University College of Osteopathy

Individual Enquiry

Research Paper 2020

Title: How effective are manual therapies and non-manual interventions in the treatment of functional voice disorders? A Scoping Review.

Author: Chloe Spencer, AVCM, BA, MSc

Supervisor: Dr Hilary Abbey, D.Prof (Ost), MSc, PG Dip, PG Cert, DO, FHEA

The University College of Osteopathy
275, Borough High Street, London SE1 1JE

How effective are manual therapies and non-manual interventions in the treatment of functional voice disorders? A Scoping Review.

Abstract

Aims:

This scoping review aimed to assess literature published between 2009-2019 on manual therapies for non-organic voice disorders, and to draw conclusions about the effectiveness of available treatments and the recommendations for further research.

Methods:

A scoping review was chosen to allow a broad identification and examination of the available literature. Searches were completed in PubMed, OVID Emcare, OVID MEDLINE(R) ALL, AMED, the Cochrane Central Register of Controlled Trials, Researchgate and the UCO and ESO records of past dissertations. Results were filtered in two Phases and charted according to the above framework. Methodological quality was acknowledged using a hierarchy of evidence.

Results:

The search produced nine primary studies on manual therapy with or without secondary intervention. These ranged from RCTs to case series. Results showed a wide spectrum of outcome measures used to assess pain, muscle tone and voice quality with little consensus among the studies. Clinically and statistically significant positive results were shown for laryngeal manual therapy, manual circumlaryngeal therapy and TENS but only three papers detailed effect size. There were significant limitations and omissions across the selected studies and an overwhelming female bias in the study populations which were all small.

Conclusion:

The results of this scoping review show that while there is positive evidence for the effectiveness of manual therapies for functional dysphonia, future studies should aim to include more rigorous RCTs, the implications of a predominantly female study population and the precise mechanism of each intervention. Protocols should be developed for assessment of muscular tone and implementation of manual therapies.

KEYWORDS: muscle tension dysphonia, laryngeal manual therapy, manual circumlaryngeal therapy, TENS

1. Introduction

This study presents an analysis of current research literature into the treatment of non-organic voice disorders, known as functional dysphonias. Voice disorders can be complex and challenging to diagnose and treat in clinical practice because there can be multiple types, many predisposing and maintaining factors and different treatment strategies (Phillips et al., 2005; Stachler et al., 2018). Studies on the effects of manual therapy are inconsistent and wide ranging in their designs, goals and outcome measures (Mathieson, 2011; Andreassen et al., 2017), so a scoping review has been chosen as an appropriate methodology to synthesise the current understanding of the different techniques available and the quality of the evidence for their effects (Arksey and O'Malley, 2005).

1.1 Basics of voice production – the correctly functioning voice

The correctly functioning voice works in three parts: the 'power' system which consists of the lungs and lower accessory breathing muscles, the 'source' of the sound, i.e. the larynx, and the 'filter', everything above the larynx which modifies the basic sound produced when air passes over the larynx (Shewell, 2009; Steinhauer et al., 2017; Dimon and Brown, 2018; Rubin, 2018b). When any one part of this mechanism malfunctions it will produce a change in sound. For example, inadequate breath for vocalising can cause excessive quietness and an audible pharyngeal tightening at the ends of sentences (Shewell, 2009; Rubin, 2018a). Excessive muscular tension in the larynx produces a 'squeezed' or pressed quality to the voice, resulting in, among other things, vocal fatigue and a reduction in volume (Shewell, 2009; Mathieson, 2011; Harris and Moisik, 2018). Problems in the 'filter' usually result in a reduction in resonance, articulation and volume and are usually addressed by speech and language therapists (Shewell, 2009; Gates et al., 2013; Harris, 2018a).

1.2 Voice disorders

A clinical practice guideline from the American Academy of Otolaryngology suggests that dysphonia can affect up to one in thirteen adults in the United States, and can present as hoarseness, vocal fatigue, difficulty in speaking and even pain on phonation (Carding, 2003; Harris and Howard, 2018; Stachler et al., 2018). The effect of dysphonia on quality of life is significant and a US study estimated a cost to the healthcare system of $13.5bn per year (Stachler et al., 2018). No comparative literature for the cost to the UK or other European healthcare systems could be found although a Cochrane review from 2007 suggested that up to 40,000 dysphonic patients per year present to voice clinics in the UK (Ruotsalainen et al., 2007).

Voice disorders can be classified into two main types: organic and functional (Carding, 2003). Organic disorders include any pathology caused by or associated with structural abnormalities in the larynx, systemic or neurological pathologies such as cancer, Parkinson's disease and infections (Bradley, 2010; Connor and Bless, 2013). Other disorders include those caused by functional or behavioural changes, which are often grouped together under the umbrella term 'muscle tension dysphonia' (MTD) (Bradley, 2010; Behlau et al., 2015) and can be subdivided into two further groups: primary and secondary (Mathieson, 2011; Garaycochea et al., 2019). Primary MTD is the result of muscular tension in the absence of any underlying pathology, and secondary MTD is muscular tension as a result of compensation for underlying pathologies (Mathieson, 2011; Harris, 2018a). The focus of this scoping review will be on interventions for primary MTD which is most suited to treatment with manual therapy (Rubin et al., 2000; Mathieson, 2011; Lieberman, 2018).

1.3 Manual therapies used for voice disorders

Treatment for MTD is predominantly carried out in voice clinics by a multidisciplinary team (Harris and Howard, 2018; Awad et al., 2019), and can range from voice therapy, exercises and manual therapy to psychological support and education (Mathieson, 2011; Andreassen et al., 2017; Stachler et al., 2018). Manual therapy techniques available include laryngeal manual therapy (LMT), manual circumlaryngeal therapy (MCT), laryngeal manipulation and transcutaneous electrical nerve stimulation (TENS). Other methods, often used in conjunction with manual therapy include breathing exercises, vocal hygiene education and vocal facilitating techniques based on speech therapy (Van Lierde et al., 2010; Aghadoost et al., 2019).

The aim of manual therapy is to relax the perilaryngeal musculature using a combination of direct massage of structures including the supra- and infra-hyoid muscles, sternocleidomastoid (SCM) muscles and the area surrounding the hyoid, with depression and displacement of the larynx to effectively stretch the hypertonic musculature (Rubin et al., 2000; Mathieson et al., 2009; Mathieson, 2011; Kennard et al., 2015). Early papers (1993, 1997) by Roy and Leeper showed rapid improvements in vocal function with manual therapy, but these studies used small study populations, new, unvalidated techniques and variable validity of outcome measures (Roy and Leeper, 1993; Roy et al., 1997). Jacob Lieberman is an osteopath recognised internationally for his work on laryngeal manipulation (Kennard et al., 2015; Cardoso et al., 2017). Unfortunately, there are no clinical trials of his specific techniques in the available literature. However, his work informs the basis of the manual laryngeal therapy methods used in the selected studies (Mathieson et al., 2009; Reimann et al., 2016; Siqueira et al., 2017; Conde et al., 2018).

1.4 Osteopathic relevance

Osteopaths are often part of the multi-disciplinary voice clinic (Harris & Howard, 2018) and may also see and treat patients with voice disorders privately. The British Voice Association (British Voice Association, 2019), British Association of Performing Arts Medicine (BAPAM, 2019) and the Osteopathic Performing Arts Care Association (OPACA, 2019) all support the use of manual therapy for the voice. Even where a patient is not presenting with a vocal dysfunction as the primary reason for a consultation, the effect of dysphonia on quality of life may be important (Ramos et al., 2018), and is something that osteopaths should be able to assess and treat (Lieberman, 2018). Clear guidelines for the assessment and treatment of voice disorders using osteopathic techniques is therefore undoubtedly indicated.

2. Methods

The breadth of the research question was most suited to a scoping review, to enable a wide analysis of the heterogenous literature on the subject (Arksey and O'Malley, 2005; Levac et al., 2010). The methodology for this scoping review was based on Arksey & O'Malley's five-step methodological framework (2005, see Appendix I for a summary) which was further refined by Levac et al (2010) and Daudt et al (2013). The resulting steps below also take into account the PRISMA-ScR Checklist guidelines from the PRISMA Extension for Scoping Reviews (Tricco et al., 2018) (Appendix II).

Step 1: identifying the research question

This scoping review was conducted to answer the following research questions:

1. How effective are laryngeal manual therapy techniques used to treat functional and behavioural dysphonia?
2. How effective are they in comparison to other non-manual approaches?
3. What recommendations can be drawn from the results?

Step 2: identifying relevant studies

The search strategy was conducted using the following databases: PubMed, OVID Emcare, OVID MEDLINE(R) ALL, AMED, the Cochrane Central Register of Controlled Trials, Researchgate and the UCO and ESO records of past dissertations. Searches were restricted to those databases accessible without payment (CINAHL could not be included). Databases were searched from 2009 up to September 2019 (within the last 10 years to cover the newest research). A combination of the following terms was used (MESH and non-MESH):

Database	Search Terms (MESH & non-MESH)
PubMed	“laryngeal manual therapy” AND “dysphonia”; “circumlaryngeal” and “dysphonia”; “laryngeal manual therapy” AND “voice disorders”; “functional dysphonia”; “behavioural dysphonia”; “manual therap*” AND “dysphonia”; “manual therap*” AND “dysphoni*”; “laryngeal” AND “voice”; “laryngeal” AND “vocal”
OVID Emcare, OVID MEDLINE(R) ALL, AMED	“laryngeal manual therapy” AND “dysphonia”; “circumlaryngeal” and “dysphonia”; “laryngeal manual therapy” AND “voice disorders”; “functional dysphonia”; “behavioural dysphonia”; “manual therap*” AND “dysphonia”; “manual therap*” AND “dysphoni*”
Cochrane	“laryngeal manual therapy”; “dysphonia”
ResearchGate	“laryngeal manual therapy” AND “dysphonia”; “circumlaryngeal” and “dysphonia”; “laryngeal manual therapy” AND “voice disorders”; “functional dysphonia”; “behavioural dysphonia”;
UCO, ESO	“dysphonia”; “laryngeal”; “voice”; “vocal”; “MTD”; “LMT”

Table 1: Search Terms

Additional papers were found through hand searching the Journal of Voice and the references from selected and non-selected papers. Search results were exported to Excel to further filter and remove duplicates. Online searches were completed between September 2019 and January 2020.

Step 3: study selection, inclusion and exclusion criteria

For inclusion, the primary language of the study had to be English (unless an English translation was available), human-based, within the last 10 years and had to be a primary study on laryngeal manual therapy for dysphonia (functional, behavioural or muscle tension).

