Exploring the factor structure of the Glasgow Children's Benefit Inventory (GCBI): new recommendations for reporting results

Haytham Kubba MBBS MPhil MD FRCS

Paediatric Otolaryngology

Royal Hospital for Children

1345 Govan Road

Glasgow G51 4TF

Scotland

William M Whitmer PhD

Hearing Sciences (Scottish Section)

University of Nottingham

Glasgow Royal Infirmary

10- 16 Alexandra Parade

Glasgow G31 2ER

Scotland

Correspondence to Mr Kubba at:

E: hkubba@nhs.net

T: +44 141 4516598

ABSTRACT

Objective

Patient-reported outcomes can be useful for reporting benefit from non-life-saving interventions, but often they report a single overall score, which means that much information on the specific areas of benefit is lost. Our aim was to perform a new factor analysis on the Glasgow Children's Benefit Inventory (GCBI) to create subscales reflecting domains of benefit. Further aims were to assess the internal consistency of the GCBI, and to develop guidelines for reporting both a total score and sub-scales in future studies.

Methods

We collected four existing datasets of GCBI data from children who have undergone tonsillectomy, ventilation tube insertion, pinnaplasty and submucous diathermy to the inferior turbinates. We performed exploratory factor analysis with principal axis factoring with varimax rotation, we sought redundancy in question items, and we measured internal consistency.

Results

Using the combined dataset of 772 cases, we found four factors which accounted for 64% of the variance and which we have labelled "Psycho-social", "Physical health", "Behaviour" and "Vitality". Subscale results varied in predictable ways depending on the nature of the intervention. Cronbach's alpha was 0.928. Item-total correlations were high, and no item could be deleted to improve alpha. Floor effects were apparent for various questions but were not consistent between different interventions.

Conclusions

The GCBI contains a range of questions which each add value in different clinical interventions. We can now make recommendations for reporting the results of the GCBI and its four new subscales.

Keywords

Quality of life; Outcomes research; Children; Paediatrics; Tonsillectomy; Ventilation tubes; Pinnaplasty; Turbinate surgery

INTRODUCTION

Patient-reported outcome measures have become increasingly popular in recent years as clinicians have sought to demonstrate the benefits of their treatments in ways that are meaningful to patients, themselves and healthcare administrators, and which go beyond crude measures such as complication rates and deaths.
The Glasgow Children's Benefit Inventory was developed in 2004 [1] and has been used in at least 50 clinical studies since (unpublished data: review of 85 Medline citations, January 2020). It is generic in scope, with no reference to any specific disease process, body part or symptom. It is designed for use in children of any age, and to be used retrospectively to rate benefit (or harm) after any medical intervention or treatment has occurred. It was designed to be sensitive to change after an intervention, something which can be difficult to measure with traditional before-and-after questionnaire studies [2], where the differences between individuals are often larger than any effect of the intervention.

The original study began with a review of the published literature, a review of existing instruments and some semi-structured parent interviews to identify a list of potential items for inclusion in the instrument. A draft instrument was piloted on a small number of families, and the instrument modified in line with parent feedback. The final instrument was then posted to the parents of 1,777 children who had undergone tonsillectomy or ventilation tube insertion over the preceding 3 years. The instrument score correlated with parental satisfaction, and objective measures of technical success (residual sore throats, ear infections and reported hearing difficulties).

It has been used primarily in otolaryngology, including studies on such diverse topics as correction of prominent ears [3], turbinate surgery [4], cochlear implantation [5], bone anchored hearing aids [6], tonsillectomy [7], and drooling of saliva [8], but has also been used in studies of constipation and faecal incontinence

[9,10], robot-assisted pyeloplasty [11] and secondary vaginoplasty for disorders of sexual development [12]. It has been translated into German [13-16], Italian [17], Dutch [18,19], Spanish [11], Portuguese [20], Swedish [21], Greek [22], Turkish [23], Russian [24] and Mandarin Chinese[25].

Although there is some value in reporting a total score from a measure such as the GCBI, it is possible to provide more detailed information on the specific areas of patient benefit by reporting sub-scales. These can be produced using subsets of questions that relate to specific areas of benefit. In the original description of the GCBI, a statistical technique called principal component analysis (PCA) with varimax rotation was used to derive four factor-based scores: Emotion, Physical Health, Learning, and Vitality. These factors accounted for 62% of the total variance.

