

The 14-item Paediatric Throat Disorders Outcome Test: a valid, sensitive, reliable, parent-reported outcome measure for paediatric throat disorders

C HOPKINS, J FAIRLEY*, M YUNG†, I HORE, S BALASUBRAMANIAM*, M HAGGARD‡

Abstract

Objectives: We modified and abbreviated a pre-existing research questionnaire, the Tonsil and Adenoid Health Status Instrument, to make it suitable for rapid completion as a disease-specific, health-related quality of life research tool for children with tonsil and adenoid disease in the UK. We determined the main psychometric properties of the resulting 14-item Paediatric Throat Disorders Outcome Test.

Design, setting and participants: Pre- and post-operative questionnaires were completed by the parents of children with throat disorders referred to two large hospitals. We included children with recurrent tonsillitis and with obstructive sleep apnoea. A separate cohort of healthy children of comparable age range was also studied.

Main outcome measures: The test's internal consistency and responsiveness were analysed and its construct validity documented via known-group differences.

Results: A total of 126 completed questionnaires were received from the hospital referral group. The children's mean age was 6.5 years (range one to 16). The 40 unaffected children were well matched in age to the study population (mean 6.1 years, range two to 15). Cronbach's α coefficient for the pre-operative assessment total score was 0.84. The test–retest reliability coefficient for the total score was 0.98, indicating very high reproducibility. The 14-item Paediatric Throat Disorders Outcome Test discriminated well between children known to suffer with throat problems and a group of healthy controls ($p < 0.0001$; $t = 24.016$). Six months after surgical intervention, parentally reported questionnaire scores had improved (i.e. were lower) ($p < 0.0001$; $t = 7.01$). The standard effect size (i.e. change in mean divided by baseline standard deviation) for children for whom post-operative questionnaires were completed was 1.53; this is very large.

Conclusions: The 14-item Paediatric Throat Disorders Outcome Test is an appropriate, disease-specific, parent-reported outcome measure for children with throat disorders, for which we have demonstrated internal consistency, reliability, responsiveness to change and two forms of construct validity.

Key words: Outcomes Assessment; Patient Outcomes; Tonsillitis; Tonsillectomy

Introduction

The use of patient-reported outcome measures is rapidly growing in studies of clinical effectiveness and quality of care. By 2010, all National Health Service (NHS) hospitals in England will be required to submit such data for a range of indicator procedures.¹ Despite obvious problems in the reliability of very brief measures, there is even a suggestion that these outcomes may influence how much hospitals are paid for a procedure.² Whatever the details of application, patient-reported outcome measures are likely to play an increasingly important role in the evaluation of care for ENT conditions.

Sore throats are said to account for 35 million days lost from school or work per annum in the UK.³ In excess of 30 000 tonsillectomies were performed on children aged two to 14 years in English NHS hospitals during the 2005–2006 financial year (Hospital Episode Statistics (HES) data).⁴ Seventy to 80 per cent of these procedures were undertaken as management for recurrent sore throats, the remainder for obstructive symptoms.⁵ In the same period, the Chief Medical Officer for England, Sir Liam Donaldson, highlighted geographical variation in tonsillectomy rates which he claimed was 'unacceptable'.⁶ (For example, data from The Scottish National Tonsil Audit showed that rates of tonsillectomy in

From the ENT Departments of Guy's and St Thomas' Hospital, London, *William Harvey Hospital, Ashford, and †Ipswich Hospital, and the ‡Department of Experimental Psychology, University of Cambridge, UK.
Accepted for publication: 14 September 2009. First published online 6 January 2010.

childhood varied from less than four in 10 000 in Forth Valley to almost 10 in 10 000 in Dumfries and Galloway.)

About 8000 tonsillectomies were allegedly unnecessary; amounting to more than six million pounds of wasted expenditure and unjustified patient morbidity. The National Institute for Health and Clinical Excellence has been requested to issue guidance on ineffective treatments, and has reported a lack of evidence to support tonsillectomy. This has resulted in some Primary Care Trusts refusing to commission tonsillectomy for any childhood throat disorders, unfortunately without any attempt thus far to discriminate between throat disorder severity or frequency.

Impact of childhood throat disorders on quality of life, and current evidence base for tonsillectomy

Stewart *et al.* have demonstrated that tonsil and adenoid disease has a significant impact on quality of life (QoL) and health status in affected children, using both disease-specific and generic QoL measures.⁷ However, there are few well designed randomised, controlled trials assessing tonsillectomy benefit in this patient population. Four trials, published 25 years ago or more, used the number of episodes of sore throat and number of days of illness as the main outcome measures.^{8–11} These trials had problems with recruitment and potential bias, but overall they showed a small reduction in the frequency of sore throats in groups undergoing tonsillectomy, compared with those undergoing conservative management. A recent Cochrane Review concluded that adeno-tonsillectomy is effective in reducing both the number of episodes of sore throat and days with sore throat, with improvement being more marked in those most severely affected.¹² A number of non-randomised trials have also demonstrated benefit from surgery. Improvements on all subscales of the Glasgow Children's Benefit Inventory were found in a group of 191 children undergoing tonsillectomy.¹³ The recently published 'Tonsillitis Outcomes: Toward Reaching Evidence in Adults and Tots' study prospectively followed children undergoing tonsillectomy, and demonstrated benefit using both disease-specific and global QoL instruments.¹⁴ The 'North of England Study of Tonsillectomy and Adeno-tonsillectomy in Children' study, a randomised, controlled trial of tonsillectomy, is due to report in 2010.¹⁵ This study aims to recruit 400 children, randomised to receive either surgical or conservative management. The main outcome measure will be the number of self-reported episodes of sore throats, but participants will also complete a generic QoL instrument, the Pediatric Quality of Life Inventory (PedsQL).