Inclusion	Exclusion
• Primary language: English (or English translation available • Studies on humans • Within 10 years (2009 – present) • Primary studies • Primary studies on manual therapies for muscle tension dysphonia (MTD), functional dysphonia (FD), behavioural dysphonia (BD) • Primary studies on combined therapies (including manual therapy) for MTD, FD, BD • Unpublished studies	• Non-english language studies (except where a full translation is available • Secondary studies (systematic reviews, qualitative studies or opinion pieces) • Studies on organic dysphonia • Studies on asymptomatic populations (except where used as a control with symptomatic patients)

Table 2: Inclusion & Exclusion Criteria

Further filtering involved selecting only primary studies, then selecting studies which were focused to the research questions. These included only studies which assessed the effects of laryngeal manual therapies either alone or in combination with other approaches, either manual or non-manual. The population of studies was restricted to dysphonic patients, so studies on asymptomatic individuals (except where used as a control) were rejected.

Owing to this being a student project, it was not possible to use a panel of researchers to screen and select papers as recommended in the scoping review framework (Arksey and O'Malley, 2005; Levac et al., 2010; Daudt et al., 2013). Therefore, whilst acknowledging this factor as a potential limitation, the researcher was responsible for the entire selection process which was conducted in two phases: Phase I identified all papers matching the search terms by title and abstract, which were then filtered for duplicates using an Excel spreadsheet. Phase II further filtered the results by title and abstract into studies on laryngeal manual therapy (LMT) or manual circumlaryngeal

therapy (MCT) with or without a comparative intervention, and then into primary studies only.

Step 4: charting the data

Data was extracted using headings similar to those suggested by Arksey & O'Malley (2005) but modified to engage with the research question of this review (Arksey and O'Malley, 2005). The limitations of this study prevented secondary testing of the data extraction table, but this would ideally have been done with a small number of randomly selected papers from phase II (Levac et al., 2010; Daudt et al., 2013; Tricco et al., 2018).

In line with the methodology and purpose of scoping reviews, the quality of each study was not formally assessed using named checklists to assess methodological quality and risk of bias (Arksey and O'Malley, 2005) but the relative quality of the studies was noted in reference to the hierarchy of research evidence (figure 2, below) (Evans, 2003; Hoppe et al., 2009; Greenhalgh, 2014). A limitation of a scoping review is that while allowing a broader range of literature to enable a more generalised study of the subject, there is some risk of the results of scoping reviews being less applicable or relevant owing to the omission of formal scoring by checklists (Levac et al., 2010).

Step 5: collating, summarising and reporting the results

Following the data extraction and charting, the studies were classified into themes or patterns (Arksey and O'Malley, 2005; Levac et al., 2010). An initial synthesis was performed to organise the data and identify patterns. Questions which arose from detailed analysis of the information are detailed in the results (Popay et al., 2006).

Ethical Approval

Ethical approval for this scoping review was granted by the University College of Osteopathy Research Ethics Committee.

3. Results

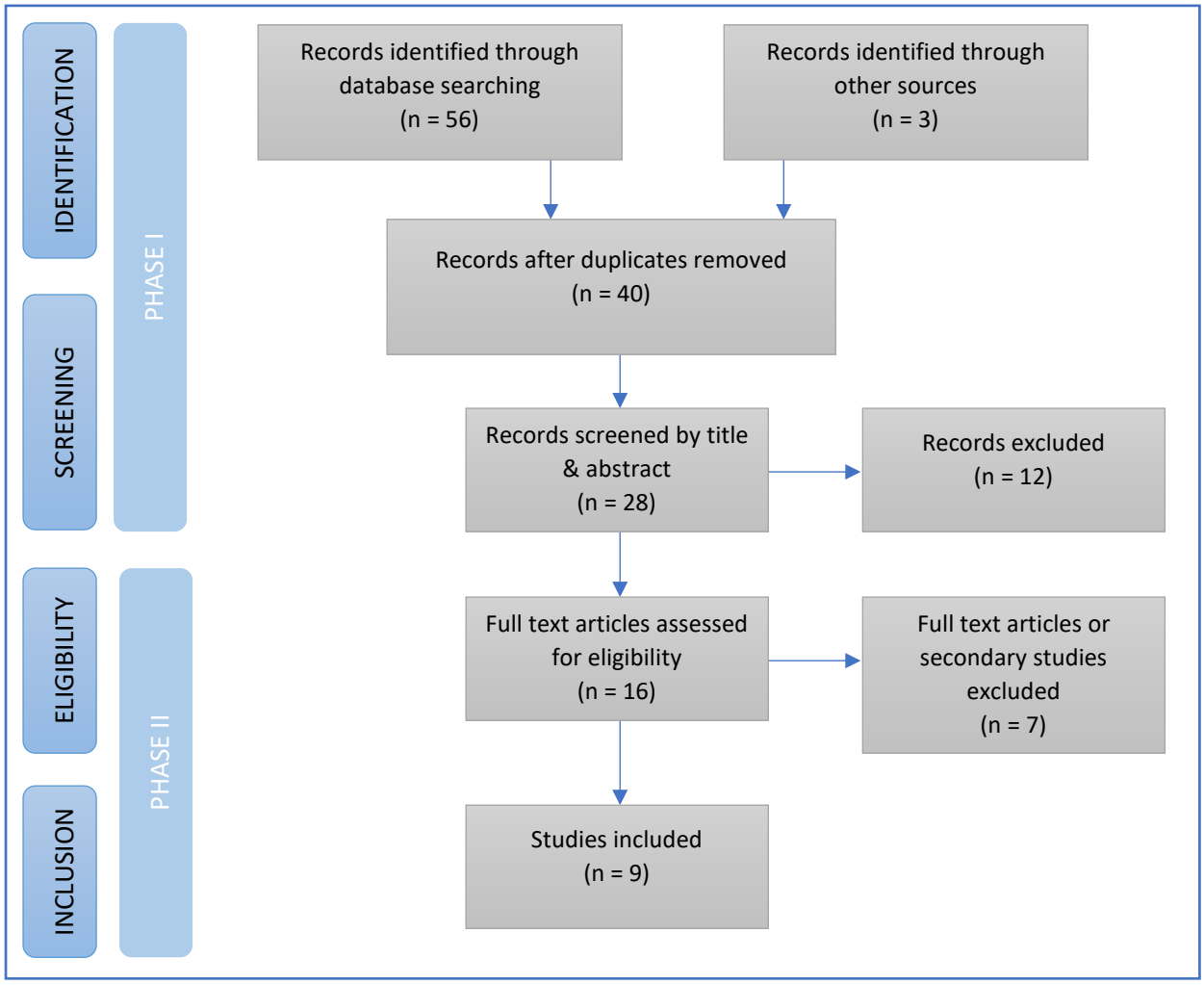

Figure 1: PRISMA flow diagram for selected papers

After Phase I, 16 papers were identified. Of those that were filtered out due to not being primary studies, two were systematic reviews, one a Cochrane review (Ruotsalainen et al., 2007), and one a systematic review with meta-analysis (Ribeiro et al., 2018). Both scored highly on the AMSTAR 2 checklist ((Shea et al., 2017)

Appendix II)), although the Cochrane review was completed in 2007 and assessed studies prior to the selected time period for this paper. Ribiero et al's 2018 review was highly specific, selecting only three papers of which only two were subject to meta-analysis. This further supports the issue of heterogeneity of the available literature. Nine papers matched the selection criteria after Phase II filters were applied. These papers were all primary studies of laryngeal manual therapies, with or without secondary comparative interventions.

The selected studies were assessed for methodological quality using the hierarchy of evidence (Evans, 2003) and ranked as can be seen in Figure 2. Four studies were randomised clinical trials (RCTs) of varying degrees of quality (Alves Silverio et al., 2015; Siqueira et al., 2017; Conde et al., 2018; Aghadoost et al., 2019), two were cohort studies of which one was a pilot study (Mathieson et al., 2009; Van Lierde et al., 2010), one was a case-control study (Reimann et al., 2016) and two case series (Tomlinson and Archer, 2015; Dehqan and Scherer, 2018). These last three study designs are usually regarded as having lower weighting in the hierarchy of evidence (Concato et al., 2000; Evans, 2003) but have been shown to have clinical value (Concato et al., 2000; Evans, 2003; Hoppe et al., 2009).

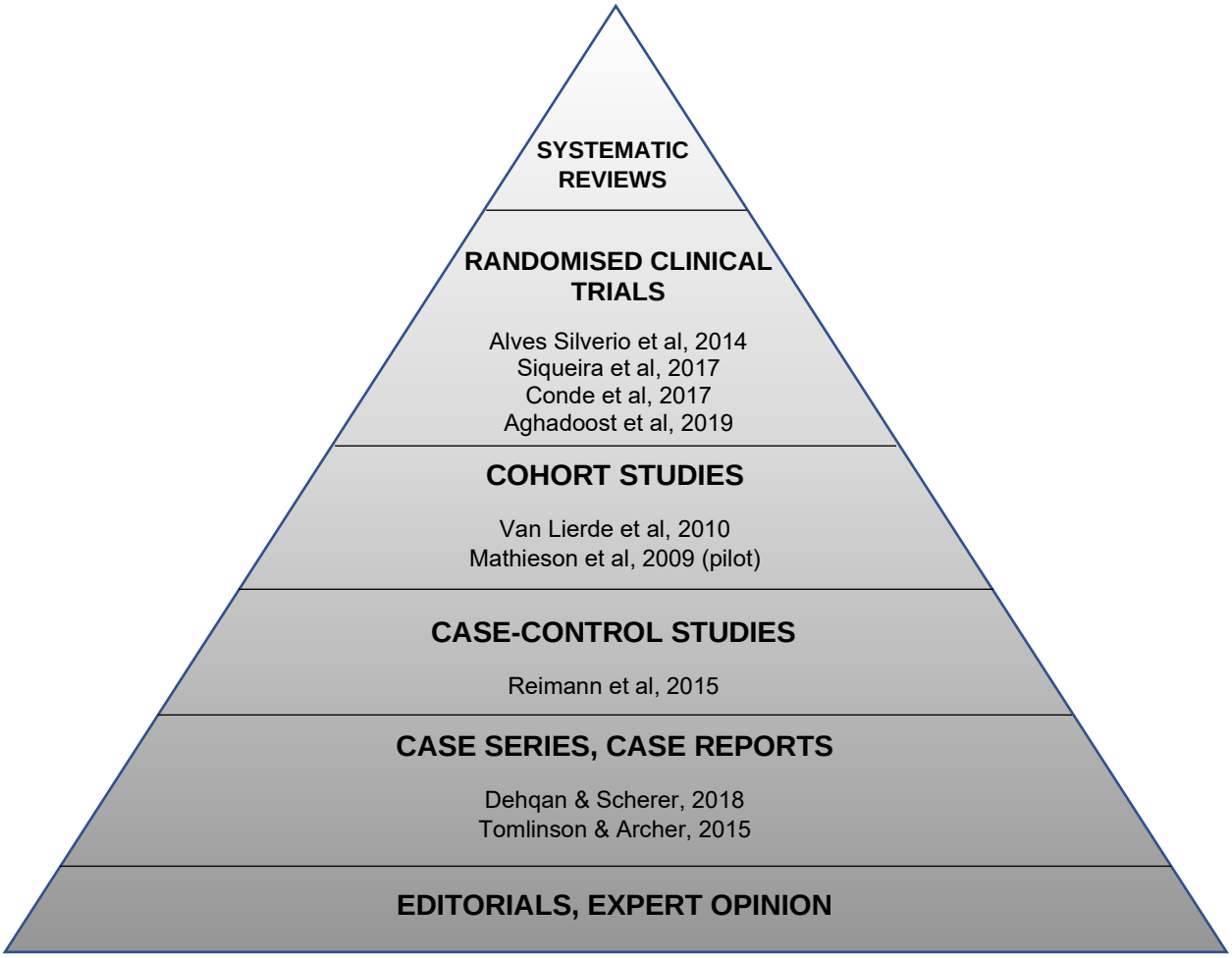

Figure 2: Hierarchy of Evidence for Research, based on Evans, 2003

Data analysis of the selected papers was performed following the Arksey & O'Malley framework for scoping reviews (2005) and the results can be seen in Table 3.

