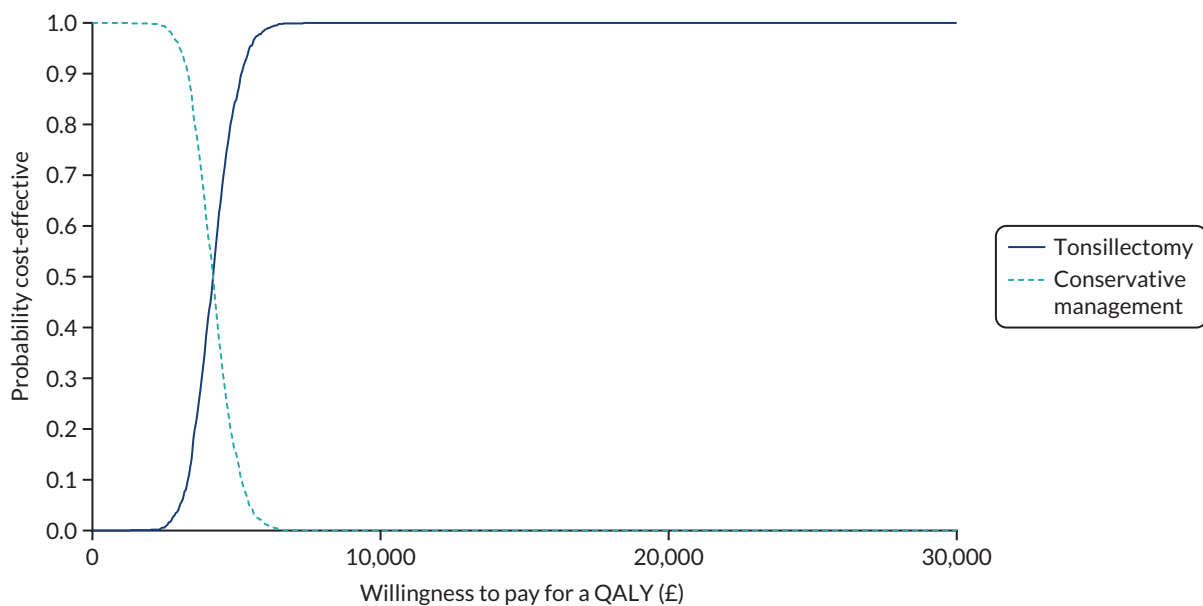


FIGURE 31 Cost-effectiveness acceptability curves for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA multiple imputation results.