Thus, it cannot yet be claimed that there is evidence of widespread material benefit from tonsillectomy as currently applied, although it remains probable that there is benefit in appropriately selected cases. In contrast, the National Prospective Tonsillectomy Audit has highlighted the risks associated with elective tonsillectomy, with 5 per cent of patients reporting

complications and 0.9 per cent requiring a return to theatre for arrest of haemorrhage.¹⁶

It is desirable to have some constraint on the severity of symptoms for which tonsillectomy is undertaken, and some check on the general satisfactoriness of outcome. Therefore, a disease-specific instrument is needed to measure health-related QoL in children suffering throat disorders.

Here, we report the further development and psychometric validation of a disease-specific, health-related QoL instrument for use in childhood throat disorders, the 14-item Paediatric Throat Disorders Outcome Test. This is a modification of a pre-existing instrument, the Tonsil and Adenoid Health Status Instrument.¹⁷

What is the Tonsil and Adenoid Health Status Instrument?

Although there are several research instruments currently available which assess global QoL in children, global instruments are likely to lack sensitivity for conditions such as childhood throat disorders. Disease-specific health status instruments are usually more sensitive to differences or changes in health status due to one organ system, compared with generic instruments, and their results are easier to interpret.¹⁸ Average benefits too small to show through in generic measures may nevertheless be important to patients and healthcare commissioners. A literature search identified three instruments which measured disease-specific QoL in children with throat disorders, but only one was intended for children with recurrent or chronic infections. The Tonsil and Adenoid Health Status Instrument is claimed to be valid, reliable and responsive to clinical change. Because tonsil and adenoid disease affects children over a wide age range and with different communication skills, the primary caregiver completes the questionnaire as a proxy for the affected child.¹⁷

The Tonsil and Adenoid Health Status Instrument was used as the main outcome measure in the 'Tonsillitis Outcomes: Toward Reaching Evidence in Adults and Tots' study.¹⁴ This was a multicentre, prospective, observational outcomes study. Ninety-two children (mean age (standard deviation (SD)) 10.6 (3.4) years) were enrolled, with follow up available for 58 children at six months and 38 children at one year. The children showed significant improvement in all subscales of the Tonsil and Adenoid Health Status Instrument, including airway and breathing, infection, healthcare utilisation, cost of care, eating, swallowing (all $p < 0.001$), and behaviour ($p = 0.01$). This was a small study, with outcome data on only a modest proportion of recruits (63 per cent (58/92)), so possibly subject to biased attrition.

For our own research, we preferred to build on the most promising existing instrument, thus making best use of previous developmental effort, rather than start again 'from scratch'. The Clinical Audit and Practice Advisory Group of ENT-UK therefore reviewed the Tonsil and Adenoid Health Status Instrument. Our expert panel concluded that this instrument would need modification for the UK

population. This modification, and the revalidation of the resulting instrument, is presented here.

Methods

Item content

A panel of eight experienced ENT surgeons reviewed the Tonsil and Adenoid Health Status Instrument to assess face and content validity, and thought some questionnaire items to be unsuitable for a UK population. Notably, the item regarding the direct cost of care and medications has less relevance in the UK, where all medical care and prescriptions are free of charge for children. The indirect costs of caring for a sick child are relevant, but would not be measured by this item. However, a question regarding missed school days would assess the impact on both the child and their caregiver, and so was substituted. A questionnaire item regarding 'strep throat' begs the question of microbiological testing; also, this is not a term commonly used in the UK or understood by parents here. Severity of infections would be indicated by other items assessing acute and chronic infections, making the 'strep throat' item redundant. In addition, the panel felt ear infections were important in relation to adenoid hypertrophy and infection, so added a question on these.

The Tonsil and Adenoid Health Status Instrument was thus modified to a 16-item prototype questionnaire, in which the range of scores was zero to 80, with lower scores implying a better health-related QoL. The 14-item Paediatric Throat Disorders Outcome Test (Appendix 1) was the result of further item reduction, described and justified below. Two items from the 16-item prototype were deleted (i.e. 'your child not gaining weight as expected' and 'behaviour problems at school or home'), resulting in the final, 14-item questionnaire. The reliability and construct validity of this modified instrument were reassessed.

Recruitment of participants

In order to aid recruitment of children with a broad range of disease severity, two centres were invited to participate: one was a specialist paediatric hospital accepting both secondary and tertiary referrals, the other a large district general hospital treating children over three years of age. The parents of children referred for assessment of throat disorders in both participating centres were asked to complete the 16-item prototype questionnaire. The questionnaire also contained a free text box in which parents were asked to mention any important symptoms not included in the questionnaire.

