AJA

Research Article

Randomized Controlled Trial of a Novel Device for Tinnitus Sound Therapy During Sleep

Sarah M. Theodoroff,[a,b] Garnett P. McMillan,[a] Tara L. Zaugg,[a] Megan Cheslock,[a] Chan Roberts,[a] and James A. Henry[a,b]

Purpose: The aim of this study was to determine if a customized stimulus from the Otoharmonics Levo System reduces tinnitus perceptions and reactions for people with bothersome tinnitus.
Method: Sixty participants were randomized to 1 of 3 groups that used sound therapy devices during sleep that differed in their acoustic stimulus: (a) tinnitus-matched (TM), (b) noise stimulus (NS), and (c) bedside sound generator (BSG). Outcome measures were the Tinnitus Functional Index (TFI), numeric rating scale of tinnitus loudness, and tinnitus loudness match. A Bayesian hierarchical model was fit to estimate the differences in treatment efficacy among groups.
Results: Average tinnitus reactions and perceptions improved across treatment groups. We are at least 87% certain that treatment with TM or NS reduces mean TFI compared to treatment with BSG, with an estimated relative efficacy of 4.5–5 points greater reduction. We are at least 95% certain that treatment with TM results in greater reduction in mean numeric rating scale (NRS) of tinnitus loudness compared to the other groups, with an estimated relative efficacy of about 0.75 points greater reduction.
Conclusions: This study offers some support for greater average improvement in reactions to tinnitus with TM or NS devices compared to the BSG device. The TM group, compared to the BSG and NS groups, showed a greater reduction in ratings of tinnitus loudness on the NRS on average.
Supplemental Material: https://doi.org/10.23641/asha.5545759

Tinnitus is the perception of sound that has no source outside of the head. It is most typically associated with exposure to loud noise, which can also cause hearing loss (Axelsson & Barrenäs, 1992; Penner & Bilger, 1995). A direct correlation exists between degree of hearing loss and prevalence of tinnitus—the likelihood of incurring tinnitus increases with a greater degree of hearing loss (Coles, 2000). In general, tinnitus can occur as the result of noise damage, blast exposures, head and neck trauma or pathology, drugs or medications, and other medical conditions (e.g., acoustic neuroma, cardiovascular and cerebrovascular disease, hyper- and hypothyroidism; Hoffman & Reed, 2004; Meikle, Creedon, & Griest, 2004). Tinnitus can result in psychological reactions (e.g., anxiety, sadness, distress), attention and concentration difficulties, problems with sleeping, and overall reduced quality of life (Cima, Crombez, & Vlaeyen, 2011; Crönlein et al., 2016; Dobie, 2003; Erlandsson & Hallberg, 2000).

Research has shown that, regardless of the initial injury associated with tinnitus onset (e.g., cochlear damage), the continued perception of tinnitus is generated by neural activity within the central auditory system (Eggermont, 2003). Although the exact neural mechanism giving rise to the percept is still unknown (Eggermont, 2015), the various neurophysiological models of tinnitus are helpful in providing the foundation for different therapies to be developed.

A commonly used therapy for tinnitus is sound stimulation. One rationale for using sound stimulation is to induce neural plastic changes in the central auditory system to counteract the maladaptive changes thought to be linked to the tinnitus percept (Shore, Roberts, & Langguth, 2016). Another proposed benefit of using sound stimulation is

[a]VA RR&D National Center for Rehabilitative Auditory Research, VA Portland Health Care System, OR
[b]Department of Otolaryngology–Head & Neck Surgery, Oregon Health & Science University, Portland
Correspondence to Sarah M. Theodoroff: sarah.theodoroff@va.gov

Editor-in-Chief: Sumitrajit Dhar
Editor: Owen Murnane
Received March 3, 2017
Revision received May 2, 2017
Accepted July 18, 2017
https://doi.org/10.1044/2017_AJA-17-0022

Disclosure: *This research was funded by the Baker Group, LLP, who manufactured the Levo System device (Portland VA Research Foundation #387004) and by the Department of Veterans Affairs, Rehabilitation Research & Development (RR&D) Service (C9247S and C9230C) at the VA Portland Health Care System in Portland, Oregon. Once agreement was reached as to study design, all study procedures and data analyses were conducted independent of any company influence.*

that it may help induce habituation, resulting in noticing the tinnitus less often and reduced reactions to tinnitus (McKenna, 2004).

There are many different methods and products offering sound therapy for tinnitus patients that are primarily to be used when the individual is awake (Hoare, Searchfield, El Refaie, & Henry, 2014). Pedemonte, Testa, Diaz, and Suarez-Bagnasco (2014) examined the use of sound stimulation during sleep. Study participants matched their tinnitus percept as closely as possible to an acoustic stimulus, which was played back to them during sleep to induce changes in neural activity and networks associated with information processing. This study provided the basis for the development of a commercially available product that was designed and manufactured by the Otoharmonics Corporation (Levo System, version 2.2.6).

Objective

Evidence-based research should guide the interventions recommended for people distressed by their tinnitus. Randomized controlled trials (RCTs) that are properly conducted are the most important source for providing such evidence (Keech, Gebski, & Pike, 2007). This study was an RCT designed to determine if an acoustic stimulus mimicking the tinnitus perception delivered during sleep from the Otoharmonics Corporation's Levo System reduces tinnitus-related distress and/or perceived loudness of tinnitus during awake hours for people who experience bothersome tinnitus. More specifically, the goal was to evaluate the efficacy of the Levo System, a custom in-ear device designed to be used with a tinnitus-matched stimulus, compared to (a) the same Levo System but with a noise stimulus chosen by the participant from a limited range of options and (b) the Marsona 1288 bedside sound generator (manufactured by Marpac) for reducing tinnitus-related distress and/or perceived loudness of tinnitus, measured using the Tinnitus Functional Index (TFI; Meikle et al., 2012) total score, numeric rating scale (NRS) of tinnitus loudness, and tinnitus loudness match (LM) in dB sensation level (SL) at 1 kHz.

Method

Recruitment and Screening

Participants were recruited from individuals who previously participated in research projects at the National Center for Rehabilitative Auditory Research and gave their consent to be contacted for future studies. Participants were also recruited from the surrounding community via research flyers posted in public locations and advertisements in local newspapers and online. Initial screening was done over the telephone to ensure as much as possible that callers who were invited for an assessment were suitable candidates. Both veterans and nonveterans were recruited.

Interested callers were asked screening questions to determine (a) if they experienced tinnitus that was either intermittent or constant, (b) the duration and temporal characteristics of their tinnitus using the Tinnitus Screener (Henry, Griest, Austin, et al., 2016), and (c) how much the tinnitus affected their quality of life using the Tinnitus and Hearing Survey (THS; Henry, Griest, et al., 2015). When tinnitus and hearing loss co-occur, it is common for people to misattribute the tinnitus as the cause of the hearing problem rather than the hearing loss (Ratnayake, Jayarajan, & Bartlett, 2009; Zaugg, Schechter, Fausti, & Henry, 2002). The THS is a 10-item questionnaire with three sections evaluating the degree to which auditory complaints (tinnitus, hearing problem, and sound tolerance problem) have been present over the past week, with response options ranging from "not a problem" to "very big problem." Candidates were required to score at least a 6 for Section A of the THS, which addresses tinnitus-specific problems. If candidates scored high enough on the THS, then additional screening questions were asked to determine their noise exposure, hearing aid use, and whether or not they were currently engaged in tinnitus treatment/management. Candidates who passed the telephone screening were invited to the National Center for Rehabilitative Auditory Research for evaluation to determine eligibility.