The verification and validation of questionnaires has evolved a lot over recent decades and there is agreement that PCA may not be the best way to uncover the complex, hidden pattern of how the question responses group together. When the aim is to reveal the sub-groups of questions that tend to correlate with each other in some way there are better, more robust techniques available (for example, exploratory factor analysis). We cannot be sure, therefore, that the published factor structure for the GCBI is correct. Given the frequency of use of the GCBI in research and clinical contexts, it is critical that the psychometric properties of the instrument are sound.

In addition, no guidance was given as to how the instrument should be used and reported. As a result, published articles vary widely as to how the GCBI data are presented (e.g., total scores given as median and range vs mean and standard deviation, scores given as averages for individual questions with no total score, factors used in some but not all studies).

The aim of this study was to use existing datasets from the original validation of the GCBI[1] and from other clinical studies on pinnaplasty [3] and submucous diathermy of the inferior turbinates [4] to perform and then test a new factor analysis. By using data from studies of very different clinical interventions, we can investigate the stability of the factor structure in different conditions. Further aims of the study were to assess the internal consistency of the item/factor structure, and to develop guidelines for reporting the GCBI in future studies using both a total score and sub-scales.

METHODS

Ethical considerations

This study was performed using existing data from three previous studies, each of which had approval from their local research ethics committee. As no new data was generated and all data were anonymised, no new ethical issues arose for this study and no new ethical committee approval was required. All three of these studies were conducted in our department so the full datasets were already held in secure storage and available to us.

Datasets

1. Tonsillectomy and ventilation tube insertion

The original 2004 study [1] describing the GCBI was done using the completed questionnaires from a cohort of 670 children aged 1-15 years (median 6) of whom 452 had undergone tonsillectomy and 218 had undergone ventilation tube insertion between 2 and 5 years previously

2. Submucous diathermy of the inferior turbinates (SMDT)

The GCBI was used by Montgomery in 2011 [4] in a study of 47 children aged 3-14 years (median 8) who had undergone SMDT 35-75 months previously as a treatment for rhinitis. 35 children (76%) underwent SMDT as their only procedure, while 11 had concomitant procedures including tonsillectomy (12, 25%), insertion of ventilation tubes (9, 19%) and adenoidectomy (1, 2%).

3. Pinnaplasty

The GCBI was used by Fraser in 2016 [3] in a study of 91 children undergoing cosmetic pinnaplasty for prominent ears. They were aged 4-16 years (median 12) and they completed the GCBI 3-69 months after surgery (median 37 months).

Statistical methods

Appendix 1 contains a glossary of technical terms in psychometrics relating to factor analysis and the detailed analysis of questionnaire data.

Sensitivity was investigated by looking for question items where at least 75% of participants answered "no change" (suggesting a question item that is uninformative) within each of the four datasets (tonsillectomy, ventilation tubes, pinnaplasty and SMDT). All four datasets combined were used to test for internal consistency with item-total correlations and Cronbach's alpha.

Exploratory analysis was done on the combined dataset using principal axis factoring with varimax rotation and Kaiser normalisation. We selected for factors with an eigenvalue greater than 1 and for question items with a factor loading of at least 0.4 [26]. Alternative analyses using other factor extraction methods (such as unweighted least squares) and non-orthogonal (correlated) rotations (such as oblimin and promax) were no more informative, in that they produced similar factors but explained less of the overall variance.

Finally, subscale scores were calculated for each of the four clinical datasets to investigate their usefulness in distinguishing areas of benefit from differing interventions.

RESULTS

Internal consistency

Considering the combined dataset of all four clinical conditions, there were 772 complete GCBI records with no missing answers, out of a total of 808 (95.5%). Cronbach's alpha for the internal consistency of the 24 items was very high at 0.928. None of the 24 question items could be deleted to improve Cronbach's alpha further. Item-total correlations are shown in **Table 1**.

Item relevance

Question items in a change questionnaire have no value if everyone responds "no change". We therefore sought out question items where 75% or more of responses were the same within each dataset.

