The natural history of subjective tinnitus in adults: a systematic review and meta-analysis of 'no-intervention' periods in controlled trials

Running title: Natural history of tinnitus

John S Phillips[1] FRCS(ORL-HNS) (Corresponding author)
Don J McFerran[2] MA, FRCS
Deborah A Hall[3,4*] BSc, PhD
Derek J Hoare[3,4*] BSc, MRes, PhD

[1]Norfolk & Norwich University Hospital NHS Foundation Trust, Colney Lane Norwich NR4 7UY

[2] Colchester Hospital University NHS Foundation Trust, Turner Road, Colchester, Essex, CO4 5JL

[3] NIHR Nottingham Hearing Biomedical Research Unit, 113 The Ropewalk, Nottingham, NG1 5DU

[4]Otology and Hearing Group, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham, NG7 2UH

*these authors made equal contribution

Correspondence:
Mr John S Phillips
Department of Otolaryngology
Norfolk & Norwich University Hospital
Colney Lane
Norwich, Norfolk
NR4 7UY, United Kingdom
Tel.: +44 1603 286770
Fax.: +44 1603 287288
Email: john.phillips@mac.com

Source of Funding:
This work was supported by a British Society of Audiology Applied Research Grant awarded to the authors.

Financial Disclosures and Conflicts of Interest:
JP has paid consultancies with Otonomy. DM has received payment for trials of drugs for tinnitus (GSK and Autifony) and has received payment for providing tinnitus talks to healthcare professionals and the public. DM has received royalties for books on tinnitus. DAH has paid consultancies with NeuroMod, Desyncra, and Otonomy. DAH and DJH are funded by the NIHR Biomedical Research Unit Programme. The views expressed in this article are those of the authors and not necessarily those of the NIHR, the NHS, or the Department of Health.

The natural history of tinnitus: a systematic review and meta-analysis of controlled trials

ABSTRACT

Objectives

Tinnitus is a prevalent condition but little has been published regarding the natural history of the condition. One technique for evaluating the long-term progression of disease is to examine what happens to participants in the no-intervention control arm of a clinical trial. The aim of this study was to examine no-intervention or waiting-list data reported in trials in which participants on the active arm received any form of tinnitus intervention.

Data sources

CINAHL, PsychINFO, EMBASE, ASSIA, PubMed, Web of Science, Science Direct, EBSCO Host and Cochrane.

Methods

Inclusion criteria followed the PICOS principles: Participants: adults with tinnitus; Intervention: none; Control: any intervention for alleviating tinnitus; Outcomes: a measure assessing tinnitus symptoms using a multi-item patient-reported tinnitus questionnaire. Secondary outcome measures included multi-item patient-reported questionnaires of mood and health-related quality of life and measures that quantified change in tinnitus loudness; Study design: randomized controlled trials or observational studies utilizing a no-intervention or waiting-list control group. Data were extracted and standardized mean difference was calculated for each study to enable meta-analysis.

Results

The evidence strongly favored a statistically significant decrease in the impact of tinnitus over time, though there was significant heterogeneity and clinical significance cannot be interpreted. Outcome data regarding secondary measures did not demonstrate any clinically significant change.

Conclusions

Participants allocated to the no-intervention or waiting-list control arm of clinical trials for a tinnitus intervention show a small but significant improvement in self-reported measures of tinnitus with time; the clinical significance of this finding is unknown. There is however considerable variation across individuals. These findings support previous work and can cautiously be used when counselling patients.

Key Words

Tinnitus; Natural history; Outcomes; Control; Waiting List

Level of evidence

NA

INTRODUCTION

Part of the counselling provided to tinnitus patients by practitioners involves reassurance that both the perceived loudness of the tinnitus sounds and the emotional symptoms of tinnitus generally improve with time. Although this may be true, data to support the validity of this statement and to quantify any improvement in symptoms have been poorly presented in the literature. There are a small number of longitudinal studies of tinnitus which give some support to the suggestion that tinnitus impact lessens with time.[1,2,3] However, participants in these studies could access healthcare services for their symptom and it is therefore difficult to ascertain whether any change is natural improvement with time or treatment effect. One technique used to study what happens to symptoms over time among people receiving no treatment is to examine the outcome of participants on a no-intervention or waiting-list control arm of clinical trials and this methodology has a long pedigree of usage in the field of mental health.[4,5,6] By amalgamating the control groups of multiple trials, meta-analysis of the outcome is viable. A limited study of what happens to patients with tinnitus while on a no-intervention and waiting-list control group has previously been undertaken[7] but this was restricted to studies that had incorporated cognitive behavior therapy as the active arm of the trial. Restricting participants to those willing to embrace psychological therapies for their tinnitus potentially produces a study population that is not representative of the wider tinnitus population. The aim of the current study was to expand that original work by looking at people with tinnitus who had been allocated to a no-intervention or waiting-list control group in the context of a trial evaluating any form of tinnitus therapy.