When considering surgical intervention for recurrent acute tonsillitis, clinicians followed the Scottish Intercollegiate Guidelines Network guidelines.¹⁹ Children with recurrent sore throats were to receive tonsillectomy if they reported more than five tonsillitis episodes per year which prevented normal activity, with this incidence having persisted for at least one year. A diagnosis of obstructive sleep apnoea (OSA) was made on the basis of a clear history and supportive clinical findings, or on the basis of a sleep study

in equivocal cases. Children who did not meet indications for tonsillectomy were offered a period of 'watchful waiting'. The surgeon recorded the diagnosis and treatment offered.

Over a period of two months in one of the participating centres, consenting parents were asked to complete a second copy of the questionnaire two weeks after their child's pre-operative out-patient appointment, in order to assess test-retest reliability. Parents were contacted by letter, sent a copy of the questionnaire and asked to return it in a stamped, addressed envelope, provided there had been no change in their child's health.

Parents of children in one of the participating centres were also contacted six months after their initial visit and asked to complete a further questionnaire. In addition to the single occasion status ratings captured by the 16-item prototype questionnaire, all post-operative questionnaires contained an extra section of transition ratings comparing pre- and post-operative health and health-related QOL problems on a five-point scale (with 1 = much better, 2 = a little better, 3 = about the same, 4 = a little worse and 5 = much worse). Such transition ratings (i.e. patient ratings of the extent to which their health has improved or deteriorated) are extremely common in clinical practice and clinical research, but are prone to expectancy bias.²⁰

The 16-item prototype questionnaires were also given to parents of a healthy cohort of children not known to be suffering from tonsillitis, otitis media with effusion or OSA. These participants were contacted through a school near one of the hospitals. Children in year one, of the same mean age (6.5 years) as the study population, were targeted. Parents agreeing to take part were asked to complete a 16-item prototype questionnaire not only for the index child but also for the child's siblings, in an attempt to achieve a similar age range as the target population (i.e. one to 16 years). An information sheet was provided and participation was voluntary. Parents were asked if their children had been diagnosed with recurrent tonsillitis or OSA, and were excluded if responding positively to either. Only data for the 14 items contained in the final version of the instrument were used for the final stages of analysis.

Ethical approval

The relevant ethical committee (National Research Ethics Service) was contacted about the study. Their formally recorded view was that the study did not require formal ethical approval, as the instrument was being used to audit current practice and the content did not differ materially from questions that would be widely used in normal practice.

Imputation of missing data

In cases where some answers to 16-item prototype questionnaire items were missing, a total score was imputed from the mean of completed items, provided that over 50 per cent of items had been completed. This was crude but adequate for many purposes, and preferable to large scale loss of participants; it was

also consistent with scoring practices for other patient-based outcome measures (e.g. the 36-item short-form health survey (SF-36)).²¹ Imputation (that is, substitution of a total most likely from the pattern of available items, using inter-item or item-total regression) deals simply but effectively with missing values, and provides a result that is usually more powerful than that derived after excluding cases with incomplete data.²² Imputed values were not included in principle component and factor analyses.

Item reduction

Classical psychometric techniques were used to explore the optimum number of items in the final version of the questionnaire, and unsatisfactory question items were identified for removal. Data derived from the 126 children referred with throat disorders were subjected to principal components analysis (using the first unrotated factors) to judge inclusion for the total score, and to factor analysis with Varimax orthogonal rotation to judge inclusion for the subscores. In order to be retained, an item had to (1) load at least 0.35 on the first principal component, (2) load at least 0.40 on a factor representing a score to be supported, and (3) not cross-load excessively (i.e. the difference between the two rotated loadings had to be more than 0.20). The first and third of these criteria are in partial conflict, as an item cross-loading on two rotated main factors will also load well on the first principal component. Clinical expert opinions and the parents' comments were also considered to ensure important items were not excluded in light of the statistical findings.

Data analysis

Once item reduction had produced the optimum number of items, all subsequent analysis was performed for the 14-item Paediatric Throat Disorders Outcome Test. In order to evaluate this questionnaire, the internal consistency, test-retest reliability and known-groups validity (one form of construct validity) were analysed.

Reliability

Reliability was assessed by test-retest reproducibility and internal consistency. The latter contains validity aspects as well, and refers to the inter-correlation of individual items with others and hence with the total. Except when aggregation over heterogeneity is a declared aim, there should be homogeneity of items in the scale. This is measured using Cronbach's α coefficient, which is in effect the aggregate intercorrelation between questionnaire items.²³ The conventionally accepted score for Cronbach's α coefficient is ≥ 0.7 . Cronbach's α coefficients for the two factor solutions were obtained from the scores of the items associated with each factor.

Test-retest reliability measures the stability of an instrument over time with repeated testing, and is assessed by administering the instrument to respondents on two different occasions and examining the correlation between scores. Like Cronbach's α

coefficient, the conventional threshold is 0.70. A test-retest reliability coefficient was obtained by correlating the total scored in the same way for matched responses in the initial and repeat questionnaires.

