Capturing tinnitus

Audiological assessment and tailored sound-based intervention

Jose López Santacruz

Contents

1

Introduction

1.1. General Introduction

Tinnitus is often defined as the perception of a sound in the ear or in the head whose source is not outside of the body. We distinguish between two main types of tinnitus: subjective and objective. Subjective tinnitus is only perceived by the patient, and it is initiated in the auditory system and exhibited somewhere along the auditory pathway (Henry et al., 2020). Objective tinnitus involves a mechanical source in the body that results in an acoustic signal, registered in the cochlea. In this thesis, I will refer to the subjective tinnitus only.

Tinnitus affects around 15% of European adults, becoming increasingly apparent in the older population (Biswas et al., 2022). Hearing loss is the prevailing comorbidity associated with tinnitus, with a significant odds ratio of 8.5 observed in the Dutch population when it comes to the development of tinnitus (Schubert et al., 2021). In parallel to the increased prevalence of hearing loss with age, also tinnitus prevalence increases with age. Despite the exponential growth of tinnitus research over the past two decades, our understanding of its underlying mechanisms remains incomplete. Furthermore, finding an objective measure or biomarker for this condition remains a significant challenge.

The most common pathophysiological model of tinnitus suggests that auditory deprivation serves as a trigger, inducing alterations in the central auditory and extra-auditory networks that result in the tinnitus percept (Henton and Tzounopoulos, 2021). This model assumes that there is a certain degree of hearing impairment, and consequently, if the hearing function is restored, the tinnitus process can potentially be reverted, at least partially. This is the main reason why sound therapies (such as hearing aids) are among the most common tinnitus treatments in today's clinical practice. But the existing literature lacks evidence for proving their effectiveness in tinnitus treatment.

Current challenges in tinnitus research go beyond the lack of objective measures and effective treatments for tinnitus. Even the therapies that have proved to be effective (like hearing aids or counseling), have shown positive results in only some patients. Tinnitus is heterogeneous, so that it is manifested differently in each patient and there might exist different tinnitus subtypes (Beukes et al., 2021). All things considered, it is unlikely to find a universal solution for tinnitus. It is more probable that different therapies will achieve certain benefits for different patients. In the following section, I'll introduce the different assessment methods and treatments that were explored in this thesis, which will be described in more detail in the next chapters.

1.2. Tinnitus assessments and treatments

1.2.1. Auditory Brainstem Responses (ABRs): a biomarker for tinnitus?

Auditory Brainstem Responses (ABRs) are a well-known method for assessing hearing loss, especially in infants, who are not able to undergo regular tone

audiometry. This technique uses acoustical signals as stimuli and measures the synchronized neural response along the auditory pathway. Due to the low signal-to-noise ratio of the measurement, it's necessary to average hundreds or thousands of responses to obtain the so-called ABRs.

Among the 7 waves that comprise the ABR, only waves I, III and V are reliably and routinely detected in humans (Burkard et al., 2007). In clinical practice, the diagnosis usually consists in identifying these three waves and comparing their amplitudes and latencies to the normative values, in addition thresholds are often determined (Milloy et al., 2017).

As mentioned earlier, tinnitus probably induces alterations in the central auditory system. Therefore, ABRs have the potential to capture these changes and help us to better understand how tinnitus is generated. However, the existing literature on ABRs and tinnitus provides mixed results (Milloy et al., 2017). Tinnitus heterogeneity and the differences in equipment, stimulus type, presentation levels or recording filters used in these studies could explain the different outcomes. Nevertheless, the literature suggests that the differences in latency and amplitude of the ABR waves are only revealed using specific stimuli (Eggermont, 2019; Hofmeier et al., 2018). Finding the appropriate stimulus might be essential to verify whether ABRs are a suitable instrument in tinnitus research.