Pattern analysis identified several questions common to the selected studies:

- What different measures are used within and between the selected studies? Why is there such a range of measures used?
- Are interventions that focus on improving voice quality assessed differently from interventions that focus on reducing pain or physical/muscular symptoms?
- Why is there a predominance of female patients in the selected studies?

- What different treatment times and follow ups are used, and is the number of treatments and the length of treatment time and follow up significant?
- What limitations and omissions did the studies have in common and how important were they to the results?

Following this questioning process, and using the guidance in Popay et al (2006), the findings were reviewed and grouped into the following themes for discussion:

- Outcome measures
- Mechanism of interventions
- Female gender predominance in the literature
- Number of interventions and follow-up period
- Limitations, omissions and recommendations

Table 3: Data Charting & Analysis

Authors & Study Type	Sample & Aims	Intervention and Follow-up	Outcome measures	Statistically Significant Results	Conclusions & Further Research
Siqueira et al (2017) Randomised clinical trial	**N=20 (20F)** symptomatic: randomly allocated 10 x intervention 1; 10 x intervention 2; no control **Aim:** To verify and compare the effect of TENS and LMT on laryngeal DDK in dysphonic women	TENS or LMT 12 x 20 minute ttt, x 2 per week **FU:** 6 weeks	DDK rate: average DDK rate, repetition rate, SD of DDK period; DDK period variation coefficient; disturbances of DDK period; DDK intensity peak variation coefficient	Significant results post-LMT ttt in period (p=0.041), period variation coefficient (p=0.040) and peak variation coefficient (p=0.032). No significant results post-TENS ttt in any measure.	Conclusion: LMT provides greater regularity of movement but TENS doesn't affect it. No effect size stated, small sample size, no other outcome measures (e.g. pain, mm tension, other dysphonia measure). Recommendations: more investigation into neuromotor behaviour of vf in different mass lesions, effect of different exx and ttt types, long term effect of interventions.
de Cassisa Macedo Conde et al (2017) Randomised clinical trial	**N=30 (30F)** symptomatic; randomly allocated 15 TENS and 15 LMT. No control **Aim:** To verify the immediate effect of low-frequency TENS and LMT in MSK pain, voice quality & self reported signs in dysphonic women.	TENS: 20 minute single session, no vocalisation, pt supine; LMT: 20 minute single session, no vocalisation, pt in chair. 1 x 20 minutes ttt **FU:** Immediate	MSK Pain questionnaire, vocal assessment (recorded), auditory-perceptual analysis (3 double blinded specialists in SLP), Acoustic analysis (f0, jitter, shimmer, NHR); p≤0.05	TENS: 'significant' decreased pain post/ant neck, shoulders, u/l back, masseter (p=0.012-0.043); LMT: 'significant' decreased pain in post. neck, shoulders, l back, temporal area (p=0.012-0.028); instability significantly improved after TENS (p=0.031); Strain significantly improved after LMT (p=0.001); no differences in acoustic parameters between LMT/TENS;	Conclusion: TENS better immediate results than LMT. Follow up of 1 week or more needed. Small sample size, some justification for female only-population, no effect size given, no referral to contextual factors. Recommendations: further acoustic analysis/research into action at glottis/vocal tract; adjust timings of LMT & TENS ttts; placebo/control group needed.
Alves Silverio et al (2014) Randomised clinical trial	**n=20 (20F)** randomly allocated Group 1: TENS (10) Group 2: LMT (10) Dysphonia and bilateral vocal fold nodules	TENS & LMT (modified Mathieson technique with no vocalisation) 20 min, 2 x week	Vocal & laryngeal symptoms: MSK pain (NMSQ) & VAS; vocal quality - auditory perceptual analysis (3	TENS: 'high pitch/effort to speak' p=0.023, lower frequency of pain in post. neck/shoulders p=0.033-0.038, AP analysis showed only improvement in 'strain'; LMT produced improvement in 'sore throat'	Conclusions: TNS & LMT used in conjunction addresses more issues. Unvalidated outcome measures for VQ, no objective measurement of

	Aim: Compare effects of TENS and LMT on vocal/laryngeal symptoms, pain & voice quality	12 sessions **FU:** 6 weeks	blinded specialist judges); p≤0.05	p=0.045 and significantly lower pain in ant. neck (pain intensity in post. neck also reduced) p=0.019. No significant change in acoustic parameters in either ttt.	contextual factors, small sample size, no control, modification of ttt time. Recommendations: RCTs, application of voice hygiene/education alongside TENS/LMT, also need for speech therapy/vocal training and guidelines.
Aghadoost et al (2019) Randomised clinical trial	**N=16 (16F)** symptomatic teachers; mean age 38.6yrs; randomly allocated 8: VFT and 8: MCT; no control **Aim:** To compare the effect of VFT and MCT in teachers with MTD	MCT; VFT: chewing, respiration training, yawn-sigh, open-mouth, loudness variation, glottal fry, chant talk 10 x 45 minutes ttt, x 2 per week **FU:** 5 weeks	Vocal Handicap Index (VHI - voice-related quality of life); Dysphonia Severity Index (DSI): highest frequency, lowest intensity, max phonation time, jitter.	Effect sizes large in all of the following measures: Within group: VHI physical: MCT = 0.92, VFT = 0.73 VHI emotional: MCT = 0.92, VFT = 0.83 DSI: MCT = 0.38, VFT = 0.90 Between group: VHI physical: MCT = 0.51, VFT = 0.31 VHI emotional: MCT = no data, VFT = 0.44 DSI: MCT = 0.30, VFT = 0.33	Conclusions: Both ttts showed significant improvements in VHI and DSI. Physical improvement greatest on VHI after MCT and on DSI after VHI. Both ttts need to be used together to address both outcome measures Recommendations: larger sample size, combined vs individual techniques on teachers with MTD, blind study.
Van Lierde et al (2010) Cohort Study	**N=10 (4F, 6M)** 18-65, dysphonic + increased laryngeal mm tension; no control **Aim:** To measure the effectiveness of two treatment techniques - vocalisation with abdominal breath (V&AB) support and MCT.	V&AB: single ttt, 45 mins, info, identification, breathing without phonation, breathing with phonation; MCT (Aronson & Roy): hyoid, thyroid cartilage, larynx with sustained vowels (no manipulation/ reposturing) 1 x 45 minutes ttt **FU:** immediate	DSI (dysphonia severity index): max phonation time (MPT), highest frequency, lowest intensity, jitter; p=0.05	Intake -post-ttt 2: voice intensity p=0.05, shimmer p= 0.05, DSI p=0.003 Intake -post-ttt 1&2: aerodynamic MPT p=0.05, voice intensity p=0.007, voice frequency (high) p=0.05, jitter p= 0.05, shimmer p=0.05, DSI p=0.001	Conclusions: Greatest difference from intake to after MCT (p<0.001), and from abdominal support to after MCT (p=0.003). Limitations: small sample size, no effect size calculated, no consideration of contextual factors, no control, no randomisation. Recommendations: research into precise mechanism of MCT effect on vocal quality; duration of positive effects of MCT, is there a late effect response to breathing support training? EMG recordings.

Mathieson et al (2009) Cohort Study (pilot study)	**n=10 (8F)** All symptomatic **Aim:** Assess acoustic and outcome measures for the evaluation of LMT methods	LMT only: bimanual circular massage of SCM, hyoid and supralaryngeal mm, manual depression of larynx 1 x 20 min ttt **FU**: immediate and after 1 week	Self-reported VTD, Relative Average Perturbation (RAP), Formant frequency, Noise-to-harmonics (NHR) ratio, Soft-phonation index (SPI), Perturbation irregularity (PI), Muscle resistance	RAP significant difference p=0.02 Formant frequency = inconclusive Effect sizes: F_2 mean/variance = 0.344/0.203 (large); RAP = 0.450 (large); NHR = 0.242 (large); SPI = 0.262 (large), PI = 0.274 (large); DQx 1&2 = 0.184 (large) VTD: significant change in symptom frequency & severity post LMT (imm & 1 week) effect size = 0.222-0.749 (large) Palpatory changes large effect size = 0.894-0.988	Conclusions: VTD scale was a useful evaluation tool for measuring positive effects of LMT, which were clinically significant acc. to effect sizes. Recommendations: Pilot study, identified that further investigation needed into formant frequency post-ttt, effect of forced lowering (vs raising) of larynx in small number of dysphonic pts; identification of dysphonic subgroups - high and low held larynxes. Also, palpatory evaluation protocol needed, larger sample sizes, better descriptive terminology for non-English speakers
Reimann et al (2015) Case Control	**n=30 (24F)** 18-45 years Dysphonic Group (DG) n=15, symptomatic; Control Group (CG) n=15 asymptomatic **Aim:** Assess effect of LMT on pain, vocal quality, physical sensations	LMT only: 5 mins massage on SCMs & suprahyoid 3 mins massage on SCM & suprahyoid 2 mins manual sliding of larynx & displacement mvt of thyroid region 1 x 20 min ttt **FU**: Immediate	Pain: MSK pain questionnaire; Voice Quality (VQ): Auditory-perceptual analysis (3 double blind specialists in SLP); Acoustic analysis incl. f0, jitter, shimmer & NHR	DG: significant reduction of pain in temporal, larynx, post. neck, UEX, upper/lower back hip/thigh P=0.005-0.036 DG: Increased roughness post ttt; Jiitter sig. reduction (p=0.033) DG reported better self-report sensations post-ttt (p=0.016-0.039 for larynx & articulation) No diff. in auditory-perceptual analysis post-ttt; speech analysis	Conclusion: LMT reduces pain in dysphonic pts; although 'roughness' reported, msk/pain sensations still improved so still a valid technique. No control, small sample size, multiple outcome measures (some unvalidated). Recommendations: controlled, randomised, blinded studies are needed to assess LMT in association with other interventions, the imm. effects of LMT on roughness, length of ttt, longer FU.
Dehqan et al (2018) Case Series	**n=28 (28F)** 18-40 years Primary MTD (min 6mth prior to diagnosis) **Aim:** Verify 6 months effects of MCT in the ttt of MTD	MCT w phonation 15 x 30 min ttts 30 min, 3 per week, Same practitioner **FU:** Immediate and after 6 months	Pre/post-ttt recordings of sustained vowels, selected sentences and connected speech samples AP & acoustic analysis	Significant improvements in all measures except fundamental frequency F0 F1 p=0.008, ES n2=0.24 (large); 6 mth p=0.02; Jitter p=0.001, ES n2=0.33 (large); 6 mth p=0.004; Shimmer p=0.008, ES n2=0.36 (large); 6 mth p=0.009; HNR p=0.006, ES n2=0.25, 6 mth p=0.02	Conclusion: MCT can lead to positive clinically significant changes in ttt of primary MTD. Case series so lower evidence level, small sample, no control, some unvalidated outcome measures, no objective means of measuring laryngeal positioning; results dependent on clinicians' expertise.

				Subjective CAPE-V ratings improved in all patients (p=0.001). Changes were sustained at 6-months	Good measurement & reporting of effect sizes. Recommended the development of protocols for using MCT to manage MTD
Tomlinson and Archer (2015) Case Series	**N=9 (9F)** symptomatic **Aim:** To determine whether physical therapy, exercise and stress management would be beneficial in reducing excess MTD in a symptomatic sample	Contract-relax home exx, csp stretches, self LMT (Mathieson modified); ergonomic & postural education; 30 mins laryngeal, csp, scapula, TMJ, tongue , resp mm and hip flexor MT; 20 mins exx (undefined); 10 mins education stress management & relaxation; 9 x 60 minutes ttt (split), x 2 per week **FU** = 9 weeks FU	Numerical rating scale (NRS), Patient-Specific Functional Scale (PSFS), Voice Handicap Index (VHI), csp/tmj ROM by goniometric measurement.	8/9 pain free after intervention, 9: improved PFSF (7 clinically meaningful **no statistical analysis performed**); 3 clinically meaningful VHI improvement; 9: increased csp flex, tmj lateral flex/jaw opening; 8/9 inc. csp ext & rot	Conclusion: vague statement that physical therapy might be valuable for MTD patients. Severe limitations reduce the quality of this study: case series, no statistical analysis performed, small selective sample, no consideration of contextual factors (acknowledged), no follow up, generalised findings, no blinding. Recommendations: RCT needed to further test hypothesis.

4. Discussion

The aims of this scoping review were to assess the effectiveness of a range of manual therapies in the treatment of muscle tension dysphonia, and to determine where further research should be directed. The results in Table 3 show that there was evidence of clinically and statistically significant benefits from LMT, MCT and TENS for reduction of pain, improvement of dysphonic symptoms and voice quality, with some variations depending on the outcomes being assessed. It can be seen in studies which used two interventions (LMT/MCT and TENS, LMT/MCT and vocal facilitating techniques or breathing techniques) that results showed larger effect sizes where interventions were combined than when used in isolation (Van Lierde et al., 2010; Alves Silverio et al., 2015; Aghadoost et al., 2019). Studies which focused solely on one treatment modality (LMT, MCT) showed some mixed results with both pain reduction and improvements in voice quality and dysphonic symptoms (Mathieson et al., 2009; Dehqan and Scherer, 2018) but with one study showing a negative result for voice quality (Reimann et al., 2016).

4.1 Outcome Measures

Outcome	Outcome Measure (Tool/Scale)	Validation	Paper
Pain	MSPQ, NRS, NMSQ, VAS	Validated	Tomlinson 2014, Silverio 2014, Reimann 2015, Conde 2017
Voice Quality	AP analysis, formant frequency	Unvalidated	Mathieson 2009, Silverio 2014, Reimann 2015, Conde 2017
	CAPE-V, MSP	Validated	Siqueira 2017, Dehqan 2018
Dysphonia (specific)	VTD, DSI	Validated	Mathieson 2009 (unvalidated), Van Lierde 2010, Aghadoost 2019
Muscular Tone/ROM	Goniometric evaluation	Validated	Tomlinson 2014
	Palpation	Unvalidated	Mathieson 2009
Voice-related QoL	VHI, PSFS	Validated	Tomlinson 2014, Aghadoost 2019

Table 4: Outcome Measures

The papers selected for this study use a variety of outcome measures, both validated and unvalidated, to assess pain, voice quality, muscular tone and range of motion, and voice-related quality of life (see Table 4). While all the selected papers sought to quantify the effectiveness of laryngeal manual therapies, with or without comparative therapies, there are several different approaches. Measurement of pain and/or muscular tension using a validated outcome measure was only carried out in five of the nine papers. Muscle Tension Dysphonia implies impairment of voice due to muscular tension. Four studies sought to analyse the effects of LMT/MCT on the voice but failed to use a validated tool to assess muscular tone or pain/symptoms before or after treatment (Van Lierde et al., 2010; Siqueira et al., 2017; Dehqan and Scherer, 2018; Aghadoost et al., 2019).

It can be seen from the results that there is a wide variation in the scales or tools used to evaluate voice quality before and after treatment. Validated tools to measure voice quality include the GRBAS protocol (grade, roughness, breathiness, asthenia, and strain) (Hirano, 1981), CAPE-V (Consensus Auditory Perceptual Evaluation—Voice) (Nemr et al., 2012), Cepstral Spectral Index of Dysphonia (CSID) (Awan et al., 2016), Acoustic Voice Quality Index (AVQI) (Faham et al., 2019) among others (Kreiman and Gerratt, 2010; Nemr et al., 2012; Faham et al., 2019). Most of these have been demonstrated to be both effective and show a general consensus (Kreiman and Gerratt, 2010), with CAPE-V, a development of the GRBAS scale (Shewell, 2009) being the most widely adopted tool internationally (Chen et al., 2018; Khoramshahi et al., 2018; de Almeida et al., 2019; Ertan-Schlüter et al., 2019). It is surprising therefore that only one study uses CAPE-V and none use the GRBAS scale.

Only two studies assess voice related quality of life with a validated outcome measure (Mathieson et al., 2009; Tomlinson and Archer, 2015). Since MTD can have a strong

behavioural component (voice misuse/abuse, anxiety, stress levels, profession, bereavement) the lack of assessment of these factors in the selected studies is a notable omission, and will be discussed further at a later point in this review. Voice related quality of life can have a demonstrable impact on MSK pain levels (Hogikyan and Sethuraman, 1999; Ramos et al., 2018), so is another important measure which is missing from the majority of the selected studies.

4.2 Mechanism of Interventions

The interventions used in the selected studies range in invasiveness from breathing techniques to deep palpation. The intention of these different interventions varies across the studies, as has been discussed with reference to the outcome measures. The common factor to all treatments is the assessment of voice before and after treatment – another reason why a standardised assessment method would be helpful in future studies.

LMT and MCT differ only in that in MCT the patient phonates during treatment. The clinical reasoning for this difference is that the practitioner is able to assess changes to the voice during treatment can therefore adjust the treatment accordingly (Mathieson, 2011). Justification for postponing phonation to after treatment is to allow maximum relaxation of the perilaryngeal musculature and the patient to phonate at an optimally balanced, pain free point (Mathieson et al., 2009; Mathieson, 2011). Only two studies assessed muscular tone prior to treatment and it is acknowledged that palpatory evidence is both subjective and lacking in standardisation of protocol (Mathieson, 2011; Woźnicka et al., 2017; Davidson et al., 2020).

A 2017 study by Woźnicka, based on the work of Lesley Mathieson and Jacob Lieberman (an osteopath specialising in laryngeal manipulation) aimed to create a

protocol for palpatory diagnosis of dysphonia (Rubin et al., 2000; Mathieson et al., 2009; Woźnicka et al., 2017). The Laryngeal Manual Therapy Palpatory Evaluation Scale (LMTPE) has not been widely adopted in the literature to date, but would be a useful addition to future studies combining assessment of muscular tone and voice quality.

Transcutaneous electrical nerve stimulation (TENS) is another technique aimed at reducing muscular tension. None of the three selected studies which compared TENS to LMT or MCT attempted to assess muscular tone through palpation, but relied on patient reported symptoms of discomfort, stiffness or pain before and after treatment (Alves Silverio et al., 2015; Siqueira et al., 2017; Conde et al., 2018). A 2019 Cochrane systematic review found that the effectiveness of TENS for the treatment of chronic neck pain found little or no evidence for TENS over sham treatment (Martimbianco et al., 2019), and similarly a 2015 review found that TENS treatment for acute pain showed only tentative evidence (Johnson et al., 2015). The continued use of TENS in clinical studies, despite the lack of evidence for its effectiveness, suggests that greater attention should be directed in future studies to the use of clinically proven interventions.

Vocal facilitating techniques (VFT), otherwise known as voice therapies, are already widely used in voice clinics with good results (Craig et al., 2015; Harris, 2018a; Awad et al., 2019; LeBorgne and Donahue, 2019). However, recent studies indicate the need for a wider range of techniques to incorporate the multi-dimensional nature of MTD, which supports the study by Aghadoost et al (2019). Voice therapy conducted by speech and language therapists (SLTs) aims through vocal and physical exercises to minimise constriction in the larynx, optimise flow of air and improve resonance in the articulators (Harris & Howard, 2018). The study by Aghadoost et al (2019) shows that

voice therapy is effective against severity of dysphonia, but that an extra element is needed to address the physical symptoms (i.e. MCT), the conclusion being that these therapies used together are most effective. Since this study used only a small sample of symptomatic female patients and the intervention was carried out by a student, the results must be regarded with caution. A reasonable progression would be for more robust RCTs to assess the combined elements of VFT and LMT/MCT and to inform the construction of a standardised protocol.

4.3 Female Gender Predominance in the Study Populations

The combined populations of the nine selected studies showed a large female predominance (159 out of 173 = 91.91%). Of these only 12 women made up an asymptomatic control in one study. The very high predominance of women in dysphonia studies has not been widely acknowledged and the literature on gender bias in the general dysphonic population is minimal (Hunter et al., 2011; Korn et al., 2018). There are some studies on gender and dysphonia among teachers and these all show a higher proportion of female teachers reporting vocal problems (van Houtte et al., 2012; Korn et al., 2018; Abou-Rafée et al., 2019). It is important to note that teaching appears to be vocally a high-risk occupation, particularly for women, and this has been examined in the papers mentioned above.

Other vocally high-risk occupations mentioned in the literature include singers, actors, fitness instructors, clergy, hospitality workers and telemarketers (van Houtte et al., 2012; Benninger et al., 2017; Remacle et al., 2017; Korn et al., 2018; Phyland and Miles, 2019). Two factors would account for the female bias: a higher proportion of female workers in these industries (Phyland and Miles, 2019) and the gender differences which may act as risk factors (Hunter et al., 2011). This second point has

implications for those treating dysphonic voices in that a complete understanding of the differences between male and female voices, both anatomically and physiologically, is important.

Anatomical and physiological factors which have significant impact on the female voice include the size of the larynx, which in adult females is approximately 20% smaller than in adult males (Hunter et al., 2011). This difference alone means that the female fundamental frequency (f0) is higher than that of males (225Hz (female) vs 120Hz (male) (Howard, 2018)) resulting in more frequent oscillations of the vocal folds and consequently a higher risk of injury (Hunter et al., 2011; Howard, 2018). The post-pubertal changes in the female endocrine system also affect the voice. The monthly menstrual cycle in women not taking contraceptive medication causes fluctuations in the oestrogen and progesterone levels (Hari Kumar et al., 2016; Pavela Banai, 2017). Oestrogen causes a retention of fluid, laryngeal oedema, which affects the mucosal cover of the vocal folds, resulting in vocal fatigue and loss of higher harmonics (Hunter et al., 2011; Hari Kumar et al., 2016; Harris, 2018b). Conversely, progesterone causes a drying effect which can lead to hoarseness and difficulty phonating (Hari Kumar et al., 2016). The impact of contraceptive medication is important but outside the scope of this study.

The monthly cycle can also have a significant psychological impact, creating mood swings, anxiety, adding to perceived stress levels and reducing quality of life (Armour et al., 2019a, 2019b). This monthly cycle of both physical and psychological factors, combined with the high vocal risk of the occupations in which women are more prevalent, could account for the high predominance of women in the studies. It is surprising therefore that even in the selected studies which used female only populations (five out of nine), all excluded women post-menopause to eliminate the

variable of muscular tone changes, but only one excluded patients who were currently menstruating (Aghadoost et al., 2019). Increased focus on the effects of the menstrual cycle on MTD in future studies would add to the understanding of this factor.

The occupational vocal risk for women in the workplace can range from a need for constant projection and stressful work environments, to a perceived inequality based on voice (van Houtte et al., 2012; Remacle et al., 2017; Neemuchwala, 2018; Phyland and Miles, 2019). Owing to the smaller female larynx there is a higher tendency for women to over-project in loud situations, such as fitness studios or noisy classrooms, known as the Lombard Effect (Bottalico et al., 2018). Only two of the nine selected studies in this review include occupation as a variable (Siqueira et al., 2017; Aghadoost et al., 2019). Given the high predominance of dysphonic women and the various vocal load risk of female-predominant occupations, this is a factor which merits closer scrutiny.

4.4 Intervention Length & Follow-up Period

Results in Table 3 show, like the outcome measures, a wide heterogeneity between the papers in terms of intervention time, number of sessions and follow-up time. The shortest intervention time of 20 minutes is used by those studies selecting LMT as one of the interventions and is based on Mathieson et al's 2009 study (but increasing the time to 20 minutes instead of 10). Only one study gives no justification for doubling the treatment time (Reimann et al., 2016)– in other cases it is in order to match the comparative intervention (TENS) (Alves Silverio et al., 2015; Siqueira et al., 2017; Conde et al., 2018). As previously stated, the research into the application of TENS shows at best tentative evidence for its effectiveness, therefore further research is

warranted into the optimum treatment time for both TENS and LMT in order to produce a rigorous comparative test.

The follow-up times selected by the studies show a discrepancy in the aims of the papers. Mathieson et al's 2009 study shows a large effect size for both the immediate effects of the intervention and the effect after one week (Mathieson et al., 2009). Aghadoost et al (2019) and Dehqan & Scherer (2018) also show large effect sizes after 5 weeks – 6 months of treatment. Where the effect size is not given, the significance (p value) of the results are used as a measure and these are, with small exceptions (Reimann et al., 2016) shown to be significant in most of the selected studies. However, recent research has shown a lack of reporting of effect sizes in biomedical literature which limits the clinical significance of the results (Lantz, 2013; Karadaghy et al., 2017; Vila et al., 2017). Of the selected studies in this review, six were performed within the last 5 years and of those only two describe effect size.

A significant omission in all the selected studies is any justification for the length of treatment plan or the follow up, with the exception of Dehqan & Scherer (2018), where 6 months is selected as 'long term'. Given that the effects of interventions can be seen to be significant after 1 treatment, none of the selected studies explains why a course of 10 or 12 treatments was chosen. The studies which selected 9 or 15 treatments were both case series studies and therefore ranked as weak evidence. However, further explanation of the selected treatment times in the RCTs is warranted to strengthen their evidential value.

4.5 Limitations, Omissions & Recommendations

Several limitations were common to all nine studies and acknowledged. These included small sample sizes, uncertainty over the exact therapeutic mechanism of some interventions and a lack of protocols or guidelines for the interventions. Some of these may be more influential to the results of the studies than others. With one exception (Reimann et al., 2016), all the selected studies used small, symptomatic populations with no asymptomatic control group and of the clinical trials (RCTs) two were non-randomised (Van Lierde et al., 2010; Alves Silverio et al., 2015). In a normal hierarchy of evidence, the RCTs would be second only to systematic reviews, but the lack of control and randomisation places the value of the results somewhat lower, although still above the other study designs. (Evans, 2003; Greenhalgh, 2014; Hohmann et al., 2018).

Recommendations of the selected studies included the need for rigorous randomised controlled trials (Tomlinson and Archer, 2015; Reimann et al., 2016), the development of protocols for interventions and better training and guidelines for therapists delivering interventions (Alves Silverio et al., 2015; Dehqan and Scherer, 2018). Among the selected studies the proficiency of the therapist delivering the intervention was not widely specified, or was noted to be a student (Aghadoost et al., 2019). The experience of the therapist has been shown to have a measurable effect on patient outcomes, in terms of palpation, treatment delivery and communication skills (Testa and Rossettini, 2016; Rossettini et al., 2018). Therefore, the omission of this information weakens the evidential value of some of the selected studies.

There is evidence to show that contextual factors such as the patient's own expectations, emotions and memories alongside the clinical setting, specific therapist

and therapy administered are hugely influential in the positive or negative outcomes of a treatment (Testa and Rossettini, 2016; Kollbrunner and Seifert, 2017; Rossettini et al., 2018). The selected studies fail to acknowledge contextual factors such as the validation of the patient's symptoms, the therapist-patient alliance, the effect of the clinical setting and the emotional effects of the treatment. A 2017 study by Kollbrunner and Seifert aimed to bring attention to this omission and clearly stated that the psychological element of MTD must be more rigorously addressed alongside other more easily measurable interventions (Kollbrunner and Seifert, 2017).

The effect of the language used by the therapist is also important and not acknowledged in any of these studies. A full discussion of placebo and nocebo effects of language is outside the scope of this study, but it is important to note that all interactions in the selected studies took place within a clinical setting, and as part of treatment trials. There is substantial evidence to show that the language used by practitioners can have both positive and negative effects on the patient's experience and the outcome of the treatment (Richter et al., 2010; Corsi et al., 2019). None of the studies mention the language used when communicating with the patients or describe the patients' emotional response to the interventions. This could have been a significant factor in the positive or negative outcomes of the treatments.

5. Conclusion

The available literature on the effectiveness of manual therapies for MTD shows a wide variety of approaches and assessment methods. All the selected studies showed significant positive effects over different time periods, but this information is given in a range of formats, and clinical significance can only really be inferred where effect size is stated (Sullivan and Feinn, 2012; Karadaghy et al., 2017). The value of these results is hampered by the heterogeneity of the study designs, low levels of evidence and methodological problems which make replication of the studies and comparison of the findings difficult. In general, the evidence for the use of LMT and MCT is reasonably strong, while the evidence for TENS remains strong but less widely used and not as clinically significant as that of LMT/MCT.

Standardisation of outcome measures in future studies is critical to creating a homogenous body of research enabling comparison of interventions. The scoping review has identified that there is still confusion among researchers over the target for treatment (pain, muscular tone, physical and acoustic voice quality) and therefore consensus should be reached before the appropriate outcome measure can be selected. Similarly, protocols or standardisations for the assessment of pre-intervention muscular tone (for example the LMTPE (Woźnicka et al., 2017)) should be more widely researched, validated and adopted. Behavioural aspects of MTD should also be further researched and a standardised measurement or assessment performed as part of future studies.

The overwhelming predominance of female subjects in the selected studies shows a gap in the research into why women are more at risk and why they present to voice clinics more frequently than men. Further studies on this subject would lead to a better understanding of the gender differences between men and women which might affect

the outcomes of interventions in different ways. Contextual factors, which remain unacknowledged or unaddressed in the selected studies, would also benefit from further exploration.

Other recommendations for future research identified by this scoping review include randomised controlled trials with more rigorous methodology than those under review, and a greater emphasis on measurement of effect size in order to determine clinical significance of interventions.

Words: 5,238

6. References

Abou-Rafée, M., Zambon, F., Badaró, F., Behlau, M., Abou-Rafée, M., Zambon, F., Badaró, F. and Behlau, M. (2019) Vocal fatigue in dysphonic teachers who seek treatment, *CoDAS*, 31 (3). DOI:10.1590/2317-1782/20182018120.

Aghadoost, S., Jalaie, S., Khatoonabadi, A. R., Dabirmoghaddam, P. and Khoddami, S. M. (2019) A Study of Vocal Facilitating Techniques Compared to Manual Circumlaryngeal Therapy in Teachers With Muscle Tension Dysphonia, *Journal of Voice*. DOI:10.1016/j.jvoice.2019.06.002.

Alves Silverio, K. C., Brasolotto, A. G., Thaís Donalonso Siqueira, L., Carneiro, C. G., Fukushiro, A. P. and Roberto de Jesus Guirro, R. (2015) Effect of Application of Transcutaneous Electrical Nerve Stimulation and Laryngeal Manual Therapy in Dysphonic Women: Clinical Trial, *Journal of Voice*, 29 (2), p. 200–208. DOI:10.1016/j.jvoice.2014.06.003.

Andreassen, M. L., Litts, J. K. and Randall, D. R. (2017) Emerging techniques in assessment and treatment of muscle tension dysphonia:, *Current Opinion in Otolaryngology & Head and Neck Surgery*, 25 (6), p. 447–452. DOI:10.1097/MOO.0000000000000405.

Arksey, H. and O'Malley, L. (2005) Scoping studies: towards a methodological framework, *International Journal of Social Research Methodology*, 8 (1), p. 19–32. DOI:10.1080/1364557032000119616.

Armour, M., Ee, C. C., Naidoo, D., Ayati, Z., Chalmers, K. J., Steel, K. A., Manincor, M. J. de and Delshad, E. (2019a) Exercise for dysmenorrhoea, *Cochrane Database of Systematic Reviews*, (9). DOI:10.1002/14651858.CD004142.pub4.

Armour, M., Parry, K., Al-Dabbas, M. A., Curry, C., Holmes, K., MacMillan, F., Ferfolja, T. and Smith, C. A. (2019b) Self-care strategies and sources of knowledge on menstruation in 12,526 young women with dysmenorrhea: A systematic review and meta-analysis, *PloS One*, 14 (7). DOI:10.1371/journal.pone.0220103.

Awad, R., Shamil, E., Gibbins, N., Aymat, A. and Harris, S. (2019) From Voice Clinic to Operating Room: Are We Out of Tune?, *Journal of Voice*. DOI:10.1016/j.jvoice.2018.12.016.