Candidates were paid $20 with cash or a prepaid debit card for attending the initial appointment, whether or not they were enrolled. At the initial appointment, candidates first underwent informed consent followed by the Health Insurance Portability and Accountability Act authorization, administered by study personnel. All participants provided informed consent prior to any procedures being performed; the VA Portland Health Care System's Institutional Review Board reviewed and approved the study protocol (IRB 3631).

Inclusion Criteria

To be included, candidates were required to (a) have a score of ≥ 25 (out of a maximum score of 100) on the TFI; (b) pass the Blessed Orientation-Memory-Concentration Six-Item Test (Katzman et al., 1983), a screening test for dementia; (c) have sufficient hearing to enable perception of the acoustic stimulus at required levels (no single audiometric threshold of > 70 dB HL, 0.25–8 kHz); (d) report the presence of tinnitus for at least 6 months; and (e) be English speaking and capable and willing to complete all aspects of the study.

Exclusion Criteria

Candidates were excluded if they (a) reported they "never," "rarely," or "sometimes" hear noises in their ears or head; (b) had a hearing threshold of > 70 dB HL at any tested frequency between 0.25 and 8 kHz; (c) had conductive hearing loss defined as an air–bone gap of 15 dB at more than two frequencies in one ear; (d) were unable to read and respond appropriately to instructions that appeared on the computer screen and/or to perform the auditory-related procedures; (e) were exposed to loud noise as part of employment or recreational activities (e.g., construction work or factory work requiring use of hearing protection, target shooting); (f) had a behavioral flag in their medical record (at the discretion of study staff); (g) had mental,

emotional, or health conditions that would preclude full study participation; or (h) reported already using one of the study devices.

Outcome Measures

Efficacy of intervention was evaluated with the TFI, NRS, and LM at 1 kHz. NRS and LM are measures of tinnitus perception, whereas the TFI measures reactions to or effects of tinnitus.

TFI

The TFI is a 25-item self-report questionnaire that has documented validity both for scaling the negative impact of tinnitus and for measuring treatment-related changes in effects of tinnitus (*responsiveness*; Meikle et al., 2012). The total score for the TFI ranges from 0 to 100, with higher scores indicating greater problems with tinnitus. The TFI has excellent internal consistency (Cronbach's $\alpha = .97$) and high test–retest reliability ($r = .86$). The authors of the TFI estimated that a 13-point decrease on the TFI is likely to reflect a change that feels meaningful to an individual (Meikle et al., 2012).

NRS

The NRS consists of a line marked with equal intervals labeled 0–10 (0 = *no tinnitus*, 10 = *very loud tinnitus*; Johnson, 2005), as distinguished from a visual analog scale, which requires marking a position along a continuous line anchored between two end-points. Participants were instructed to draw a vertical line at any point between the anchors 0 and 10, indicating the loudness of their tinnitus at that moment. This was completed in the sound-attenuated suite before any audiometric or psychoacoustic testing to ensure that auditory stimulation would not affect the tinnitus perception. NRS scores were obtained by measuring the distance from 0 to the vertical line. That value was divided by the total distance of the line and rounded to the nearest tenth.

LM

Numerous studies have provided results suggesting that tinnitus loudness is significantly underestimated when LMs are obtained at the tinnitus frequency. For this reason, LMs were obtained and reported at 1 kHz, as recommended by Goodwin and Johnson (1980). The "tinnitus ear" was designated as the ear in which the tinnitus was perceived to be loudest. If the tinnitus was perceived "in the head" or to be symmetrical, then the tinnitus ear was selected as the ear with poorer hearing. All stimuli used for LM testing were presented to the ear contralateral to the tinnitus ear. First, a hearing threshold in dB HL at 1 kHz in 2-dB steps was found. Next the 1-kHz tone was presented at an intensity of 10 dB above the threshold (i.e., 10 dB SL). The participant was instructed to tell the audiologist if the intensity of the tone needed to be made louder or softer to match the loudness of the tinnitus in the opposite ear. The participant continued to direct the clinician to make changes in the intensity of the stimulus until it was perceived to be the same loudness as the tinnitus. LMs at 1 kHz were reported in dB SL (difference between LM and hearing threshold).

NRS and LM testing were performed to capture any changes in tinnitus perception; these measurements are two different approaches yielding subjective judgments about tinnitus loudness.

Procedures

Initial Visit

After the informed consent and the Health Insurance Portability and Accountability Act authorization forms were signed, candidates completed a case history form, the Tinnitus Screener, THS, TFI, and the Blessed Orientation-Memory-Concentration Six-Item Test. This was followed by NRS testing, standard audiologic evaluation (audiometry and immittance testing), and LM testing. Audiologic testing was performed by a study audiologist in a sound-attenuated suite, using conventional audiologic instrumentation (audiometer and immittance system). Hearing thresholds were obtained using pure-tone air- and bone-conduction testing.

Qualified participants were enrolled and randomized into three treatment groups with equal probability: (a) Levo System with a tinnitus-matched stimulus (TM group), (b) Levo System with a noise stimulus (NS group; white noise and/or band noise), or (c) Marsona 1288 Sound Conditioner/Tinnitus Masker (bedside sound generator device; BSG group).

Participants randomized to the TM and NS groups had earmold impressions made, which were sent to a lab to make custom-fit earbuds; the earbuds are specially designed earphones, which present the in-ear acoustic stimuli and therefore are an essential component of the Levo System.

The dispense appointment for participants in the TM and NS groups was not scheduled until the earbuds arrived on-site. Participants randomized to receive the BSG device were scheduled for their dispense appointment. All participants had outcomes measured prior to randomization (baseline) and at all follow-up visits.

Levo System Devices

The study audiologists were trained by an Otoharmonics Corporation representative to dispense the TM and NS sound therapy devices according to company guidelines. Participants randomized to the TM group were fit according to company guidelines. Participants in the TM and NS groups received the standard company instructions on care and maintenance of the sound therapy devices and custom-fit earbuds. Participants randomized to the TM group created a tinnitus "sound print" using company software to identify the sound(s) that most closely matched the sound(s) of their tinnitus. They were asked to listen to the sound print every night and to adjust its volume settings to match the loudness of the tinnitus (different from the LM at 1 kHz performed in clinic). Participants randomized to the NS group were instructed to pick a sound

(either white noise, band noise, or a combination of noise stimuli) that was comfortable. Specific instructions were "pick a sound that is comfortable for you, something that you could comfortably listen to all night. It is fine to find something that is soothing, but it is not necessary for the sound to be soothing. It should not be annoying." After self-performing the LM with the device, they adjusted the volume of the device to a comfortable level prior to using it during sleep.

Marsona Device

Participants randomized to the BSG group were instructed on how to use the sound therapy device (including care and maintenance) during sleep and received the same instructions on "how to choose a sound" as received by participants in the TM and NS groups. Specifically, they were instructed to select a sound that was comfortable to listen to all night and that should not be annoying in any way. Participants in the BSG group were allowed to select any or a combination of the various sounds the device generated. In addition, because the BSG device delivers sounds in the environment and not at ear level, participants were told if they shared a bedroom to feel free to include their partner in the conversation about what sound to listen to throughout the night.