For the ventilation tubes dataset (n=215), frequent "no change" responses were found for q5 (liveliness: 75.9%), q7 (food: 77.0%), q8 (self-conscious: 76.4), q11 (embarrassment: 81.0%) and q20 (self-care: 89.8%).

For the tonsillectomy dataset (n=450), frequent "no change" responses were found for q3 (behaviour: 77.8%), q8 (self-conscious: 87.8%), q9 (family harmony: 88.7%), q11 (embarrassment: 91.8%), q12 (easily distracted: 85.2%), q15 (concentration: 77.2%), q17 (self-esteem: 76.1%), q19 (confidence: 81.4%) and q20 (self-care: 91.1%).

For the pinnaplasty dataset (n=91), frequent "no change" responses were found for q3 (behaviour: 75.8%), q6 (sleep: 83.5%), q7 (food: 92.3%), q9 (family harmony: 83.5%), q12 (easily distracted: 87.9%), q13 (learning: 83.5%), q14 (time off school: 83.5%), q15 (concentration: 85.7%), q20 (self-care: 83.5%), q22 (catches colds: 87.9%), q23 (visits to doctor: 85.7%) and q24 (need for medication: 91.2%).

Interestingly, there was also a frequent positive response for q1 (overall life) with 76.9% reporting "much better".

For the SMDT dataset (n=47), frequent "no change" responses were found for q9 (family harmony: 85.1%), q12 (easily distracted: 76.6%) and q20 (self-care: 76.6%).

Question 20 (self-care) was the only item which had a response of "no change" for at least 75% of respondents in all four datasets.

Factor analysis

Using the complete combined dataset of 772 cases, a principal axis factoring with varimax rotation produced four factors which accounted for 64% of the variance. The Kaiser-Meyer-Olkin measure of sampling adequacy [27] was satisfactory at 0.930 as was Bartlett's test of sphericity (p<0.001). The items loading on to each factor are shown in **Table 2** and a scree plot is shown in **Figure 1**. There were between 5 and 8 items loading onto each factor. Question 20 (self-care) did not adequately load onto any of the factors. We have chosen to name the four factors "Psycho-social", "Physical health", "Behaviour" and "Vitality" based on item content and their resemblance to the original and translated factors [1,13,20].

Subscale scores

On the basis of the factor loadings shown in **Table 2**, we calculated subscale scores by adding the scores for each of the question items listed and re-scaling on the same -100 to +100 scale as the total score:

- Psycho-social subscale: questions 8, 11, 17, 18, 19
- Physical health subscale: 1, 14, 22, 23, 24
- Behaviour subscale: 3, 4, 9, 12, 13, 15, 16
- Vitality subscale: 2, 5, 6, 7, 10, 21

The results for the total score and these subscale scores are shown for each of the four clinical datasets in **Figure 2** and **Table 3**. Although the magnitude and range of overall benefit shown for each of the interventions is similar, it is clear that pinnaplasty has a large benefit in the Psycho-social subscale, while tonsillectomy and ventilation tubes show greatest benefit in the Physical health subscale.

DISCUSSION

It may seem to a clinician that this is just obscure statistical analysis with esoteric discussion of eigenvalues and orthogonal rotations, but the value of this study is in allowing us to capture more detail about how children derive benefit from our interventions. A quality-of-life outcome measure such as the GCBI can produce a single, overall score that is easy for everyone to understand, on a scale from -100 (maximum harm) to +100 (maximum benefit) with zero as no change. This overall score has been considered useful for comparative studies. However, the starting point for calculating that score is a set of 24 questions about various aspects of day-to-day functioning, and all that rich detail is lost in a single overall score. By calculating statistically appropriate sub-scales we can present some of that detail and provide a summary of benefit in key areas in a way that is much more informative for the clinician.

The interventions studied here are very different in their intentions (preventing recurrent infection, improving hearing, improving breathing and changing cosmesis). We feel that this is one of the strengths of our study, because by including children of varying ages and with varying pathologies, we are better able to demonstrate the wider generalisability of the GCBI as a measure. The range of interventions studied means that we cannot expect that they will all produce "benefit" in the same way. Even though their overall benefit, as shown in the total GCBI score, is similar, it is the subscales that show the intuitive differences in how these treatments affect patients, such as tonsillectomy having its largest effect on physical health and pinnaplasty having its largest effect on the psychosocial.