1.2.2. Minimum Masking Levels (MMLs): a signature of tinnitus heterogeneity?

Currently, sound-based therapies, in combination with education and counseling, seem to be the most effective approach for treating tinnitus (Hoare et al., 2014). Being exposed to certain acoustic stimuli such as masking sounds can potentially diminish or temporarily alleviate the symptoms of tinnitus in patients (Feldmann, 1971; Hazell and Wood, 1981). This is the case for hearing aids, which are able to enhance the environmental sound to mask the tinnitus percept. Patients tend to benefit the most from acoustic treatments such as hearing aids when their tinnitus is easily masked (Vernon, 1977). Therefore, assessing the feasibility of masking could potentially predict the effectiveness of a sound-based therapy in a particular patient. Moreover, further exploring psychoacoustic measures that assess masking could shed light on tinnitus diagnosis and, potentially, provide a better understanding on tinnitus subtyping for tailor-made treatments.

The minimum masking level (MML) is the minimum intensity of a masking sound needed to effectively mask the tinnitus. It is a method used for measuring tinnitus intrusiveness and the acceptance of masking (Andersson, 2003). There are only a few studies on this topic in the tinnitus literature, most of which tried to discriminate tinnitus subtypes (Feldmann, 1971; Mitchell, 1983; Tyler and Conrad-Armes, 1984b; Zagólski and Stręk, 2014). Although MML is commonly used as a psychoacoustical measure to screen tinnitus patients, there is a lack of evidence on the relationship between hearing thresholds and MML.

As stated above, there is a significant heterogeneity observed among tinnitus patients (Langguth et al., 2017; Van den Berge et al., 2017). Patients' tinnitus

differs in numerous aspects, such as perception (as in laterality of tinnitus, pitch or sound type), causal risk factors, distress and treatment responses (Cederroth et al., 2019). Cluster analyses on psychoacoustical data (such as hearing thresholds and MMLs) have the potential to identify groups that may not be immediately apparent through superficial examination of these variables. These analyses can be useful for tailoring a treatment to the members of a particular cluster. Cluster analyses can also serve as a tool for better understanding the relation between tinnitus and psychoacoustic measures, and potentially helping to personalize treatments.

1.2.3. The Tinnitus Functional Index (TFI): assessing the functional consequences of tinnitus

Tinnitus has a subjective nature, meaning that it is a highly individualized experience that currently cannot be objectively measured or quantified. For this reason, self-reported questionnaires are the most common assessment tool for measuring tinnitus severity in patients. Several questionnaires have been developed to evaluate the impact of tinnitus on patients' quality of life (Hall et al., 2016; Meikle et al., 2008). However, most of these questionnaires were not designed to measure the effectiveness of treatment outcomes (Kamalski et al., 2010). For this reason, Meikle et al. (2012) developed the Tinnitus Functional Index (TFI), a tinnitus questionnaire that is able to assess the impact of tinnitus and is also responsive to treatment effects. The TFI contains 25 questions organized in several subscales of factors, covering different tinnitus domains, such as intrusiveness, sense of control, sleep and quality of life.

To date, the TFI is a standard questionnaire in tinnitus practice, and it has been translated to more than 20 languages and used in more than 22 countries. In order to be used in a clinical setting in another country„ the questionnaire has to be translated and validated in a new clinical population. By achieving similar reliability measures as the original questionnaire, a translated version is considered validated and can be used as a clinical tool.

1.2.4. Tinnitus pitch matching: can we "measure" the tinnitus frequency?

The concept of tinnitus pitch refers to the perceived frequency or tonal quality of the sound experienced by a patient. In the context of sound-based therapies, there might exist a relationship between the tinnitus pitch and the effectiveness of a treatment. This has motivated the development of treatments like the tailor-made notch noise training (Stracke et al., 2010), the notch filter amplification (Marcrum et al., 2021) or the harmonic sound therapy (Mahboubi et al., 2012), among others.