Follow-Up Visits

Participants returned about 3–6 weeks following their enrollment date to receive their respective devices, then 1–2 weeks later to check device adjustment and usage. During the follow-up visits, participants randomized to the TM group had their tinnitus sound print adjusted as needed. At each follow-up visit (i.e., Visits 2–4), outcome testing included the Tinnitus Screener (Visit 2 only), TFI, NRS, THS, audiometric testing (0.25–8 kHz), and tinnitus testing (i.e., LM at 1 kHz). Participants used their device for about 3 months, at which time they attended the final study visit (i.e., Visit 4) to complete outcome testing and a follow-up audiologic evaluation. They also answered questions from an exit interview to determine their subjective impressions of using the device. This ended their participation in the study, and they were allowed to keep their device. At this time, participants in the NS group were informed that they had been using the sound therapy device differently than how it is used commercially. These participants were given the option (if keeping the device) to have it reprogrammed and create a tinnitus sound print or keep it programmed to the noise stimulus they had been listening to, whichever they preferred.

Counseling

All participants received general educational counseling about tinnitus using the company's counseling booklet. Topics covered by the study audiologists included tinnitus basics, hearing basics, how the brain works (e.g., neuroplasticity and habituation), and options for tinnitus management. Participants in the BSG and NS groups received the same counseling as those in the TM group, with the exception of the last page of the counseling booklet, which was only shown to participants in the TM group because it addressed specific information about the acoustic stimulus the TM group received.

Participant Safeguards

Safeguards were established to minimize risk of exposure to high levels of noise: The Levo System device will not allow the health care professional to set the treatment duration for more than 8 hr, as it is limited in the software design. Also, the suggested guidelines indicate that the treatment duration should be limited to less than 8 hr once the sound intensity exceeds 65 dB SPL. Specifically, the company guidelines recommend that when the device is used at 80 dB SPL, the treatment duration should be limited to 4 hr, and when used at maximum output (84.9 dB SPL), 2 hr should be the limit. The hours of use are preset as part of the process of programming the device in the clinic. Once the designated hours of use are achieved, the device stops delivering the sound therapy. The device has data-logging capabilities, and at every follow-up visit, data were downloaded and reviewed to determine if participants were using the device as instructed.

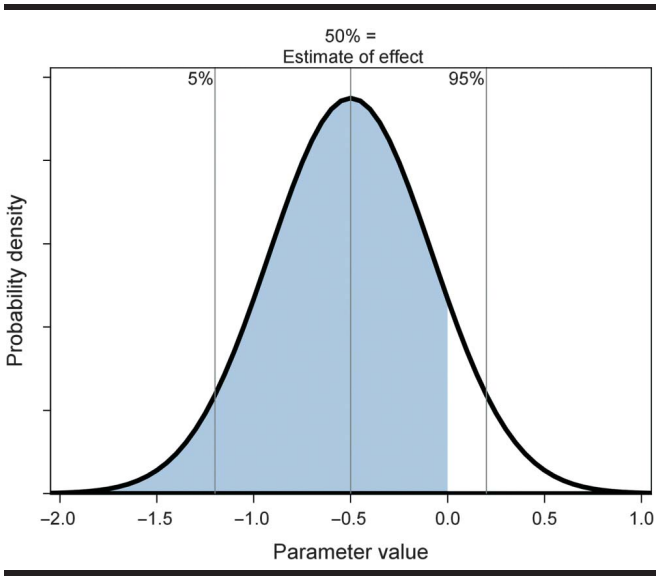
In spite of these safeguards, to ensure that participants did not receive excessive daily exposure to sound, any candidate who was exposed to loud noise as part of their job or recreational activity was excluded from participation. In addition, audiometric testing was performed at every visit to verify hearing thresholds had not changed from baseline.

Additional safeguards included otoscopic evaluations at every follow-up visit to inspect participants' ears for any sign of irritation from the custom-fit earbuds. If irritation was reported or detected, the custom-fit earbuds were remade to resolve the discomfort.

Statistical Considerations

A Bayesian approach to the data analysis was used in this study (Spiegelhalter, Abrams, & Myles, 2004). Briefly, a Bayesian analysis is distinguished from classical approaches by defining probability as a measure of belief in particular propositions. Figure 1 displays an example probability distribution expressing beliefs about a particular parameter (x-axis), which may correspond to, for example, the effects of treatment with TM compared to BSG. The 50% reference line denotes the estimated value of the parameter, and the 5th and 95th percentiles express 90% Bayesian confidence intervals. These percentiles define the range in which we are 90% certain that the true parameter value lies. This distribution contains all information at our disposal about the parameter value, including prior studies, substantive knowledge in particular clinical settings, and new data. New measurements are combined with this probability distribution via Bayes theorem to define a new "posterior" probability distribution of the parameter values. Because the posterior distribution describes everything that we know about a particular parameter, we can compute, for example, the probability that the parameter is less

Figure 1. Example probability distribution with 5% and 95% designating the 90% Bayesian confidence interval and the 50% value as the best "estimate" of the parameter value.



than zero by taking the area under the curve below zero, as indicated by the shaded region.

In many aspects of data analysis a Bayesian approach is indistinguishable from classical methods. However, Bayesian analysis forgoes esoteric constructs such as *p* values in favor of intuitive probability statements about treatment effect sizes and relative effects. For example, a classical analysis would test the one-sided null hypothesis that the true difference in treatment effects is greater than or equal to zero and reject this hypothesis with a *p* value of less than .05. With a Bayesian analysis, one simply computes the posterior probability that the true difference in treatment effects is less than zero, as shown in Figure 1.

The goal of this study was to evaluate the efficacy of each device for relieving tinnitus perceptions and/or tinnitus distress measured using the TFI score, NRS, and LM at 1 kHz. It is useful to consider these effects from clinical as well as research perspectives. On the one hand, it is of great benefit to know how, for example, an in-ear sound therapy device that is matched as closely as possible to an individual's tinnitus percept (i.e., the TM sound therapy device) affects perceived loudness of tinnitus after 3 months of nighttime use. Alternatively, a clinician would benefit more from an estimate of the expected improvement in reactions to tinnitus in a new patient seeking treatment for chronic tinnitus. We present results of this analysis for both research and clinical "consumers," in terms of the average effects of treatment with each sound therapy device on each outcome (a research focus), as well as the predicted benefits to a new patient of treatment with each particular sound therapy device on each outcome (a clinical focus).