The subscales are made up of question items that seem to make sense as coherent concepts (such as Psycho-social and Physical health) but it is important to note that they are derived mathematically in an abstract way from the data.

This process (factor analysis) shows the underlying domains of benefit that we can't see directly but which reflect the way questions group together statistically. Also shown here is the high degree to which all the questions correlate together, known as internal consistency and measured by Cronbach's alpha. The high value for alpha shows that it is reasonable to interpret an overall score as relating to a single over-arching concept which is coherent mathematically and interpretable clinically as "benefit".

The subscales reported here are broadly similar but do differ to some degree from those in the original description of the GCBI [1] and from the factor analyses reported in translated versions of the GCBI [12,19]. The method used here is more appropriate to an exploratory analysis with the aim of discovering the underlying structure of the data (principal axis factoring rather than the principal component analysis used in the original study). In addition, the analysis reported in this current study uses data from children undergoing a broader range of interventions which may have benefit in different domains not represented in the original study. Furthermore, the original description of the GCBI did not include any instructions on how to calculate or present the results of the subscales and therefore they have been little used in research studies to date. For these reasons, we believe the current study is more informative.

We have examined the sensitivity of the question items making up the GCBI for the first time. It is tempting to remove all questions where many people answer "no change" on the basis that they are not generating useful information. Care must be taken when a question shows such an effect in a study on one clinical condition, however: the same question may be very pertinent to a different clinical condition. We can see this in the questions about embarrassment (q11), self-esteem (q17) and self-consciousness (q8) which have frequent responses of "no change" after tonsillectomy but which are very pertinent after cosmetic pinnaplasty. We have considered deleting question 20, "Has your child's operation affected his/her ability to care for himself/herself as well as you think they should, such as washing, dressing and using the toilet?" on the basis that it has a response of "no change" in the four conditions studied here and that it does not load significantly onto any of the four factors. We should note that, while this question may not be pertinent to otolaryngological conditions, it may be very pertinent in other health conditions. The GCBI has found use in a variety of general surgical and urological conditions [8-11] including faecal incontinence, and at least one of these studies has shown non-neutral responses to question 20 [11]. For that reason (and because the extra burden of one question is small, and its effect on the total score negligible) we provisionally recommend keeping the question in place until further data become available on its usefulness in measuring outcomes outwith otolaryngology.

CONCLUSIONS: RECOMMENDATIONS FOR SCORING AND REPORTING THE RESULTS OF THE GCBI

1. The main score should be calculated by taking the mean of the 24 question item scores and re-scaling on a range -100 (maximum harm) to +100 (maximum benefit) with zero as no change.
2. Four subscale scores should be calculated the same way, using the following subsets of questions, each also re-scaled on a range of -100 to +100:
   - Psycho-social subscale: questions 8, 11, 17, 18, 19
   - Physical health subscale: 1, 14, 22, 23, 24
   - Behaviour subscale: 3, 4, 9, 12, 13, 15, 16
   - Vitality subscale: 2, 5, 6, 7, 10, 21
3. The data are rarely normally distributed, so it is not appropriate or meaningful to report only the mean and standard deviation. The median and range, as text and/or boxplot, are easiest to interpret and most appropriate given the expected positive skew of the data. These should be used for the total GCBI score and the four subscale scores (Psychosocial, Physical health, Behaviour and Vitality).

ACKNOWLEDGEMENTS

The authors are very grateful for Ms Lyndsay Fraser and Ms May Yaneza-Parr for providing the data on pinnaplasty and submucous diathermy of the turbinates.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. WMW was supported by the Medical Research Council [grant number MR/S003576/1]; and the Chief Scientist Office of the Scottish Government.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest

DATA ACCESS

The primary data are available from the corresponding author on reasonable request.