Validity

The known-groups validity of the 14-item Paediatric Throat Disorders Outcome Test was examined via an unmatched case-control contrast between the clinical and unaffected population groups, using the standardised effect size. (As the unaffected group was smaller, it was less suited to providing a stable reference SD, so we took the mean difference divided by pooled SD.) The associated statistical significance is given by the unpaired *t*-test and one-way analysis of variance (ANOVA). This was done both for the total score and the two factor-based scores. We also hypothesised that children not meeting the Scottish Intercollegiate Guidelines Network criteria for tonsillectomy (and so subject to watchful waiting) would give lower scores on the 14-item Paediatric Throat Disorders Outcome Test than those considered candidates for surgery. Furthermore, we investigated whether there was a difference in scores in children referred with recurrent acute tonsillitis, OSA or both indications for treatment.

Responsiveness (i.e. can the new scale detect significant clinical change over time?)

Responsiveness was assessed by examining the 16-item questionnaire scores before and after surgery in order to measure the treatment effect size (i.e. the mean change score divided by the SD of scores at baseline). (Note that the SD of the change scores themselves is one contributor to statistical significance, but because of the general pre-post correlation it gives an inflated impression of the magnitude of change; thus, scaling by an absolute SD, either baseline or pooled, is more meaningful and generalisable.) By convention, an effect size of >0.2 is considered small, >0.5 moderate and >0.8 a 'large' effect.²⁵ The associated statistical significance comes from the related-sample *t*-test.

All study analyses were performed using the Stata version 7.0 software package (Statacorp; College Station, Texas, USA).

Results

Patient characteristics

One hundred and twenty-six completed questionnaires were received from the hospital referral group. The children's mean age was 6.5 years, and age range one to 16 years. The 40 unaffected children were well matched in age to the study population (mean 6.1 years; range two to 15).

Missing data

The number of missing items of data was small. Responses were missing from six questionnaires (4.8 per cent of those received); on these questionnaires, two to six items (i.e. 14 to 42 per cent of items) were

unanswered. Given this small number of missing responses, our simple method for imputing missing values avoided bias and slightly increased power for the psychometric issues addressed.

Item reduction and factor analysis

Using Stata statistical software, principal components analysis was performed using the 16-item prototype questionnaire data obtained for the children referred with throat disorders. This analysis gave 35, 55, 62 and 69 per cent of the cumulative variance explained by one, two, three and four factors, respectively. The respective Eigen values were 4.63, 2.29, 0.74 and 0.53. Extracting two factors was highly justified by the sharp drop after two (from 2.29 to 0.74) per extra factor. Importantly, the two-factor solution also grouped the items into two highly interpretable sub-sets (see Discussion below). It is possible that, with a larger sample and item set, further factors could be justified in a comprehensive questionnaire for this condition. However, as the target number of items (approximately 15 items maximum) could reliably support only two factors, factors three and four were abandoned. Rotated factor three received high loadings (more than 0.45 in the three-factor solution) from only two items (problems with weight and eating); this was relevant to a later decision. Two items (problems with weight gain, and poor behaviour at school) were then deleted. The two-factor solution is given by the loadings in Table I, both before item reduction ($K = 16$) and after ($K = 14$). The loadings for all 14 chosen items remained within 1 per cent, further affirming the stability of the two-factor solution. The deleted item pair had high uniqueness (greater than 0.79), and one item also had a poor contribution to the total (first principal component loading, unrotated), of less than 0.38. After rotation, neither of these two items met jointly (in the 16-item, two-factor solution) the loading and cross-loading criteria distinguished above, which the retained 14 items did. Table II shows that most differences between factors in items' loadings exceeded 0.5; there were three slightly cross-loading items (regarding ear ache, eating and sleepiness), but deleting these also would have impaired general reliability.

All further analyses relate to the reduced, 14-item instrument.

Internal consistency

The Cronbach's α coefficient for pre-operative assessment using the 14-item Paediatric Throat Disorders Outcome Test total score was 0.84. Given that the change in item content from the original instrument was minimal, this high internal consistency provides further assurance that the modified instrument measured the same underlying constructs as the original instrument. Cronbach's α coefficients were 0.86 and 0.80 for the two factor scores (factors one and two, respectively).

Test-retest reliability

Sixteen parents completed and returned a second questionnaire two weeks after their child's initial

out-patient assessment. The test-retest reliability coefficient for the total score was 0.98, indicating very high reproducibility. For the separate factors, the coefficients were 0.94 for the total of infective items and 0.96 for obstructive items.

Known-groups comparisons

Table II presents the mean pre-operative 14-item Paediatric Throat Disorders Outcome Test scores by key groups. The mean score in children referred with throat disorders was 30.2 (95 per cent confidence interval (CI) 27.8–32.7), with a range of one to 60. The distribution is shown in Figure 1.