We estimate the effects of treatment on each outcome using a multilevel model with sources of variation in the observed outcomes due to treatment group effects, time of measurement, outcome scale, participant effects, and interactions among these. Let y_i^k denote the kth outcome measurement on the ith observation, $k = 1, 2, 3$, indicating the NRS, LM 1 kHz (dB SL), and TFI score measurements, respectively. Raw data are given in Supplemental Material S1. Associated with each y_i^k is a subject index s and treatment group index $g = 1$ (BSG), 2 (NS), and 3 (TM), and a binary indicator v_i identifying whether the ith observation is from baseline ($v_i = 0$) or posttreatment ($v_i = 1$). The y_i^k have a normal likelihood so that

$$y_i^k \sim N\left(\delta_{s[i]}^k + \alpha_{g[i]}^k \cdot v_i, \sigma\right). \tag{1}$$

(All normal distributions are parameterized in terms of the mean and standard deviation.) In this particular notation, $s[i]$ denotes the sth subject on whom the ith observation was made, and $g[i]$ denotes the gth treatment group to which the ith observation is attached. The $\delta_{s[i]}^k$ are participant-specific random effects on the kth outcome measure centered at a participant-specific, across-measure mean ξ_s plus the overall population mean on the kth response, denoted μ^k, so that

$$\begin{gathered}\delta_s^k \sim N\left(\mu^k + \xi_s, \tau_\delta\right) \text{ and} \\ \xi_s \sim N\left(0, \tau_\xi\right).\end{gathered} \tag{2}$$

Note that, according to this model, the baseline mean of the kth response is equal to μ^k regardless of treatment group, which is reasonable because participants were not randomized to a treatment group until after the baseline measurement was taken. The average change to the kth response occurring after intervention are described by the a_g^k terms, which are also normal random variables, such that

$$a_g^k \sim N\left(\theta^k + \varphi_g, \tau_a\right). \tag{3}$$

The mean components are also normal:

$$\begin{gathered}\theta^k \sim N(\theta_0, \tau_\theta), \\ \varphi_g \sim N\left(0, \tau_\varphi\right), \text{ and} \\ \theta_0 \sim N(-0.1, 2).\end{gathered} \tag{4}$$

One can think of the θ^k as the "participation" effect on the kth response, whereas φ_g is the overall effect of the gth treatment group on all outcomes. The a_g^k are interaction effects between the kth outcome and gth treatment group on the mean response at follow-up visit. The difference in treatment efficacy between group g and group $\acute{g}$ on the kth outcome measure is given by $a_g^k - a_{\acute{g}}^k$.

All measurements were standardized to a mean of zero and standard deviation of 1, so that the μ^k are given a $N(0,\tau_\mu)$ prior. All of the variance components, τ and σ, are given half-normal prior with scale parameter = 0.5. These give a 90% prior interval on the contrasts between any two treatment groups of 0 ± 4 points on the NRS scale, 0 ± 18 dB LM SL, and 0 ± 37 points on the TFI

scale. We also fit the model using a half-normal priors with scale = 2 on the variance components with no appreciable effect on the results.

Posterior distributions of the parameters a_g^k and contrasts $a_g^k - a_{\tilde{g}}^k$ summarize our understanding of the effects of each sound therapy device, alone or in contrast, on tinnitus perceptions and reactions. We predict the benefits to a new patient by simulating new observations from the normal likelihood that generates the y_i^k in our model. From this effort, we can predict the probability of any improvement in a new patient under each treatment modality, as well as the magnitude of expected benefit.

Computation

The model was fit using SAS software Version 9.4. The model-fitting code is given in the Supplemental Material S1. The No-U-Turn Sampler was used for evaluating the joint posterior distribution. Three chains with random, dispersed starting values were run for 5,000 iterations following a 1,000-iteration burn-in period, and convergence was evaluated with Gelman–Rubin diagnostics. All parameters had diagnostics of 1.004 or less. Model fit was evaluated by plotting the y_i^k and posterior predictive interquartile range for each observation, and no gross deviations due to lack of fit were found.

Results

A total of 454 individuals were prescreened over the telephone. After learning more about the nature of the study, 64 decided to not continue the screening process. After administering the Tinnitus Screener and THS, 260 callers were deemed ineligible. The remaining 130 prequalified to attend the initial in-clinic screening visit. After 104 individuals attended the in-clinic screening visit, the study reached its enrollment goal of 60 participants. Figure 2 displays the number of individuals screened, consented, enrolled, and randomized into each group.

Randomization Error

Two participants were incorrectly randomized by study personnel and were therefore not included in the statistical analyses. This resulted in $n = 19$ in the TM group, $n = 19$ in the NS group, and $n = 20$ in the BSG group, for a total of $N = 58$.

Demographic Information

The majority of the participants were male ($n = 39$; female, $n = 19$) with a mean age of 59.7 years (range = 30–72). Participants' age as a function of group assignment and other baseline demographic information is displayed in Table 1.

Raw Data

Figure 3 shows the sample data collected in this study. Each panel corresponds to a treatment group (columns) and outcome measure (rows). The *x*-axis denotes the time point of measurement, and the *y*-axis denotes the possible score on the respective measurement scale. Black circles are individual participant data with lines connecting each participant's observed data at baseline and end of treatment. The thick red line is the sample mean with standard errors. Sample data show fairly consistent changes over time across treatment groups, although the TM group appears to improve slightly better on the NRS scale. There is little difference in changes on the LM scale across groups. Last, TFI score appears to change fairly consistently across groups.

Model-Based Results

Posterior distributions of the parameters estimating the mean change from baseline in each outcome for each treatment group are shown in Figure 4. Recall that in Bayesian analysis knowledge about the parameters (in this case change from baseline) is expressed as probability distributions. Each point in Figure 4 is the 50th percentile of the posterior probability distribution of the average change from baseline; the thick bars are the posterior interquartile range, and the thin lines are the posterior 90% intervals. The reference line at zero indicates no change from baseline; positive values indicate worsening of symptoms; negative values indicate improvement. Note that the change from baseline is expressed in standard deviations to facilitate comparisons across outcome measures. The posterior estimated change on the original scale is shown along the *x*-axis for each treatment group.

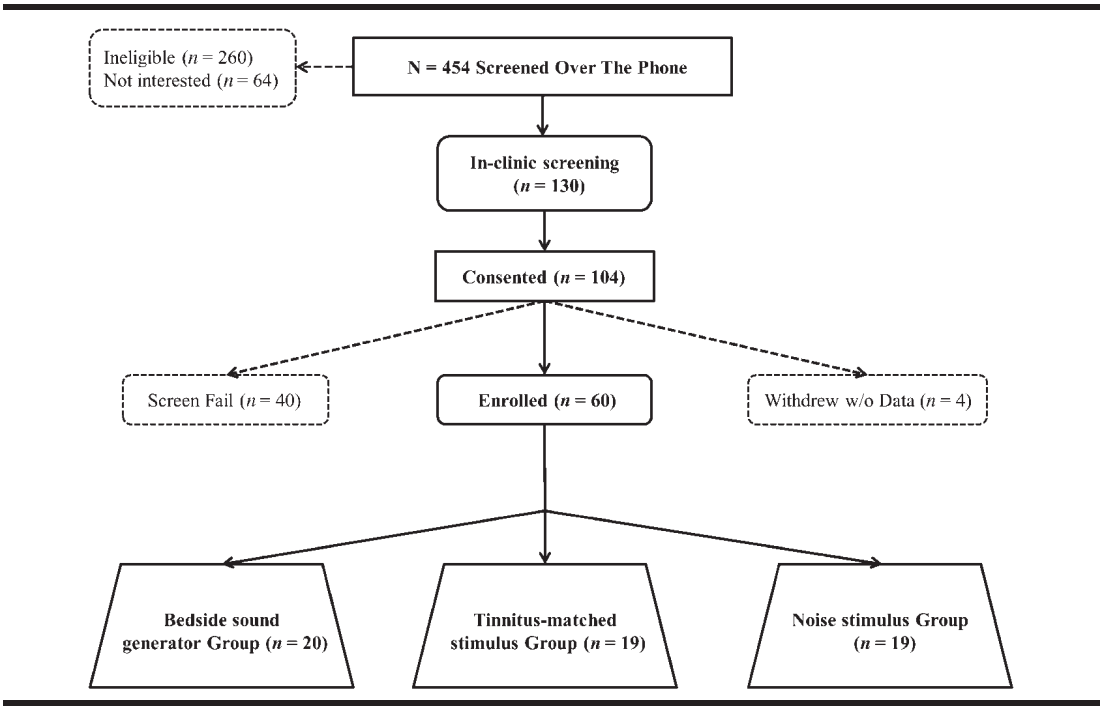
Figure 4 shows considerable variation in the change from baseline across outcome measures and comparatively less variability among treatment groups within each outcome measure. It is apparent that all groups showed some average improvement from baseline on all outcomes, though there is a nonnegligible probability (> 10%) that average NRS got worse for the BSG and NS groups. There is little variation among groups in improvement on the LM at 1 kHz (3.6–4 dB SL reduction). Average change in the TFI after 3 months of sound therapy during sleep reveals a reduction of 16 points in the BSG group, 20.6 points in the NS group, and 21.2 points in the TM group. There is little evidence of any difference between the TM and NS groups.