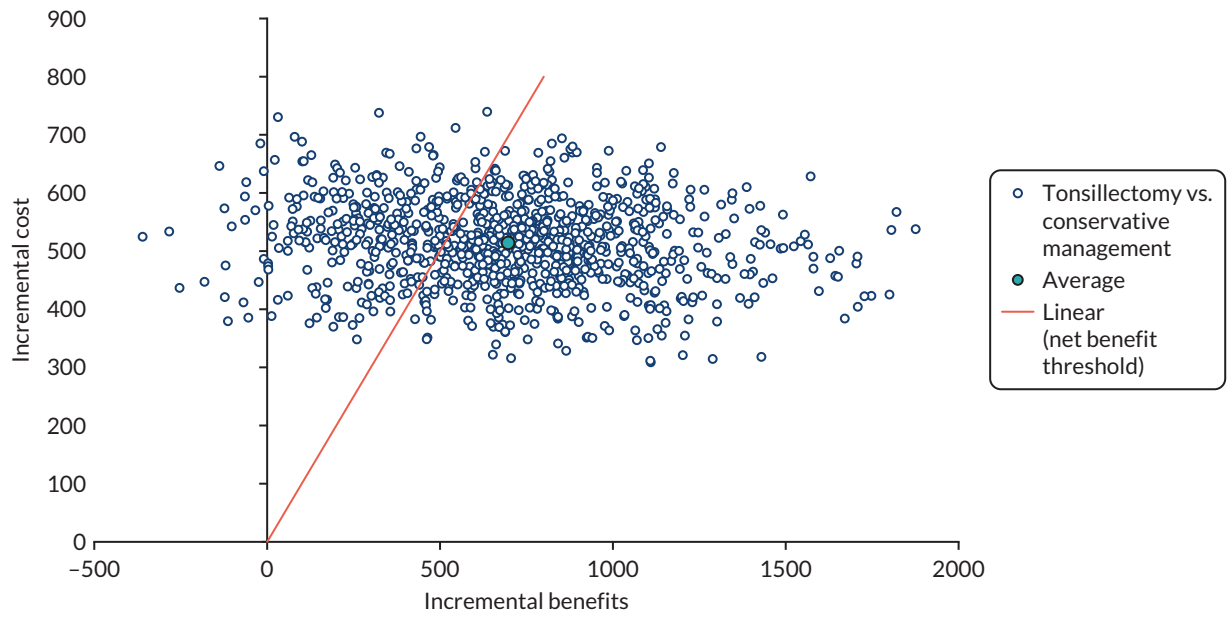


FIGURE 32 Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CBA multiple imputation results.

TABLE 62 Cost-utility analysis sensitivity analysis: tonsillectomy costs estimated using microcosting

Investigation strategy	Cost (£) (95% CI) ^a	Incremental cost (£) (95% CI) ^b	Effect (95% CI) ^a	Incremental effect (95% CI) ^b	ICER (£)	Probability that tonsillectomy is cost-effective for different threshold values for society's WTP for an additional QALY				
						£0	£10,000	£20,000	£30,000	£50,000
Outcome: QALYs – results (sensitivity analysis – microcosting)										
Conservative management (costs, n = 217; outcomes, n = 215)	830 (734 to 926)		1.44 (1.41 to 1.47)			1.00	0.00	0.00	0.00	0.00
Tonsillectomy (costs, n = 231; outcomes, n = 229)	1247 (1161 to 1333)	417 (288 to 545)	1.55 (1.52 to 1.57)	0.118 (0.09 to 0.14)	3534	0.00	1.00	1.00	1.00	1.00

a Point estimates are based on the unadjusted analysis (costs, n = 448; QALYs, n = 444).

b Incremental estimates are based on the adjusted analysis (n = 444).

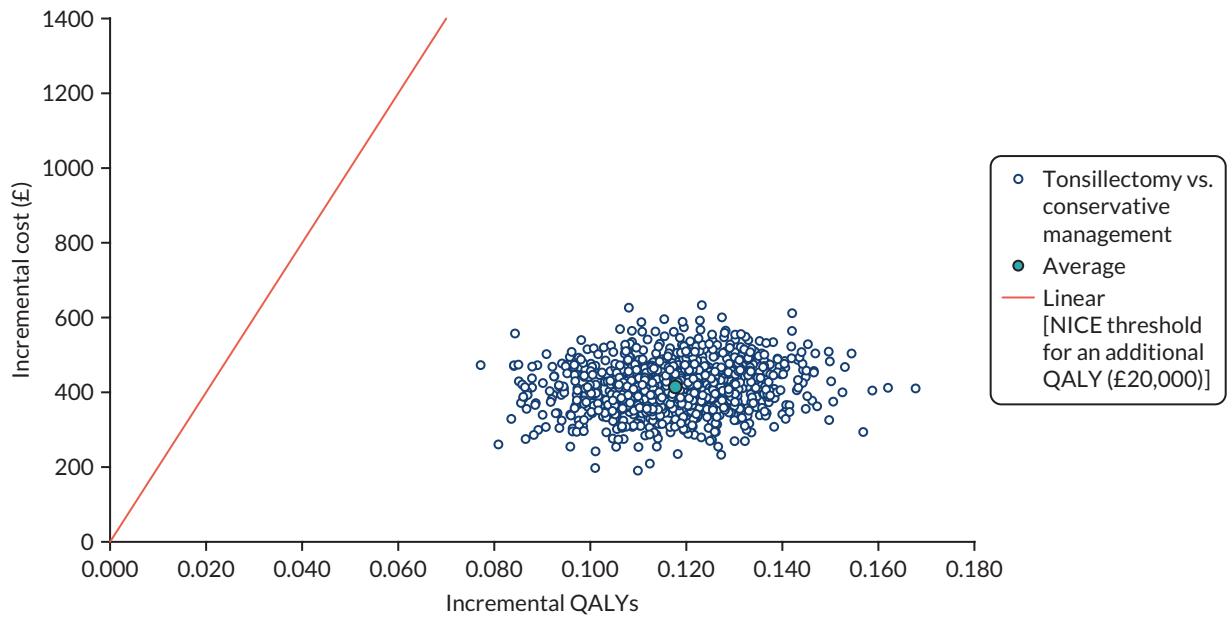


FIGURE 33 Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (tonsillectomy costs were estimated using microcosting).

