Is voice therapy an effective treatment for dysphonia? A randomised controlled trial

Citation for published version:

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
BMJ

Publisher Rights Statement:

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Is voice therapy an effective treatment for dysphonia?
A randomised controlled trial

Kenneth MacKenzie, Audrey Millar, Janet A Wilson, Cameron Sellars, Ian J Deary

Abstract

Objectives To assess the overall efficacy of voice therapy for dysphonia.
Design Single blind randomised controlled trial.
Setting Outpatient clinic in a teaching hospital.
Participants 204 outpatients aged 17-87 with a primary symptom of persistent hoarseness for at least two months.
Interventions After baseline assessments, patients were randomised to six weeks of either voice therapy or no treatment. Assessments were repeated at six weeks on the 145 (71%) patients who continued to this stage and at 12-14 weeks on the 133 (65%) patients who completed the study. The assessments at the three time points for the 70 patients who completed treatment and the 63 patients in the group given no treatment were compared.
Main outcome measures Ratings of laryngeal features, Buffalo voice profile, amplitude and pitch perturbation, voice profile questionnaire, hospital anxiety and depression scale, clinical interview schedule, SF-36.
Results Voice therapy improved voice quality as assessed by rating by patients (P = 0.001) and rating by observer (P < 0.001). The treatment effects for these two outcomes were 4.1 (95% confidence interval 1.7 to 6.6) points and 0.82 (0.50 to 1.13) points. Amplitude perturbation showed improvement at six weeks (P = 0.005) but not on completion of the study. Patients with dysphonia had appreciable psychological distress and lower quality of life than controls, but voice therapy had no significant impact on either of these variables.
Conclusion Voice therapy is effective in improving voice quality as assessed by self rated and observer rated methods.

Introduction

Many patients have transient, self limiting changes in voice, but those who have been hoarse for more than three weeks need specialist assessment to exclude underlying laryngopharyngeal pathology. Once conditions that need surgery have been excluded, patients are usually referred to a speech and language therapist for voice therapy. Up to 40 000 patients with dysphonia are referred for voice therapy annually in the United Kingdom. At the time of referral, many patients with vocal dysfunction have entered a vicious cycle in which psychological factors exacerbate voice pathology and poor voice quality adversely affects psychological wellbeing. The relation between these factors is complex, and the relative influence of each factor varies from individual to individual.

No study has yet examined the overall effectiveness of voice therapy for dysphonia in terms of either changes in voice quality or changes in psychological distress or laryngoscopic findings. We aimed to examine the efficacy of voice therapy in patients with dysphonia and to identify those patients for whom voice therapy might be most beneficial.

Participants and methods

We recruited consecutive outpatients attending the department of otorhinolaryngology and head and neck surgery of Glasgow Royal Infirmary with a primary complaint of dysphonia (hoarseness) present for a minimum of two months and without any relevant organic pathology (for example, polyp, papilloma, tumour, vocal cord palsy) or need for surgery.

The inclusion criteria were age greater than 16 years, motivation to resolve the voice problem, and willingness to enter into regular voice therapy sessions. The exclusion criteria were previously treated dysphonia, neurological disease, or upper aerodigestive tract malignancy; marked hearing impairment; acid reflux; multiple medical complaints; professional voice user unwilling to go on voice rest; patients referred for surgery; patients with persistent hoarseness of less than two months; and patients with an underlying laryngopharyngeal pathology. These conditions were chosen so that voice therapy could not influence the response to surgery.