Table 2 shows 5th, 50th, and 95th percentiles of the posterior distribution of change from baseline contrasts between treatment groups on each outcome measure. Contrasts are expressed as mean differences between groups on the original measurement scale. The last column in Table 2 shows the posterior probability that the first group listed in the contrast column showed greater improvement than the second group listed.

Contrasts Among Treatment Groups on Tinnitus Perception

The NRS measure shows at least 95% certainty (95% for TM vs. BSG; 97% for TM vs. NS) that the TM sound therapy device generated greater reduction in average subjective ratings of tinnitus loudness. This effect was on the

Figure 2. Flowchart of the number of individuals screened, consented, enrolled, and randomized into each group.



order of about 0.7- or 0.8-point greater improvement compared to the other groups. As seen in the 50% column in Table 2, we estimate that the TM group showed about a 0.75-point greater reduction in perceived loudness on the NRS than the other two groups (90% confidence interval [−1.4, −0.0] for TM vs. BSG and [−1.7, −0.1] for TM vs. NS; these correspond to roughly [−1.5, −0.1] across both contrasts). The treatment groups were virtually indistinguishable in their effects on the LM at 1 kHz scale, with no contrast greater than 0.5 dB SL.

Table 1. Baseline demographics and tinnitus outcomes of the study sample.

Baseline features		Group BSG	NS	TM	All
N		20	19	19	58
Gender					
Male	*n*	14	12	13	39
	%	70	63	68	67
Female	*n*	6	7	6	19
	%	30	37	32	33
Age	*M*	62.3	59.6	57.0	59.7
	Min	54	35	30	30
	Max	72	72	70	72
NRS	*M*	6.6	5.8	6.4	6.3
	Min	2	2	3	2
	Max	10	10	9	10
LM 1 kHz	*M*	20.8	20.4	14.3	18.6
	Min	8	6	2	2
	Max	40	66	26	66
TFI	*M*	64.8	57.8	62.7	61.8
	Min	30	33	28	28
	Max	98	82	90	98

Note. BSG = bedside sound generator; NS = noise stimulus; TM = tinnitus-matched; NRS = numeric rating scale of tinnitus loudness; LM = tinnitus loudness match; TFI = Tinnitus Functional Index.

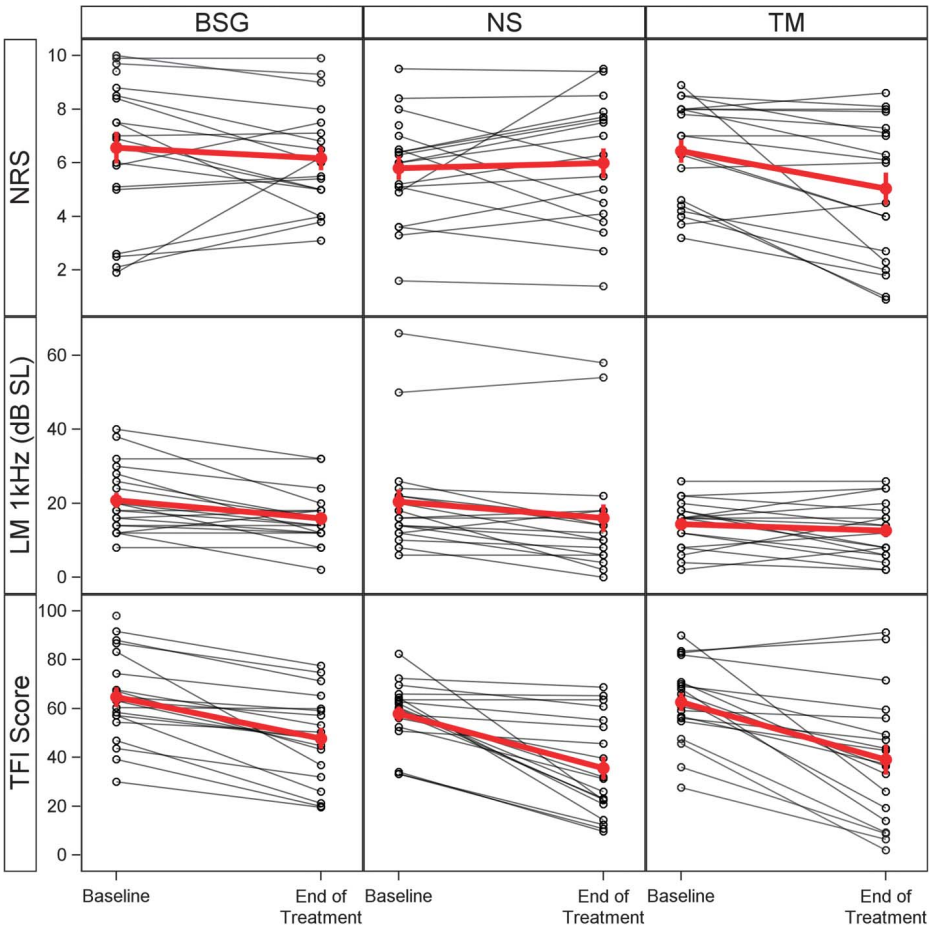
Contrasts Among Treatment Groups on Tinnitus Reactions

In terms of average change in tinnitus-related distress, the model reveals an 87% and 91% posterior probability that NS and TM groups, respectively, improved average TFI more so than the BSG group (Table 2). We estimate that the TM and NS groups reduced tinnitus reactions by, on average, 4.5–5 points more on the TFI as compared to the BSG group (90% confidence interval [−12.0, 0.9] for TM vs. BSG and [−11.9, 1.6] for NS vs. BSG; these correspond to roughly [−12, 1] across both contrasts). The TM and NS groups were essentially indistinguishable from each other in their effects on the TFI: With 56% posterior probability, the TM group was better than NS group by about 0.6 points.

Predicted Benefits From Each Sound Therapy Device

The Bayesian analysis also facilitates predictions about the efficacies of each treatment option for a new patient. These predictions are made under the assumption that a new patient is drawn from the same pool of people with tinnitus distress from whom the study participants were drawn. Table 3 shows the posterior predicted distribution of change from baseline under each treatment regimen for each outcome. Also shown are the posterior probabilities of any improvement from baseline. Note that predictions for an individual patient are intrinsically less precise than

Figure 3. Outcomes observed in this study. Each panel corresponds to a treatment group (columns) and outcome measure (rows). The *x*-axis denotes the time point of measurement, and the *y*-axis denotes the possible score on the respective measurement scale. BSG = bedside sound generator; NS = noise stimulus; TM = tinnitus-matched; NRS = numeric rating scale of tinnitus loudness; LM = tinnitus loudness match; TFI = Tinnitus Functional Index.



estimates of average treatment effects in a population of patients. Accordingly, the posterior probabilities of improvement are considerably lower than seen in Figure 4 and give a more tempered view of the expected effects of these sound therapy devices on tinnitus.

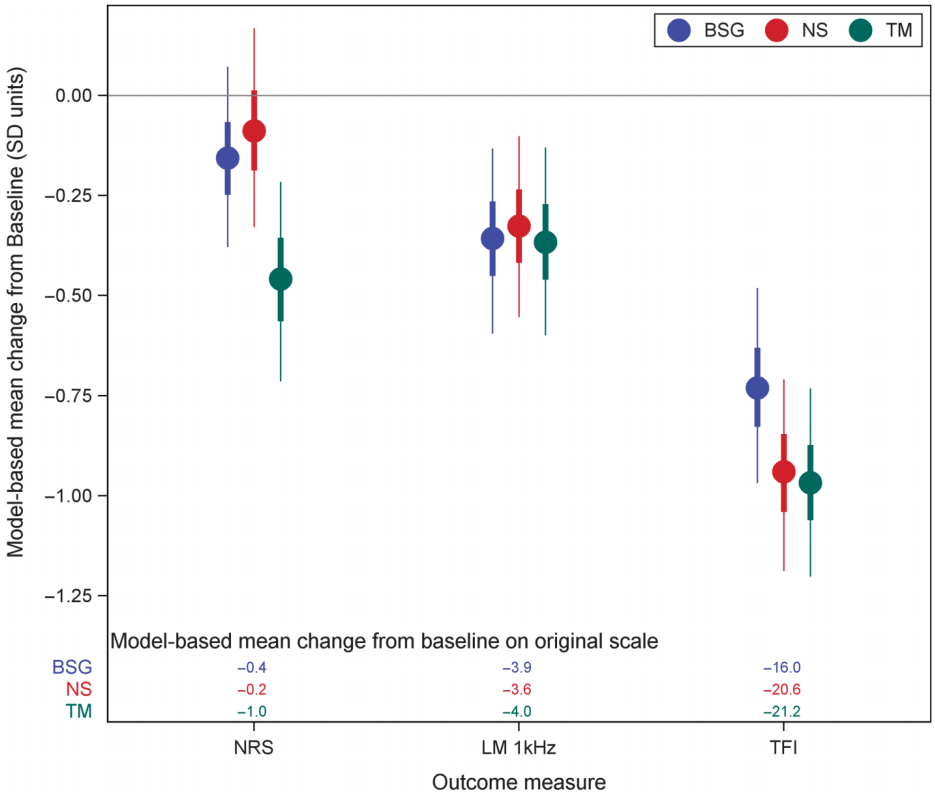
The predicted chance of improvement on subjective ratings of tinnitus loudness (i.e., NRS) is a little better than a coin toss in the BSG or NS groups and somewhat better for the TM group (74% chance of improvement). The chance of improvement on the 1-kHz LM is at least 67% (NS group) with little expected difference among groups. Finally, there is at least 84% certainty that any of the study devices will improve TFI outcomes, although there is little basis for preferring any particular sound therapy device. The last column in Table 3 shows the predicted probability of improving by at least 13 points on the TFI, which is considered enough of a reduction to be a meaningful improvement to an individual (Henry, Griest, Thielman, et al., 2016; Meikle et al., 2012). These last probabilities are considerably lower than that of any improvement, because they require greater changes for what an individual would consider a meaningful reduction in tinnitus distress.

Discussion

Using some form of sound therapy to manage tinnitus perceptions and reactions has been employed for centuries and is documented in writings dating back to the Roman Empire, the Middle Ages, and the Renaissance, to name a few (Heller, 2003). Current-day approaches to manage the effects of tinnitus still favor using some form of sound therapy, often via a medical device (for a review, see Hoare et al., 2014). Comparing and contrasting the numerous sound-based approaches that exist is difficult because of the lack of standardized methods to assess tinnitus perceptions and reactions and the limited number of RCTs performed evaluating the clinical effectiveness of sound therapy for tinnitus (Hobson, Chisholm, & El Refaie, 2010, 2012; Landgrebe et al., 2012).

The goal of this RCT was to evaluate the efficacy of using a specific in-ear sound therapy device during sleep

Figure 4. Posterior distributions of the mean change from baseline in each group on each outcome α_g^k. The *y*-axis shows estimates in standardized units to facilitate comparison across outcomes. The values along the *x*-axis show the posterior median of the average change from baseline on the original scale. Points are the posterior estimates of α_g^k thin bars or the posterior 90% interval, and thick bars are the posterior interquartile range. BSG = bedside sound generator; NS = noise stimulus; TM = tinnitus-matched; NRS = numeric rating scale of tinnitus loudness; LM = tinnitus loudness match; TFI = Tinnitus Functional Index.



for 3 months (i.e., TM group) for relieving tinnitus perceptions and/or distress compared to two control sound therapy devices (NS and BSG groups). The interpretations of the results are discussed in terms of what we know to be true based on (a) average effects of treatment with each sound therapy device on each outcome and (b) predicted benefits of treatment with each particular sound therapy device on each outcome.

Table 2. Model-based estimates of the contrasts in treatment effects between groups.

Outcome measure	Contrast	5%	50% (estimate)	95%	Probability
NRS	NS–BSG	−0.5	0.1	0.8	36%
	TM–BSG	−1.4	−0.7	−0.0	95%
	TM–NS	−1.7	−0.8	−0.1	97%
LM 1 kHz	NS–BSG	−2.7	0.3	3.6	43%
	TM–BSG	−3.3	−0.1	3.4	52%
	TM–NS	−3.6	−0.4	2.8	59%
TFI	NS–BSG	−11.9	−4.4	1.6	87%
	TM–BSG	−12.0	−5.1	0.9	91%
	TM–NS	−6.7	−0.6	6.0	56%

Note. Column headers "5%," "50% (estimate)," and "95%" correspond to the posterior 5th, 50th, and 95th percentiles of the posterior distribution of contrasts $a_g^k - a_{\bar{g}}^k$. Negative values indicate greater improvement in the first group listed in the contrast column. "Probability" denotes the posterior probability that the first listed group gives better improvement than the second listed group. NRS = numeric rating scale of tinnitus loudness; NS = noise stimulus; BSG = bedside sound generator; TM = tinnitus-matched; LM = tinnitus loudness match; TFI = Tinnitus Functional Index.

Table 3. Predicted improvements in a new patient by treatment option.

Outcome	Group	Predicted change from baseline			Probability of improvement	Probability of improvement > 13 points
		5%	50% (estimate)	95%		
NRS	BSG	–3.1	–0.4	2.4	58%	
	NS	–2.9	–0.2	2.6	54%	
	TM	–3.8	–1.1	1.6	74%	
LM 1 kHz	BSG	–17.1	–3.9	9.4	69%	
	NS	–16.7	–3.4	9.7	67%	
	TM	–16.9	–3.9	9.3	69%	
TFI	BSG	–42.5	–15.9	10.4	84%	57%
	NS	–46.7	–20.6	5.8	90%	68%
	TM	–47.5	–21.0	5.3	91%	69%

Note. NRS = numeric rating scale of tinnitus loudness; BSG = bedside sound generator; NS = noise stimulus; TM = tinnitus-matched; LM = tinnitus loudness match; TFI = Tinnitus Functional Index.

To evaluate efficacy, three outcome measures were used: (a) a questionnaire psychometrically validated to examine treatment responsiveness (TFI score), (b) a measure of self-rated tinnitus loudness (NRS), and (c) a psychoacoustic LM at 1 kHz. The test battery used in this RCT is consistent with the recommendations of Newman, Sandridge, and Jacobson (2014) by capturing aspects of the tinnitus percept as well as treatment-related change.

Average Effects of Treatment

Tinnitus reactions and perceptions improved across treatment groups. We found in our sample at least 85% posterior certainty that the TM and NS groups improved more than the BSG group on the TFI. The analysis also showed at least 95% posterior certainty that treatment with the in-ear TM sound therapy device results in greater improvement on the NRS than the other two groups. The treatment groups are virtually indistinguishable in their effects on the LM at 1 kHz, with no contrast greater than half a dB.

Predicted Benefit

It is important to remember that discussing expectations of what a new patient would experience is different from discussing the mean treatment effects of the sound therapy devices used in this study. Even though these are fundamentally different concepts, it is possible to discuss the results of this study in terms of what a new patient could expect when selecting a sound therapy device (for tinnitus) to be used during sleep, at least to the extent that a new patient responds in the same way as a participant drawn from this RCT.

Results on each of the outcome measures independently have value in assisting clinicians to help patients make an informed decision about selecting a sound therapy device. For patients most interested in learning how to manage their reactions to tinnitus, these "new patients" are very likely to experience reduced tinnitus distress with any of the sound therapy devices evaluated in this RCT.

For patients focused on the loudness of their tinnitus, our model predicts the best chances for the most improvement on self-rated tinnitus loudness, but not subjective LMs, are with the in-ear TM sound therapy device. The terms "improvement" or "reduced perception" used to describe the effects on tinnitus perception are not synonymous with elimination of the perception or a "cure" and should be interpreted only as some reduction in the subjective 0–10 rating of tinnitus loudness. It is not possible to know if this change in subjective rating of tinnitus loudness would be considered meaningful to a patient.

Unlike the TFI metric, there are no established estimates for how much change is needed on either the NRS or LM at 1 kHz to be considered meaningful improvement to an individual. Evaluating "treatment-related change" is not straightforward. Applying this concept to tinnitus is challenging because there is no widespread acceptance of what type of, or degree of, change in tinnitus perceptions would be considered meaningful improvement for a patient receiving treatment. It is possible for changes in tinnitus perceptions or reactions to occur, but they may not be equated with meaningful improvement in the eyes of the person seeking treatment. A variety of possible changes following tinnitus interventions can occur, and the difficulty lies in evaluating what type or degree of change other than suppressing the tinnitus (i.e., finding a cure) would constitute enough of a treatment-related change to be considered beneficial to a tinnitus patient (Theodoroff, Griest, & Folmer, 2017).

Limitations

Because of the nature of the sound therapy devices used in this study, blinding was not possible nor was it feasible to create a true placebo device. All participants knew up front that two groups would be receiving custom earbuds for devices to be used at ear level, whereas the third group was to use a sound therapy device that played sounds through a BSG. It is possible that participants who received custom earbuds to be used with the in-ear devices (TM and NS groups) were influenced by receiving

something "special," and the degree that might have contributed to the outcomes of this study is unknown.

The BSG and NS groups served as control groups, but not in the same manner as a placebo-controlled group; therefore, it is not possible to quantify how much of a placebo effect might also have contributed to these outcomes (Kaptchuk, 2002). When interpreting these findings, it is important to recognize that this study used self-reported outcomes. Just the act of participating in a clinical trial can have a positive effect on outcomes, regardless of the treatment received (for more information on placebo effects and controlling for expectations, see Potter, Mallinckrodt, & Detke, 2014).

It is not uncommon for individuals with tinnitus to report fluctuations in their tinnitus perceptions, and no established standard exists to assess the reliability of subjective ratings of tinnitus loudness. In an effort to create a standard, Henry, Fausti, Flick, Helt, and Ellingson (2000) provided guidelines to address reliability in tinnitus psychoacoustic LMs at various frequencies and also noted that there are two primary sources of variability in obtaining consistent tinnitus psychoacoustic measures over time: (a) variability in the tinnitus sensation itself and (b) measurement error that occurs due to changes in equipment, testing procedures, interpretation, and normal response variability that occurs regardless of whether the tinnitus percept is stable or not.

Finally, this study was not designed to look at sustainability of improvement. No measurements were performed at any time point after the "end of treatment" to determine how long any reported benefit continued. Future prospective studies involving the TM device should include assessments after 3 months to investigate how long possible benefit is sustained. Another limitation is the unknown generalizability of these results to tinnitus patients seen in the clinic. All participants in the current study were given their respective devices free of charge within the context of a research study. In a clinical setting, tinnitus patients may have different expectations than in a research study and would likely take into account the cost of different sound therapy device options and consider the cost-to-benefit ratio of each. The economic value compared to expected benefit of a therapy device is something that would commonly be discussed and factored into the decision-making process and is not something this study can address (Newman & Sandridge, 2012).

Conclusions

This study evaluated the efficacy of a unique tinnitus sound therapy device (TM) that is used only while sleeping. An RCT was conducted to compare outcomes between treatment with TM and two control treatments: (a) sham stimulus (NS) and (b) standard of care (BSG). The sham stimulus allowed for insights into how much of the effectiveness of the TM device might be attributed to selecting a stimulus that mimics the tinnitus percept. The standard-of-care group was intended to represent typical tinnitus sound therapy during sleep through use of a commonly used BSG.

Results from this study offer some support for greater average improvement in reactions to tinnitus with TM or NS devices compared to the BSG device. The TM group, compared to the BSG and NS groups, showed a greater reduction in ratings of tinnitus loudness on the NRS on average. The magnitude of these relative effects and the extent to which they generalize to other clinical environments and patient populations require additional study. Future controlled trials are needed to determine if these results are replicable and to evaluate additional variables and patient factors that would inform clinical practice.

Acknowledgment

This research was funded by the Baker Group, LLP, who manufactured the Levo System device (Portland VA Research Foundation 387004), and by the Department of Veterans Affairs, Rehabilitation Research & Development Service by a Program Award to the National Center for Rehabilitative Auditory Research (C9230C) and by a research career scientist award granted to Dr. James A. Henry (C9247S) at the VA Portland Health Care System in Portland, Oregon. These contents do not represent the views of the U.S. Department of Veterans Affairs, Department of Defense, or the U.S. Government.

References

Axelsson, A., & Barrenäs, M.-L. (1992). Tinnitus in noise-induced hearing loss. In A. L. Dancer, D. Henderson, R. J. Salvi, & R. P. Hamnernik (Eds.), *Noise-induced hearing loss* (pp. 269–276). St. Louis, MO: Mosby-Year Book.

Cima, R., Crombez, G., & Vlaeyen, W. S. (2011). Catastrophizing and fear of tinnitus predict quality of life in patients with chronic tinnitus. *Ear & Hearing*, *32,* 634–641.

Coles, R. R. A. (2000). Medicolegal issues. In R. S. Tyler (Ed.), *Tinnitus handbook* (pp. 399–417). San Diego, CA: Singular.