Awan, S. N., Roy, N., Zhang, D. and Cohen, S. M. (2016) Validation of the Cepstral Spectral Index of Dysphonia (CSID) as a Screening Tool for Voice Disorders: Development of Clinical Cutoff Scores, *Journal of Voice: Official Journal of the Voice Foundation*, 30 (2), p. 130–144. DOI:10.1016/j.jvoice.2015.04.009.

BAPAM: health advice for musicians, actors, singers, dancers and all performing artists (2019), *BAPAM*. Available from: https://www.bapam.org.uk/ [Accessed 23 May 2019].

Behlau, M., Madazio, G. and Oliveira, G. (2015) Functional dysphonia: strategies to improve patient outcomes, *Patient Related Outcome Measures*, 6, p. 243–253. DOI:10.2147/PROM.S68631.

Benninger, M. S., Holy, C. E., Bryson, P. C. and Milstein, C. F. (2017) Prevalence and Occupation of Patients Presenting With Dysphonia in the United States, *Journal of Voice*, 31 (5), p. 594–600. DOI:10.1016/j.jvoice.2017.01.011.

Bottalico, P., Codino, J., Cantor-Cutiva, L. C., Marks, K., Nudelman, C. J., Skeffington, J., *et al.* (2018) Reproducibility of Voice Parameters: The Effect of Room Acoustics and Microphones, *Journal of Voice*. DOI:10.1016/j.jvoice.2018.10.016.

Bradley, P. J. (2010) Voice Disorders: Classification, in: Anniko, M., Bernal-Sprekelsen, M., Bonkowsky, V., Bradley, P. J., and Iurato, S. (eds.) *Otorhinolaryngology, Head and Neck Surgery*. Berlin, Heidelberg: Springer, pp. 555–562.

British Voice Association (2019), *Who's Who in the Voice Clinic*. Available from: https://www.britishvoiceassociation.org.uk/voicecare_voice-clinic-whos-who.htm [Accessed 10 May 2019].

Carding, P. (2003) Voice pathology in the United Kingdom, *BMJ*, 327 (7414), p. 514–515. DOI:10.1136/bmj.327.7414.514.

Cardoso, R., Meneses, R. F. and Lumini-Oliveira, J. (2017) The Effectiveness of Physiotherapy and Complementary Therapies on Voice Disorders: A Systematic Review of Randomized Controlled Trials, *Frontiers in Medicine*, 4, p. 45. DOI:10.3389/fmed.2017.00045.

Chen, Z., Fang, R., Zhang, Y., Ge, P., Zhuang, P., Chou, A. and Jiang, J. (2018) The Mandarin Version of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) and Its Reliability, *Journal of Speech, Language, and Hearing Research: JSLHR*, 61 (10), p. 2451–2457. DOI:10.1044/2018_JSLHR-S-17-0386.

Concato, J., Shah, N. and Horwitz, R. I. (2000) Randomized, Controlled Trials, Observational Studies, and the Hierarchy of Research Designs, *The New England Journal of Medicine*, 342 (25), p. 1887–1892.

Conde, M. de C. M., Siqueira, L. T. D., Vendramini, J. E., Brasolotto, A. G., Guirro, R. R. de J. and Silverio, K. C. A. (2018) Transcutaneous Electrical Nerve Stimulation (TENS) and Laryngeal Manual Therapy (LMT): Immediate Effects in Women With Dysphonia, *Journal of Voice*, 32 (3), p. 385.e17-385.e25. DOI:10.1016/j.jvoice.2017.04.019.

Connor, N. P. and Bless, D. M. (2013) Functional and organic voice disorders, in: Cummings, L. (ed.) *The Cambridge Handbook of Communication Disorders*. Cambridge: Cambridge University Press, pp. 321–340.

Corsi, N., Emadi Andani, M., Sometti, D., Tinazzi, M. and Fiorio, M. (2019) When words hurt: Verbal suggestion prevails over conditioning in inducing the motor nocebo effect, *The European Journal of Neuroscience*, 50 (8), p. 3311–3326. DOI:10.1111/ejn.14489.

Craig, J., Tomlinson, C., Stevens, K., Kotagal, K., Fornadley, J., Jacobson, B., Garrett, C. G. and Francis, D. O. (2015) Combining voice therapy and physical therapy: A novel approach to treating muscle tension dysphonia, *Journal of Communication Disorders*, 58, p. 169–178. DOI:10.1016/j.jcomdis.2015.05.001.

Daudt, H. M., van Mossel, C. and Scott, S. J. (2013) Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework, *BMC Medical Research Methodology*, 13 (1), p. 48. DOI:10.1186/1471-2288-13-48.

Davidson, M. J., Nielsen, P. M. F., Taberner, A. J. and Kruger, J. A. (2020) Is it time to rethink using digital palpation for assessment of muscle stiffness?, *Neurourology and Urodynamics*, 39 (1), p. 279–285. DOI:10.1002/nau.24192.

de Almeida, S. C., Mendes, A. P. and Kempster, G. B. (2019) The Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) Psychometric Characteristics: II European Portuguese Version (II EP CAPE-V), *Journal of Voice: Official Journal of the Voice Foundation*, 33 (4), p. 582.e5-582.e13. DOI:10.1016/j.jvoice.2018.02.013.

Dehqan, A. and Scherer, R. C. (2018) Positive Effects of Manual Circumlaryngeal Therapy in the Treatment of Muscle Tension Dysphonia (MTD): Long Term Treatment Outcomes, *Journal of Voice*, p. S0892199718301280. DOI:10.1016/j.jvoice.2018.07.010.

Dimon, T. and Brown, G. D. (2018) *Anatomy of the voice: an illustrated guide for singers, vocal coaches, and speech therapists*. Berkeley, California: North Atlantic Books.

Ertan-Schlüter, E., Demirhan, E., Ünsal, E. M. and Tadıhan-Özkan, E. (2019) The Turkish Version of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V): A Reliability and Validity Study, *Journal of Voice*. DOI:10.1016/j.jvoice.2019.05.014.

Evans, D. (2003) Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions, *Journal of Clinical Nursing*, 12 (1), p. 77–84. DOI:10.1046/j.1365-2702.2003.00662.x.

Faham, M., Laukkanen, A.-M., Ikävalko, T., Rantala, L., Geneid, A., Holmqvist-Jämsén, S., Ruusuvirta, K. and Pirilä, S. (2019) Acoustic Voice Quality Index as a Potential Tool for Voice Screening, *Journal of Voice*. DOI:10.1016/j.jvoice.2019.08.017.

Garaycochea, O., Navarrete, J. M. A., del Río, B. and Fernández, S. (2019) Muscle Tension Dysphonia: Which Laryngoscopic Features Can We Rely on for Diagnosis?, *Journal of Voice*, 33 (5), p. 812.e15-812.e18. DOI:10.1016/j.jvoice.2018.04.015.

Gates, R., Arick Forrest, L. and Obert, K. (2013) *The Owner's Manual to the Voice: A Guide For Singers And Other Professional Voice Users*. Oxford; New York: Oxford University Press.

Greenhalgh, T. (2014) *How to read a paper: the basics of evidence-based medicine*. Fifth edition. Chichester, West Sussex: John Wiley & Sons Inc.

Hari Kumar, K. V. S., Garg, A., Ajai Chandra, N. S., Singh, S. P. and Datta, R. (2016) Voice and endocrinology, *Indian Journal of Endocrinology and Metabolism*, 20 (5), p. 590–594. DOI:10.4103/2230-8210.190523.

Harris, S. (2018a) Speech therapy for dysphonia, in: Harris, T. and Howard, D. M. (eds.) *The Voice Clinic Handbook*. 2nd ed. Oxford: Compton Publishing Ltd., pp. 83–158.

Harris, T. (2018b) Drugs and the pharmacological treatment of dysphonia, in: Harris, T. and Howard, D. M. (eds.) *The Voice Clinic Handbook*. 2nd ed. Oxford: Compton Publishing Ltd., pp. 251–268.

Harris, T. and Howard, D. M. (eds.) (2018) *The Voice Clinic Handbook*. 2nd ed. Oxford: Compton Publishing Ltd.

Harris, T. and Moisik, S. (2018) Laryngopharyngeal mechanisms in normal function and dysfunction, in: Harris, T. and Howard, D. M. (eds.) *The Voice Clinic Handbook*. 2nd ed. Oxford: Compton Publishing Ltd., pp. 55–80.

Hirano, M. (1981) *Clinical examination of voice*. London: Springer London, Limited.

Hogikyan, N. D. and Sethuraman, G. (1999) Validation of an instrument to measure voice-related quality of life (V-RQOL)., *Journal of Voice : Official Journal of the Voice Foundation*, 13 (4), p. 557–569. DOI:10.1016/S0892-1997(99)80010-1.

Hohmann, E., Feldman, M., Hunt, T. J., Cote, M. P. and Brand, J. C. (2018) Research Pearls: How Do We Establish the Level of Evidence?, *Arthroscopy: The Journal of Arthroscopic & Related Surgery*, 34 (12), p. 3271–3277. DOI:10.1016/j.arthro.2018.10.002.

Hoppe, D. J., Schemitsch, E. H., Morshed, S., Tornetta, P. I. and Bhandari, M. (2009) Hierarchy of Evidence: Where Observational Studies Fit in and Why We Need Them, *JBJS*, 91 (Supplement_3), p. 2–9. DOI:10.2106/JBJS.H.01571.

Howard, D. M. (2018) Practical Voice Measurement, in: Harris, T. and Howard, D. M. (eds.) *The Voice Clinic Handbook*. 2nd ed. Oxford: Compton Publishing Ltd., pp. 319–378.

Hunter, E. J., Smith, M. E. and Tanner, K. (2011) Gender differences affecting vocal health of women in vocally demanding careers, *Logopedics, Phoniatrics, Vocology*, 36 (3), p. 128–136. DOI:10.3109/14015439.2011.587447.

Johnson, M. I., Paley, C. A., Howe, T. E. and Sluka, K. A. (2015) Transcutaneous electrical nerve stimulation for acute pain, *The Cochrane Database of Systematic Reviews*, (6). DOI:10.1002/14651858.CD006142.pub3.

Karadaghy, O. A., Hong, H., Scott-Wittenborn, N., Sinha, P., Suko, J., Tait, S., Wamkpah, N. S., Kallogjeri, D. and Piccirillo, J. F. (2017) Reporting of Effect Size

and Confidence Intervals in JAMA Otolaryngology-Head & Neck Surgery, *JAMA Otolaryngology-- Head & Neck Surgery*, 143 (11), p. 1075–1080. DOI:10.1001/jamaoto.2017.1504.

Kennard, E. J., Lieberman, J., Saaid, A. and Rolfe, K. J. (2015) A Preliminary Comparison of Laryngeal Manipulation and Postural Treatment on Voice Quality in a Prospective Randomized Crossover Study, *Journal of Voice*, 29 (6), p. 751–754. DOI:10.1016/j.jvoice.2014.09.026.

Khoramshahi, H., Khatoonabadi, A. R., Khoddami, S. M., Dabirmoghaddam, P. and Ansari, N. N. (2018) Responsiveness of Persian Version of Consensus Auditory Perceptual Evaluation of Voice (CAPE-V), Persian Version of Voice Handicap Index (VHI), and Praat in Vocal Mass Lesions with Muscle Tension Dysphonia, *Journal of Voice*, 32 (6), p. 770.e21-770.e30. DOI:10.1016/j.jvoice.2017.08.008.

Kollbrunner, J. and Seifert, E. (2017) Encouragement to Increase the Use of Psychosocial Skills in the Diagnosis and Therapy of Patients With Functional Dysphonia, *Journal of Voice*, 31 (1), p. 132.e1-132.e7. DOI:10.1016/j.jvoice.2015.11.021.

Korn, G. P., Park, S. W., Pontes, A. A. de L. and Pontes, P. (2018) Vocal Symptoms and Associated Risk Factors between Male and Female University Teachers, *International Archives of Otorhinolaryngology*, 22 (3), p. 271–279. DOI:10.1055/s-0037-1606604.

Kreiman, J. and Gerratt, B. R. (2010) Perceptual Assessment of Voice Quality: Past, Present, and Future, *Perspectives on Voice and Voice Disorders*, 20 (2), p. 62–67. DOI:10.1044/vvd20.2.62.

Lantz, B. (2013) The Large Sample Size Fallacy, *Scandinavian Journal of Caring Sciences*, 27 (2), p. 487–492. DOI:10.1111/j.1471-6712.2012.01052.x.

LeBorgne, W. D. and Donahue, E. N. (2019) Voice Therapy as Primary Treatment of Vocal Fold Pathology, *Otolaryngologic Clinics of North America*, 52 (4), p. 649–656. DOI:10.1016/j.otc.2019.03.009.

Levac, D., Colquhoun, H. and O'Brien, K. K. (2010) Scoping studies: advancing the methodology, *Implementation Science*, 5 (1), p. 1–9. DOI:10.1186/1748-5908-5-69.

Lieberman, J. (2018) Osteopathic medicine in the management of muscle tension dysphonia, in: Harris, T. and Howard, D. M. (eds.) *The Voice Clinic Handbook*. 2nd ed. Oxford: Compton Publishing Ltd., pp. 159–198.

Martimbianco, A. L. C., Porfírio, G. J., Pacheco, R. L., Torloni, M. R. and Riera, R. (2019) Transcutaneous electrical nerve stimulation (TENS) for chronic neck pain, *The Cochrane Database of Systematic Reviews*, 12. DOI:10.1002/14651858.CD011927.pub2.

Mathieson, L. (2011) The evidence for laryngeal manual therapies in the treatment of muscle tension dysphonia:, *Current Opinion in Otolaryngology & Head and Neck Surgery*, 19 (3), p. 171–176. DOI:10.1097/MOO.0b013e3283448f6c.

Mathieson, L., Hirani, S. P., Epstein, R., Baken, R. J., Wood, G. and Rubin, J. S. (2009) Laryngeal Manual Therapy: A Preliminary Study to Examine its Treatment Effects in the Management of Muscle Tension Dysphonia, *Journal of Voice*, 23 (3), p. 353–366. DOI:10.1016/j.jvoice.2007.10.002.

Neemuchwala, P. (2018) Psychogenic factors in dysphonia, in: Harris, T. and Howard, D. M. (eds.) *The Voice Clinic Handbook*. 2nd ed. Oxford: Compton Publishing Ltd., pp. 233–250.

Nemr, K., Simões-Zenari, M., Cordeiro, G. F., Tsuji, D., Ogawa, A. I., Ubrig, M. T. and Menezes, M. H. M. (2012) GRBAS and Cape-V Scales: High Reliability and Consensus When Applied at Different Times, *Journal of Voice*, 26 (6), p. 812.e17-812.e22. DOI:10.1016/j.jvoice.2012.03.005.

Osteopathic Performing Arts Care Association (2019), *OPACA*. Available from: https://www.opaca.co.uk/ [Accessed 23 May 2019].

Pavela Banai, I. (2017) Voice in different phases of menstrual cycle among naturally cycling women and users of hormonal contraceptives, *PLoS ONE*, 12 (8). DOI:10.1371/journal.pone.0183462.

Phillips, P. S., Carlson, E. and Chevretton, E. B. (2005) Does a specialist voice clinic change ENT clinic diagnosis?, *Logopedics Phoniatrics Vocology*, 30 (2), p. 90–93. DOI:10.1080/14015430500233286.

Phyland, D. and Miles, A. (2019) Occupational voice is a work in progress: active risk management, habilitation and rehabilitation, *Current Opinion in Otolaryngology & Head and Neck Surgery*, 27 (6), p. 439–447. DOI:10.1097/MOO.0000000000000584.

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K. and Duffy, S. (2006, April) *Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: A Product from the ESRC Methods Programme*. Lancaster University, Lancaster, United Kingdom.

Ramos, A. C., Floro, R. L., Ribeiro, V. V., Brasolotto, A. G. and Silverio, K. C. A. (2018) Musculoskeletal Pain and Voice-related Quality of Life in Dysphonic and Non-dysphonic Subjects, *Journal of Voice: Official Journal of the Voice Foundation*, 32 (3), p. 307–313. DOI:10.1016/j.jvoice.2017.05.019.

Reimann, A. P., Siqueira, L. T. D., Rondon, A. V., Brasolotto, A. G. and Silverio, K. C. A. (2016) Efeito imediato da terapia manual laríngea em indivíduos disfônicos, *CoDAS*, 28 (1), p. 59–65. DOI:10.1590/2317-1782/20162015089.

Remacle, A., Petitfils, C., Finck, C. and Morsomme, D. (2017) Description of patients consulting the voice clinic regarding gender, age, occupational status, and diagnosis, *European Archives of Oto-Rhino-Laryngology*, 274 (3), p. 1567–1576. DOI:10.1007/s00405-016-4332-7.

Ribeiro, V. V., Pedrosa, V., Silverio, K. C. A. and Behlau, M. (2018) Laryngeal Manual Therapies for Behavioral Dysphonia: A Systematic Review and Meta-analysis, *Journal of Voice*, 32 (5), p. 553–563. DOI:10.1016/j.jvoice.2017.06.019.

Richter, M., Eck, J., Straube, T., Miltner, W. H. R. and Weiss, T. (2010) Do words hurt? Brain activation during the processing of pain-related words, *Pain*, 148 (2), p. 198–205. DOI:10.1016/j.pain.2009.08.009.

Rossettini, G., Carlino, E. and Testa, M. (2018) Clinical relevance of contextual factors as triggers of placebo and nocebo effects in musculoskeletal pain, *BMC Musculoskeletal Disorders*, 19 (1), p. 27. DOI:10.1186/s12891-018-1943-8.

Roy, N., Bless, D. M., Heisey, D. and Ford, C. N. (1997) Manual circumlaryngeal therapy for functionaldysphonia: An evaluation of short- and long-term treatment

outcomes, *Journal of Voice*, 11 (3), p. 321–331. DOI:10.1016/S0892-1997(97)80011-2.

Roy, N. and Leeper, H. A. (1993) Effects of the manual laryngeal musculoskeletal tension reduction technique as a treatment for functional voice disorders: Perceptual and acoustic measures, *Journal of Voice*, 7 (3), p. 242–249. DOI:10.1016/S0892-1997(05)80333-9.

Rubin, J. S. (2018a) Mechanisms of respiration (the bellows), in: Harris, T. and Howard, D. M. (eds.) *The Voice Clinic Handbook*. 2nd ed. Oxford: Compton Publishing Ltd., pp. 43–54.

Rubin, J. S. (2018b) The structural anatomy of the larynx and supraglottic vocal tract: a review, in: Harris, T. and Howard, D. M. (eds.) *The Voice Clinic Handbook*. 2nd ed. Oxford: Compton Publishing Ltd., pp. 3–18.

Rubin, J. S., Lieberman, J. and Harris, T. M. (2000) Laryngeal manipulation, *Otolaryngologic Clinics of North America*, 33 (5), p. 1017–1034. DOI:10.1016/s0030-6665(05)70261-9.

Ruotsalainen, J. H., Sellman, J., Lehto, L., Jauhiainen, M. and Verbeek, J. H. (2007) Interventions for treating functional dysphonia in adults, *Cochrane Database of Systematic Reviews*. DOI:10.1002/14651858.CD006373.pub2.

Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., *et al.* (2017) AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both, *BMJ*, p. j4008. DOI:10.1136/bmj.j4008.

Shewell, C. (2009) *Voice work: art and science in changing voices*. Chichester, West Sussex: John Wiley & Sons Inc.

Siqueira, L. T. D., Silverio, K. C. A., Brasolotto, A. G., Guirro, R. R. de J., Carneiro, C. G. and Behlau, M. (2017) Efeitos da terapia manual laríngea e da estimulação elétrica nervosa transcutânea (TENS) na diadococinesia laríngea em mulheres disfônicas: estudo clínico randomizado, *CoDAS*, 29 (3). DOI:10.1590/2317-1782/20172016191.

Stachler, R. J., Francis, D. O., Schwartz, S. R., Damask, C. C., Digoy, G. P., Krouse, H. J., *et al.* (2018) Clinical Practice Guideline: Hoarseness (Dysphonia) (Update), *Otolaryngology–Head and Neck Surgery*, 158 (1_suppl), p. S1–S42. DOI:10.1177/0194599817751030.

Steinhauer, K., McDonald Klimek, M. and Estill, J. (2017) *The Estill Voice Model: Theory & Translation*. Pittsburgh, Pennsylvania: Estill Voice International.

Sullivan, G. M. and Feinn, R. (2012) Using Effect Size—or Why the P Value Is Not Enough, *Journal of Graduate Medical Education*, 4 (3), p. 279–282. DOI:10.4300/JGME-D-12-00156.1.

Testa, M. and Rossettini, G. (2016) Enhance placebo, avoid nocebo: How contextual factors affect physiotherapy outcomes, *Manual Therapy*, 24, p. 65–74. DOI:10.1016/j.math.2016.04.006.

Tomlinson, C. A. and Archer, K. R. (2015) Manual Therapy and Exercise to Improve Outcomes in Patients With Muscle Tension Dysphonia: A Case Series, *Physical Therapy*, 95 (1), p. 117–128. DOI:10.2522/ptj.20130547.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., *et al.* (2018) PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation, *Annals of Internal Medicine*, 169 (7), p. 467. DOI:10.7326/M18-0850.

van Houtte, E., Claeys, S., Wuyts, F. and van Lierde, K. (2012) Voice disorders in teachers: Occupational risk factors and psycho-emotional factors, *Logopedics Phoniatrics Vocology*, 37 (3), p. 107–116. DOI:10.3109/14015439.2012.660499.

Van Lierde, K. M., Bodt, M. D., Dhaeseleer, E., Wuyts, F. and Claeys, S. (2010) The Treatment of Muscle Tension Dysphonia: A Comparison of Two Treatment Techniques by Means of an Objective Multiparameter Approach, *Journal of Voice*, 24 (3), p. 294–301. DOI:10.1016/j.jvoice.2008.09.003.

Vila, P. M., Townsend, M. E., Bhatt, N. K., Kao, W. K., Sinha, P. and Neely, J. G. (2017) The P Value Problem in Otolaryngology: Shifting to Effect Sizes and Confidence Intervals, *Otolaryngology--Head and Neck Surgery*, 156 (6), p. 978–980. DOI:10.1177/0194599816677735.

Woźnicka, E., Niebudek-Bogusz, E., Morawska, J., Wiktorowicz, J. and Śliwińska-Kowalska, M. (2017) Laryngeal manual therapy palpatory evaluation scale: A preliminary study to examine its usefulness in diagnosis of occupational dysphonia, *Medycyna Pracy*, 68 (2), p. 179–188. DOI:10.13075/mp.5893.00463.

Acknowledgements:

My sincere thanks to Dr Hilary Abbey for her patience, time, support and knowledge, and her willingness to delve into such a specialist topic.

Thanks also to Dr Guy Bunce for proof-reading this paper, and his and my family's unwavering emotional and financial support through the last five years. For second proofing and sense checking, thanks to Kathryn Procter.

To Dr Jenevora Williams and Debbie Winter of Voice Workshop my thanks, I think, for helping to set me on this path, and Lucy Scott, Gayle Hocking and Stephen King for their faith in my future abilities.

Lastly, thanks to fellow members of my part-time cohort for their unfailing ability to generate laughter and tears at exactly the right moment.

Appendix I: Summary of Arksey & O'Malley (2005) Methodological Framework for Scoping Reviews

1. **Identifying the research question:** starting with wide definitions for study population, interventions or outcomes, to ensure breadth of coverage in the search, and then setting parameters based on the scope and volume of references generated.
   - Levac et al., (2010): maintaining a broad search strategy with clearly defined concepts and their continuous refinement

2. **Identifying relevant studies:** as comprehensively as possible identifying primary studies (published and unpublished) and reviews suitable for answering the central research question. Adopting a strategy that involves searching for research evidence via different sources.
   - Armstrong et al., (2011): From a practical point of view, decisions have to be made at the outset about the coverage of the review in terms of time span and languages.

3. **Study selection:** unlike systematic reviews, inclusion and exclusion criteria are developed post hoc, once familiarity with the literature has been gained
   - Daudt et al., (2013); Levac et al., (2010): using multidisciplinary expertise and group consultation within the scoping team to inform and guide the definition of the search criteria and clinical applicability of data for extraction

4. **Charting the data:** data synthesis and interpretation may adopt a narrative or descriptive approach in place of a more systematic data extraction or analytic method.
   - Armstrong et al., (2011): allowing for post-hoc development of inclusion/exclusion criteria and data synthesis in terms of the value yielded by qualitative or quantitative analysis of results.

5. **Collating, summarising and reporting the results:** emphasis is not placed on the "weight of evidence" nor on evaluating the quality of evidence, but an analytic or thematic framework to guide the narrative account of existing literature is recommended.

6. **Consultation exercise:** although this is an optional step, this is recommended as a useful contribution, where "contributors to the consultation provided additional references about potential studies to include in the review as well as valuable insights about issues relating to the effectiveness and cost-effectiveness of services that the scoping review alone would not have alerted us to".
   - Daudt et al., (2013): An additional, parallel element is also described regarding the use of a 'consultation exercise' to inform and validate findings from the main scoping review. Whilst consultation might be viewed as an optional component of the scoping study framework, it greatly enhanced our work, a view confirmed by other researchers.

Table. PRISMA-ScR Checklist

Section	Item	PRISMA-ScR Checklist Item
Title	1	Identify the report as a scoping review.
Abstract		
Structured summary	2	Provide a structured summary that includes (as applicable) background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.
Introduction		
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.
Methods		
Protocol and registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.
Eligibility criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.
Information sources*	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.
Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.
Selection of sources of evidence†	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.
Data charting process‡	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.
Data items	11	List and define all variables for which data were sought and any assumptions and simplifications made.
Critical appraisal of individual sources of evidence§	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate).
Summary measures	13	Not applicable for scoping reviews.
Synthesis of results	14	Describe the methods of handling and summarizing the data that were charted.
Risk of bias across studies	15	Not applicable for scoping reviews.
Additional analyses	16	Not applicable for scoping reviews.
Results		
Selection of sources of evidence	17	Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.
Characteristics of sources of evidence	18	For each source of evidence, present characteristics for which data were charted and provide the citations.
Critical appraisal within sources of evidence	19	If done, present data on critical appraisal of included sources of evidence (see item 12).
Results of individual sources of evidence	20	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.
Synthesis of results	21	Summarize and/or present the charting results as they relate to the review questions and objectives.
Risk of bias across studies	22	Not applicable for scoping reviews.
Additional analyses	23	Not applicable for scoping reviews.
Discussion		
Summary of evidence	24	Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups.
Limitations	25	Discuss the limitations of the scoping review process.
Conclusions	26	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.
Funding	27	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.

JBI = Joanna Briggs Institute; PRISMA-ScR = Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews.
* Where *sources of evidence* (see second footnote) are compiled from, such as bibliographic databases, social media platforms, and Web sites.
† A more inclusive/heterogeneous term used to account for the different types of evidence or data sources (e.g., quantitative and/or qualitative research, expert opinion, and policy documents) that may be eligible in a scoping review as opposed to only studies. This is not to be confused with *information sources* (see first footnote).
‡ The frameworks by Arksey and O'Malley (6) and Levac and colleagues (7) and the JBI guidance (4, 5) refer to the process of data extraction in a scoping review as data charting.
§ The process of systematically examining research evidence to assess its validity, results, and relevance before using it to inform a decision. This term is used for items 12 and 19 instead of "risk of bias" (which is more applicable to systematic reviews of interventions) to include and acknowledge the various sources of evidence that may be used in a scoping review (e.g., quantitative and/or qualitative research, expert opinion, and policy documents).

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

1. Did the research questions and inclusion criteria for the review include the components of PICO?

For Yes:
- ☐ Population
- ☐ Intervention
- ☐ Comparator group
- ☐ Outcome

Optional (recommended)
- ☐ Timeframe for follow-up

- ☐ Yes
- ☐ No

2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?

For Partial Yes:
The authors state that they had a written protocol or guide that included ALL the following:
- ☐ review question(s)
- ☐ a search strategy
- ☐ inclusion/exclusion criteria
- ☐ a risk of bias assessment

For Yes:
As for partial yes, plus the protocol should be registered and should also have specified:
- ☐ a meta-analysis/synthesis plan, if appropriate, *and*
- ☐ a plan for investigating causes of heterogeneity
- ☐ justification for any deviations from the protocol

- ☐ Yes
- ☐ Partial Yes
- ☐ No

3. Did the review authors explain their selection of the study designs for inclusion in the review?

For Yes, the review should satisfy ONE of the following:
- ☐ *Explanation for* including only RCTs
- ☐ OR *Explanation for* including only NRSI
- ☐ OR *Explanation for* including both RCTs and NRSI

- ☐ Yes
- ☐ No

4. Did the review authors use a comprehensive literature search strategy?

For Partial Yes (all the following):
- ☐ searched at least 2 databases (relevant to research question)
- ☐ provided key word and/or search strategy
- ☐ justified publication restrictions (e.g. language)

For Yes, should also have (all the following):
- ☐ searched the reference lists / bibliographies of included studies
- ☐ searched trial/study registries
- ☐ included/consulted content experts in the field
- ☐ where relevant, searched for grey literature
- ☐ conducted search within 24 months of completion of the review

- ☐ Yes
- ☐ Partial Yes
- ☐ No

5. Did the review authors perform study selection in duplicate?

For Yes, either ONE of the following:
- ☐ at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include
- ☐ OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer.

- ☐ Yes
- ☐ No

6. Did the review authors perform data extraction in duplicate?

For Yes, either ONE of the following:

- ☐ at least two reviewers achieved consensus on which data to extract from included studies
- ☐ OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer.

☐ Yes
☐ No

7. Did the review authors provide a list of excluded studies and justify the exclusions?

For Partial Yes:

- ☐ provided a list of all potentially relevant studies that were read in full-text form but excluded from the review

For Yes, must also have:

- ☐ Justified the exclusion from the review of each potentially relevant study

☐ Yes
☐ Partial Yes
☐ No

8. Did the review authors describe the included studies in adequate detail?

For Partial Yes (ALL the following):

- ☐ described populations
- ☐ described interventions
- ☐ described comparators
- ☐ described outcomes
- ☐ described research designs

For Yes, should also have ALL the following:

- ☐ described population in detail
- ☐ described intervention in detail (including doses where relevant)
- ☐ described comparator in detail (including doses where relevant)
- ☐ described study's setting
- ☐ timeframe for follow-up

☐ Yes
☐ Partial Yes
☐ No

9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?

RCTs

For Partial Yes, must have assessed RoB from

- ☐ unconcealed allocation, *and*
- ☐ lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all-cause mortality)

For Yes, must also have assessed RoB from:

- ☐ allocation sequence that was not truly random, *and*
- ☐ selection of the reported result from among multiple measurements or analyses of a specified outcome

☐ Yes
☐ Partial Yes
☐ No
☐ Includes only NRSI

NRSI

For Partial Yes, must have assessed RoB:

- ☐ from confounding, *and*
- ☐ from selection bias

For Yes, must also have assessed RoB:

- ☐ methods used to ascertain exposures and outcomes, *and*
- ☐ selection of the reported result from among multiple measurements or analyses of a specified outcome

☐ Yes
☐ Partial Yes
☐ No
☐ Includes only RCTs

10. Did the review authors report on the sources of funding for the studies included in the review?

For Yes

- ☐ Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies

☐ Yes
☐ No

11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?

RCTs

For Yes:

- ☐ The authors justified combining the data in a meta-analysis
  - ☐ AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present.
  - ☐ AND investigated the causes of any heterogeneity

☐ Yes
☐ No
☐ No meta-analysis conducted

For NRSI

For Yes:

- ☐ The authors justified combining the data in a meta-analysis
  - ☐ AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present
  - ☐ AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available
  - ☐ AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review

☐ Yes
☐ No
☐ No meta-analysis conducted

12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?

For Yes:

- ☐ included only low risk of bias RCTs
- ☐ OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect.

☐ Yes
☐ No
☐ No meta-analysis conducted

13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?

For Yes:

- ☐ included only low risk of bias RCTs
- ☐ OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results

☐ Yes
☐ No

14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?

For Yes:

- ☐ There was no significant heterogeneity in the results
- ☐ OR if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review

☐ Yes
☐ No

15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?

For Yes:

- ☐ performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias

☐ Yes
☐ No
☐ No meta-analysis conducted

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

For Yes:

- ☐ The authors reported no competing interests OR
- ☐ The authors described their funding sources and how they managed potential conflicts of interest

☐ Yes
☐ No

To cite this tool: Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V, Kristjansson E, Henry DA. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017 Sep 21;358:j4008.