REFERENCES

1. Kubba H, Swan IRC, Gatehouse S. The Glasgow Children's Benefit Inventory: a new instrument to assess changes in children's quality of life after an intervention Annals of Otology Rhinology and Laryngology 2004; 113(12): 980-6
2. Gatehouse S (1999) Glasgow Hearing Aid Benefit Profile: Derivation and validation of a client-centered outcome measure for hearing aid services. *Journal of the American Academy of Audiology* 1999; 10: 80-103.
3. Fraser L, James HL, Wynne DM, MacGregor FB, Morrissey MSC, Clement WA, Kubba H Validation of a family-centred outcome questionnaire for pinnaplasty: a cross-sectional pilot study Clinical Otolaryngology 2016; 41(5): 472-480
4. Montgomery JA, Sadiq HA, Kubba H Long-term follow-up of children after submucosal diathermy to the inferior turbinate for rhinitis. *International Journal of Pediatric Otorhinolaryngology* 2011; 75(3): 387-90
5. Speaker RB, Robertson J, Simoes-Franklin C, Glynn F, Walshe P, Viani L. Quality of life outcomes in cochlear implantation of children with profound and multiple learning disability, Cochlear Implants International, 2018; 19(3): 162-166.
6. McDermott AL, Williams J, Kuo M, Reid A, Proops D. Quality of Life in Children Fitted With a Bone-Anchored Hearing Aid. Otology & Neurotology. 30(3):344-349, 2009
7. Wood JM, Harris PK, Woods CM, McLean SC, Esterman A, Carney AS. Quality of life following surgery for sleep disordered breathing: subtotal reduction adenotonsillectomy versus adenotonsillectomy in Australian children. ANZ Journal of Surgery. 2011; 81(5): 340-4
8. Chanu NP, Sahni JK, Aneja S, Naglot S Four-duct ligation in children with drooling American Journal of Otolaryngology. 2012; 33(5) 604-7

9. Lu PL, Koppen IJN, Orsagh-Yentis DK, et al. Sacral nerve stimulation for constipation and fecal incontinence in children: Long- term outcomes, patient benefit, and parent satisfaction. Neurogastroenterol Motil. 2018;30:e13184.
10. Strine AC, Misseri R, Szymanski KM, Kaefer M, Rhee AC, Hillier K, Rink RC, Cain MP. Assessing health related benefit after reconstruction for urinary and fecal incontinence in children: a parental perspective. The Journal of Urology, 2015; 193: 2073–2078.
11. Freilich DA, Penna FJ, Nelson CP, Retik AB, Nguyen HT. Parental Satisfaction After Open Versus Robot Assisted Laparoscopic Pyeloplasty: Results From Modified Glasgow Children's Benefit Inventory Survey. Journal of Urology 2010; 183(2): 704-708
12. Willihnganz-Lawson KH, Isharwal S, Lewis JM, Sarafoglou K, Boisclair-Fahey A, Shukla AR Secondary vaginoplasty for disorders for sexual differentiation: Is there a right time? Challenges with compliance and follow-up at a multidisciplinary center. Journal of Pediatric Urology 2013: 9: 627-633
13. Schwentner I, Schwentner C, Schmutzhard J, Radmayr C, Grabher G, Sprinzl G, Höfer S. Validation of the German Glasgow children's benefit inventory. Journal of Evaluation in Clinical Practice 2007; 13(6): 942-946
14. Schwentner I, Schmutzhard J, Schwentner C, Abraham I, Höfer S, Sprinzl GM. The impact of. adenotonsillectomy on children's quality of life. Clinical Otolaryngology. 2008; 33(1): 56-59
15. Papadopulos NA; Niehaus R; Keller E; Henrich G; Papadopoulos ON; Staudenmaier R; Kovacs L; Herschbach P; Pototschnig H; Machens H-G. The Psychologic and Psychosocial Impact of Otoplasty on Children and Adults Journal of Craniofacial Surgery. 26(8):2309-2314, 2015
16. Braun T, Berghaus A. Ohranlegung und Lebensqualität HNO (2015) 63: 28
17. Marsella P, Scorpecci A, Vallarino MV, Di Fiore S, Pacifico C. Sophono in Pediatric Patients: The Experience of an Italian Tertiary Care Center. Otolaryngology-Head and Neck Surgery 2014; 151 (2): 328-332