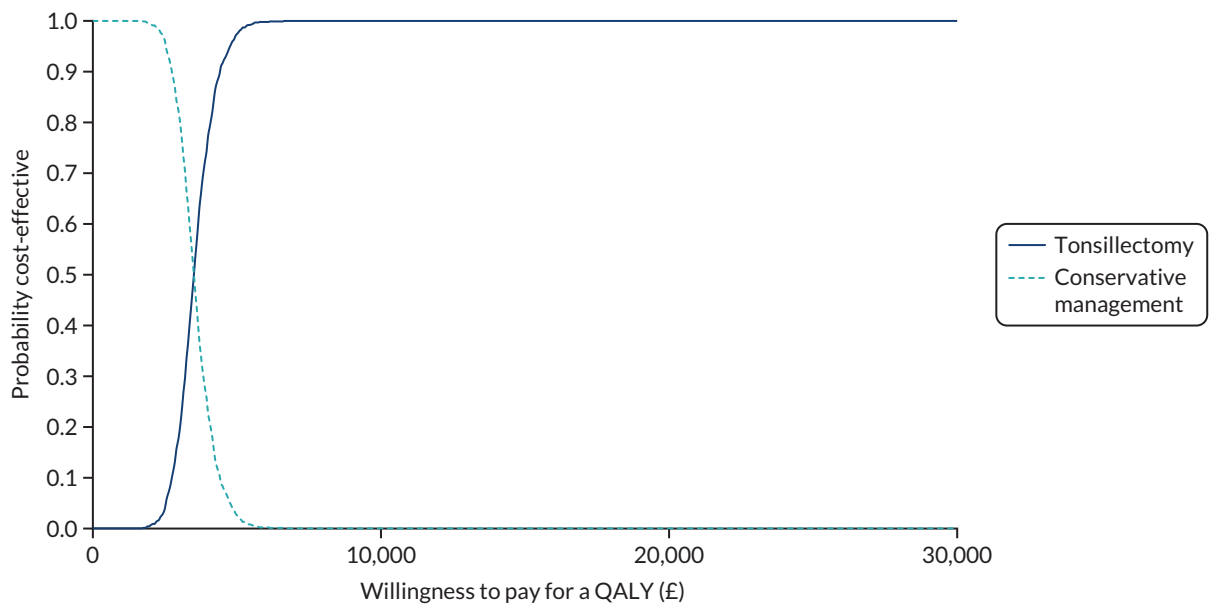


FIGURE 34 Cost-effectiveness acceptability curves for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (tonsillectomy costs were estimated using microcosting).

TABLE 63 Cost-utility analysis sensitivity analysis: participant costs

Investigation strategy	Cost (£) (95% CI) ^a	Incremental cost (£) (95% CI) ^b	Effect (95% CI) ^a	Incremental effect (95% CI) ^b	ICER (£)	Probability that tonsillectomy is cost-effective for different threshold values for society's WTP for an additional QALY					
						£0	£10,000	£20,000	£30,000	£50,000	
Outcome: QALYs – results (sensitivity analysis – participant costs)											
Conservative management (costs, n = 220; outcomes, n = 215)	4914 (4185 to 5644)		1.444 (1.42 to 1.46)			0.02	0.00	0.00	0.00	0.00	
Tonsillectomy (costs, n = 233; outcomes, n = 229)	4033 (3857 to 4479)	-885 (-1735 to 35)	1.551 (1.53 to 1.57)	0.118 (0.09 to 0.14)	Dominant	0.98	1.00	1.00	1.00	1.00	1.00

a Point estimates are based on the unadjusted analysis (costs, n = 453; QALYs n = 444).

b Incremental estimates are based on the adjusted analysis (n = 444).