The mean score in the unaffected children was 0.78 (95 per cent CI 0.04–1.50), the median score was zero and the range was zero to 11. The 14-item Paediatric Throat Disorders Outcome Test discriminated well between children known to suffer with throat problems and a group of healthy controls ($p < 0.0001$, $t = 24.016$). The disease effect size was very large, 1.58 SD, corresponding to an overlap between these two groups of only seven cases out of a total of 166 (with a cut-off score of 10 points giving a sensitivity of 94.4 per cent and a specificity of 97.5 per cent).

Thirteen children failed to meet the Scottish Intercollegiate Guidelines Network criteria, as they had only infrequent episodes of sore throat, a symptom duration of less than a year, or no history or supportive findings of OSA. The mean score of children who were listed for surgical intervention was higher than that of those selected for watchful waiting ($p = 0.0003$; $t = -3.69$), and the effect size was again substantial, on the boundary of 'large' (0.78 SD).²⁴

Lower scores were reported for children considered for tonsillectomy for obstructive symptoms, compared with those with recurrent acute tonsillitis, while the highest pre-operative scores were reported for children with both conditions (one-way ANOVA, $f = 23.71$, two degrees of freedom, $n = 113$; $p < 0.0001$).

Responsiveness at six months

Of the 52 questionnaires sent out to parents six months after their children had undergone tonsillectomy or adeno-tonsillectomy, 36 were returned. The mean 14-item Paediatric Throat Disorders Outcome Test score in these children was 34.9 pre-operatively (95 per cent CI 29.0–40.8) and 9.0 post-operatively (95 per cent CI 1.68–16.32). The paired t -test showed a statistically significant decrease in parent-reported 14-item Paediatric Throat Disorders Outcome Test scores six months after surgical intervention ($p < 0.0001$; $t = 7.01$). The effect size (i.e. change in mean divided by baseline SD) for children of parents completing the post-operative questionnaires was 1.53; this is very large.

There were few additional symptoms volunteered by parents in the questionnaire free text box. Parents reported problems with vomiting (two responses), high temperatures (three), cough

TABLE I
ROTATED FACTOR LOADINGS FOR THE TWO QUESTIONNAIRE VERSIONS

Item	Rotated factor loadings for T-16		Rotated factor loadings for T-14	
	Factor 1	Factor 2	Factor 1	Factor 2
Snoring	0.01	0.68	0.01	0.69
Apnoea	-0.02	0.68	-0.02	0.66
Visits to GP	0.76	0.08	0.76	0.08
Phone to GP	0.61	0.09	0.61	0.10
Antibiotics <2 wks	0.84	0.16	0.83	0.16
Antibiotics >2 wks	0.64	0.17	0.64	0.19
Ear infections	0.42	0.21	0.42	0.24
Infections <2 wks	0.70	0.00	0.70	-0.01
Infections >2 wks	0.61	0.31	0.61	0.31
Mouth breathing	0.12	0.81	0.12	0.82
Noisy breathing	0.13	0.79	0.14	0.81
Poor weight gain	0.27	0.37	Deleted	Deleted
Problems eating	0.21	0.44	0.19	0.37
Poor behaviour	0.18	0.36	Deleted	Deleted
Disturbed sleep	0.21	0.46	0.21	0.45
Missing school	0.71	-0.09	0.70	-0.10

T-16 = 16-item prototype questionnaire; T-14 = 14-item Paediatric Throat Disorders Outcome Test; GP = general practitioner; wks = weeks

(four), drooling (two) delayed speech (two) and eczema (one). The relatively low level of additional comments in any symptom area (maximum four of 126; 3 per cent) suggests that all important aspects of the disease were already addressed by the 14-item Paediatric Throat Disorders Outcome Test (and hence by its source questionnaire, the Tonsil and Adenoid Health Status Instrument). However, the free text response box was retained in the final instrument for three reasons: parent involvement, openness to further improvement once a large database was available, and notification of clinical symptoms possibly relevant to diagnosis and management of the individual child.

Discussion

Validity of 14-item Paediatric Throat Disorders Outcome Test

Validity comes in many types and stages, and accrues with use. Concurrent (criterion) validity can be

relevant to short forms but only when there is an accepted long form with which to correlate scores. This was not applicable for the 14-item Paediatric Throat Disorders Outcome Test, so other types of validity had to be pursued. Here, the initial construct validity justifying further use came from the large case-control difference and the ready interpretation of the two-factor solution. We demonstrated good internal consistency for both the total score and the factor scores. Test-retest reliability was high. Known-group comparisons demonstrated discriminant and construct types of validity. Children with symptomatic throat disorders had higher reported symptom scores than healthy school children. Of the children who were referred, those meeting clinical guidelines as candidates for tonsillectomy (either in terms of infection frequency or obstructive symptom severity) had higher reported scores than those who did not. Finally, as expected, children with both infectious and obstructive indications for tonsillectomy had higher symptom scores than children with only one indication. Responsiveness was also demonstrated.

The original Tonsil and Adenoid Health Status Instrument was subjected to a number of tests of validity. The initial 37 items were reduced to a final 18. These were subjected to factor analysis (yielding four subscales), and assessment of test-retest reliability and internal consistency. Construct validity was determined by: correlating scores with a global child health instrument; known-group comparisons; and comparing subscale scores with antibiotic use and polysomnography results. Our results indicated that, in these areas, the 14-item Paediatric Throat Disorders Outcome Test compares favourably with the Tonsil and Adenoid Health Status Instrument.