The following research questions were posited:

1. During a period of no-intervention or waiting-list, what changes occur in self reported measure of tinnitus?

2. During a period of no-intervention or waiting-list, what changes occur in self-reported measures of tinnitus-related problems of mood and quality of life?
3. During a period of no-intervention waiting, what changes occur in perceived tinnitus loudness?

The first research question was our primary outcome measure and our second and third research questions were our secondary outcome measures.

METHODS

Study registration

Details of the proposed study eligibility criteria, information sources, search strategy, selection and data collection processes, as well as data synthesis methods were registered at PROSPERO, the international database of prospectively registered systematic reviews (PROSPERO 2013:CRD42013003334). Reporting of the review has been conducted using the criteria recommended by Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).[8] Presentation of the meta-analysis complies with MOOSE Guidelines for Meta-Analyses and Systematic Reviews of Observational Studies.[9]

Study selection

In the protocol registered in PROSPERO, the condition of interest was referred to as 'watchful waiting'. Because this term implies some degree of symptom monitoring which was not necessarily evident in the records found and because our study selection strategy did not necessarily seek to exclude study designs in which a group was not anticipating receiving an intervention, we refer instead to this group throughout as 'no-intervention or waiting-list control'. Inclusion criteria were formed using the Participants, Intervention, Control, Outcomes, and Study designs (PICOS) strategy.[10] These are; Participants: adults with tinnitus; Intervention: no intervention or waiting list control; Comparator: any intervention for tinnitus; Outcomes, primary: one or more tinnitus-specific measures using a multi-item patient-reported questionnaire; Study design: randomized controlled trial or observational study with a control group involving no intervention.

Studies that were not available in English were also excluded as we did not have the resources to translate them. Records that had not been through a peer-review process (grey literature) were excluded as a quality control measure.

The search was not explicitly time limited, but the first multi-item patient-reported tinnitus-related questionnaire was published in 1988.[11] Hence, no clinical trials meeting our inclusion criteria would have been published prior to this date. For the purposes of the review, adult was defined as aged 16 or older.

Appropriate outcome measures

Eligible studies were those reporting at least one patient-reported outcome relating to tinnitus, measured using a multi-item patient-reported tinnitus-specific questionnaire with scores that were reported both before and after the time period corresponding to the intervention for the active comparator group. Examples of acceptable measurement instruments are shown in Supplemental Table S1. This is not an exhaustive list and, if encountered, other tools were considered. Outcomes that were considered as a secondary question in this review were those multi-item patient-reported questionnaires of mood and quality of life, and tools for estimating a change in tinnitus percept, namely loudness, with scores that were reported both before and after the time period corresponding to the intervention for the comparator group. Such assessments were not prerequisites for study inclusion, but where such information was available it was extracted and analyzed. In a change to the study design as registered in PROSPERO, we did not investigate the change in audiological or physiological outcome measures as secondary questions.

Appropriate study design

Eligible study designs were randomized controlled trials in which adult participants were allocated to a no-intervention control group receiving no support. Observational studies in which there was a no-intervention group were also eligible. Cross-over designs were included if a no-intervention period preceded an active intervention comparator and data from the pre-intervention period could be separately extracted.

Search strategy

A systematic search of the literature was conducted by DJH to identify relevant articles from 8 literature search platforms; CINAHL, PsychINFO, EMBASE, ASSIA, PubMed, Web of Science, Science Direct, and EBSCO Host. For each database the search was run using the Boolean search term: tinnitus AND waiting OR wait* OR waiting-list OR watchful OR observation. For interest, a sample search strategy (generated by PubMed) in executing the search is given in Supplemental Table S2.