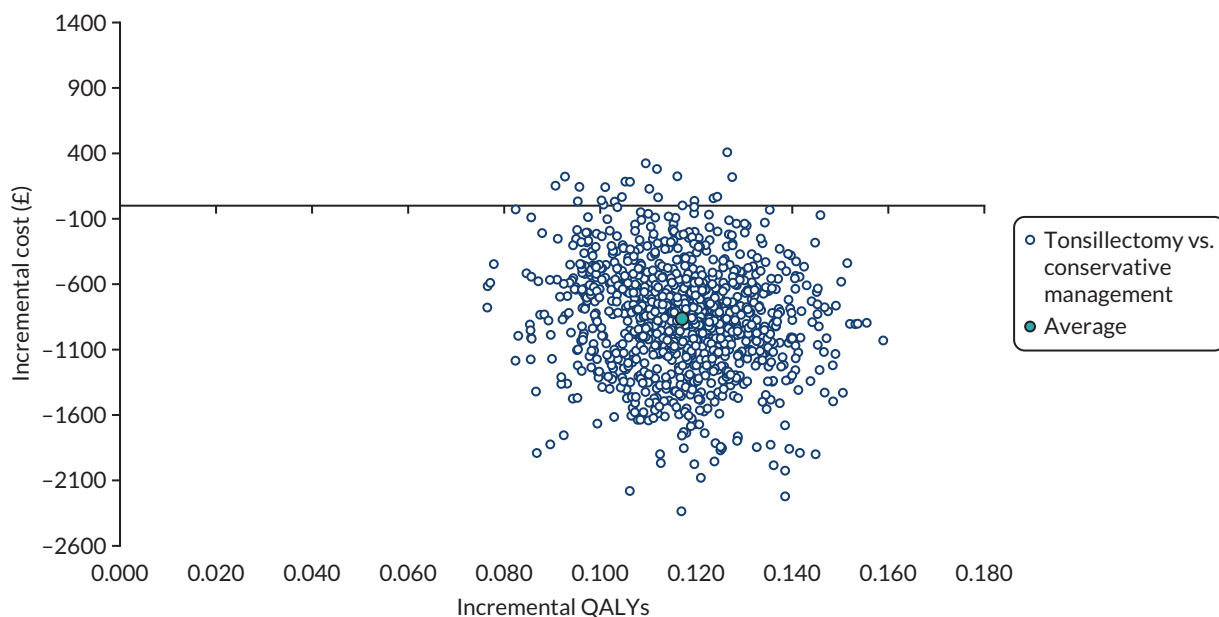


FIGURE 35 Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (costs and QALYs were estimated using multiple imputation and participant costs have been included with total costs).