Clinicians recognise two main indications for childhood tonsillectomy: recurrent sore throat attributed to infection, and airway obstruction due to adeno-tonsillar hypertrophy. Here, eight questionnaire items load coherently onto an 'infection' factor: frequent medical visits, frequent phone calls

TABLE II

14-ITEM PAEDIATRIC THROAT DISORDERS OUTCOME TEST SCORES BY KEY GROUPS

Group	Pts (n)	Score (mean (95% CI))	SD
Healthy pts	40	0.78 (0.04-1.5)	2.3
Pts referred with throat disorders	126	30.2 (27.8-32.7)	14.8
<i>Management plan</i>			
Pts undergoing surgery	113	31.9 (29.4-34.3)	14.3
Pts undergoing 'watchful waiting'	13	17.3 (10.4-24.1)	11.3
<i>Diagnostic category</i>			
OSA	25	20.0 (15.6-24.6)	11.1
Recurrent acute tonsillitis	47	29.9 (26.5-33.4)	11.7
OSA & recurrent acute tonsillitis	41	39.8 (36.1-43.5)	11.7

Pts = patients; CI = confidence interval; SD = standard deviation; OSA = obstructive sleep apnoea

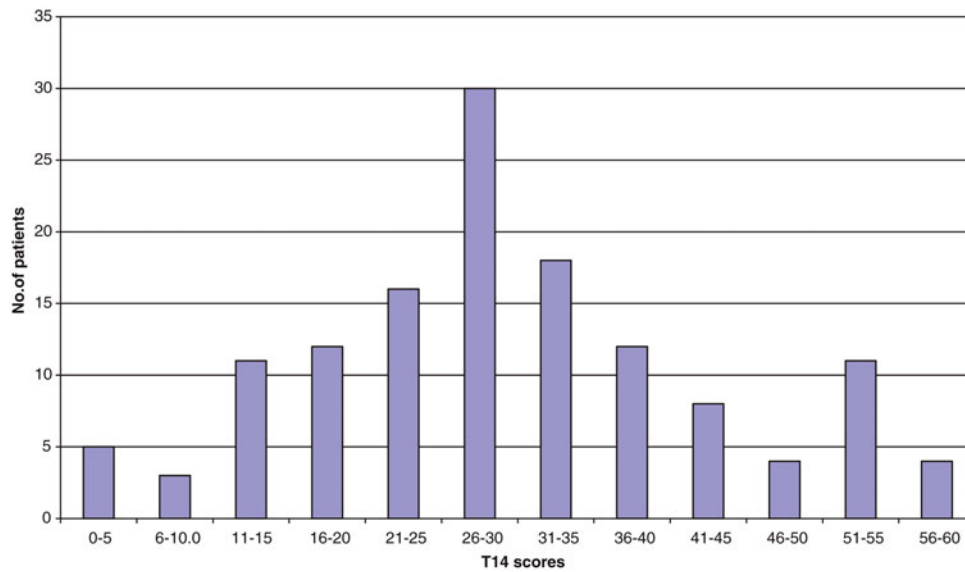


FIG. 1

Distribution of T14 Scores (skew 0.18, kurtosis 2.52).

for medical advice, antibiotics for less than two weeks, antibiotics for more than two weeks, frequent ear infections, short term throat infections, chronic throat infections and frequent school absence. The remaining six questionnaire items are similarly coherent, but relate to airway obstruction: snoring, apnoea, mouth breathing, noisy breathing, problems with eating and daytime sleepiness. These are exactly the items that load onto the second, 'obstruction' factor. The basis for the separation of factors is that recurrent infection can occur independently of tonsil size, while large tonsils can cause a separate set of obstructive symptoms. Patients may require referral for either infection or obstructive indications. The underlying aetiological variation is independent enough to dissociate the two sets of signs and symptoms and so allow two factors to emerge. A tendency to obstruction may only become relevant if there is some infection. Nevertheless, the sum of all these items (as given in the total score and justified by the first unrotated component) is also meaningful as a gross severity metric. Thus, the 14-item Paediatric Throat Disorders Outcome Test may be used as a single score, and/or the two subscores may be calculated, and repeat multiple measures of either can be produced throughout the treatment course.

Clinical utilisation

The 14-item Paediatric Throat Disorders Outcome Test is intended to measure the severity of throat disorders affecting the QoL of afflicted children, and also to act as an outcome measure. Its purpose is not primarily diagnostic, and any use of cut-off values on (sub)scores to guide management decisions would have to await randomised evidence on what best predicts benefit. For the present, we believe the Scottish Intercollegiate Guidelines Network guidelines should continue to be used to guide decisions as to

whether surgical intervention is warranted in recurrent acute tonsillitis.