In addition, handsearching of the reference lists of all articles returned from the search was undertaken, and articles published by shortlisted authors were screened to identify any relevant articles which may not have been returned by the initial database searches. Cochrane and other relevant systematic reviews were searched. In October 2015, hand searches were conducted of articles published in issues since April 2013 of the pre-specified journals (Acta Otolaryngologica; Ear and Hearing; Hearing Research; Journal of Psychosomatic Medicine; Psychosomatic Research; International Journal of Audiology; International Tinnitus Journal; Laryngoscope; Otolaryngology Head and Neck Surgery; Otology and Neurology, and PLOS ONE). Finally, the data collection form associated with an independent systematic review of clinical trials of tinnitus published between July 2006 and March 2015 was searched.[12]

Data management

All identified records were saved into a Microsoft Excel master file where records were tracked through the screening and data collection process by a unique study identification number. A simple system of record annotation was implemented to capture reasons for exclusion. Two authors (JSP and DJM) independently assessed the search results to identify

studies for inclusion in the review and extracted the relevant data. Any discrepancies in study selection or data extraction were resolved in discussion with a third author (DAH or DJH). JSP was the data guarantor.

Data extraction

Data extracted included study design, participants (demographics, baseline characteristics), context of waiting (waiting-list for crossover or no-intervention), comparator, outcomes measures used, study findings, and conclusions. A data extraction form was developed and piloted for the purpose. Where data was missing or unclearly reported, an attempt was made to contact the relevant corresponding author of the study; the most common problem was that the results had been presented graphically and numerical data for the meta-analysis could not be extracted. Supplemental Table S3 provides a summary of 18 study records for which we sought clarification or additional information. Of those, only three did not reply; six did reply but were unable to provide the data requested.

Risk of bias (quality) assessment

Risk of bias assessment was guided by Higgins *et al.*[13] and was conducted by three authors (JP, DAH, DJH) on those study records included in the meta-analysis. The following terminology was specified: (1) Selection bias refers to how participants were allocated to the intervention arms of the trial and was assessed according to two criteria, namely sequence generation for the randomization process and allocation concealment to ensure that the schedule of random assignments prevented advance knowledge about the forthcoming allocations; (2) Attrition bias refers to how participants withdrew from any trial and was assessed by identifying incomplete outcome data; (3) Detection bias refers to how the outcomes were determined and was assessed according to the blinding of participants and outcome assessors assessing patient- or clinician-reported questionnaires, respectively. In the protocol registered in PROSPERO, these three categories of risk of bias were described

as (1) Study design, (2) Compliance and drop out, and (3) Blinding. Sample size was not evaluated in this section because this is a marker of quality, not risk of bias.[14] Similarly, external validity of the study sample (i.e. specialist subgroups, e.g. occupational setting, tertiary clinic, severe tinnitus only) was not formally evaluated.

Measures of effect

From each study a Standardized Mean Difference (SMD) was calculated for every included score obtained on all tinnitus questionnaires. SMD was calculated for each post-baseline time point and was defined as the difference between the group mean questionnaire score at baseline and after *n* weeks of no-intervention waiting, divided by the pooled standard deviation. A positive SMD indicated an improvement over time. This difference was then converted to Hedges' g[15], a commonly used measure of effect which controls for the bias in effect size that might be introduced by studies with small participant sample size. The test-retest correlation between the repeated time points was set to 90% for all questionnaires. Where multiple questionnaires were used at the same time point, a mean effect size was calculated by averaging the individual effect sizes.

Meta-analysis

Mean effect sizes across studies were calculated using Comprehensive Meta-Analysis (Version 2.2.048). For the primary synthesis, the latest time-point in each study was selected, and a random effects model was run. A random effect model assumes that the true effect may vary from study to study, i.e. here it was assumed that changes in the impact of tinnitus over time are not likely a constant effect but may be influenced by study factors such as age of participants, duration of tinnitus, education level, or general health. Sensitivity analyses were conducted pooling effect sizes per time from baseline (~ 6 weeks, ~12 weeks, ~ 6 months). For all meta-analyses it was reasonably assumed that the multi-item questionnaires included showed sufficient convergent validity to be pooled, e.g.

tinnitus questionnaires are generally demonstrated to measure the same underlying construct of the everyday impact of tinnitus.