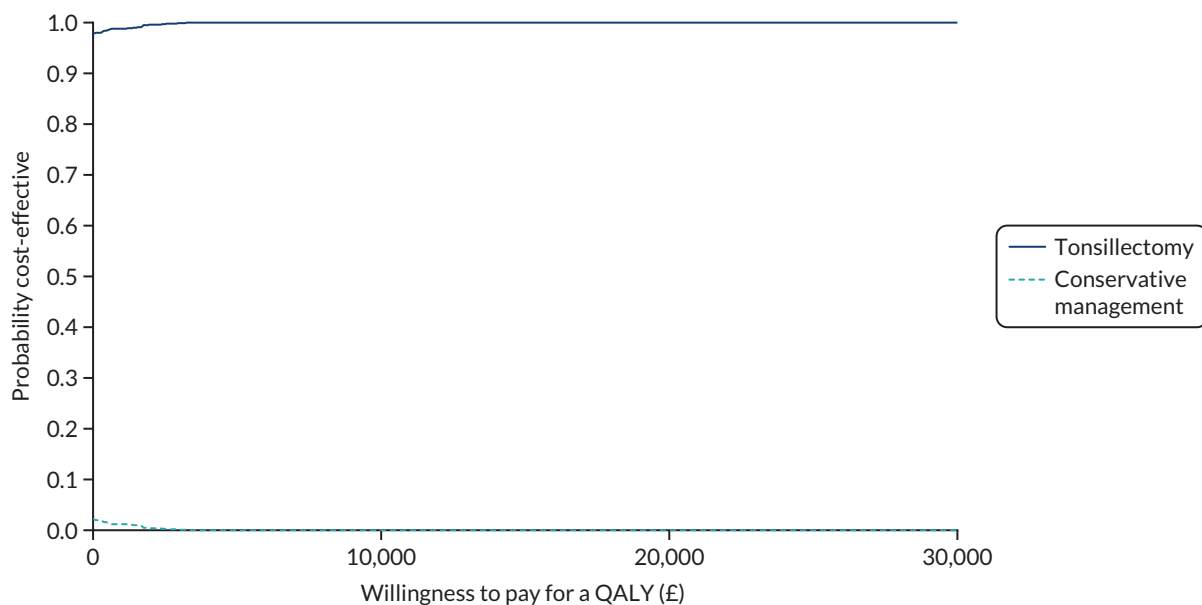


FIGURE 36 Cost-effectiveness acceptability curves for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (costs and QALYs were estimated using multiple imputation and participant costs have been included with total costs).

TABLE 64 Cost-utility analysis sensitivity analysis: healthcare resource use costs estimated based on GP linkage data

Investigation strategy	Cost (£) (95% CI) ^a	Incremental cost (£) (95% CI) ^b	Effect (95% CI) ^a	Incremental effect (95% CI) ^b	ICER (£)	Probability that tonsillectomy is cost-effective for different threshold values for society's WTP for an additional QALY				
						£0	£10,000	£20,000	£30,000	£50,000
Outcome: QALYs – results (sensitivity analysis – healthcare utilisation costs estimated using GP linkage data)										
Conservative management (costs, n = 171; outcomes, n = 215)	534 (420 to 649)		1.44 (1.42 to 1.46)			1.00	0.00	0.00	0.00	0.00
Tonsillectomy (costs, n = 187; outcomes, n = 229)	1254 (1160 to 1349)	598 (385 to 811)	1.55 (1.53 to 1.57)	0.112 (0.07 to 0.15)	5339	0.00	1.00	1.00	1.00	1.00

a Point estimates are based on the unadjusted analysis (costs, n = 358; QALYs, n = 444).

b Incremental estimates are based on the adjusted analysis (n = 197).

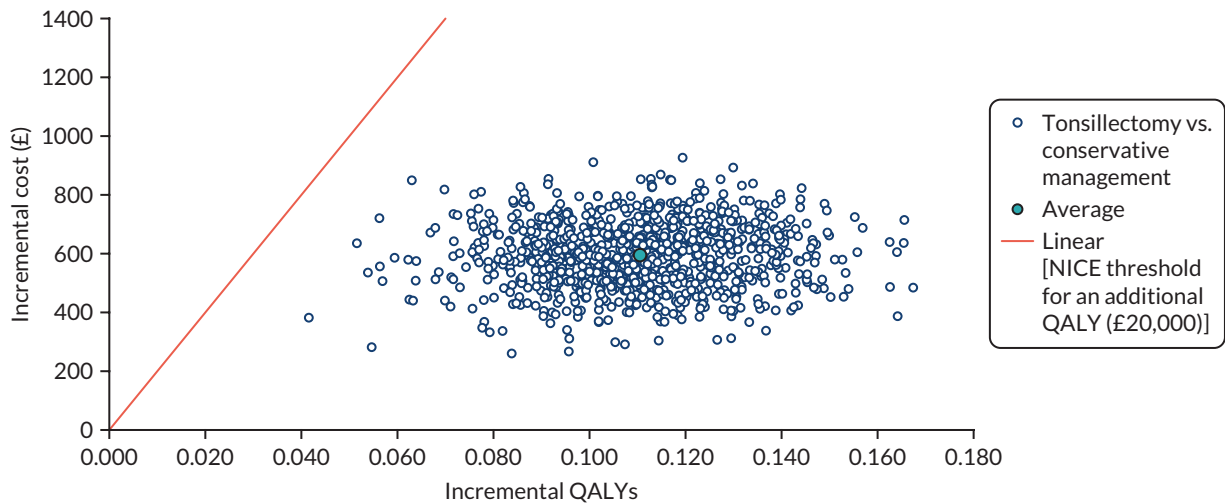


FIGURE 37 Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (QALYs were estimated using multiple imputation and healthcare utilisation costs were estimated using GP linkage data).