It was not the intention of this small study to evaluate the effectiveness of tonsillectomy, but rather to confirm the suitability of the 14-item Paediatric Throat Disorders Outcome Test, so as to facilitate such research. However, our preliminary data in a small group of patients suggests a large benefit, compared with the 'Tonsillitis Outcomes: Toward Reaching Evidence in Adults and Tots' study. We do not know how much of the benefit is due to the surgical intervention and how much may be due to increased age and conservative management. Now that the validity of the 14-item Paediatric Throat Disorders Outcome Test has been established, it is for larger, better-controlled studies to establish whether this is generalisable and explicable in part by the more highly selected caseloads seen in UK secondary care.

Necessary scope limitations for 14-item Paediatric Throat Disorders Outcome Test: recommendations for additional objective measurement of weight and height

The parents of our clinical group frequently mentioned problems with weight gain and eating, and improvement in these after treatment. It is therefore interesting that the potential third factor involved eating and weight gain. As a further factor (e.g. weight and appetite) could not be supported within such a short instrument at conventional reliability and validity, the item regarding weight gain was deleted from the original 16-item prototype questionnaire. The development of short questionnaire forms often involves such restriction of scope. Subjective reports are inexpensive yet clinically 'rich', but balanced assessment must also seek objectivity where this is feasible. We therefore recommend that children referred for tonsillitis and airway

obstruction be weighed and their weight considered in relation to their age and height. These measurements are very simple to schedule and inexpensive to make once the system is committed to using them. In this way, a low-cost, robust, objective and bias-free measure is available to complement the present outcome measure(s) before and after intervention. Adding questions to support a third factor would have added to the burden of response, potentially discouraging completeness of data.

- **The use of patient-reported outcome measures is rapidly growing in studies of clinical effectiveness and quality of care**
- **The National Institute for Health and Clinical Excellence was requested to issue guidance on ineffective treatments, and reported a lack of evidence to support tonsillectomy**
- **This study describes a modification of a pre-existing research instrument, the Tonsil and Adenoid Health Status Instrument, to make it suitable for rapid completion as a disease-specific health-related quality of life instrument in children with tonsil and adenoid disease in the UK**
- **The resulting 14-item Paediatric Throat Disorders Outcome Test is an appropriate, disease-specific, parent-reported outcome measure for children with throat disorders, for which this study demonstrated internal consistency, reliability, responsiveness to change and two forms of construct validity**

Limitations of proxy reporting

As with the Tonsil and Adenoid Health Status Instrument, the 14-item Paediatric Throat Disorders Outcome Test relies on parental rating of the impact of throat disorders on their child's QoL. When the patient is able to report his or her own QoL this is usually preferred, as information provided by proxy-respondents has been shown to differ to that reported by patients themselves. Imperfect agreement or cross-informant variance has been demonstrated in parent-reported QoL measurements for children with chronic health conditions and for healthy children.²⁵ However, many conditions, such as tonsil and adenoid disorders, affect children across a wide age range, and many will lack the insight and scale concepts required to reliably report their own health-related QoL. Moreover, it is usually the caregiver who seeks medical intervention. It is therefore appropriate to use proxy-respondents in this setting.

Application in research

An instrument with only six to eight items per construct is not an ideal research tool; the necessary brevity for the main intended applications in clinical audit and as a patient-reported outcome measure entails limited reliability. However, research

needing to cover throat disorders somewhat (but not as a sole variable of interest) can build on the present development, and use the 14-item Paediatric Throat Disorders Outcome Test as an undemanding supplementary measure, calling upon the reference data (such as limits of normal data) now likely to become available with more widespread use of this questionnaire. On a similar basis, we intend to collect further data ourselves in the quantity required to develop a more refined quantitative scoring system at the item level. Such item scaling would help to maximise the discriminative gradation and hence the potential of the two subscores, but based on only six to eight items. However, we believe that the total score from the present version of the 14-item Paediatric Throat Disorders Outcome Test, with equal item weighting and a priori scoring, offers sufficient gradation for simple and robust use as a patient-reported outcome measure, or as one score among other measures.

Conclusion

The 14-item Paediatric Throat Disorders Outcome Test questionnaire is an appropriate, disease-specific, parent-reported outcome measure for children with throat disorders, for which we have demonstrated internal consistency, reliability, responsiveness to change and two forms of construct validity.

Acknowledgement

We gratefully acknowledge Dr Michael Stewart for his work in developing the Tonsil and Adenoid Health Status Instrument, and his kind permission to modify the instrument for use in the UK population.

References

- 1 Darzi A. *High Quality Care for All: NHS Next Stage Review, Final Report*. London: Department of Health, 2008
- 2 Timmins N. NHS goes to the PROMs. *BMJ* 2008;**336**: 1464–5
- 3 Little P, Williamson I. Sore throat management in general practice. *Fam Pract* 1996;**13**:317–21
- 4 Main operations: 4 character tables. <http://www.hesonline.nhs.uk> [13 March 2009]
- 5 Blair RL, McKerrow WS, Carter NW, Fenton A. *The Scottish Tonsillectomy Audit*. Dundee: *JLO*, 1996;**110**:1–25
- 6 White C. Chief Medical Officer calls for rewards to improve clinical practice. *BMJ* 2006;**333**:216
- 7 Stewart MG, Friedman E, Sulek M, Hulka GF, Kupper-smith RB, Harrill WC *et al*. Quality of life and health status in pediatric tonsil and adenoid disease *Arch Otolaryngol Head Neck Surg* 2000;**126**:45–8
- 8 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 Eng J Med* 1984;**310**:674–83
- 9 McKee WJ. A controlled study of the effects of tonsillectomy and adenoidectomy in children. *J Br Soc Prev Med* 1963;**17**:49–69
- 10 Mawson SR, Adlington P, Evans M. A controlled study evaluation of adeno-tonsillectomy in children. *J Laryngol Otol* 1967;**81**:777–90
- 11 Roydhouse N. A controlled study of adenotonsillectomy. *Arch Otolaryngol* 1970;**92**:611–16
- 12 Burton MJ, Glasziou PP. Tonsillectomy or adenotonsillectomy versus non-surgical treatment for chronic/