18. Nelissen R, Mylanus EA, Cremers CWR, Hol MK, Snik AF. Long-term Compliance and Satisfaction With Percutaneous Bone Conduction Devices in Patients With Congenital Unilateral Conductive Hearing Loss. Otology & Neurotology. 36(5):826–833, 2015
19. Coppens-Hofman MC, Koch HH, Maassen BAM & Snik AFM. Evaluating the subjective benefit of hearing rehabilitation in adults with intellectual disability, Hearing, Balance and Communication, 2013; 11(1): 24-29
20. Dias-Vaz M, Morgado H, Severo M Estevão-Costa J. Measuring otoplasty outcome: Expanding the validity to caregivers' perspective and to Portuguese-speaking children *Clinical Otolaryngology*.2018;43:1513–1521
21. Ericsson E, Lundeborg I, Hultcrantz E. Child behavior and quality of life before and after tonsillotomy versus tonsillectomy *International Journal of Pediatric Otorhinolaryngology*. 2009; 73(9); 1254-1262
22. Fettaa M, Tsilisb NS, Segasc JV, Nikolopoulos TP, Vlastarakos PV. Functional endoscopic sinus surgery improves the quality of life in children suffering from chronic rhinosinusitis with nasal polyps. *International Journal of Pediatric Otorhinolaryngology* 100 (2017) 145-148
23. Toplu Y, Sapmaz E, Firat C et al. Clinical results and health-related quality of life in otoplasty patients using cartilage resection and suturing methods *Eur Arch Otorhinolaryngol* (2014) 271: 3147
24. Mileshina NA, Osipenkov SS, Bakhshinyan VV, Tavartkiladze GA. The impact of implantable technologies on the quality of life in the patients with hearing loss *Vestnik Otorinolaringologii* 2016; 6: 22-24
25. Fan X, Wang Y, Wang P, Fan Y, Chen Y, Zhu Y, Chen X. Aesthetic and hearing rehabilitation in patients with bilateral microtia-atresia. *International Journal of Pediatric Otorhinolaryngology* 2017; 101: 150-7
26. Stevens JP *Applied multivariate statistics for the social sciences* (2nd edition). Hillsdale, NJ: Erlbaum & Associates. 1992
27. Kaiser HL, Rice J. Little jiffy, mark IV. *Educational and Psychological Measurement* 1974; 34(1): 111-117.

APPENDIX 1: EXPLANATION OF TERMS

Item-total correlation

The extent to which a particular question varies across individual responses with the average of the rest of the question items in a questionnaire, as one would expect if all the question items refer in some way to a single underlying concept (such as "benefit" from a medical intervention)

Cronbach's alpha

A single, overall measure of the extent to which the responses to question items in a questionnaire correlate with each other. Low scores (0.6 or less) indicate poor internal consistency, while higher scores suggest that internal consistency is adequate (0.7 or higher), good (0.8 or higher) or excellent (0.9 or higher). Item-total correlation and Cronbach's alpha are related, such that removing items with poor item-total correlation tends to increase alpha.

Factor analysis

A method to simplify data, taking a large number of questions and grouping them into a smaller number of *factors*. The hypothesis is that there are a number of hidden, latent variables within the data which explain the pattern of responses to the individual question items. Analysing the correlations between the individual question items reveals the latent factors which can then be given names for convenience. As an example, in this study we believe that, within the overall concept of "benefit", there is a factor which we have chosen to call "psycho-social" which relates to self-esteem, embarrassment and self-consciousness. Interventions which target this specific area of benefit, such as cosmetic pinnaplasty, drive positive answers to certain questions within the GCBI that we have grouped together as this Psychosocial sub-scale.

Principle component analysis (PCA)

A means of reducing a large number of variables to a smaller number of variables. It is superficially similar to factor analysis: both aim to discover which question items form coherent sub-groups that are relatively independent of each other. In many cases the results of PCA and factor analysis are similar. However, the underlying aim of each is very different. PCA combines the question items in a linear way to produce a smaller number of components with the aim of explaining as much of the overall variance as possible. Factor analysis, on the other hand, is used to understand the latent constructs within the data by examining the correlations between the question items: the factors can then be combined in a linear way to understand the original question item responses.

Principle axis factoring

A method of factor analysis which aims to find the smallest number of factors accounting for the common variance of a set of variables. While there are many alternative techniques for exploratory factor analysis, this is one of the most commonly used as it makes fewer assumptions about the data being normally distributed.