The 204 patients (51 men, 153 women) who gave informed consent for inclusion were new referrals typical of patients referred for voice therapy. At entry to the study, 100 patients were randomised to voice therapy and 104 to no treatment. By completion of the study 12-14 weeks later, about a third of participants had dropped out or been excluded, leaving 70 patients in the treatment group and 63 patients in the observation group (figure). This attrition was mostly a result of patients defaulting. In addition to failure to reattend, some patients failed to complete all self report questionnaires, notably where these had to be returned by post, despite the issuing of prepaid envelopes. Another subgroup omitted a number of items within questionnaires. The voice therapy and no treatment groups were not significantly different in terms of
either rate of attrition (30% in the therapy group, 39% in the no treatment group) or characteristics of patients who dropped out (sociodemographic variables or baseline voice or psychological variables). There is thus no evidence that the attrition will have introduced bias. Analysis of each outcome was conducted only on patients with complete data.

**Measures**

**Pathophysiology**—An otolaryngologist (KMacK) used a flexible nasolaryngoscope (Olympus ENF 3) to assess four features—nodule formation, laryngitis, glottic escape, and hyperfunction of the laryngeal musculature—on a four point (0-3) rating scale. Although stroboscopic findings may facilitate the diagnosis of subtle laryngeal dysfunction, we decided not to use our stroboscope because we wanted the study to reflect routine clinical practice throughout the United Kingdom.

**Voice quality**—A digital tape recording of the patient’s reading of the phonetically balanced “rainbow” passage (a standard paragraph used in voice assessment) was analysed by a speech and language therapist blind to treatment group. Ten aspects of voice, including laryngeal tone, pitch, and resonance, were rated 0-5 on the Buffalo voice profile scale, of which the “overall rating” item was chosen as the key observer rated variable. The same therapist also extracted two key objective measures of voice quality—“jitter” (pitch perturbation) and “shimmer” (amplitude perturbation)—by using the Computerised Speech Laboratory (model 4300B; Kay Elemetrics Corp, NJ). The higher the score on these two variables the more dysphonic the voice. Patients rated their own voice quality using the validated vocal performance questionnaire, with five point scoring on 12 items (1-5, 5 = worst). The total voice profile questionnaire score was the key self reported outcome measure.

**Psychological measures**—The hospital anxiety and depression scale is a self completed questionnaire which assesses recent anxiety and depression and provides an overall score for each. The revised clinical interview schedule was conducted by a trained psychologist blind to the treatment group. The interviewer rated 14 aspects of non-psychotic psychiatric disturbance. The clinical interview schedule’s overall distress score and the hospital anxiety and depression scale’s anxiety score were the key outcome measures of psychological distress.

**Quality of life** was assessed by the SF-36, an eight dimension, extensively validated assessment of general health status.

**Intervention**

Baseline data were recorded after eligibility had been assessed and consent had been obtained by the laryngologist. The participants were then seen by one speech and language therapist (CS), who obtained a number for random allocation of the participant to either a course of voice therapy or a period of observation. Computer generated random numbers produced either rate of attrition (30% in the therapy group, 39% in the no treatment group) or characteristics of patients who dropped out (sociodemographic variables or baseline voice or psychological variables). There is thus no evidence that the attrition will have introduced bias. Analysis of each outcome was conducted only on patients with complete data.

**Treatment protocol**

The voice therapy group initially comprised 100 patients; eight patients were excluded from the study before therapy. These exclusions occurred for a range of reasons, most of which (for example, dementia) only emerged at the follow up appointment. Two patients withdrew from the study after the first visit. In addition to the eight patients who were excluded after allocation to the therapy group (two of whom withdrew), eight patients failed to attend the first therapy appointment. A further 10 patients defaulted during therapy. This meant that 74 patients completed a course of voice therapy, 70 of whom had complete data sets. Participants entering the treatment group of the study were given an appointment at which a routine voice history was taken. A recording of the “rainbow” passage was taken for possible reference in therapy. In line with current practice, and depending on clinical circumstances, a number of patients (11 out of the 74 who completed treatment) received advice on good vocal hygiene and optimal voice production and did not proceed to a fuller programme of voice therapy, although this option remained open to them until attendance for the second visit. The remainder of the patients (63) underwent voice therapy for up to six sessions (eight sessions in one case). The treatment sessions lasted 45-60 minutes. The number of sessions needed and the type of treatment carried out were determined heuristically, depending partly on the nature of the symptoms and partly on patients’ priorities. Treatment could be indirect—for example, involving discussion of issues of vocal hygiene or of lifestyle impinging on voice production. Equally, patients could be invited to practise techniques related to their vocal symptoms—for example, improving breath support for vocal production or altering vocal onset in favour of “softer” vocal attack. Because of the limited possibilities for contact with patients, therapy favouring attention to vocal symptoms predominated, but issues were incorporated relating to possible underlying psychological distress. Indeed, where these were clearly the focus of concern to the patient, a counselling approach was adopted.

The programme of voice therapy ended by mutual agreement between the treating therapist and the individual patient. Completion of therapy did not imply complete recovery at that time but rather a recognition that good vocal practice needed ongoing attention over a sustained period beyond the contact limits set for this project.
We performed a series of trials where they had improved, deteriorated, or stayed the same. We assigned patients to a category (0, 1, or 2) according to whether their score for each feature at the end of follow up analyses. We then assigned baseline values as the covariates for both the end of treatment and at the later follow up. Each statistical analysis compared the mean difference in treatment versus no treatment) as a between patients comparison. The baseline scores in the respective measure. We used general linear modelling (analysis of covariance) in SPSS 9/10 to perform the analysis. Statistical analyses compared the mean difference in the outcome variables between the groups with and without treatment. We conducted separate analyses at the end of treatment and at the later follow up. Each analysis took into account the baseline (entry to study) score for each of the outcome variables. We thus used an analysis of covariance procedure with group (treatment versus no treatment) as a between patients variable; we used people's baseline scores on the particular variable being compared as covariates. We used baseline values as the covariates for both the end of treatment and follow up analyses. Assessment of the effect of voice therapy on pathophysiological outcomes needed a categorical approach. We subtracted ratings for each pathophysiological feature for each patient at the end of treatment (visit 2) from those at baseline (visit 1); we also subtracted ratings for each feature at the end of follow up (visit 3) from those at visit 1. We then assigned patients to a category (0, 1, or 2) according to whether they had improved, deteriorated, or stayed the same. We performed a series of $\chi^2$ intergroup comparisons. We calculated treatment effects as mean differences at the relevant outcome (visits 2 and 3) controlled for baseline scores in the respective measure. We used general linear modelling (analysis of covariance) in SPSS 9/10 to perform the analysis. Results As expected, most patients in both groups were women; the groups were closely matched for age (table 1). Laryngeal features at study entry were similar in the intervention and control patients (either in the 204 patients originally recruited or in the 145 with repeat laryngoscopy at six weeks). Grade 2-3 (moderate to severe) scores were uncommon for all of the four features, and only minimal resolution of the abnormalities occurred between the two time points (table 1). The groups were well matched at entry to the study for subjective and objective voice variables (table 2). The treatment and no treatment groups differed at baseline only on the hospital anxiety and depression scale anxiety scores, which were significantly higher in the control group ($t = 2.67$, $P = 0.008$). This difference between the treatment and no treatment groups was evident in the original 204 randomised recruits and in the 133 patients who completed all three phases of the study. Both the treatment and no treatment groups had high baseline scores for anxiety on the hospital anxiety and depression scale compared with healthy controls from the scale’s normative reference data, though not necessarily with other otorhinolaryngology patients. The baseline SF-36 quality of life scores reflect a severely impaired quality of life in both the treatment and no treatment groups compared with other groups of patients. Statistical analysis

Statistical analyses compared the mean difference in the outcome variables between the groups with and without treatment. We conducted separate analyses at the end of treatment and at the later follow up. Each analysis took into account the baseline (entry to study) score for each of the outcome variables. We thus used an analysis of covariance procedure with group (treatment versus no treatment) as a between patients variable; we used people’s baseline scores on the particular variable being compared as covariates. We used baseline values as the covariates for both the end of treatment and follow up analyses.

Assessment of the effect of voice therapy on pathophysiological outcomes needed a categorical approach. We subtracted ratings for each pathophysiological feature for each patient at the end of treatment (visit 2) from those at baseline (visit 1); we also subtracted ratings for each feature at the end of follow up (visit 3) from those at visit 1. We then assigned patients to a category (0, 1, or 2) according to whether they had improved, deteriorated, or stayed the same. We performed a series of $\chi^2$ intergroup comparisons.

We calculated treatment effects as mean differences at the relevant outcome (visits 2 and 3) controlled for baseline scores in the respective measure. We used general linear modelling (analysis of covariance) in SPSS 9/10 to perform the analysis.

Results

As expected, most patients in both groups were women; the groups were closely matched for age (table 1). Laryngeal features at study entry were similar in the intervention and control patients (either in the 204 patients originally recruited or in the 145 with repeat laryngoscopy at six weeks). Grade 2-3 (moderate to severe) scores were uncommon for all of the four features, and only minimal resolution of the abnormalities occurred between the two time points (table 1).

The groups were well matched at entry to the study for subjective and objective voice variables (table 2). The treatment and no treatment groups differed at baseline only on the hospital anxiety and depression scale anxiety scores, which were significantly higher in the control group ($t = 2.67$, $P = 0.008$). This difference between the treatment and no treatment groups was evident in the original 204 randomised recruits and in the 133 patients who completed all three phases of the study. Both the treatment and no treatment groups had high baseline scores for anxiety on the hospital anxiety and depression scale compared with healthy controls from the scale’s normative reference data, though not necessarily with other otorhinolaryngology patients. The baseline SF-36 quality of life scores reflect a severely impaired quality of life in both the treatment and no treatment groups compared with other groups of patients.
To address the issue of dropout we re-ran the analyses including all patients with data at baseline. For patients with missing data at visit 2 or visit 3 we entered the baseline values. This makes the conservative assumption that there was no difference in outcome between patients who dropped out from the treatment and no treatment groups. On reanalyses of the sensitivity scores, treatment effects (points), and confidence intervals for each of the outcome measures, the results retained their significant \( P \) values and treatment effects (tables 4 and 5).

**Discussion**

This first randomised controlled trial of the efficacy of voice therapy for dysphonia has shown voice therapy to be effective in improving self rated and expert rated quality of voice. The magnitude of the observed mean improvements reflects clinically meaningful improvements in voice quality. The minimal change in laryngoscopic appearances during the study reflects the fact that many of the patients referred for non-surgical points (effect size 0.76 SD). In conventional statistical terminology these are medium to large effects. Voice therapy had an effect on only one quality of life outcome variable—mental health. This was significantly better in the treatment group at completion of treatment but not at completion of follow up.

<table>
<thead>
<tr>
<th>Test or questionnaire</th>
<th>No of patients</th>
<th>Treatment Visit 1 (baseline)</th>
<th>Treatment Visit 2 or 3</th>
<th>No of patients</th>
<th>No treatment Visit 1 (baseline)</th>
<th>No treatment Visit 2 or 3</th>
<th>F value</th>
<th>Treatment effect (points (95% CI))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychological distress:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HADS anxiety</td>
<td>67</td>
<td>6.3 (4.0)</td>
<td>5.9 (4.5)</td>
<td>63</td>
<td>8.1 (4.3)</td>
<td>7.7 (3.9)</td>
<td>0.71 (P=0.402)</td>
<td>0.39 (−0.53 to 1.31)</td>
</tr>
<tr>
<td>HADS depression</td>
<td>67</td>
<td>3.7 (3.5)</td>
<td>3.6 (3.7)</td>
<td>63</td>
<td>4.2 (3.3)</td>
<td>4.2 (3.1)</td>
<td>0.38 (P&lt;0.540)</td>
<td>0.25 (−0.55 to 1.05)</td>
</tr>
<tr>
<td>CIS total score</td>
<td>67</td>
<td>7.0 (6.9)</td>
<td>8.6 (8.2)</td>
<td>62</td>
<td>8.6 (8.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voice:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buffalo overall rating</td>
<td></td>
<td>74</td>
<td>2.9 (1.1)</td>
<td>69</td>
<td>2.9 (1.0)</td>
<td>2.5 (0.93)</td>
<td>2.33 (P=0.130)</td>
<td>0.21 (−0.01 to 0.49)</td>
</tr>
<tr>
<td>Pitch perturbation†</td>
<td>67</td>
<td>3.2 (2.4)</td>
<td>2.8 (1.5)</td>
<td>57</td>
<td>2.4 (1.4)</td>
<td>2.8 (1.5)</td>
<td>0.97 (P=0.326)</td>
<td>0.25 (−0.25 to 0.74)</td>
</tr>
<tr>
<td>Amplitude perturbation†</td>
<td>67</td>
<td>6.0 (4.2)</td>
<td>4.5 (2.6)</td>
<td>57</td>
<td>5.0 (3.3)</td>
<td>5.7 (3.8)</td>
<td>8.03 (P&lt;0.005)</td>
<td>1.6 (0.47 to 2.63)</td>
</tr>
<tr>
<td>VPQ total†</td>
<td>67</td>
<td>27.6 (8.0)</td>
<td>22.4 (5.8)</td>
<td>85</td>
<td>27.9 (7.5)</td>
<td>25.4 (8.1)</td>
<td>6.12 (P=0.015)</td>
<td>2.6 (0.51 to 4.83)</td>
</tr>
</tbody>
</table>

At completion of follow up (visit 3)

<table>
<thead>
<tr>
<th>Test or questionnaire</th>
<th>No of patients</th>
<th>Treatment Visit 1 (baseline)</th>
<th>Treatment Visit 2 or 3</th>
<th>No of patients</th>
<th>No treatment Visit 1 (baseline)</th>
<th>No treatment Visit 2 or 3</th>
<th>F value</th>
<th>Treatment effect (points (95% CI))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychological distress:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HADS anxiety</td>
<td>62</td>
<td>6.6 (4.0)</td>
<td>5.9 (3.8)</td>
<td>56</td>
<td>8.1 (4.0)</td>
<td>6.7 (4.1)</td>
<td>0.00 (P=0.969)</td>
<td>0.02 (−1.21 to 1.25)</td>
</tr>
<tr>
<td>HADS depression</td>
<td>62</td>
<td>4.2 (3.6)</td>
<td>3.8 (3.5)</td>
<td>56</td>
<td>4.2 (3.3)</td>
<td>4.3 (3.6)</td>
<td>1.08 (P&lt;0.300)</td>
<td>0.4 (−1.3 to 0.39)</td>
</tr>
<tr>
<td>CIS total score</td>
<td>67</td>
<td>7.0 (6.9)</td>
<td>7.3 (8.1)</td>
<td>62</td>
<td>8.6 (8.2)</td>
<td>7.8 (8.8)</td>
<td>0.64 (P=0.424)</td>
<td>0.89 (−1.3 to 3.06)</td>
</tr>
<tr>
<td>Voice:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buffalo overall rating</td>
<td></td>
<td>70</td>
<td>2.9 (1.1)</td>
<td>63</td>
<td>2.8 (1.0)</td>
<td>3.0 (1.0)</td>
<td>26.50 (P&lt;0.001)</td>
<td>0.82 (0.5 to 1.13)</td>
</tr>
<tr>
<td>Pitch perturbation†</td>
<td>65</td>
<td>3.1 (2.4)</td>
<td>2.7 (1.5)</td>
<td>54</td>
<td>2.4 (1.4)</td>
<td>2.2 (1.5)</td>
<td>1.58 (P=0.212)</td>
<td>0.35 (−0.20 to 0.89)</td>
</tr>
<tr>
<td>Amplitude perturbation†</td>
<td>65</td>
<td>5.8 (3.9)</td>
<td>4.1 (3.0)</td>
<td>54</td>
<td>5.1 (3.2)</td>
<td>4.5 (3.6)</td>
<td>0.90 (P=0.345)</td>
<td>0.57 (−0.62 to 1.77)</td>
</tr>
<tr>
<td>Score on VPQ‡</td>
<td>61</td>
<td>28.8 (8.2)</td>
<td>21.6 (6.3)</td>
<td>57</td>
<td>28.0 (7.5)</td>
<td>25.4 (8.5)</td>
<td>10.96 (P=0.001)</td>
<td>4.14 (1.67 to 6.61)</td>
</tr>
</tbody>
</table>

HADS=hospital anxiety and depression scale; CIS=clinical interview schedule; VPQ=vocal performance questionnaire.

†The higher the score the more dysphonic the voice.

‡The maximum score=60; the higher the score the more dysphonic the voice.
voice therapy have, by definition, relatively normal laryngeal appearances.

Psychological distress was not significantly reduced as a result of treatment. Voice therapy had a significant effect on one quality of life variable—mental health—at the end of treatment, but this was not maintained at follow up. A subgroup of patients remain psychologically distressed despite receiving treatment. Speech and language therapists often use psychological strategies but often acquire psychological training after qualification and in what has been described as an ad hoc manner. If patients with high psychological distress could be identified by screening they could be referred for psychological intervention, perhaps from a clinical psychologist.

The disconcertingly abnormal SF-36 results highlight the importance of effective vocal communication for an individual’s psychosocial wellbeing. Indeed, the level of psychological morbidity may also mainly reflect the greatly reduced quality of life in patients with dysphonia. Such interrelations underline the importance of a holistic treatment for reduction in symptoms and improvement in overall functioning.

In conclusion, this study shows that voice therapy is effective in improving self rated and observer rated measures of voice quality. However, voice therapy does...
Many patients with dysphonia are treated by voice therapy.

The effectiveness of voice therapy in a diverse group of patients is unknown.

Voice therapy is an effective treatment for dysphonia in terms of report by patients and perceptual ratings by an expert.

Psychological distress and reduction in general health status are common in patients with dysphonia but are not significantly affected by a course of voice therapy.

not significantly affect laryngeal pathophysiology or reduce the high levels of psychological distress that characterise patients with dysphonia.

We acknowledge the contributions of Nicola Bradshaw and Shonagh Scott, who were both involved as research assistants in the collection of data for this paper; Dr Martha Whitteman for running the sensitivity analyses; Catherine Dunnet, chief speech and language therapist at Glasgow Royal Infirmary, for contributions to the project and comments on a draft of the paper; and Professor Stuart Gatehouse for his comments on an earlier draft of the paper.

Contributors: KMacK initiated and coordinated the formulation of the primary study hypothesis, discussed core ideas, designed the protocol, and carried out all the voice therapy. IJD initiated and coordinated the formulation of the primary study hypothesis, discussed core ideas, designed the protocol, supervised the statistical analyses, and was principally involved in writing the paper. The guarantor for the paper is KMacK.

Funding: Scottish Office Home and Health Department (grant reference: K/RED/4/C249).

Competing interests: None declared.

What is already known on this topic

Many patients with dysphonia are treated by voice therapy. The effectiveness of voice therapy in a diverse group of patients is unknown.

What this study adds

Voice therapy is an effective treatment for dysphonia in terms of report by patients and perceptual ratings by an expert. Psychological distress and reduction in general health status are common in patients with dysphonia but are not significantly affected by a course of voice therapy.

(Accepted 8 July 2001)