Crönlein, T., Langguth, B., Pregler, M., Kreuzer, P. M., Wetter, T. C., & Schecklmann, M. (2016). Insomnia in patients with chronic tinnitus: Cognitive and emotional distress as moderator variables. *Journal of Psychosomatic Research, 83,* 65–68.

Dobie, R. A. (2003). Depression and tinnitus. *Otolaryngologic Clinics of North America, 36,* 383–388.

Eggermont, J. J. (2003). Central tinnitus. *Auris Nasus Larynx, 30*(Suppl.), S7–S12.

Eggermont, J. J. (2015). Tinnitus and neural plasticity (Tonndorf lecture at XIth International Tinnitus Seminar, Berlin, 2014). *Hearing Research, 319,* 1–11.

Erlandsson, S. I., & Hallberg, L. R. M. (2000). Prediction of quality of life in patients with tinnitus, *British Journal of Audiology, 34,* 11–20.

Goodwin, P. E., & Johnson, R. M. (1980). The loudness of tinnitus. *Otolaryngologic Clinics of North America, 90,* 353–359.

Heller, A. J. (2003). Classification and epidemiology of tinnitus. *Otolaryngologic Clinics of North America, 36*(2), 239–248.

Henry, J. A., Fausti, S. A., Flick, C. L., Helt, W. J., & Ellingson, R. M. (2000). Computer-automated clinical technique for tinnitus quantification. *American Journal of Audiology, 9,* 36–49.

Henry, J. A., Griest, S., Austin, D., Helt, W., Gordon, J., Thielman, E., … Carlson, K. (2016). Tinnitus Screener: Results from first 100 participants in epidemiology study. *American Journal of Audiology, 25*(2), 153–160.

Henry, J. A., Griest, S., Thielman, E., McMillan, G., Kaelin, K., & Carlson, K. F. (2016, April). Tinnitus Functional Index: Development, validation, outcomes research, and clinical application. *Ear and Hearing, 334,* 58–64.

Henry, J. A., Griest, S., Zaugg, T. L., Thielman, E., Kaelin, C., Galvez, G., & Carlson, K. F. (2015). Tinnitus and Hearing Survey: A screening tool to differentiate bothersome tinnitus from hearing difficulties. *American Journal of Audiology, 24*(1), 66–77.

Hoare, D. J., Searchfield, G. D., El Refaie, A., & Henry, J. A. (2014). Sound therapy for tinnitus management: Practicable options. *Journal of the American Academy of Audiology, 25,* 62–75.

Hobson, J., Chisholm, E., & El Refaie, A. (2010). Sound therapy (masking) in the management of tinnitus in adults. *The Cochrane Database of Systematic Reviews, 12,* CD006371.

Hobson, J., Chisholm, E., & El Refaie, A. (2012). Sound therapy (masking) in the management of tinnitus in adults. *The Cochrane Database of Systematic Reviews, 11,* CD006371.

Hoffman, H. J., & Reed, G. W. (2004). Epidemiology of tinnitus. In J. B. Snow (Ed.), *Tinnitus: Theory and management* (pp. 16–41). Lewiston, NY: BC Decker.

Johnson, C. (2005). Measuring pain. Visual analog scale versus numeric pain scale: What is the difference? *Journal of Chiropractic Medicine, 4*(1), 43–44.

Kaptchuk, T. J. (2002). The placebo effect in alternative medicine: Can performance of a hearing ritual have clinical significance? *Annals of Internal Medicine, 136,* 817–825.

Katzman, R., Brown, T., Fuld, P., Peck, A., Schechter, R., & Schimmel, H. (1983). Validation of a short orientation-memory concentration test of cognitive impairment. *American Journal of Psychiatry, 140*(6), 734–739.

Keech, A., Gebski, V., & Pike, R. (Eds.). (2007). *Interpreting and reporting clinical trials. A guide to the CONSORT statement and the principles of randomised controlled trials* (pp. 1–181). Sydney, Australia: MJA Books.

Landgrebe, M., Azevedo, A., Baguley, D., Bauer, C., Cacace, A., Coelho, C., … Langguth, B. (2012). Methodological aspects of clinical trials in tinnitus: A proposal for an international standard. *Journal of Psychosomatic Research, 73,* 112–121.

McKenna, L. (2004). Models of tinnitus suffering and treatment compared and contrasted. *Audiological Medicine, 2,* 41–53.

Meikle, M. B., Creedon, T. A., & Griest, S. E. (2004). *Tinnitus archive* (2nd ed.). Retrieved from http://www.tinnitusarchive.org/

Meikle, M. B., Henry, J. A., Griest, S. E., Stewart, B. J., Abrams, H. B., McArdle, R., … Vernon, J. A. (2012). The Tinnitus Functional Index: Development of a new clinical measure for chronic, intrusive tinnitus. *Ear and Hearing, 33*(2), 153–176.

Newman, C. W., & Sandridge, S. A. (2012). A comparison of benefit and economic value between two sound therapy tinnitus management options. *Journal of the American Academy of Audiology, 23*(2), 126–138.

Newman, C. W., Sandridge, S. A., & Jacobson, G. P. (2014). Assessing outcomes of tinnitus intervention. *Journal of the American Academy of Audiology, 25,* 76–105.

Pedemonte, M., Testa, M., Diaz, M., & Suarex-Bagnasco, D. (2014). The impact of sound on electroencephalographic waves during sleep in patients suffering from tinnitus. *Sleep Science, 7*(3), 143–151.

Penner, M. J., & Bilger, R. C. (1995). Psychophysical observations and the origin of tinnitus. In J. A. Vernon & A. R. Møller (Eds.), *Mechanisms of tinnitus* (pp. 219–230). Needham Heights, MA: Allyn & Bacon.

Potter, W. Z., Mallinckrodt, C. H., & Detke, M. J. (2014). Controlling placebo response in drug development: Lessons learned from psychopharmacology. *Pharmaceutical Medicine, 28*(2), 53–65.

Ratnayake, S. A., Jayarajan, V., & Bartlett, J. (2009). Could an underlying hearing loss be a significant factor in the handicap caused by tinnitus? *Noise Health, 11*(44), 156–160.

Shore, S. E., Roberts, L. E., & Langguth, B. (2016, March). Maladaptive plasticity in tinnitus-triggers, mechanisms and treatment. *Nature Reviews Neurology, 12*(3), 150–160.

Spiegelhalter, D. J., Abrams, K. R., & Myles, J. P. (2004). Basic concepts from traditional statistical analysis (Chap. 2). In D. J. Spiegelhalter, K. R. Abrams, & J. P. Myles (Eds.), *Bayesian approaches to clinical trials and health-care evaluation* (pp. 9–48). West Sussex, England: Wiley.

Theodoroff, S. M., Griest, S. E., & Folmer, R. L. (2017). Transcranial magnetic stimulation for tinnitus: Using the Tinnitus Functional Index to predict benefit in a randomized controlled trial. *Trials, 18*(1), 64.

Zaugg, T. L., Schechter, M. A., Fausti, S. A., & Henry, J. A. (2002). Difficulties caused by patients' misconceptions that hearing problems are due to tinnitus. In R. Patuzzi (Eds), *Proceedings of the Seventh International Tinnitus Seminar* (pp. 226–228). Perth, Australia: Physiology Deptartment, University of Western Australia.