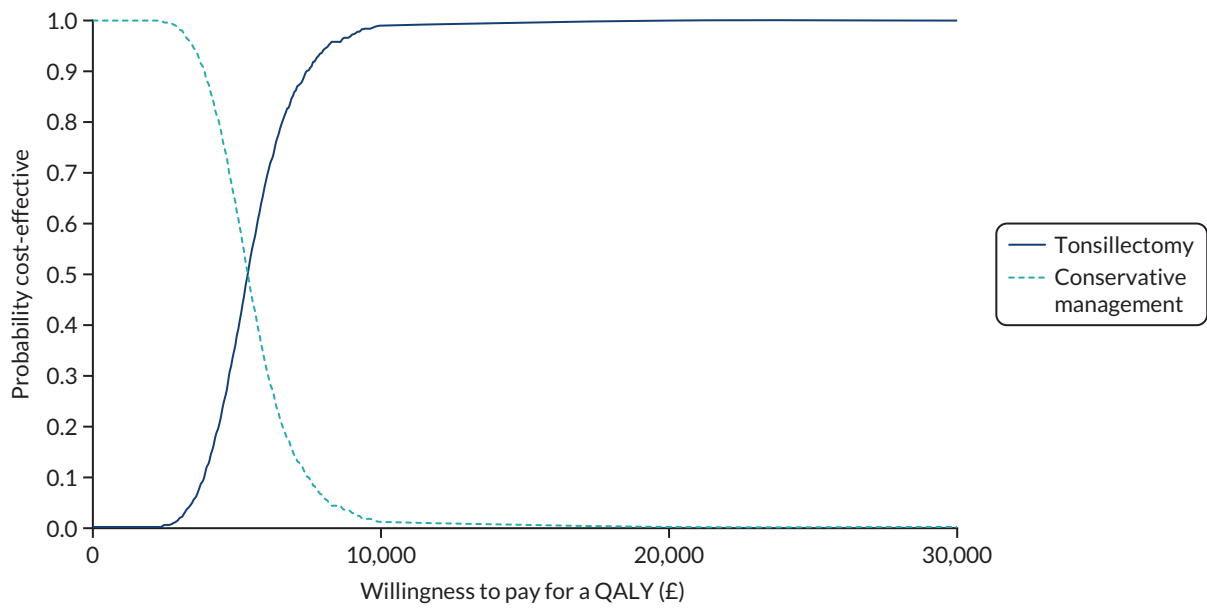


FIGURE 38 Cost-effectiveness acceptability curves for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (QALYs were estimated using multiple imputation and healthcare utilisation costs were estimated using GP linkage data).

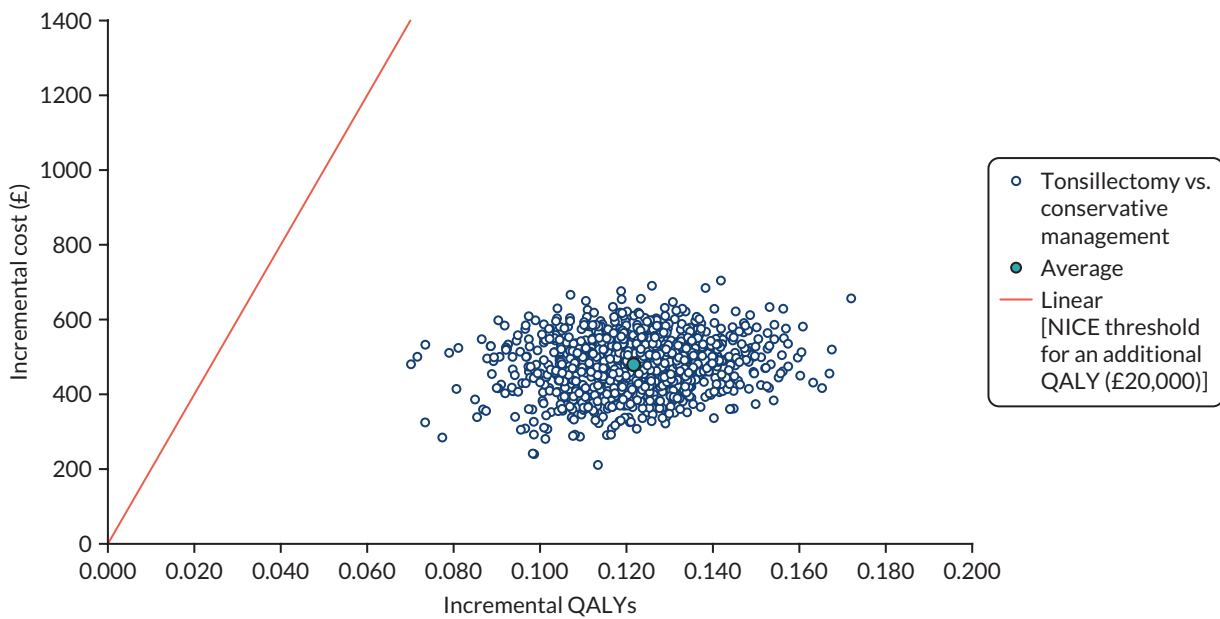


FIGURE 39 Cost-effectiveness plane for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (QALYs were estimated using multiple imputation and include utility decrements associated with a sore throat episode).

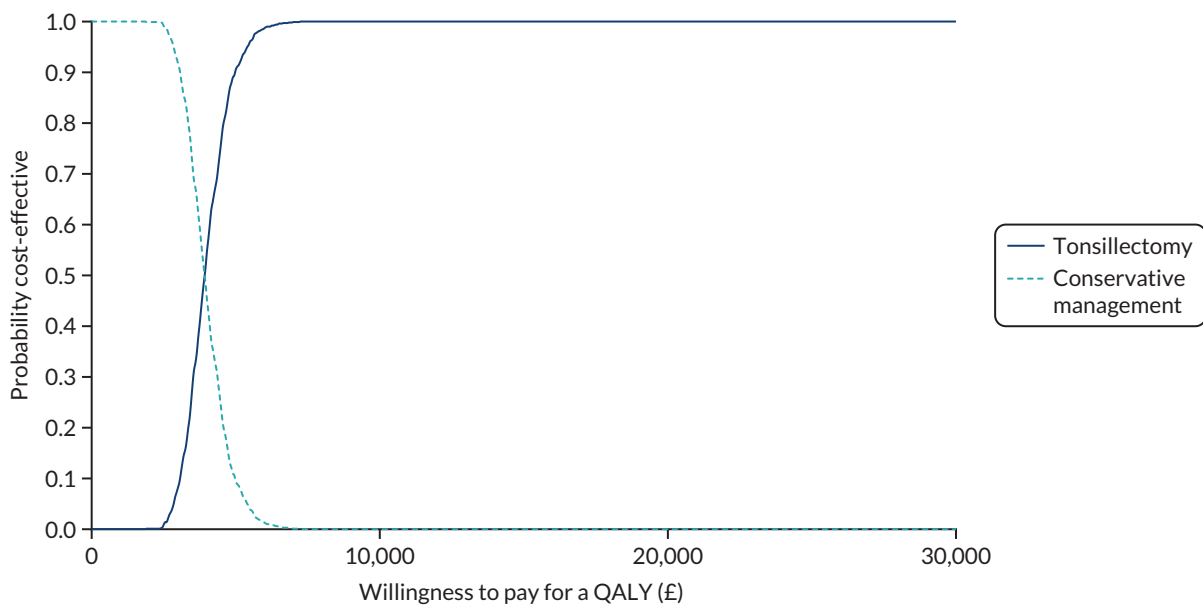


FIGURE 40 Cost-effectiveness acceptability curves for tonsillectomy vs. conservative management using the adjusted bootstrapped CUA results (QALYs were estimated using multiple imputation and include utility decrements associated with a sore throat episode).

EME
HSDR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

*This report presents independent research funded by the National Institute for Health and Care Research (NIHR).
The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the
Department of Health and Social Care*

Published by the NIHR Journals Library