Orthogonal and oblique rotations

Factors are described as *oblique* if they correlate with each other and *orthogonal* if they do not. When performing exploratory factor analysis and PCA, part of the process involves "rotating" the potential factor solutions to find the best fit to the data. Various methods for rotating the data exist, each with advantages and disadvantages. *Varimax* rotation tends to produce a small number of orthogonal factors with each question item loading onto one factor at most, by maximising the sum of the variances of the squared loadings. However, it is often the case in medicine and the social sciences that factors may correlate with each other to some degree, since they will all reflect aspects of one over-arching concept such as "benefit". It is often appropriate, therefore, to use oblique rotations such as *Promax* or *Oblimin* for exploratory factor analysis. In practice, if factors are sufficiently robust they will appear in a similar way regardless of the technique used.

Eigenvalue

The amount of variance explained by a factor is called its *eigenvalue*. If the factor explains more of the variance than the question item on its own then the eigenvalue will be greater than 1 and the factor is likely to be of use.

Scree plot

A graphical representation of the eigenvalues for each of the factors identified in an exploratory factor analysis. The factors in the initial steeper part of the curve are the most important to retain as they explain the majority of the variance.

Factor loading

The factor loading is the degree to which the factor influences the response to a particular question item, expressed as the correlation coefficient between the item response and the factor. Factor loadings closer to 1 indicate that the underlying factor is a very strong driver of the response to the question item, whereas factor loadings closer to zero indicate that the effect is weak. Items that load strongly onto a particular factor can be grouped together to create a sub-scale score.

Kaiser-Meyer-Olkin measure of sampling adequacy

This is an estimate of the proportion of variance that might be explained by underlying factors: a value approaching 1 indicates that factor analysis might be useful.

Bartlett's test of sphericity

This tests the hypothesis that all the variables in the dataset are completely unrelated and therefore unsuitable for factor analysis: p-values less than 0.05 indicate that the data are suitable for factor analysis.

TABLE 1

Internal consistency analysis of the GCBI using the complete dataset from all four clinical conditions combined (n=772). Cronbach's alpha was 0.928.

GCBI question item	Item-total correlation	Cronbach's alpha if item deleted
q1 - overall life	.590	.924
q2 - things they do	.635	.924
q3 - behaviour	.570	.925
q4 - progress & development	.634	.924
q5 - liveliness	.659	.923
q6 - sleep	.573	.925
q7 - food	.515	.926
q8 - self consciousness	.479	.926
q9 - family harmony	.587	.925
q10 - fun with friends	.662	.923
q11 - embarrassment	.436	.927
q12 - easily distracted	.563	.925
q13 - learning	.597	.924
q14 - time off school	.500	.926
q15 - concentration	.611	.924
q16 - irritability	.689	.923
q17 - self esteem	.570	.925
q18 - happiness	.673	.923
q19 - confidence	.527	.925
q20 - self care	.486	.926
q21 - leisure	.621	.924
q22 - catches colds	.500	.926
q23 - visits to doctor	.530	.926
q24 - need for medication	.533	.926

TABLE 2

Factor loadings for the 24 items of the GCBI. Factor analysis was performed using principal axis factoring with varimax rotation, selecting for factors with an eigenvalue greater than 1. Factor loadings of at least 0.4 are shown.

	Factor			
	1	2	3	4
	"Psycho-social"	"Physical health"	"Behaviour"	"Vitality"
q1 - overall life		.518		
q2 - things they do				.443
q3 - behaviour			.529	
q4 - progress & development			.532	
q5 - liveliness				.694
q6 - sleep		.401		.515
q7 - food				.511
q8 - self consciousness	.807			
q9 - family harmony			.598	
q10 - fun with friends				.508
q11 - embarrassment	.805			
q12 - easily distracted			.726	
q13 - learning			.593	
q14 - time off school		.707		
q15 - concentration			.745	
q16 - irritability			.489	
q17 - self esteem	.785			
q18 - happiness	.556			
q19 - confidence	.765			
q20 - self care				
q21 - leisure				.498
q22 - catches colds		.753		
q23 - visits to doctor		.917		

q24 - need for medication		.893		

TABLE 3

Total GCBI score and subscale scores for the four clinical datasets (tonsillectomy, ventilation tubes, pinnaplasty and SMDT, each given as the median score (range in brackets).

Intervention	N	Total	Psycho-social	Physical health	Behaviour	Vitality
Tonsil	452	+31.3 (-43.8 to +100)	+10 (-70 to +100)	+90 (-50 to +100)	+7.1 (-100 to +100)	+33.3 (-41.7 to +100)
Vent tubes	216	+22.9 (-27.1 to +100)	+10 (-50 to +100)	+60 (-40 to +100)	+21.4 (-21.4 to +100)	+16.7 (-25 to +100)
Pinnaplasty	91	+25.0 (-25.0 to +100)	+80 (-100 to +100)	+20 (-40 to +100)	+7.1 (-14.3 to +100)	+16.7 (-16.7 to +100)
SMDT	47	+18.8 (-64.6 to +93.8)	+10 (-60 to +100)	+30 (-90 to +100)	0 (-50 to +92.9)	+16.7 (-58.3 to +100)

FIGURE 1

Scree plot showing the eigenvalues of the factors extracted in the exploratory factor analysis. The majority of the variance is in the first 2 factors, with 4 factors having an eigenvalue greater than 1.

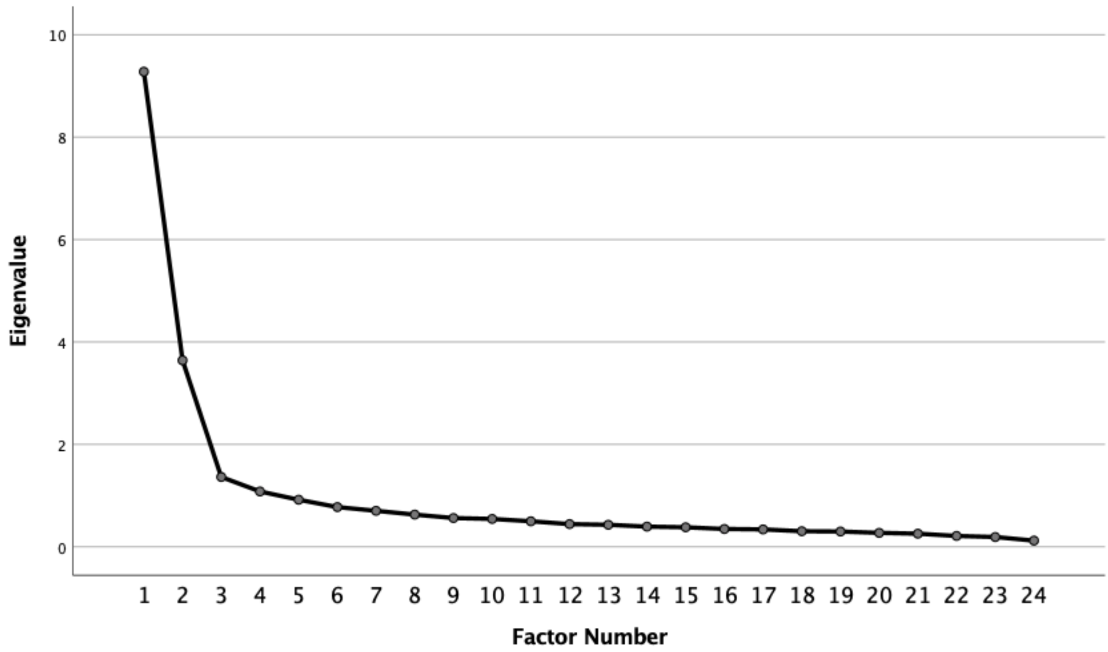

FIGURE 2

Total GCBI score and subscale scores for each of the four clinical datasets studied (tonsillectomy 450 cases, ventilation tubes 215 cases, pinnaplasty 91 cases and SMDT 47 cases). All are presented on a scale from -100 (maximum harm) to +100 (maximum benefit) with zero being no change. The boxplots show the median as the thick horizontal line in the centre of the box, with the interquartile range as the edges of the box. The whiskers show 1.5 times the interquartile range. Outliers beyond 1.5 times the interquartile range are shown as circles and extreme values beyond 2 times the interquartile range are shown as asterisks.