- recurrent acute tonsillitis. *Cochrane Database Syst Rev* 2008;**4**:1291–1326
- 13 Schwentner I, Schmutzhard J, Schwentner C, Abraham I, Höfer S, Sprinzel GM. The impact of adenotonsillectomy on children's quality of life. *Clin Otolaryngol* 2008;**33**:56–9
- 14 Goldstein NA, Stewart MG, Witsell DL, Hannley MT, Weaver EM, Yueh B *et al.* Quality of life outcomes after tonsillectomy – results from the TO TREAT study. *Otolaryngol Head Neck Surg* 2005;**133**:109–10
- 15 Bond J, Wilson J, Eccles M, Vanoli A, Steen N, Clarke R *et al.* Protocol for north of England and Scotland study for tonsillectomy and adeno-tonsillectomy in children (NESSTAC). *BMC - Ear, Nose and Throat Disorders, Biomed Central* 2006;**6**:13
- 16 National Prospective Tonsillectomy Audit. Royal College of Surgeons of England. 2005. London. ISBN 1-904096-02-06
- 17 Stewart M, Friedman E, Sulek M. Validation of an outcomes instrument for tonsil and adenoid disease. *Arch Otolaryngol Head Neck Surg* 2001;**127**:29–35
- 18 Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989;**27**:S217–32
- 19 *Management of Sore Throat and Indications for Tonsillectomy SIGN Publication No. 34.* 1999
- 20 Guyatt GH, Norman GR. A critical look at transition ratings. *J Clin Epidemiol* 2002;**55**:900–8
- 21 Ware JE, Kosinski M, Dewey JE. *How to Score Version of SF-36 Health Survey.* Lincoln, Rhode Island: Quality-Metric Incorporate, 2000
- 22 Roth PL, Switzer FS 3rd, Switzer D. Missing data in multi-item scales: a Monte Carlo analysis of missing data techniques. *Organizational Research Methods* 1999;**2**:211–32
- 23 Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;**16**:297–306
- 24 Cohen J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. New Jersey: Lawrence Erlbaum, 1988
- 25 Levi RB, Drotar D. Health-related quality of life in childhood cancer: discrepancy in parent–child reports. *Int J Cancer* 1999;**12**:58–64

Address for correspondence:
Ms Claire Hopkins,
Carmay, Chelsfield Lane,
Orpington BR6 7RR, UK.

E-mail: clairehopkins@yahoo.com

Ms C Hopkins takes responsibility for the integrity
of the content of the paper.
Competing interests: None declared

Appendix 1. 14-Item Paediatric Throat Disorders Outcome Test

Below you will find a list of symptoms and problems that may be caused by your child's throat problems. We would like to know more about these, and would appreciate your answering the following questions to the best of your ability. There are no right or wrong answers, and only you can provide us with this information.

Date of completion:

Considering how severe the problem is when your child experiences it and how often it happens over the past six months for your child, please rate each item below on how 'bad' it is by circling the number that corresponds with how you feel using this scale. If a certain question is not a problem for your child, please circle '0'. Please try not to miss any questions.

Item	No problem	Very mild problem	Mild or slight problem	Moderate problem	Severe problem	Problem as bad as it could be
Snoring loudly during sleep	0	1	2	3	4	5
Irregular or stopped breathing (apnoea) during sleep	0	1	2	3	4	5
Many visits to the family doctor or A&E department	0	1	2	3	4	5
Many phone calls to the doctor or NHS Direct	0	1	2	3	4	5
Taking antibiotics over and over for less than 2 weeks at a time	0	1	2	3	4	5
Taking antibiotics for more than 2 weeks straight	0	1	2	3	4	5
Frequent ear ache or ear infections	0	1	2	3	4	5
Repeated short-term throat infections that last less than 2 weeks	0	1	2	3	4	5
Constant, or chronic, throat infections that last more than 2 weeks	0	1	2	3	4	5
Breathing through the mouth during the day	0	1	2	3	4	5
Noisy breathing during the day	0	1	2	3	4	5
Problems with poor appetite, or poor eating habits (choking on food etc)	0	1	2	3	4	5
Missing school days due to sore throats	0	1	2	3	4	5
Daytime sleepiness	0	1	2	3	4	5

Please enter any further important symptoms that occur as a result of your child's throat problems that we have missed from the list above, and give each a rating from 0 to 5 like the ones already listed: