Received: 29 January 2020 | Revised: 1 June 2020 | Accepted: 5 June 2020

DOI: 10.1111/coa.13596

ORIGINAL ARTICLE

WILEY

The effectiveness of voice therapy on voice-related handicap: A network meta-analysis

Ben Barsties v. Latoszek[1,2] | Christopher R. Watts[3] | Katrin Neumann[1]

[1]Department of Phoniatrics and Pediatric Audiology, University Hospital Münster, Westphalian Wilhelm University, Münster, Germany

[2]Speech-Language Pathology, SRH University of Applied Health Sciences, Düsseldorf, Germany

[3]Harris College of Nursing & Health Sciences, Texas Christian University, Fort Worth, TX, USA

Correspondence
Ben Barsties v. Latoszek, Speech-Language Pathology, SRH University of Applied Health Sciences, Graf-Adolf-Straße 67, 40210 Düsseldorf, Germany.
Email: benjamin.latoszek@srh.de

Abstract

Background: Treatment approaches for voice therapy are diverse, yet their differential effects are not well understood. Evaluations of treatment effects across approaches are important for clinical guidance and evidence-based practice.

Objective of review: To quantify the evidence of treatment effectiveness on the outcome measure Voice Handicap Index with the 30-items (VHI-30) from existing randomised controlled/clinical trials (RCT) of voice therapy using the statistical approach of a network meta-analysis (NMA) with a random effects model.

Type of review: Meta-analysis.

Search strategy: We searched in MEDLINE (PubMed, 1950 to 2019), Embase (1974 to 2019) and Science Citation Index (1994 to 2019) using five key terms. The inclusion criteria were reports of randomised controlled/clinical trials (RCTs) published in English or German which evaluated the effectiveness of a specific voice therapy treatment using VHI-30 as an outcome measure in adult participants with non-organic or organic voice disorders. Studies were excluded if participants had been diagnosed with neurological motor speech disorders or who were vocally healthy. Furthermore, no medical, pharmacological or instrumental (eg voice amplification) treatments were considered.

Evaluation method: The primary outcome variable was VHI-30 with a score from 0 to 120. The pre-post treatment change in VHI-30 scores was an average score of 13 points related to various VHI-30 test-retest results.

Results: We retrieved 464 publications (ie with duplicates) and included 13 RCTs, which evaluated nine interventions, in the final analysis. The most effective intervention with a significant and clinically relevant effect was Stretch-and-Flow Phonation (SFP) (mean pre-post difference −28.37, 95% confidence interval [CI], −43.05 to-13.68). Resonant Voice (RV), the Comprehensive Voice Rehabilitation Program (CVRP) and Vocal Function Exercises (VFE) also demonstrated significant improvements.

Conclusions: Of the nine voice interventions identified with the present NMA, SFP, RVT, CVRP, and VFE effectively improved VHI-30 scores from pre- to post-treatment. SFP proved to be the most significant and clinically relevant treatment. Further contributions of high-quality intervention studies are needed to support evidence-based practice in vocology.

1 | INTRODUCTION

Voice problems longer than 1 week affect one in 13 adults annually and present a substantial burden and impact on quality of life.[1,2] Current prevalence rates for dysphonia diagnosis based on insurance claims are approximately 1.7% of the population.[2] Overall prevalence rates might be much higher, because many patients with voice problems do not seek medical care.[2] The treatment effectiveness for voice therapy is typically measured by several outcome variables such as pre- to post-treatment changes in endoscopic, auditory-perceptual, acoustic and aerodynamic assessments. Various voice diagnostic protocols also recommend as a standard dimension of the assessment battery the patient's self-assessment of the perceived handicap.[3-8]

The Voice Handicap Index (VHI) is an instrument designed to assess self-perceived voice-related handicap using a 30-item questionnaire (VHI-30).[9] It consists of three subscales with statements relating to physical, functional and emotional domains. Each domain includes 10 statements evaluated on a five-point Likert scale. The VHI-30 total score ranges from 0 to 120. The VHI is a popular clinical tool used to quantify the impact of a voice disorder on daily voice-related activities. More than twenty VHI-30 validation studies in different languages have been published.[10] Further studies of the VHI-30 evaluated the potential reduction of questions,[11-13] its application to various populations with voice disorders,[14-22] the relationship to objective measurements,[17,23-25] test-retest differences[10,26-28] and its applicability as screening tool.[28] The VHI has been used by numerous studies to measure the effect of voice therapy; its sensitivity treatment-associated voice changes have established it as a useful voice assessment tool.[29-32]

Dysphonia has different aetiologies[33,34] that determine the primary approach to treatment. Treatment options can include voice therapy,[35-45] phonosurgery[46,47] or pharmacological[48] treatment. The present paper focuses on studies reporting the effectiveness of voice therapy as the primary intervention approach. In surveying the different systematic reviews of voice therapy effectiveness,[35-45] only three researcher groups performed a meta-analysis on different treatment approaches (ie Vocal Function Exercises, Laryngeal Manual Therapy, and indirect and combined voice therapy).[36,42,43]

There exist several barriers for conducting a meta-analysis of voice therapy effectiveness. First, there is a lack of consistency of outcome measurements across studies to evaluate a treatment effect. Second, many studies used small-scale uncontrolled observational study designs with the inclusion of only small samples or specific populations. Third, there are differences in timing, frequency, and intensity of treatment or home practice. However, to compare multiple treatments for a given medical or healthcare condition in one analysis, a network meta-analysis (NMA) can be used. The NMA includes both direct comparisons of interventions within randomised controlled trials and indirect comparisons across trials based on a common comparator.[49] This statistical approach might be useful for meta-analyses of voice therapy effectiveness because different voice treatments could be compared with each other if the same outcome measure is reported in the same way across studies.

The aim of the present study was to perform a NMA for estimating treatment effectiveness and establish a ranking for different voice therapy approaches for dysphonia based on randomised controlled/clinical trial (RCT) studies. While the use of objective outcome parameters from acoustic and aerodynamic voice quantification and laryngeal imaging would be desirable, it is often limited by the availability and comparability of software and hardware, data acquisition, and processing methodology, which also affects the comparability of studies. Therefore, the VHI was chosen as the outcome measure for this NMA because it is one of the most commonly used measures in voice therapy research due to its simplicity and ease of use without instruments or secondary assessments by others.

Keypoints

- We compared the effectiveness of various voice treatments using the Voice Handicap Index 30.
- To our knowledge, this is the first network meta-analysis on the treatment effectiveness of dysphonia.
- Five from nine voice treatments resulted in a significant improvement of VHI-30 scores.
- Stretch-and-Flow Phonation has been identified by our network meta-analysis as the most effective intervention.

2 | METHODS

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension statements for reporting of systematic reviews incorporating NMA of healthcare interventions.[50] A systematic literature search was performed to locate studies in electronic databases. We searched in MEDLINE (PubMed, 1950 to 2019), Embase (1974 to 2019) and Science Citation Index (1994 to 2019). References from published studies were also reviewed. Potential articles were identified by title and abstract. In these publications, the specific voice therapy approach per RCT had to be evaluated with several comparisons between subsets of treatments to be further considered in the NMA. In addition, treatment effectiveness had to be measured with the results of the VHI-30. To identify the studies to be included, the following search terms were used: randomised controlled trial, randomised clinical trial, randomised sham-controlled trial, voice therapy and voice handicap index.

Published studies were included that evaluated the effectiveness of treatments that targeted non-organic or organic voice disorders in adults and adolescents 16 years or older with a pre-post treatment design. The outcome measure for the effectiveness of treatments was evaluated using the total score of the VHI-30. Clinically relevant critical difference scores of the total VHI-30 score have been interpreted as meaningful when the difference is minimally 13 points. This score corresponds to an average cut-off score from various test-retest studies of VHI total scores ranging from 8 to 18.[10,26-28] Included studies minimally reported the number of subjects per group, the mean results between pre- and post-treatment outcomes and standard deviation (SD) values of the differences of change or p-values of the pre-post outcomes. Finally, scientific reports were considered in English and German languages.

We excluded studies in which any of the subjects had been diagnosed with a neurological motor speech disorders (eg Parkinson's disease) or involved subjects who were vocally healthy (eg for prevention) or singers without voice disorders. Furthermore, we did not include studies that used medical or pharmacological treatments in participant groups. Additionally, studies were excluded if adequate descriptions of the voice therapy approaches were not provided or the voice therapy protocol of the groups had not a primary single approach. Finally, studies that incorporated instrumentation in the application of voice therapy (eg voice amplification) were excluded as well.

2.1 | Risk of bias assessment

To assess risk of bias of the included studies, the RoB 2 tool was used.[51] The following domains were evaluated to conclude an overall risk of bias (ie low, some concerns or high): randomisation process, deviations from intended interventions, missing outcome data, measurement of the outcome and selection of the reported result.

2.2 | Statistical analysis

Where possible, mean pre-post differences (MD) of VHI-30 total scores and their SDs for each treatment arm in each study were directly extracted from the publications. Where standard deviations were missing, p-values of the pre-post MD were used for calculating them. For random effects NMA, R package netmeta from the open statistical programming environment R was used.[52,53] Results were presented as MD between pre-post differences with 95% confidence intervals (CI). Furthermore, we ranked interventions based on a quantity, called P-score, which is a critical appraisal of ranking that can be considered as a frequentist analogue to surface under the cumulative ranking curve (SUCRA) of the Bayesian concept without need for resampling methods.[54] It is a simple analytical method, which is based on the frequentist point estimates and their standard errors. P-scores produce a ranking on a scale from 0 to 1, where 0 means worst and 1 means best and are based on both size and uncertainty of the effects. This metric can be interpreted as measuring the mean degree of confidence that one treatment is better than a comparable treatment.

3 | RESULTS

3.1 | Study characteristics

Figure 1 shows the details of exclusion and inclusion of studies using a flow chart. Table 1 presents the characteristics of the thirteen included RCT's involving of 389 subjects in intervention groups and 51 subjects in control groups. Thus, the total number of sample sizes for intervention groups ranged from 8 to 37 while sample sizes for control groups ranged from 8 to 20 per group. The control groups included only voice-disordered subjects without any preceding intervention. In total, nine different voice therapy approaches were reported, namely the following: Comprehensive Voice Rehabilitation Program (CVRP), Lip Trill (LT), Resonant Voice (RV), Respiratory Muscle Training (RMT), Stretch-and-Flow Phonation (SFP), Tube Phonation in Air (TPA), Vocal Function Exercises (VFE), Vocal Hygiene (VH) and Water Resistance Therapy (WRT).

The risk of bias was evaluated for all thirteen studies, and the results showed that an overall low risk of bias was reported.

Figure 2 illustrates the pairwise comparisons between the nine interventions and control groups in a network graph. The nodes (red circles) in this graph represent voice therapy approaches reported in the included studies. Lines represent direct comparisons of interventions. Numbers at the lines represent the number of included studies for the given comparison.

3.2 | Control group versus treatment groups

The results of the NMA showed considerable heterogeneity (I^2 = 42.9) and are presented as a forest plot in Figure 3. Four interventions (SFP, RV, CVRP and VFE) manifested significant outcome effects on measures of VHI-30 as evidenced by confidence intervals which did not cross the null line in the forest plot. These interventions had also mean pre-post MD values above equal the threshold of clinical relevant (SFP −28.37; RV = −21.84; CVRP = −22.01; VFE = −13.00). SFP was the only intervention which demonstrated statistically significant effects where the low end of the confidence interval was greater than the −13 VHI-30 difference threshold for clinical relevance (MD = −28.37, 95% CI −43.05 to 13.68).

According to the P-score ranking, SFP resulted in the highest outcome (P-score = 0.95) that markedly exceeded all other values (Figure 3).

4 | DISCUSSION

To the best of our knowledge, this is the first NMA to assess the treatment effectiveness of different voice therapy approaches in subjects with dysphonia reported in RCT study designs. The quality of the included studies had low risk of bias. We used the VHI-30 as the primary outcome variable, as it is a frequently used voice

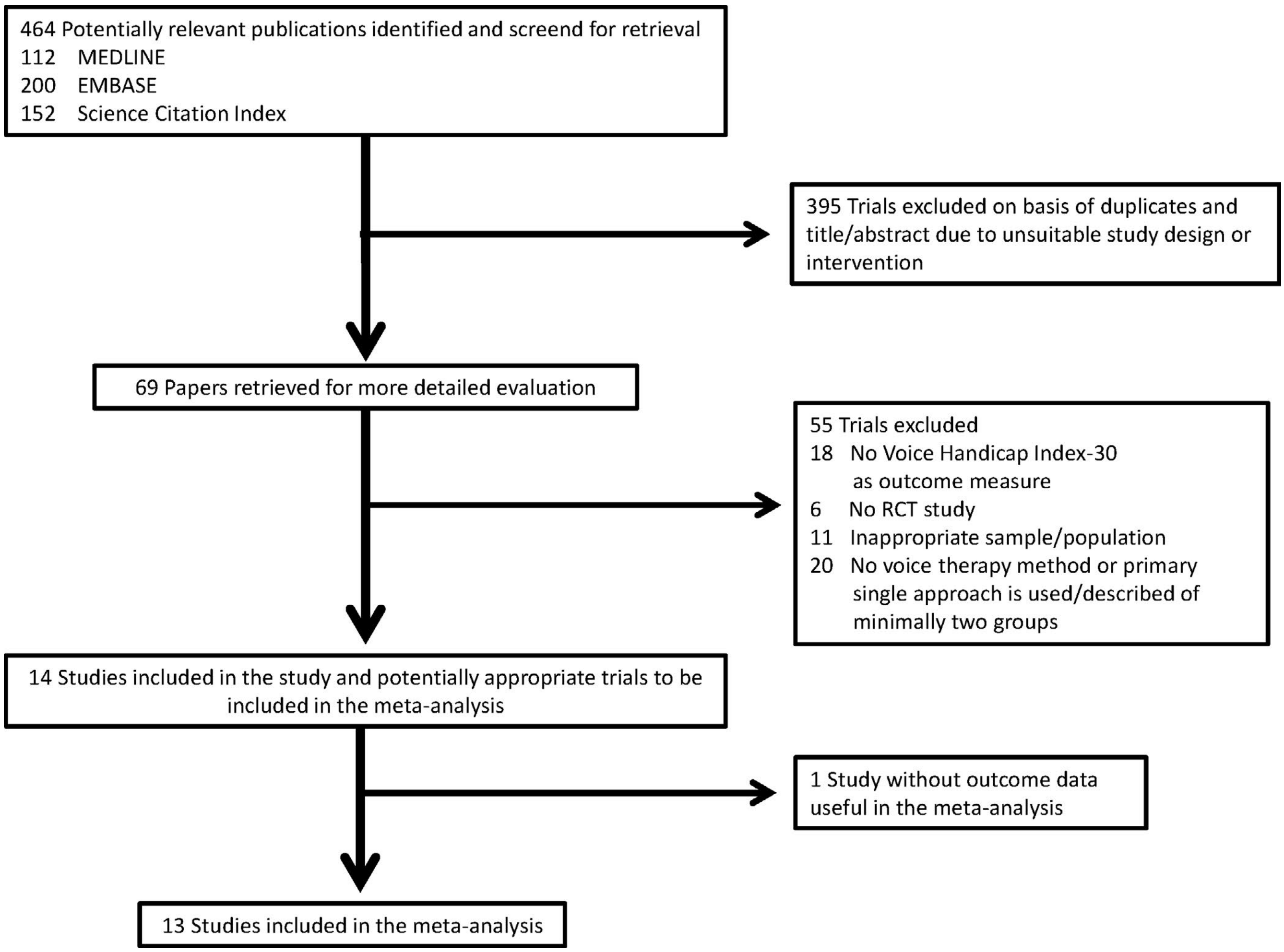

FIGURE 1 Quality of reporting of meta-analyses statement trial flow

assessment tool and measures the impact of treatment from the perspective of the patient. SFP manifested significant (ie P-score, and 95%-CI results) and clinically relevant (ie VHI-30 difference score > 13 points) treatment outcomes and placed it as the superior treatment approach across all those compared. The RV, CVRP and VFE approaches also demonstrated statistically significant improvements on VHI-30 scores (ie MD score ≥ 13 points and a confidence interval which did not cross the null line), in which VFE was the most evaluated voice therapy method. Interestingly, three of these four approaches share a common framework in that they each have a hierarchical structure with a physiological concept. For example, the SFP approach aims to control airflow and laryngeal control with increasing complexity moving from a voiceless airflow over an exaggerated airflow during connected speech with elongated vowels, to connected speech with normal articulation and vowel production with the perception of an easy, effortless airflow. The primary perceptual target of SFP is airflow movement throughout the vocal tract, at each level of the hierarchy. It has been applied to individuals with both non-organic and organic voice disorders.[55,56]

The RV approach has been investigated in individuals with hyperfunctional (muscle tension) or phonotraumatic voice disorders.[39] Like SFP, the RV approach moves across a framework from low complexity to high complexity (eg continuous speech) stimuli. The working of RV is based on an easy voice production with vibratory sensations in facial bones that reflects a relatively high-intensity glottal source spectrum yielding a loudness that is easily heard and intense oral air pressure variations that result in the vibratory sensations.[57]

In case of the VFE approach, the perceptual targets include low intensity phonation with a resonant production. The treatment stimuli consist of four core exercises which are produced in multiple repetitions and sets for a specified number of weeks. Over time, the number of sets and specific exercises can be tapered to the needs of the individual. A previous meta-analysis of VFE[43] showed comparable results to our NMA with mean VHI improvements of −11 and −13.00, respectively. This meta-analysis included four studies to calculate the effect size.[43] Both meta-analyses showed that treatment effectiveness as measured by VHI can be expected in patients with both non-organic or organic voice disorders.

The concept of CVRP contains an eclectic voice therapy approach in a six-week application. This approach was mainly developed for behavioural dysphonia. It is model of care, which relates to a consolidation of experiences from the Brazilian Larynx Institute

TABLE 1 Characteristics of RCTs included in the NMA

Study ID	Subjects	Primary single interventions (number of participants of each group; duration of treatment in weeks)	Outcome measures
Roy et al (2001)[66]	58 voice-disordered teachers with non-organic voice disorder	1. No intervention (19; 6) 2. VFE (19; 6) 3. VH (20; 6)	1. Self-evaluation (VHI, self-designed teacher questionnaire)
Roy et al (2002)[67]	29 voice-disordered teachers with non-organic voice disorder	1. VH (15; 6) 2. No intervention (14; 6)	1. Self-evaluation (VHI, voice severity self-rating scale) 2. Acoustics (jit and shim, NHR)
Roy et al (2003)[68]	39 voice-disordered teachers with non-organic voice disorder	1. RV (19; 6) 2. RMT (20; 6)	1. Self-evaluation (VHI, voice severity self-rating scale)
Behrman et al (2008)[69]	62 women with phonotrauma (pre-nodules but moreover non-organic voice disorders)	1. RV (31; 6) 2. VH (31; 6)	1. Self-evaluation (VHI)
Kapsner-Smith et al (2015)[70]	20 subjects with heterogenic voice disorders (eg non-organic dysphonia, VF polyp, VF nodules, VF oedema, VF granuloma, VF paralysis, and VF ulcer)	1. No intervention of each 5 subjects from group B and C who had delayed interventions (10; 6) 2. VFE (10; 6) 3. TPA (10; 6)	1. Self-evaluation (VHI, interview questions) 2. Perceptual Judgement (CAPE-V)
Watts et al (2015)[55]	20 subjects with non-organic voice disorders	1. SFP (10; 6) 2. VH (10; 6)	1. Self-evaluation (VHI) 2. Acoustics (CPP) 3. Aerodynamics (MPT, S/Z Ratio)
Pedrosa et al (2016)[71]	72 voice professionals with non-organic voice disorder	1. VFE (35; 6) 2. CVRP (37;6)	1. Self-evaluation (VHI, VRQOL) 2. Laryngeal stroboscopic findings 3. Perceptual Judgement (overall voice quality)
Guzman et al (2017)[72]	20 subjects with non-organic voice disorders	1. WRT (10; 6) 2. TPA (10; 6)	1. Self-evaluation (VHI, voice severity self-rating scale) 2. Perceptual judgement (overall voice quality) 3. Electroglottography (CQ) 4. Aerodynamics (glottal airflow, subglottal pressure, glottal resistance, PTP)
Kao et al (2017)[64]	19 subjects with unilateral adductor vocal fold paralysis within 6 mo of initial diagnosis	1. VFE (10; 12.4) 2. VH (9; 12.4)	1. Self-evaluation (VHI) 2. Laryngeal stroboscopic findings 3. Perceptual judgement (CAPE-V) 4. Acoustics (jit and shim, NHR) 5. Aerodynamics (MPT, PQ, PTP)
Rojas et al (2017)[73]	16 subjects with presbyphonia	1. TPA (8; 6) 2. WRT (8; 6)	1. Self-evaluation (VHI)
La Mantia et al (2018)[74]	19 subjects with primary laryngeal cancer treated with curatively intended radiotherapy	1. VFE (10; 6) 2. VH (9; 6)	1. Self-evaluation (VHI, EORTC QLQ-H&N35) 2. Laryngeal stroboscopic findings with high-speed parameters 3. Perceptual judgement (GRBAS) 4. Acoustics (jit and shim) 5. Aerodynamics (MPT, MFR)
Meerschman et al (2019)[62]	35 subjects with heterogenic voice disorders who were referred for voice therapy	1. LT (9; 3) 2. WRT (9; 3) 3. TPA (9; 3) 4. No intervention (8; 3)	1. Self-evaluation (VHI, VTDS) 2. Acoustics (AVQI, DSI) 3. Perceptual judgement (GRBAS)
Watts et al (2019)[56]	21 subjects with predominantly non-organic voice disorders	1. SFP (12; 6) 2. RV (9; 6)	1. Self-evaluation (VHI) 2. Acoustics (AVQI, CPPS) 3. Perceptual judgement (CAPE-V)

Abbreviations: AVQI, acoustic voice quality index; CPP, cepstral peak prominence; CPPS, Cepstral Peak Prominence Smoothed; CQ, Contact Quotient; CVRP, Comprehensive Voice Rehabilitation Program; DSI, dysphonia severity index; EORTC QLQ-H&N35, Health-Related Quality of Life questionnaire Head and Neck module; jit, Jitter; LT, Lip Trill; MFR, mean flow rate; MPT, maximum phonation time; NHR, Noise to Harmonics Ratio; PQ, Phonation Quotient; PTP, Phonation Threshold Pressure; RCT, Randomised controlled/clinical trial; RMT, Respiratory Muscle Training; RV, Resonant Voice; S/Z ratio, Ratio of the durations for which a person can sustain the sounds “s” and “z,” VRQOL, Voice-Related Quality of Life; shim, Shimmer; VA, Voice Amplification; VF, Vocal Folds; VFE, Vocal Function Exercises; VH, Vocal Hygiene; VHI, Voice Handicap Index; VTDS, Vocal Tract Discomfort Scale; WRT, Water Resistance Therapy.

in the 1970s, the Ambulatories of Larynx and Voice of Universidade Federal de São Paulo, and Center for the Study of Voice in the 1990s.[58] Main aspects of the programme are voice orientation (eg vocal hygiene and habits), vocal psychodynamics (eg analysis of the impact of altered voices from professional, social and emotional viewpoints) and vocal training (eg body posture, physiological voice techniques such as various semi-occluded vocal tract and vocal facilitating exercises, resonation, articulation and the combination of all these subsystems).[58] This kind of holistic concept is very common in voice therapy without a specific therapy method known from clinical practice[59] and research[36] being identifiable.

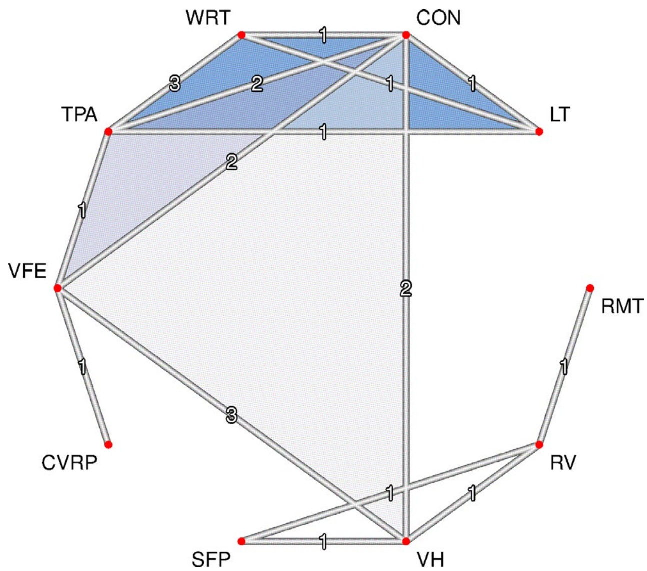

FIGURE 2 Network graph with focus on the comparison of CON and different voice treatments (8 interventions compared with CON by direct estimates and 14 interventions compared with CON by indirect estimates) from a multiple pairwise meta-analysis. Coloured polygons represent multi-arm studies (dark blue = four-arm study, light blue/light grey = three-arm studies). CON, Control Group; SFP, Stretch-and Flow Phonation; RV, Resonant Voice; CVRP, Comprehensive Voice Rehabilitation Program; VFE, Vocal Function Exercises; RMT, Respiratory Muscle Training; WRT, Water Resistance Therapy; LT, Lip Trill; VH, Vocal Hygiene; and TPA, Tube Phonation in Air

The weakest treatment effects were reported for TPA and VH. An explanation of the weak performance of TPA could be that this method was mostly tested in dysphonic patients with heterogenic voice disorders. With the exception to VFE, other methods were mostly verified in subjects with non-organic voice disorders. Thus, the treatment outcomes could be compromised by the range of voice disorders within a study, because some voice disorders may require more intensive or longer lasting interventions than do others. In our NMA, VFE showed better VHI-30 results than TPA, which also mostly included heterogeneous voice disorders in the intervention groups. In a direct comparison between VFE and TPA, it remains unclear whether either of the two approaches could be important for patient care.

VH is an indirect voice therapy concept (ie focus on managing the aspects which contribute to the voice problem) that is typically utilised for some benign vocal fold lesions.[60] However, the concept of VH is controversial because the treatment effectiveness is dependent on various factors such as the quality of the vocal hygiene education programme.[60] Furthermore, the treatment effectiveness is lower in VH than in combined VH and direct voice therapy or in direct voice therapy that does not incorporate VH.[36,61]

4.1 | Caveats, limitations and future directions

Limitations of the present meta-analysis concern the generalisability of its results, but also provide a direction for future research. The following section is divided into two categories (ie representativeness of the current meta-analysis and methodological limitation):

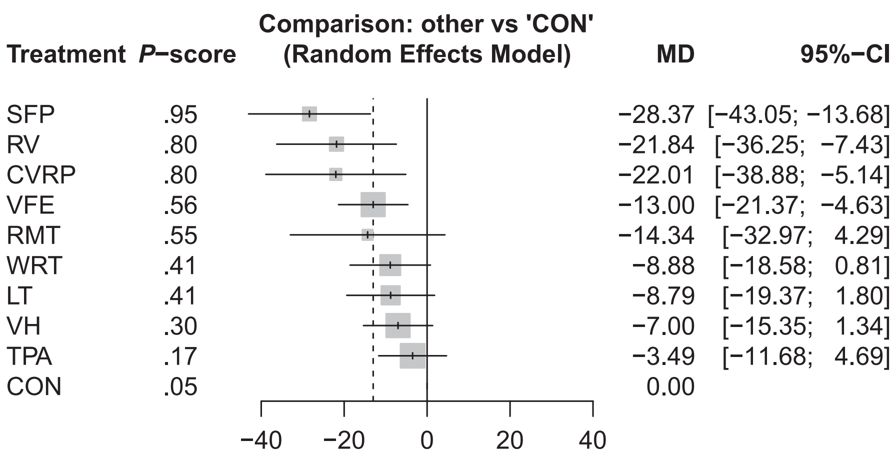

FIGURE 3 Forest plot representing the treatment effect sizes of the VHI-30 by specific interventions in comparison with the control group sorted at the order of treatment rankings of the P-score. The dashed line represents the equivalence region for the total VHI-30 score at −13 points. CON, Control Group; SFP, Stretch-and Flow Phonation; RV, Resonant Voice; CVRP, Comprehensive Voice Rehabilitation Program; VFE, Vocal Function Exercises; RMT, Respiratory Muscle Training; WRT, Water Resistance Therapy; LT, Lip Trill; VH, Vocal Hygiene; TPA, Tube Phonation in Air; MD, Mean difference; and CI, Confidence interval

4.1.1 | Representativeness of the current meta-analysis

First, there was considerable heterogeneity of studies in the present NMA (I^2 = 42.9%), which may be due to publication bias, selection biases across studies regarding kind of dysphonia and differences in responsiveness to change of the VHI-30 scale. Publication bias was not explicitly statistically tested but overall a low number of subjects were included in the treatment groups, which could explain the high 95% CI in the forest plots. Although non-organic voice disorders were considered in the studies, five from thirteen RCTs included heterogenic or organic voice disorder populations, which could explain variances in VHI-30 responsiveness. Across all included studies, the treatment duration was mostly six weeks. Only Meerschmanet al.[62] reported a duration of three weeks for voice therapy. However, the most evaluated interventions in the present study required a lower duration for intervention than a common voice therapy which had an average of 9.25 weeks.[63] Just Kao et al[64] used a higher duration with an average of 12.4 weeks.

Second, multidimensional assessments are needed for the diagnosis and measurement of treatment outcomes in voice disorders. Therefore, a consistent standardisation of diagnostic voice assessments or compliance with already established standards is important for the performance of both voice therapies and high-quality comparative studies on the outcome of different voice therapies.[36,38,43] Over the last two decades, attempts were undertaken to attain consensus and recommendations for voice measurements of laryngeal imaging (eg stroboscopy),[3,4,8,65] auditory-perceptual judgement (eg CAPE-V [Consensus Auditory-Perceptual Evaluation–Voice], and GRBAS [grade, roughness, breathiness, asthenia, strain] scales),[3,7,8] acoustics (eg cepstral peak prominence, dysphonia severity index and acoustic voice quality index),[7,8,65] aerodynamics (eg maximum phonation time, glottal airflow rate and subglottal air pressure)[3,7,8,65] and self-evaluation (eg VHI).[7]

Third, several voice therapy approaches were not included in the present meta-analysis with regard to those that were analysed in previous systematic reviews on voice therapy.[35-45] This was due to the fact that currently no high-quality studies were available for those approaches, or studies used voice treatment outcome measures other than VHI, or omitted essential statistical data for a meta-analysis.

Fourth, the search strategy for relevant scientific reports for the meta-analysis included only two languages. Thus, potentially relevant publications in other languages were omitted.

4.1.2 | Methodological limitations

First, more studies are needed to investigate voice therapy approaches for homogenous groups of voice disorders. The present study included mostly subjects with non-organic dysphonia. To evaluate the treatment effects of specific approaches on other voice disorders which are commonly encountered in clinical practice, for example vocal nodules and vocal fold paralysis, further investigations are needed.[33,34]

Second, the evaluation of intensity and frequency of voice therapy approaches should be further investigated, as issues surrounding treatment dose are poorly understood in the voice treatment literature.

Third, the main outcome variable used in this study was the VHI-30. Some studies have used different versions of the VHI, and these were not comparable to the outcome of the VHI-30. As a result, those studies had to be excluded for the present meta-analysis.

5 | CONCLUSION

The above-stated limitations notwithstanding, the current NMA revealed SFP, RV, CVRP, and VFE interventions were effective for improving VHI-30 total scores. SFP had the greatest effect. Other voice therapy approaches did not result in significant or clinically relevant treatment effects on VHI-30 scores. Thus, the presented NMA reduced the number of effective interventions from a larger pool of voice therapy approaches for individuals with predominantly non-organic dysphonia, based on the self-evaluation of a perceived vocal handicap. On the one hand, the available information could be particularly useful for clinicians who use different approaches to dysphonia treatment, in order to better assess the method-specific expectations for the improvement of voice-related quality of life. On the other hand, this meta-analysis underscored the need of additional high-quality intervention studies with large samples of treated patients, multi-dimension assessments and homogeneous groups of voice disorders to better inform clinical practice on significant constituents of voice therapies and associated outcomes.

ACKNOWLEDGEMENT

The authors thank Dr Gerta Rücker (Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center—University of Freiburg, Germany) for her support in network meta-analysis statistics.

CONFLICT OF INTEREST

None to declare.

ORCID

Ben Barsties v. Latoszek https://orcid.org/0000-0002-0086-8163

REFERENCES

1. Bhattacharyya N. The prevalence of voice problems among adults in the United States. *Laryngoscope*. 2014;124:2359-2362.
2. Benninger MS, Holy CE, Bryson PC, Milstein CF. Prevalence and occupation of patients presenting with dysphonia in the United States. *J Voice*. 2017;31:594-600.
3. Dejonckere PH, Bradley P, Clemente P, et al. A basic protocol for functional assessment of voice pathology, especially for

investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). *Eur Arch Otorhinolaryngol*. 2001;258:77-82.
4. Kaszuba SM, Garrett CG. Strobovideolaryngoscopy and laboratory voice evaluation. *Otolaryngol Clin North Am*. 2007;40:991-1001.
5. Rodríguez-Parra MJ, Adrián JA, Casado JC. Voice therapy used to test a basic protocol for multidimensional assessment of dysphonia. *J Voice*. 2009;23:304-318.
6. Cohen W, Wardrop A, Wynne DM, Kubba H, McCartney E. Development of a minimum protocol for assessment in the paediatric voice clinic. Part 2: subjective measurement of symptoms of voice disorder. *Logoped Phoniatr Vocol*. 2012;37:39-44.
7. Ziethe A, Patel R, Kunduk M, Eysholdt U, Graf S. Clinical analysis methods of voice disorders. *Curr Bioinform*. 2011;6:270-285.
8. Boominathan P, Samuel J, Arunachalam R, Nagarajan R, Mahalingam S. Multi parametric voice assessment: Sri Ramachandra University protocol. *Indian J Otolaryngol Head Neck Surg*. 2014;66:246-251.
9. Jacobson BH, Johnson A, Grywalski C, et al. The Voice Handicap Index (VHI): development and validation. *Am J Speech Lang Pathol*. 1997;6:66-70.
10. Herbst CT, Oh J, Vydrová J, Švec JG. DigitalVHI–a freeware open-source software application to capture the Voice Handicap Index and other questionnaire data in various languages. *Logoped Phoniatr Vocol*. 2015;40:72-76.
11. Rosen CA, Lee AS, Osborne J, Zullo T, Murry T. Development and validation of the voice handicap index-10. *Laryngoscope*. 2004;114:1549-1556.
12. Nawka T, Verdonck-de Leeuw IM, De Bodt M, et al. Item reduction of the voice handicap index based on the original version and on European translations. *Folia Phoniatr Logop*. 2009;61:37-48.
13. Feige K, Strauss A, Strauss G. Voice handicap index-18 as an instrument for assessing subjective voice impairment between voice pre- and post-treatment. *J Voice*. 2019. http://dx.doi.org/10.1016/j.jvoice.2019.04.005. [Epub ahead of print].
14. Moerman M, Martens JP, Dejonckere P. Application of the Voice Handicap Index in 45 patients with substitution voicing after total laryngectomy. *Eur Arch Otorhinolaryngol*. 2004;261:423-428.
15. Wingate JM, Ruddy BH, Lundy DS, et al. Voice handicap index results for older patients with adductor spasmodic dysphonia. *J Voice*. 2005;19:124-131.
16. Evans E, Carding P, Drinnan M. The voice handicap index with post-laryngectomy male voices. *Int J Lang Commun Disord*. 2009;44:575-586.
17. Schindler A, Mozzanica F, Vedrody M, et al. Correlation between the Voice Handicap Index and voice measurements in four groups of patients with dysphonia. *Otolaryng Head Neck Surg*. 2009;141:762-769.
18. Sanuki T, Yumoto E, Kodama N, et al. Long-term Voice Handicap Index after type II thyroplasty using titanium bridges for adductor spasmodic dysphonia. *Auris Nasus Larynx*. 2014;41:285-289.
19. Karlsen T, Heimdal JH, Grieg AR, Aarstad HJ. The Norwegian Voice Handicap Index (VHI-N) patient scores are dependent on voice-related disease group. *Eur Arch Otorhinolaryngol*. 2015;272:2897-2905.
20. Crawley BK, Dehom S, Thiel C, et al. Assessment of clinical and social characteristics that distinguish presbylaryngis from pathologic presbyphonia in elderly individuals. *JAMA Otolaryngol Head Neck Surg*. 2018;144:566-571.
21. Tafiadis D, Helidoni ME, Chronopoulos SK, et al. The Hellenic Voice Handicap Index of different laryngeal mass lesions: a receiver-operating characteristic analysis. *J Voice*. 2019; 22:S0892.
22. Kovatch KJ, Reyes-Gastelum D, Hughes DT, et al. Assessment of voice outcomes following surgery for thyroid cancer. *JAMA Otolaryngol Head Neck Surg*. 2019;145(9):823.
23. Lopes LW, da Silva JD, Simões LB, et al. Relationship between acoustic measurements and self-evaluation in patients with voice disorders. *J Voice*. 2017;31:119.e1-119.e10.
24. Pommée T, Maryn Y, Finck C, Morsomme D. The acoustic voice quality index, version 03.01, in French and the voice handicap index. *J Voice*. 2018. https://doi.org/10.1016/j.jvoice.2018.12.008. [Epub ahead of print].
25. Valentino WL, Park J, Alnouri G, et al. Diagnostic value of acoustic and aerodynamic measurements in vocal fold movement disorders and their correlation with laryngeal electromyography and Voice Handicap Index. *J Voice*. 2019. https://doi.org/10.1016/j.jvoice.2019.10.008. [Epub ahead of print].
26. Hakkesteegt MM, Wieringa MH, Gerritsma EJ, Feenstra L. Reproducibility of the Dutch version of the Voice Handicap Index. *Folia Phoniatr Logop*. 2006;58:132-138.
27. Van Gogh CD, Mahieu HF, Kuik DJ, Rinkel RN, Langendijk JA, Verdonck-de Leeuw IM. Voice in early glottic cancer compared to benign voice pathology. *Eur Arch Otorhinolaryngol*. 2007;264:1033-1038.
28. Niebudek-Bogusz E, Kuzańska A, Woznicka E, Sliwinska-Kowalska M. Assessment of the voice handicap index as a screening tool in dysphonic patients. *Folia Phoniatr Logop*. 2011;63:269-272.
29. Rosen CA, Murry T, Zinn A, Zullo T, Sonbolian M. Voice handicap index change following treatment of voice disorders. *J Voice*. 2000;14:619-623.
30. Hengen J, Peterson M, McAllister A. Patient characteristics and intervention effect as measured by Voice Handicap Index. *Logoped Phoniatr Vocol*. 2017;42:93-98.
31. Moore J, Greenberg C, Thibeault SL. Predictors of six-month change in the Voice Handicap Index in a treatment-seeking population. *J Voice*. 2017;31:41-47.
32. Joshi A, Dave VJ, Bradoo R, Sapkale D. Evaluation of validity of Voice Handicap Index among Indian Population. *Int J PhonosurgLaryngol*. 2017;7:59-62.
33. De Bodt M, Van den Steen L, Mertens F, et al. Characteristics of a dysphonic population referred for voice assessment and/or voice therapy. *Folia Phoniatr Logop*. 2015;67:178-186.
34. Martins RH, do Amaral HA, Tavares ELM, et al. Voice disorders: etiology and diagnosis. *J Voice*. 2016;30:761.e1-761.e9.
35. Speyer R. Effects of voice therapy: a systematic review. *J Voice*. 2008;22:565-580.
36. Ruotsalainen J, Sellman J, Lehto L, Verbeek J. Systematic review of the treatment of functional dysphonia and prevention of voice disorders. *Otolaryngol Head Neck Surg*. 2008;138:557-565.
37. Eastwood C, Madill C, Mccabe P. The behavioural treatment of muscle tension voice disorders: a systematic review. *Int J Speech Lang Pathol*. 2015;17:287-303.
38. Walton C, Conway E, Blackshaw H, Carding P. Unilateral vocal fold paralysis: a systematic review of speech-language pathology management. *J Voice*. 2017;31:509.e7-509.e22.
39. Yiu EM, Lo MC, Barrett EA. A systematic review of resonant voice therapy. *Int J Speech Lang Pathol*. 2017;19:17-29.
40. Desjardins M, Halstead L, Cooke M, Bonilha HS. A systematic review of voice therapy: what "effectiveness" really implies. *J Voice*. 2017;31:392.e13-392.e32.
41. da Cunha PG, de Oliveira LI, DalboscoGadenz C, Cassol M. Effects of voice therapy on muscle tension dysphonia: a systematic literature review. *J Voice*. 2018;32:546-552.
42. Ribeiro VV, Pedrosa V, Silverio KCA, Behlau M. Laryngeal manual therapies for behavioral dysphonia: a systematic review and meta-analysis. *J Voice*. 2018;32:553-563.
43. Angadi V, Croake D, Stemple J. Effects of vocal function exercises: a systematic review. *J Voice*. 2019;33:124.e13-124.e34.
44. Barsties v. Latoszek B. Evidenced-based voice therapy programs for the treatment of dysphonia: a systematic literature review. [German] *Sprache-Stimme-Gehör*. 2020; 44:16-22.

45. Desjardins M, Bonilha HS. The impact of respiratory exercises on voice outcomes: a systematic review of the literature. *J Voice*. 2019; 25:S0892.
46. Sataloff RT, Hawkshaw MJ, Divi V, Heman-Ackah YD. Voice surgery. *Otolaryngol Clin North Am*. 2007;40:1151-1183.
47. Reiter R, Hoffmann T. Phonomicrosurgery - a retrospective analysis of 400 cases. [German] *Laryngorhinootologie*. 2017;96:597-606.
48. Schönweiler R. An update of medication in voice treatment and side effects on voice relating to medication. [German] *Sprache-Stimme-Gehör*. 2020;44:23-28.
49. Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research Synthesis Methods*. 2012;3:80-97.
50. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med*. 2015;162:777-784.
51. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898.
52. Rücker G, Krahn U, König J, Efthimiou O, Schwarzer G. netmeta: network meta-analysis using frequentist methods. 2019. https://github.com/guido-s/netmeta and http://meta-analysis-with-r.org. Accessed November 26, 2019.
53. R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2019. https://www.R-project.org/ (Accessed 26 November 2019)
54. Rücker G, Schwarzer G. Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Med Res Methodol*. 2015;15:58.
55. Watts CR, Hamilton A, Toles L, Childs L, Mau T. A randomized controlled trial of stretch-and-flow voice therapy for muscle tension dysphonia. *Laryngoscope*. 2015;125:1420-1425.
56. Watts CR, Hamilton A, Toles L, Childs L, Mau T. Intervention outcomes of two treatments for muscle tension dysphonia: a randomized controlled trial. *J Speech Lang Hear Res*. 2019;62:272-282.
57. Casper JK, Murry T. Voice therapy methods in dysphonia. *Otolaryngol Clin North Am*. 2000;33:983-1002.
58. Behlau M, Pontes P, Vieira VP, Yamasaki R, Madazio G. Presentation of the comprehensive vocal rehabilitation program for the treatment of behavioral dysphonia. *Codas*. 2013;25:492-496.
59. Burg I, Meier B, Nolte K, et al. Selection of voice therapy methods. results of an online survey. *J Voice*. 2015;29:776.e1-776.e6.
60. Hosoya M, Kobayashi R, Ishii T, et al. Vocal hygiene education program reduces surgical interventions for benign vocal fold lesions: a randomized controlled trial. *Laryngoscope*. 2018;128:2593-2599.
61. Rodríguez-Parra MJ, Adrián JA, Casado JC. Comparing voice-therapy and vocal-hygiene treatments in dysphonia using a limited multidimensional evaluation protocol. *J CommunDisord*. 2011;44:615-630.
62. Meerschman I, Van Lierde K, Ketels J, et al. Effect of three semi-occluded vocal tract therapy programmes on the phonation of patients with dysphonia: lip trill, water-resistance therapy and straw phonation. *Int J Lang CommunDisord*. 2019;54:50-61.
63. De Bodt M, Patteeuw T, Versele A. Temporal variables in voice therapy. *J Voice*. 2015;29:611-617.
64. Kao YC, Chen SH, Wang YT, et al. Efficacy of voice therapy for patients with early unilateral adductor vocal fold paralysis. *J Voice*. 2017;31:567-575.
65. Patel RR, Awan SN, Barkmeier-Kraemer J, et al. Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association expert panel to develop a protocol for instrumental assessment of vocal function. *Am J Speech Lang Pathol*. 2018;27:887-905.
66. Roy N, Gray SD, Simon M, et al. An evaluation of the effects of two treatment approaches for teachers with voice disorders: a prospective randomized clinical trial. *J Speech Lang Hear Res*. 2001;44:286-296.
67. Roy N, Weinrich B, Gray SD, et al. Voice amplification versus vocal hygiene instruction for teachers with voice disorders: a treatment outcomes study. *J Speech Lang Hear Res*. 2002;45:625-638.
68. Roy N, Weinrich B, Gray SD, et al. Three treatments for teachers with voice disorders: a randomized clinical trial. *J Speech Lang Hear Res*. 2003;46:670-688.
69. Behrman A, Rutledge J, Hembree A, Sheridan S. Vocal hygiene education, voice production therapy, and the role of patient adherence: a treatment effectiveness study in women with phonotrauma. *J Speech Lang Hear Res*. 2008;51:350-366.
70. Kapsner-Smith MR, Hunter EJ, Kirkham K, Cox K, Titze IR. A randomized controlled trial of two semi-occluded vocal tract voice therapy protocols. *J Speech Lang Hear Res*. 2015;58:535-549.
71. Pedrosa V, Pontes A, Pontes P, Behlau M, Peccin SM. The effectiveness of the comprehensive voice rehabilitation program compared with the vocal function exercises method in behavioral dysphonia: a randomized clinical trial. *J Voice*. 2016;30:377.e11-377.e19.
72. Guzman M, Jara R, Olavarria C, et al. Efficacy of Water Resistance Therapy in subjects diagnosed with behavioral dysphonia: a randomized controlled trial. *J Voice*. 2017;31:385.e1-385.e10.
73. Rojas GVE, Carrijo MR, Ricz LNA, Pereira PM. Voice handicap index in elderly people submitted to the voice program for presbyphonia. *Int Arch Otorhinolaryngol*. 2017;21:83.
74. La Mantia I, Cupido F, Andaloro C. Vocal function exercises and vocal hygiene combined treatment approach as a method of improving voice quality in irradiated patients for laryngeal cancers. *Acta MedicaMediterranea*. 2018;34:517-523.

How to cite this article: Barsties v. Latoszek B, Watts CR, Neumann K. The effectiveness of voice therapy on voice-related handicap: A network meta-analysis. *Clin Otolaryngol*. 2020;45:796–804. https://doi.org/10.1111/coa.13596