RESULTS

Literature searches were conducted in December 2013 and updated in October 2015. The initial database searches yielded 902 records. A further 23 eligible records were identified through hand searches (see Figure 1 for PRISMA flow chart). Following screening, 50 records were retrieved for full text review. Twenty-five studies were included[16-40] (Table 1) and 21 studies were suitable for inclusion in meta-analyses.[16-18,21-26,28,29,31-40] Among these, one record reported two independent control groups which were retained as independent groups.[18] Ross et al (2007) also used multiple control groups depending on tinnitus duration and this enabled two independent control groups to be used for the meta-analysis.[35] Thus, the meta-analyses examined up to 23 no-intervention or waiting-list control groups from 21 study records.

Missing data

Data queries were satisfactorily answered for eight study records, and partly answered for one other (See Supplemental Table S3).

Data synthesis

The period spent on the no-intervention or waiting-list period varied from 1-52 weeks, with an average of 12 weeks. Information about the individual percentage and effect size of change in tinnitus severity, as measured by tinnitus questionnaire score is provided in Table 2. Two studies (Fackrell et al., 2016, Krick et al., 2015) were excluded from the meta-analysis because the interval between assessments for most or all patients was as little as 7 days.[20,30]

Caffier et al. (2006) was excluded as numerical data were not sufficiently available.[19] Jakes et al. (1992) was excluded[27] from meta-analysis as their tinnitus outcome questionnaire was the Tinnitus Effects Questionnaire[11] which does not yield a global score.

Across the remaining studies, over the longest period reported, there was a small decrease in global tinnitus of 2.3% indicating a trend for improvement over time. How clinically meaningful that is cannot be interpreted; although it was assumed that tinnitus questionnaires measure the same construct of the everyday impact of tinnitus, clinically meaningful change scores on those questionnaires differ. Strikingly, no study demonstrated statistically significant worsening of tinnitus over time. There was, however, considerable heterogeneity across studies. Reports of changes in depression, anxiety, quality of life, and tinnitus loudness were few and not significant.

Risk of bias assessment

A summary of the risk of bias of the 21 study records that were included in the meta-analysis is shown in Table 3. Low risk of bias was achieved on 51% of occasions. Six studies had a high or unclear risk of bias on two of the criteria[22,23,34,37,38], while one study had a high or unclear risk of bias on all three criteria.[36] Detection bias was the most poorly reported. Support for judgement concerning selection bias, attrition bias and detection bias is provided in Supplemental Tables S4, S5 and S6 respectively.

Effects over time on global tinnitus

Twenty-three study groups (788 participants) in 21 study records reported changes in tinnitus over time. Effect sizes (Hedge's *g*) for the maximum interval within studies ranged -.17 to .55. The primary meta-analysis pooled data across studies using the longest timeframe reported in each study record, irrespective of the absolute length of time (23 study groups, *M* = 12 weeks, range = 4-52 weeks). There was significant heterogeneity

across studies; $Q(df=22) = 112.97$, $p < 0.001$, $I^2 = 80.53$. In a random effects model the mean effect size was statistically significant in favor of tinnitus improving; Hedge's $g = .122$ (95% CI = .055 to .188), $p < 0.001$ (Figure 2).

For 14 study groups, tinnitus questionnaire data at up to two months from baseline were available ($M = 6.6$ weeks, range = 2-8 weeks). There was significant heterogeneity; $Q(df=13) = 48.7$, $p < 0.001$, $I^2 = 73.29$. In a random effects model the mean effect size was statistically significant towards improvement in tinnitus; Hedge's $g = .097$ (95% CI = .019 to .176), $p = 0.015$. The pattern was the same for ten study groups reporting questionnaire data up to four months ($M = 10.6$ weeks, range = 9-12 weeks, n = 238). There was significant heterogeneity; $Q(df=8) = 51.6$, $p < 0.001$, $I^2 = 82.57$. In a random effects model the mean effect size was still significant; Hedge's $g = .154$ (95% CI = .027 to .281), $p = 0.018$. Longer term effects (measured in seven study groups) were not significant however ($M = 29$ weeks, range = 17-52, n = 256); Hedge's $g = .112$ (95% CI = -.013 to .236), $p = 0.079$.

Effects over time on depression

Eight studies (301 participants) reported changes in depression questionnaire scores, using Hospital Anxiety Depression Scale-Depression[41] or Beck Depression Index[42] questionnaires over intervals ranging 6-26 weeks ($M = 16.2$). Henry et al. (1998) measured BDI score at three time intervals (baseline, 26 and 52 week later). Hedge's g across the eight studies ranged .469 to .182, with one study favoring a worsening and two studies favoring an improvement in scores over time (Figure 3). The pooled effect size across all studies (using the 26 week measure from Henry et al., 1996) was positive but not significant (Hedge's $g = .006$ [95% CI = -.045 to .057], $p = 0.828$) indicating no significant change in depression over time.

Effects over time on generalized anxiety

Five studies (161 participants) reported changes in anxiety questionnaire scores, using the Hospital Anxiety Depression Scale-Anxiety (HADS-A; Zigmond and Snaith 1983) over intervals ranging 6-12 weeks (*M* = 8). Hedge's *g* across studies ranged .089 to .206, with one study favoring an improvement in scores over time (Figure 4). The pooled effect size across all studies was positive but not significant (Hedge's *g* = .058 [95% CI = -.012 to .127], *p* = 0.104) indicating no significant change in anxiety over time. Andersson et al. (2002)[16] additionally measured "fear of anxiety-related somatic sensation" using the Anxiety Sensitivity Index[43] noting a slight improvement over time; mean score reduced from 19.1 (SD = 12.7) at baseline to 17.8 (SD = 12.1) at 6 weeks. In contrast, Andersson et al. (2005)[17] report an increase in Anxiety Sensitivity Index score in their waiting list control group after about 6 weeks; mean score increased from 18.9 (± 10.0) to 26.3 (± 10.5).

Effects over time on quality of life

Only two studies reported on changes in quality of life over time, with similar results. Rief et al. (2005)[34] reported on change on the Questions of Life Satisfaction questionnaire[44] ; after eight weeks their participants reported a 15.5 point (on a 100-point scale) increase in questionnaire score suggesting improvement in subjective satisfaction with quality of life, but at six-month follow-up there was no difference from baseline. Westin et al. (2011)[40] used the Quality of Life Inventory[45], and measured change at 10 weeks finding a 0.2 increase in score on a three-point scale which was not significant.

Effects over time on tinnitus loudness

In six studies, tinnitus loudness was measured using a visual analogue scale.[16,17,21,27,28,34] However, Jakes et al. (1992)[27] reported abandoning the measure during the study for several reasons including poor compliance, and Andersson et al. (2005)[17] did not report numerical

values. Of the four remaining, three used a 0-10 scale and reported a 0.8-point decrease[34], no change[16], and a 0.1-point increase[28] in scores respectively after 6-8 weeks watchful waiting. One study used a 0-4 scale and reported a reduction of <0.1 after 4 weeks.[21]

Although single item measures of tinnitus show good correlation with each other they do not measure meaningful tinnitus related constructs so these data were not subjected to meta-analysis.

DISCUSSION

This systematic review with a meta-analysis presents the most inclusive evaluation of the natural time course of tinnitus under controlled experimental conditions to date. The random effects meta-analysis gives a reliable overall summary of findings since the analysis accounts for heterogeneity and is weighted by sample size. This revealed a small but significant improvement in global tinnitus severity up to four months, but studies with longer assessment periods did not reveal any change. This finding may reflect a lack of statistical power for this subgroup analysis or an insensitivity of tinnitus questionnaire measures over longer periods. Even for the 2 and 4 month analyses, it must be cautioned that we cannot ascertain with certainty whether the small statistically significant improvement is equivalent to a clinically meaningful improvement that is noticeable to people with tinnitus. Clinical interpretation of the findings by anchoring numerical values against patient reported experience is under-reported to date.

In contrast to the small improvements in global tinnitus severity, our meta-analyses did not reveal statistically significant improvement in measures of mood. This finding contradicts that of a study by Posternak (2001) which looked at mental health conditions in isolation and found improvement while on waiting-list control groups.[4] It is possible that the null findings in the current study simply represent the relatively low number of tinnitus studies that had incorporated a measure of depression or generalized anxiety.

Although this systematic review accepted studies testing any form of tinnitus intervention, the meta-analysis was biased towards psychological interventions with 11 out of the 21 included studies testing a psychological management modality. The current study, although skewed towards psychological treatment trials adds to previous work by Hesser et al (2011)[7] because it assessed a much broader range of tinnitus experiences than this previous work. Our findings incorporated those participants enrolled into a range of tinnitus intervention

studies, namely tinnitus retraining therapy, education, auditory discrimination training, self-help using books, drug treatment and Qigong (a combination of body posture, breathing control and meditation developed in China). We believe that this inadvertent bias towards psychological interventions is in large part a reflection of the type of control group favored by trial designs assessing different types of tinnitus study. Pharmacological intervention studies will typically use a placebo medication as control, while studies assessing a device treatment such as repetitive transcranial magnetic stimulation or laser therapy will generally employ a sham treatment as control. For psychological therapies, such as cognitive behavior therapy, acceptance and commitment therapy, or mindfulness meditation, a placebo or sham psychological therapy control is unethical for the trial design and so those trials are therefore much more likely to use a no-intervention or waiting-list control. Moreover, psychological therapies present a routine therapeutic option for people with bothersome tinnitus, often with a natural waiting list for an initial appointment, and so a no-intervention control is often a straightforward pragmatic option. One limitation for our interpretation of the findings is thus that it is not clear whether tinnitus patients consenting to participate in a psychological treatment trial are representative of tinnitus patients in general, and more specifically whether they are equivalent to those consenting to join a drug trial or medical device study. Two studies drew their participants directly from US military veterans[23,24], hence participants were more likely to be male, have been exposed to a greater than average risk of noise induced hearing loss and to the psychological stress associated with military service. These limitations are mitigated to some degree by the meta-analysis which pooled findings from a wide range of studies. One further potential limitation of note is the exclusion of studies not available in English (because of limited resources), and studies that appear only in the grey literature. Whilst excluding grey literature may have introduced a publication bias, including grey literature could in itself introduce bias if the included sample

of unpublished studies was not representative of all unpublished studies. It would be interesting to explore this issue in further analyses.

CONCLUSIONS

Participants enrolled into clinical trials assessing tinnitus interventions generally demonstrate a small but statistically significant improvement in self-reported global tinnitus severity scores over time, despite receiving no intervention. This finding provides statistical evidence that tinnitus generally improves over time, albeit the effect is highly variable across individuals and how clinically meaningful the effect is cannot be interpreted at a general level. This evidence can therefore cautiously be used when counselling patients.

ACKNOWLEDGEMENTS

This work was supported by a British Society of Audiology Applied Research Grant awarded to the authors.

BIBLIOGRAPHY

1. Andersson G, Vretblad P, Larsen HC, Lyttkens L. Longitudinal follow-up of tinnitus complaints. Arch Otolaryngol Head Neck Surg. 2001; 127:175-9.

2. Nondahl DM, Cruickshanks KJ, Wiley TL, Klein R, Klein BE, Tweed TS. Prevalence and 5-year incidence of tinnitus among older adults: the epidemiology of hearing loss study. J Am Acad Audiol. 2002; 13:323-31.

3. Nondahl DM, Cruickshanks KJ, Wiley TL, Klein BE, Klein R, Chappell R, Tweed TS. The ten-year incidence of tinnitus among older adults. Int J Audiol. 2010; 49:580-5.

4. Posternak MA, Miller I. Untreated short-term course of major depression: a meta-analysis of outcomes from studies using wait-list control groups. J Affect Disord. 2001; 66:139-46.

5. Young C. What happens when people wait for therapy? Assessing the clinical significance of the changes observed over the waiting period for clients referred to a primary care psychology service. Primary Care Mental Health 2006; 4:113-119.

6. Steinert C, Stadter K, Stark R, Leichsenring F. The Effects of Waiting for Treatment: A Meta-Analysis of Waitlist Control Groups in Randomized Controlled Trials for Social Anxiety Disorder. Clin Psychol Psychother. 2016; Jul 22. doi: 10.1002/cpp.2032. [Epub ahead of print]

7. Hesser H, Weise C, Rief W, Andersson G. The effect of waiting: A meta-analysis of wait-list control groups in trials for tinnitus distress. J Psychosom Res. 2011; 70:378-384.

8. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009; 339:b2700.

9. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA. 2000; 283:2008-12.

10. Centre for Reviews and Dissemination. Systematic Reviews. CRD's guidance for undertaking reviews in health care. York, UK. York Publishing Services; 2008.

11. Hallam RS, Jakes SC, Hinchcliffe R. Cognitive variables in tinnitus annoyance. British Journal of Clinical Psychology, 1988; 27:213-222.

12. Hall DA, Haider H, Szczepek AJ, et al. Systematic review of outcome domains and instruments used in clinical trials of tinnitus treatments in adults. Trials. 2016; 17:270.

