Measuring voice outcomes: state of the science review

The Journal of Laryngology & Otology (2009), ***123***, *823–829.*

doi:10.1017/S0022215109005398

Review Article

Measuring voice outcomes: state of the science review

P N Carding, J A Wilson, K MacKenzie*, I J Deary†

Abstract

Researchers evaluating voice disorder interventions currently have a plethora of voice outcome measurement tools from which to choose. Faced with such a wide choice, it would be beneficial to establish a clear rationale to guide selection. This article reviews the published literature on the three main areas of voice outcome assessment: (1) perceptual rating of voice quality, (2) acoustic measurement of the speech signal and (3) patient self-reporting of voice problems. We analysed the published reliability, validity, sensitivity to change and utility of the common outcome measurement tools in each area. From the data, we suggest that routine voice outcome measurement should include (1) an expert rating of voice quality (using the Grade-Roughness-Breathiness-Asthenia-Strain rating scale) and (2) a short self-reporting tool (either the Vocal Performance Questionnaire or the Vocal Handicap Index 10). These measures have high validity, the best reported reliability to date, good sensitivity to change data and excellent utility ratings. However, their application and administration require attention to detail. Acoustic measurement has arguable validity and poor reliability data at the present time. Other areas of voice outcome measurement (e.g. stroboscopy and aerodynamic phonatory measurements) require similarly detailed research and analysis.

Key words: Dysphonia; Voice Outcomes; Voice Disorders; Voice Quality; Voice Handicap; Self-Reported Voice Measures

Introduction

The rise of evidence-based medicine and 'payment by results' has driven a need for sophisticated and robust clinical outcome data in all areas of medicine. Outcome measurement tools need to have established reliability, validity, sensitivity to change and utility in order to be clinically useful and to enable confident assessment of a disorder as it changes following intervention.[1] 'Reliability' refers to the internal consistency and stability of the tool,[2] free from random error[3] or unwanted variation.[4] 'Validity' is concerned with the relevance of a tool or the extent to which the instrument measures what it purports to measure.[2,3] If the reliability of a measure can be seen as its trustworthiness, then validity can be thought of as its truthfulness.[5] 'Sensitivity to change' refers to an instrument's responsiveness and ability to detect clinically important changes.[6] 'Utility' is a measure of the ease of use of a tool for both the clinician and patient. Aspects such as patient discomfort and inconvenience (e.g. the time required to complete the task) are also important here.

Several areas of voice outcome measurement have been subjected to systematic international research over the past decade. Although different voice disorders may be treated with different types of intervention (e.g. pharmacological, surgical, behavioural, mechanical or psychological, or a combination), similar voice outcome measurements can be applied to all situations.

This review concentrates on three areas of voice outcome measurement that have been subjected to extensive research: (1) perceptual rating of voice quality, (2) acoustic measurement of the speech signal and (3) patient self-reporting of voice problems. The outcome measurement tools for each area are discussed with respect to their reliability, validity, sensitivity to change and utility. It is clear that there are a number of other areas of voice outcome measurement which require similarly detailed research – for example, endoscopic laryngeal interpretation (including stroboscopy) and aerodynamic phonatory measurements. However, to date no such data have been published.

Perceptual rating of voice quality

Auditory perceptual rating of voice quality involves an expert listener judging a voice sample according to various vocal parameters, and (in most cases)

From the Department of Otolaryngology–Head and Neck Surgery, Freeman Hospital, Newcastle upon Tyne, England, the *Department of Otolaryngology–Head and Neck Surgery, Royal Infirmary, Glasgow, and the †Medical Research Council Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, Scotland, UK.
Accepted for publication: 7 January 2009. First published online 20 May 2009.

marking the extent to which the voice deviates from a perceived 'normal' range.[7] Perceptual voice quality rating is considered by all voice clinicians to be an essential outcome measure.[1] There are a number of formal voice quality rating scales available. Three of the most commonly used scales in the UK are the Buffalo Voice Profile,[8] the Vocal Profile Analysis scheme[9] and the Grade-Roughness-Breathiness-Asthenia-Strain scale.[10] An additional rating scale has recently emerged – the Consensus of Auditory Perceptual Evaluation Voice scale, which incorporates the parameters of grade, roughness, breathiness, asthenia and strain, and also allows for additional dimensions to be added.[11]

Reliability

Several studies have established good reliability for the Grade-Roughness-Breathiness-Asthenia-Strain scale in the hands of expert users.[12–14] To our knowledge, there have not been any studies examining the reliability of the Consensus of Auditory Perceptual Evaluation Voice scale, the most recent perceptual voice quality rating scale suggested by the American Speech and Hearing Association. Webb *et al.* have provided the only evidence for the comparative internal consistency, repeatability and reliability of three commonly used voice quality rating scales.[15] Webb and colleagues' study was conducted under optimal rating conditions using seven highly experienced voice clinicians. A judgement of 'overall' voice severity was the most robust rated parameter in terms of inter- and intra-rater reliability (with reliability coefficients of 0.78 and 0.81, respectively). The Grade-Roughness-Breathiness-Asthenia-Strain scale was reliable across all parameters (inter-rater reliability coefficients ranged from 0.68 to 0.70 and intra-rater reliability coefficients from 0.69 to 0.79) except strain (with an inter-rater reliability coefficient of 0.48). Almost all of the component parameters of the Buffalo Voice Profile and Vocal Profile Analysis scales were found to have either poor or moderate reliability (i.e. below 0.50).

Validity

Perceptual voice rating has strong content validity, since most patients seek help for a voice disorder based on the sound of their voice. In addition, improvement in voice quality is the outcome by which interventions are judged to be successful.[1,7] The criterion validity of perceptual voice rating (i.e. does it measure what it purports to measure?) has been demonstrated by highly significant correlations between Grade-Roughness-Breathiness-Asthenia-Strain scale ratings and self-perception and self-reporting scale scores.[16] In this particular study, the strongest correlation (Spearman's correlation coefficient of 0.32) was between the 'overall' grade of voice severity and the Vocal Performance Questionnaire total score (see below).

Sensitivity to change

To date, there has only been one quantitative study assessing the responsiveness to change (following intervention) of auditory perceptual voice quality ratings. Steen *et al.*[17] compared effect sizes of the component parameters of the Grade-Roughness-Breathiness-Asthenia-Strain scale in a cohort of 144 patients following voice therapy and phonosurgery. For subjects undergoing voice therapy, there were significant, small-to-medium effect sizes. All of the Grade-Roughness-Breathiness-Asthenia-Strain scale parameters except roughness showed moderate effect sizes (the standard deviation (SD) ranged from 0.32 to 0.57). Roughness ratings generally showed less responsiveness to change following either voice therapy or surgery (effect size SDs ranged from 0.16 to 0.29).

Utility

Perceptual voice evaluation can be quick to perform and succinct, and the results easily communicable between clinicians. It is also non-invasive, readily available and can be performed 'live' in the clinic. However, when undertaking external validation, voice samples should be recorded using high quality recording equipment (preferably in a soundproof room). It is important to note that the task requires highly trained clinicians in order to be performed adequately. A review of practice amongst UK experts in voice perception analysis concluded that the absolute minimum requirement for observer voice assessment in the clinical setting was the use of the Grade-Roughness-Breathiness-Asthenia-Strain scale.[7]

Acoustic measures of voice quality

Acoustic analysis of the voice signal involves computerised measurement of specific properties of the sound waveform as produced by the patient. For the purposes of voice outcome measurement, the three most commonly used acoustic parameters are 'jitter' (i.e. cycle-to-cycle frequency perturbation), 'shimmer' (i.e. cycle-to-cycle amplitude perturbation) and harmonics to noise ratio (an expression of aperiodic to periodic sound). In most published papers, these parameters are measured during 'steady state' vowel production.

Reliability

Steady state acoustic vowel analysis has been reported to have only moderate reliability, for both intra- and inter-system comparisons and repeated measures (i.e. within-subject) analysis.[18–20] Carding *et al.* studied a group of dysphonic patients and found that test–retest reliability (i.e. stability) coefficients were at best moderate for jitter (0.45 (95 per cent confidence interval (CI) = 0.23–0.70)) and shimmer (0.40 (95 per cent CI = 0.18–0.67)) and lower for harmonics to noise ratio (0.33 (95 per cent CI = 0.11–0.63)).[21] The intra-class correlation coefficient for reliability improved when acoustic analysis was performed on non-dysphonic or near-normal (i.e. type one) voice signals (jitter = 0.73 (95 per cent CI = 0.58–0.85), shimmer = 0.55 (95 per cent CI = 0.35–0.74) and harmonics to noise

ratio = 0.68 (95 per cent CI = 0.51–0.82)). This, however, emphasises the limited clinical application of these techniques at the present time.

Validity

Dysphonia may be defined as the degree of aperiodic sound produced by the sound source (i.e. the vibrating vocal folds).[18–21] Therefore, it may be argued that analysis of the periodicity of the sound signal may have high content validity. However, acoustic measurements of this type are only valid when applied to signals with sufficient periodic structure.[22] This could mean that at least 20 per cent of patients within a typical voice pathology population may not be analysable in this way.[21] Criterion validity has not been clearly established. Some authors (e.g. Rabinov *et al.*) have suggested that a close correlation exists between specific parameters and certain perceptual voice quality features; however, others (e.g. Carding *et al.*) have reported a less convincing and highly complex correlation.[20,21] Furthermore, many authors have debated the validity of steady state vowel analysis for the purposes of voice outcome measurement, and have argued for a more representative measure of connected speech.[23] However, more complex speech signals are inherently more difficult to analyse, and data are sparse.

Sensitivity to change

There is limited information on the comparative sensitivity of acoustic voice analysis parameters for measuring voice change. Carding *et al.* found poor-to-moderate effect sizes when assessing the sensitivity of such parameters in detecting change following treatment.[21] Following surgery, the effect sizes (SD) for this assessment were: jitter = 0.32, shimmer = 0.28 and harmonics to noise ratio = 0.34; those following voice therapy were: jitter = 0.47, shimmer = 0.34 and harmonics to noise ratio = 0.32.

Utility

Good reliability of acoustic measurement is difficult to achieve in moderately dysphonic (aperiodic) voices and is of very limited value in cases of severely dysphonic voice. The process of acquiring and analysing the speech sound signal is time-consuming (approximately one hour per patient) and requires considerable voice laboratory expertise.

Patient self-reporting

There are a number of voice-specific patient self-reporting tools reported in the literature. Most of the research activity over the past decade has concentrated on examining the Voice Handicap Index, the Vocal Performance Questionnaire and the Voice Symptom Scale.[24–26]

Reliability

Several studies have examined the comparative reliability of the Vocal Performance Questionnaire, the Voice Handicap Index and the Voice Symptom Scale.[16,17] In summary, based on assessment of 181 patients presenting with dysphonia, all three assessment tools provided excellent internal consistency (Cronbach's coefficient = 0.81–0.95) and repeatability (intra-class correlation coefficients: Voice Handicap Index total = 0.83, Vocal Performance Questionnaire = 0.75 and Voice Symptom Scale total = 0.63). For baseline measures, therefore, criteria other than reliability should direct the selection of self-reporting tools.

Validity

Patient self-reporting has high content validity since, unless patients are satisfied with their own voice, little can claim to have been achieved in treatment.

Patient self-reporting also offers an opportunity to obtain information about vocal handicap and disability, in addition to aspects of vocal quality. Furthermore, many dysphonic patients have a widely fluctuating disorder (e.g. worse at the end of the working day or the working week). Therefore, the voice that is presented to the clinician in the voice clinic may well not be representative of the overall voice performance.[27] Self-reporting tools allow the patient to give an overall voice rating, as opposed to one based solely on vocal performance on the day of consultation.[27]

Criterion validity is more difficult to prove. A central problem with many historic self-reporting tools has been the physician-centred nature of their derivation. Both the Voice Handicap Index and the Vocal Performance Questionnaire suffer from this limitation.

In this respect, the Voice Symptom Scale is considerably superior to all previous voice self-reporting tools, with 800 subjects participating in the final development of the tool.[26,28] Criterion validity is also affected by the internal component structure of the self-reporting tool. Psychometric analysis of the 800 subjects' Voice Symptom Scale responses showed three distinct subscales: impairment (15 items), emotional response (eight items) and physical symptoms (seven items).[28]

In contrast, Rosen *et al.* assessed the Voice Handicap Index and found a lack of statistically discreet subscales.[29] Further factor analysis of the Voice Handicap Index subscales revealed that only a single factor was being measured. For these reasons, a shorter, 10-item Voice Handicap Index was proposed.

Sensitivity to change

Several published studies have analysed the sensitivity of self-reporting assessment tools for measuring change following intervention.[16,17] Again, it would appear that the Vocal Performance Questionnaire, Voice Handicap Index and Voice Symptom Scale all show large effect sizes as regards sensitivity to change following either voice therapy (SD results being 1.04, 0.62 and 0.78, respectively) or surgery (SD results being 0.82, 0.72 and 1.06, respectively). In terms of sensitivity to change, the ability of the

Vocal Performance Questionnaire to demonstrate a treatment effect size of more than one (i.e. equal to the Voice Symptom Scale and somewhat higher than the Voice Handicap Index) is an impressive result for a short, 12-item questionnaire.

Utility

Deary *et al.* compared the Voice Handicap Index 10 (i.e. the shorter, 10-item version) with the Vocal Performance Questionnaire.[30] Both were found to be similar, being short, convenient, internally consistent, uni-dimensional tools used to measure the severity of a voice disorder. Furthermore, Rosen *et al.* concluded that there was no benefit to using the full (30-item) version of the Voice Handicap Index rather than the shortened Voice Handicap Index 10.[29] The use of an extended questionnaire (with a considerable risk of item redundancy) would appear to be required only for very specific reasons and requirements. In this latter case, it would appear that the Voice Symptom Scale may be most useful, since it has three discreet subscales.[28]

Discussion

When measuring outcomes, the aim is to document significant change – i.e. change that is neither random nor unimportant.[31] The established opinion is that voice outcome measurement should be multidimensional in nature.[1] We have analysed the evidence base for three common types of voice outcome measurement tools: voice quality perceptual rating, acoustic measurement of the speech signal and patient self-reporting. We suggest that the selection of voice outcome measurement tool should be based on considerations of reliability, validity, sensitivity to change and utility. Whilst our research only extended into three areas of voice assessment, we would anticipate that a similar approach to the analysis of other tools (such as laryngeal endoscopy and stroboscopy, and aerodynamic phonatory measurement) may also yield valuable clinical information.

From our research findings, we recommend that routine voice outcome measurement should include (1) an expert rating of voice quality (probably using the Grade-Roughness-Breathiness-Asthenia-Strain scale) and (2) a short self-reporting tool (either the Vocal Performance Questionnaire or the Voice Handicap Index 10). These measures have high validity, the best reported reliability to date, good sensitivity to change and excellent utility ratings. These instruments are therefore likely to provide high quality outcome information irrespective of whether the treatment choice is phonosurgery, voice therapy, pharmacological therapy or a combination of several approaches.

The obvious limitation is that, in a clinical setting, expert rating of voice quality will probably be carried out by the treating clinician. We should remember that published studies relate only to blinded, controlled, independent evaluation of voice quality by an expert rater. The effect of clinical bias and the performance of less expert raters have not been fully examined.

However, this should not prevent clinicians from applying these measures in routine practice in order to determine the effectiveness of their treatments. Furthermore, with respect to voice quality ratings, we should not forget that clinician ratings may not always correlate with patient perceptions of their own voice quality scores.[32]

The Voice Symptom Scale is certainly worth considering if a more detailed patient self-evaluation is required. The advantage of the Voice Symptom Scale over the Vocal Performance Questionnaire or the Voice Handicap Index 10 is that it includes a physical symptoms subscale.[28,30] However, whilst information on these physical symptoms may be interesting to obtain, physical symptom subscale results do not seem to correlate with vocal outcome nearly as closely as different voice measures correlate with each other. A review of the impact of surgery according to the Voice Symptom Scale impairment subscale showed an effect size of one, with a corresponding Voice Symptom Scale emotional subscale of 0.69 but a physical symptoms subscale response of only 0.43.[17] This result is perhaps predictable and indeed may be welcomed, as it suggests that these subscales may be obtaining information on an area of dysfunction which conventional strategies have yet to adequately address. Both the Voice Symptom Scale and the Vocal Performance Questionnaire are detailed in the appendices of this article.

Acoustic analysis of the speech signal would currently appear to have a limited clinical role. Reliability may be enhanced by recording and analysing multiple voice samples and averaging the results, but this is at the expense of utility.[22] Perturbation measurements of selected vowel prolongations may be greatly enhanced by following a strict recording protocol.[33] However, the value of this approach in measuring clinically useful change has yet to be established. Acoustic analysis of connected speech still appears to be in its infancy.

In a research context, there is no doubt that multidimensional analysis is best. Where high quality evidence exists, we should use it to guide our selection of the most robust voice outcome measures. However, limiting our data to that obtained by these tools only would be to the long term detriment of the development of knowledge in this area. For example, the general positive benefits of being a patient in a clinical trial mean that it would be very unwise to interpret research findings on the basis of self-reporting measures alone, however reliable they appear on statistical analysis. Clinical outcome data from laryngeal endoscopy, aerodynamic phonatory measurement and psychological impact assessment may all yield valuable data. It is however clear that these measures require considerable further attention, particularly with respect to reliability and sensitivity to change.

References

1 Carding PN. *Evaluating Voice Therapy: Evaluating the Effectiveness of Treatment*. London: Whurr, 2000

2 Anthony DM. *Understanding Advanced Statistics; a Guide for Nurses and Healthcare Researchers*. London: Churchill Livingstone, 1999
3 Armitage P, Berry G. *Statistical Methods in Medical Research*. Oxford: Blackwell Scientific Publications, 1994
4 Cronbach L. *Essentials of Psychological Testing*. London: Harper & Row, 1970
5 Schuavetti N, Metz D. *Evaluating Research in Communicative Disorders*. Boston: Allyn and Bacon, 1997
6 Fayers P, Machin D. *Quality of Life Assessment, Analysis and Interpretation*. Chichester: John Wiley & Sons, 2000
7 Carding PN, Carlson E, Epstein R, Mathieson I, Shewell C. Formal perceptual evaluation of voice quality in the United Kingdom. *Log Phon Vocol* 2000;**25**:133–8
8 Wilson D. *Voice Problems of Children*, 3rd edn. Baltimore: Williams and Wilkins, 1987
9 Laver J. *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press, 1980
10 Hirano M. *Clinical Examination of Voice*. New York: Springer-Verlag, 1981
11 Consensus of Auditory Perceptual Evaluation Voice 2001. In: http://www.asha.org [07 02 2002]
12 Dejonckere PH, Obbens C, Leeper HA, Hawkins S, Heeneman H, Doyle PC. Perceptual evaluation of dysphonia: reliability and relevance. *Folia Phoniat Logop* 1993;**45**:76–83
13 De Bodt M, Wuyts FL, Van de Heyning PH, Croux C. Test-retest of the GRBAS Scale: influence of experience and professional background on perceptual ratings of voice quality. *J Voice* 1997;**11**:74–80
14 Wuyts FL, De Bodt MS, Van de Heyning PH. Is the reliability of a visual analog scale higher than an ordinal scale? An experiment with the GRBAS Scale for the perceptual evaluation of dysphonia. *J Voice* 1999;**13**:508–17
15 Webb A, Carding P, Deary IJ, MacKenzie K, Steen IN, Wilson JA. A study of the reliability of three auditory perceptual scales for dysphonia. *Eur Arch Otorhinolaryngol* 2004;**261**:429–34
16 Webb AL, Carding PN, Deary IJ, MacKenzie K, Steen IN, Wilson JA. Optimising outcome assessment of voice interventions, I: reliability and validity of three self-reported scales. *J Laryng Otol* 2007;**121**:763–7
17 Steen IN, Webb Al, Deary IJ, MacKenzie K, Carding PN, Wilson JA. Optimising outcome assessment of voice interventions II: the sensitivity to change of self-report and observer rated measures. *J Laryng Otol* 2008;**122**:45–51
18 Gonzalez J, Cervera T, Miralles JL. Acoustic voice analysis; reliability of a set of multidimensional parameters [in Spanish]. *Acto Otorhino Espan* 2002;**53**:256–68
19 Bough D, Heur RJ, Sataloff RT, Hills JR, Carter JR. Intrasubject variability of objective voice measures. *J Voice* 1996;**10**:166–74
20 Rabinov RC, Kreiman J, Gerrart BR, Bielamonwicz S. Comparing reliability of perceptual ratings of roughness and acoustic measures of jitter. *J Speech Hear Res* 1995; **38**:26–32
21 Carding PN, Steen IN, Webb A, MacKenzie K, Deary IJ, Wilson JA. The reliability and sensitivity to change of acoustic measures of voice quality. *Clin Otol* 2004;**29**:538–44
22 Titze I. Workshop on Acoustic Voice Analysis: summary statement. Iowa, Iowa: The University of Iowa, 1995
23 Kania RE, Hartl DM, Hans S, Maeda S, Vaissiere J, Brasnu DF. Fundamental frequency histograms measured by electroglottography during speech: a pilot study for standardization. *J Voice* 2006;**20**:18–24
24 Jacobson BH, Johnson A, Grywalski C. The Voice Handicap Inventory (VHI): development and validation. *Am J Speech Lang Path* 1997;**6**:66–70
25 Carding PN, Horsley IA. An evaluation study of voice therapy in non-organic dysphonia. *Eur J Disord Commun* 1992;**27**:137–58
26 Deary IJ, Wilson JA, Carding PN, MacKenzie K. VoiSS: a patient derived voice symptom scale. *J Psychosom Res* 2003;**54**:483–9
27 Jones SM, Carding PN, Drinnan MJ. Exploring the relationship between severity of dysphonia and voice-related quality of life. *Clinical Otolaryngology*;**31**:411–17
28 Wilson JA, Webb AL, Carding PN, Steen N, MacKenzie K, Deary IJ. Comparing the Voice Symptom Scale (VoiSS) and the Voice Handicap Index: structure and content. *Otol Clin* 2004;**29**:169–74
29 Rosen CA, Lees AS, Osborne J, Zullo T, Murray T. Development and validation of the Voice Handicap Index-10. *Laryngoscope* 2004;**9**:1549–6
30 Deary IJ, Webb A, MacKenzie K, Wilson JA, Carding PN. Short, self-report voice symptom scales: psychometric characteristics of the Voice Handicap Index-10 and the Vocal Performance Questionnaire. *Head Neck Surg* 2004; **131**:232–5
31 Olswang L. *Treatment Efficacy Research; Measuring Outcomes in Speech Language Pathology*. New York: Thieme, 1998
32 Lee M, Drinnan M, Carding PN. The reliability and validity of patient self-rating of their own voice quality. *Clin Otol* 2005;**30**:357–61
33 Brockmann M, Storck C, Carding PN, Drinnan MJ. Voice loudness and gender effects on jitter and shimmer in healthy adults. *J Speech Hear Res* 2008;**51**:1152–60

Appendix 1. Vocal Performance Questionnaire

By Paul Carding, Freeman Hospital, Newcastle upon Tyne, UK

Name . . . . . Date . . . . .

Tick or circle an answer for each question.

1 How do you think your voice sounds now (compared with before your voice problems started)?
   (a) No different from usual voice
   (b) Only slightly different from usual voice
   (c) Quite different from usual voice
   (d) Very different from usual voice
   (e) Totally different from usual voice
2 Does your voice give you any physical discomfort when you talk?
   (a) No discomfort
   (b) Slight discomfort
   (c) Moderate discomfort
   (d) A lot of discomfort
   (e) Severe discomfort
3 Does your voice get worse as you talk?
   (a) Not at all – it stays the same
   (b) Occasionally when I talk
   (c) Often gets worse when I talk
   (d) Often gets a lot worse when I talk
   (e) Always gets a lot worse when I talk
4 Do you find it an effort to talk?
   (a) No effort at all
   (b) Slight effort sometimes (i.e. at the end of the day or when talking loudly)
   (c) Quite an effort sometimes
   (d) An effort most of the time
   (e) A constant effort
5 How much are you using your voice at present?
   (a) As much as I usually would
   (b) A little less than I usually would
   (c) Somewhat less than usual
   (d) A lot less than usual
   (e) Hardly at all
6 Does your voice problem stop you from doing anything that you would otherwise normally do?
   (a) Doesn't stop me doing anything that involves me using my voice
   (b) Stops me doing a few things that involve using my voice

(c) Stops me doing a lot of things that involve using my voice
(d) Stops me doing most things that involve using my voice
(e) I can hardly do anything that involves me using my voice

7 In your opinion, do you think that your voice is ever difficult to hear or understand?
(a) Not at all
(b) A little difficult
(c) Quite difficult
(d) Very difficult
(e) Extremely difficult

8 Do other people (e.g. close family) ever comment that your voice is difficult to hear or understand?
(a) No comments
(b) Occasional comments
(c) Quite often there are comments
(d) Frequent comments
(e) Very frequent comments

9 Since your voice problem started, has your voice...
(a) Improved a lot
(b) Improved a little
(c) Not improved at all
(d) Deteriorated a little
(e) Deteriorated a lot

10 Since your voice problem started, have other people (e.g. close family) commented that your voice has improved?
(a) Other people say that my voice has improved a lot
(b) Other people say that my voice has improved a little
(c) Other people say that my voice has not improved at all
(d) Other people say that my voice has got a little worse
(e) Other people say that my voice has got a lot worse

11 Would you say that the sound of your voice was...
(a) Normal
(b) Not quite normal
(c) Mildly abnormal
(d) Quite abnormal
(e) Very abnormal

12 How much do you worry about your voice problem now?
(a) Not at all
(b) Hardly at all
(c) Quite a lot
(d) A good deal
(e) Almost all of the time

Assign a value of 1 to each (a) answer, a 2 to each (b) answer and so on.
Total range of scores is therefore 12 (normal) to 60 (very severe dysfunction).
Total score.....

Appendix 2. Voice Symptom Scale

Your name.....
Your date of birth.....
Today's date.../.../....

Please circle one answer for each item
Please do not leave any blank items

1	Do you have difficulty attracting attention?	Never	Occasionally	Some of the time	Most of the time	Always
2	Do you have problems singing?	Never	Occasionally	Some of the time	Most of the time	Always
3	Is your throat sore?	Never	Occasionally	Some of the time	Most of the time	Always
4	Is your voice hoarse?	Never	Occasionally	Some of the time	Most of the time	Always
5	When talking in company do people fail to hear you?	Never	Occasionally	Some of the time	Most of the time	Always
6	Do you lose your voice?	Never	Occasionally	Some of the time	Most of the time	Always
7	Do you cough or clear your throat?	Never	Occasionally	Some of the time	Most of the time	Always
8	Do you have a weak voice?	Never	Occasionally	Some of the time	Most of the time	Always
9	Do you have problems talking on the telephone?	Never	Occasionally	Some of the time	Most of the time	Always
10	Do you feel miserable or depressed because of your voice problem?	Never	Occasionally	Some of the time	Most of the time	Always
11	Does it feel as if there is something stuck in your throat?	Never	Occasionally	Some of the time	Most of the time	Always
12	Do you have swollen glands?	Never	Occasionally	Some of the time	Most of the time	Always
13	Are you embarrassed by your voice problem?	Never	Occasionally	Some of the time	Most of the time	Always
14	Do you find the effort of speaking tiring?	Never	Occasionally	Some of the time	Most of the time	Always
15	Does your voice problem make you feel stressed and nervous?	Never	Occasionally	Some of the time	Most of the time	Always
16	Do you have difficulty competing against background noise?	Never	Occasionally	Some of the time	Most of the time	Always
17	Are you unable to shout or raise your voice?	Never	Occasionally	Some of the time	Most of the time	Always
18	Does your voice problem put a strain on your family and friends?	Never	Occasionally	Some of the time	Most of the time	Always
19	Do you have a lot of phlegm in your throat?	Never	Occasionally	Some of the time	Most of the time	Always
20	Does the sound of your voice vary throughout the day?	Never	Occasionally	Some of the time	Most of the time	Always
21	Do people seem irritated by your voice?	Never	Occasionally	Some of the time	Most of the time	Always

(Continued)

Appendix 2. *(Continued)*

22	Do you have a blocked nose?	Never	Occasionally	Some of the time	Most of the time	Always
23	Do people ask what is wrong with your voice?	Never	Occasionally	Some of the time	Most of the time	Always
24	Does your voice sound creaky and dry?	Never	Occasionally	Some of the time	Most of the time	Always
25	Do you feel you have to strain to produce voice?	Never	Occasionally	Some of the time	Most of the time	Always
26	How often do you get throat infections?	Never	Occasionally	Some of the time	Most of the time	Always
27	Does your voice 'give out' in the middle of speaking?	Never	Occasionally	Some of the time	Most of the time	Always
28	Does your voice make you feel incompetent?	Never	Occasionally	Some of the time	Most of the time	Always
29	Are you ashamed of your voice problem?	Never	Occasionally	Some of the time	Most of the time	Always
30	Do you feel lonely because of your voice problem?	Never	Occasionally	Some of the time	Most of the time	Always

For office use:
Total Voice Symptom Scale score = ...
Impairment score (items 1, 2, 4, 5, 6, 8, 9, 14, 16, 17, 20, 23, 24, 25 & 27) (maximum 60) = ...
Emotional score (items 10, 13, 15, 18, 21, 28, 29 & 30) (maximum 32) = ...
Physical score (items 3, 7, 11, 12, 19, 22 & 26) (maximum 28) = ...

Please note that the Vocal Performance Questionnaire and Voice Symptom Scale are also available in electronic format (at http://www.entuk.org/clinical_outcomes/). This website also includes information about how to score the questionnaires, as well as several supporting publications.

Address for correspondence:
Prof Paul Carding,
Professor of Speech/Voice Pathology,
Dept of Speech, Voice and Swallowing,
Otolaryngology,
Freeman Hospital,
Newcastle upon Tyne NE7 7DN, UK.

E-mail: paul.carding@nuth.nhs.uk

Professor P Carding takes responsibility for the integrity of the content of the paper.
Competing interests: None declared