13. Higgins JP, Altman DG, Gøtzsche PC. Cochrane Bias Methods Group; Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011; 18:343:d5928.

14. Higgins JPT, Altman DG. Assessing risk of bias in included studies. In: Higgins JPT and Green S, eds. Cochrane Handbook of Systematic Reviews of Interventions. West Sussex UK: John Wiley; 2008:187-241.

15. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. Psychological Methods. 1998; 3:486−504.

16. Andersson G, Stromgren T, Strom L, Lyttkens, L. Randomized controlled trial of internet-based cognitive behavior therapy for distress associated with tinnitus. Psychosom Med. 2002; 64: 810-6.

17. Andersson G, Porsaeus D, Wiklund M, Kaldo V, Larsen HC. Treatment of tinnitus in the elderly: a controlled trial of cognitive behavior therapy. International Journal of Audiology. 2005; 44:671-675.

18. Biesinger E, Kipman U, Schatz S, Langguth B. Qigong for the treatment of tinnitus: A prospective randomized controlled study. Journal of Psychosomatic Research. 2010; 69:299-304.

19. Caffier PP, Haupt H, Scherer H, Mazurek B. Outcomes of long-term outpatient tinnitus-coping therapy: psychometric changes and value of tinnitus-control instruments. Ear Hear. 2006;27:619-27.

20. Fackrell K, Hall DA, Barry JG, Hoare DJ. Psychometric properties of the Tinnitus Functional Index (TFI): Assessment in a UK research volunteer population. Hear Res. 2016; 335:220-35.

21. Henry JL, Wilson PH. The Psychological Management of Tinnitus: Comparison of a Combined Cognitive Educational Program, Education Alone and a Waiting-List Control. Int Tinnitus J. 1996; 2:9-20.

22. Henry JL, Wilson PH. An evaluation of two types of cognitive intervention in the management of chronic tinnitus. Behaviour Therapy. 1998; 27:156-66.

23. Henry JA, Loovis C, Montero M, Kaelin C. Randomized clinical trial: group counseling based on tinnitus retraining therapy. Journal of Rehabilitation Research and Development. 2007; 44:21-32.

24. Henry JA, Griest S, Thielman E, McMillan G, Kaelin C, Carlson KF. Tinnitus Functional Index: Development, validation, outcomes research, and clinical application. Hearing Research. 2016; 334:58-64.

25. Herraiz C, Hernandez FJ, Plaza G, de los Santos G. Long-term clinical trial of tinnitus retraining therapy. Otolaryngol Head Neck Surg. 2005; 133:774-779.

26. Herraiz C, Diges I, Cobo P, Aparicio JM, Toledano A. Auditory discrimination training for tinnitus treatment: the effect of different paradigms. Eur Arch Otorhinolaryngol. 2010; 267:1067-1074.

27. Jakes SC, Hallam RS, McKenna L, Hinchcliffe R. Group cognitive therapy for medical patients - an application to tinnitus. Cognitive Therapy and Research. 1992; 16:67-82.

28. Kaldo V, Cars S, Rahnert M, Larsen HC, Andersson G. Use of a self-help book with weekly therapist contact to reduce tinnitus distress: A randomized controlled trial. Journal of Psychosomatic Research. 2007; 63:195-202.

29. Kreuzer PM, Goetz M, Holl M, et al. Mindfulness-and body-psychotherapy-based group treatment of chronic tinnitus: a randomized controlled pilot study. BMC Complementary & Alternative Medicine. 2012; 12:235-242.

30. Krick CM, Grapp M, Daneshvar-Talebi J, Reith W, Plinkert PK, Bolay HV. Cortical reorganization in recent-onset tinnitus patients by the Heidelberg Model of Music Therapy. Frontiers in neuroscience. 2015; 9:49.

31. Kroner-Herwig B, Frenzel A, Fritsche G, Schilkowsky G, Esser G. The management of chronic tinnitus - Comparison of an outpatient cognitive-behavioral group training to minimal-contact interventions. Journal of Psychosomatic Research. 2003; 54:381-389.

32. Malouff JM., Noble W, Schutte NS, Bhullar N. (2010) The effectiveness of bibliotherapy in alleviating tinnitus-related distress. Journal of Psychosomatic Research. 2010; 68:245-251.

33. Neri G, De Stefano A, Baffa C et al. Treatment of central and sensorineural tinnitus with orally administered Melatonin and Sulodexide: personal experience from a randomized controlled study. Acta Otorhinolaryngologica Italica. 2009; 29:86-91.

34. Rief W, Weise C, Kley N, Martin A. Psychophysiologic treatment of chronic tinnitus: a randomized clinical trial. Psychosomatic Medicine. 2005; 67:833-838.

35. Ross UH, Lange O, Unterrainer J, Laszig R. Ericksonian hypnosis in tinnitus therapy: effects of a 28-day inpatient multimodal treatment concept measured by Tinnitus-Questionnaire and Health Survey SF-36. Eur Arch Otorhinolaryngol. 2007; 264:483-8.

36. Sadlier M, Stephens SDG, and Kennedy V. (2008) Tinnitus rehabilitation: a mindfulness meditation cognitive behavioural therapy approach. Journal of Laryngology and Otology. 2008; 122:31-37.

37. Seydel C, Haupt H, Szczepek, AJ, Klapp BF, Mazurek B. Long-term improvement in tinnitus after modified tinnitus retraining therapy enhanced by a variety of psychological approaches. Audiology and Neurotology. 2009; 15:69-80.

38. Sönmez O, Külahlı I, Vural A, Sahin MI, Aydın M. The evaluation of ozone and betahistine in the treatment of tinnitus. European Archives of Oto-Rhino-Laryngology. 2013; 270:1999-2006.

39. Weise, C, Heinecke, K , Rief, W. Biofeedback-Based Behavioral Treatment for Chronic Tinnitus: Results of a Randomized Controlled Trial. Journal of Consulting and Clinical Psychology 2008; 76:1046-1057.

40. Westin VZ, Schulin M, Hesser H, et al. Acceptance and Commitment Therapy versus Tinnitus Retraining Therapy in the treatment of tinnitus: A randomised controlled trial. Behaviour Research and Therapy 2011; 49:737-747.'

41. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand. 1983; 67:361-70.

42. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch. Gen. Psychiatry. 1961; 4: 561–571.

43. Reiss S, Peterson RA, Gurksy DM, McNally RJ. Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. Behav Res Ther. 1986; 34:1-8.

44. Henrich G, Herrschbach P. Questions on Life Satisfaction (FLZ)- a short questionnaire for assessing subjective quality of life. Euro J Psychol Assess. 2000; 16:150 –9.

45. Frisch MB, Cornell J, Villanueva M, Retzlaff PJ. Clinical validation of the Quality Of Life Inventory: a measure of life satisfaction for use in treatment planning and outcome assessment. Psychological Assessment. 1992; 4:92e101.

Figure legends

Figure 1. PRISMA flow chart

Figure 2. Meta-analysis of change in self-reported tinnitus severity over longest interval in individual studies indicating an improvement over time. Black square = effect size (Hedge's g) in that study. Black diamond = pooled effect size. The relative sample size and hence relative influence of individual studies on the pooled effect size is indicated by the size of the black square; i.e. the studies by Ross have the greatest influence on the pooled result, followed by Henry 2007 and 2015 etc.

Figure 3. Meta-analysis of change in self-reported depression shows no significant change over time. Black square = effect size (Hedge's g) in that study. Black diamond = pooled effect size. The relative sample size and hence relative influence of individual studies on the pooled effect size is indicated by the size of the black square.

Figure 4. Meta-analysis of change in self-reported anxiety shows no significant change over time. Black square = effect size (Hedge's g) in that study. Black diamond = pooled effect size. The relative sample size and hence relative influence of individual studies on the pooled effect size is indicated by the size of the black square.

Table legends

Table 1. Study characteristics

Table 2. Percentage and effect size of change in tinnitus questionnaire score. NR = baseline not reported. Negative percentage change indicates decreased questionnaire score (tinnitus improves); positive percentage change indicates increased questionnaire score (tinnitus worse).

Table 3. Risk of bias summary table

Appendices

Supplemental file 1. Table S1. Examples of outcome instruments acceptable for data collection in the present review

Supplementary file 2. Table S2. Sample search strategy using PubMed

Supplementary file 3. Table S3. Summary of those authors who were contacted for clarification or additional information

Supplementary file 4. Table S4. Support for judgement concerning selection bias which includes sequence generation for the randomization process and allocation concealment

Supplementary file 5. Table S5. Support for judgement concerning attrition bias

Supplementary file 6. Table S6. Support for judgement concerning detection bias