These sound-based therapies are often fine-tuned to the pitch of tinnitus (Hoare et al., 2014), which makes the correct pitch characterization essential for the correct application of a treatment. This process is called tinnitus pitch matching, which consists in identifying the frequency of an external sound whose pitch coincides with the predominant pitch of a patients' tinnitus. Although pitch

matching is a standard audiological assessment in tinnitus practice, its reliability is questionable since it depends on the patient's feedback and it often presents large variability between sessions (Hoare et al., 2014). But the question whether the process is simply difficult for patients or tinnitus changes between sessions still remains.

Several pitch matching methods can be found in the literature, such as the two-alternative forced-choice method (Penner and Bilger, 1992) or the likeness rating (Norena et al., 2002). In contrast to these approaches that rely on the communication between the clinician and the patient, the method of adjustment (Tyler and Conrad-Armes, 1984a) allows the patient to control the process through an interface, which can potentially result in a more precise outcome. However, tinnitus patients are often of advanced age and these procedures can be too complex. For this reason, developing pitch matching techniques that are simple, reliable and self-guided can help to better match tinnitus and, potentially, improve the effectiveness of pitch-dependent sound-based therapies.

1.2.5. Tailor-made approaches: can hearing aids help with tinnitus?

Among sound-based therapies that take into account the tinnitus pitch, it is worth mentioning the Tailor-Made Notched Music Training (TMNMT). The goal of this treatment is to reduce the spontaneous activity in neurons by enhancing lateral inhibition from those frequencies below and above the frequency of tinnitus (Pantev et al., 2012; Pape et al., 2014). These authors suggested that tinnitus might originate after the loss of lateral inhibition in specific auditory neurons. Despite these studies being focused on using music as stimulus, the same approach has been used in hearing aids (Marcrum et al., 2021). In this case, the environmental sound, shaped with the notch filter of the devices, operates as the stimulus of the treatment.

However, as it has been previously mentioned, the main function for hearing aids with regard to tinnitus is to provide masking by means of enhancing the environmental sound. Therefore, it is conceivable that further amplification at the tinnitus frequency could contribute to better masking the tinnitus percept.

These techniques are opposite in their approaches in the way they address amplification: no amplification in opposition to extra amplification. Although both methods have plausible rationales, the existing literature highlights the lack of evidence to support the effectiveness of hearing aids for treating tinnitus (Hoare et al., 2014) and, in particular, the low quality of the existing studies in the form of non-randomized controlled trials (Shekhawat et al., 2013).

Taking into account the heterogeneity of tinnitus and the fact that it is unlikely to find a one-size-fits-all solution, well designed randomized controlled trials are necessary to prove which hearing aids strategies are the most effective in tinnitus, thereby finding the tinnitus characteristics that determine the responsiveness to a specific treatment.

1.3. Thesis outline

This thesis consists of three experimental chapters and two analyses on clinical data, and aims to build knowledge on different aspects of tinnitus from the perspective of clinical practice, including measurement, assessment of tinnitus impact and treatment. With the exception of chapter 2, which investigates the mechanisms of tinnitus through electrophysiology, this thesis collects a series of chapters that address different psychoacoustic aspects of tinnitus (such as masking and pitch) and how these can be considered to tailor hearing aids according to the particularities of each patient's tinnitus.

Chapter 2 contains an experimental study on Auditory Brainstem Responses (ABRs) with the intention of finding an objective measure that could be useful in the hearing aid trial. In particular, this chapter tries to answer the following question: Is it possible to detect tinnitus in the brainstem by optimizing the ABR stimulus? And: Can we distinguish tinnitus from hearing loss using ABR?

Chapter 3 consists in a cluster analysis on audiological data (focusing on the minimum masking level) collected from a clinical population. This analysis explores the relationship between tinnitus masking and hearing thresholds with the aim of finding tinnitus subtypes. The insights on masking will be beneficial for the hearing aid trial.

In **chapter 4**, the validation of the Dutch version of the Tinnitus Functional Index is presented. The goal of this study was to translate the well-known TFI questionnaire so that it could be used in a Dutch clinical population. The validation of this questionnaire was assessed through reliability measures to ensure its adequacy as an instrument for screening tinnitus impact, obtaining similar psychometric properties to the original version.

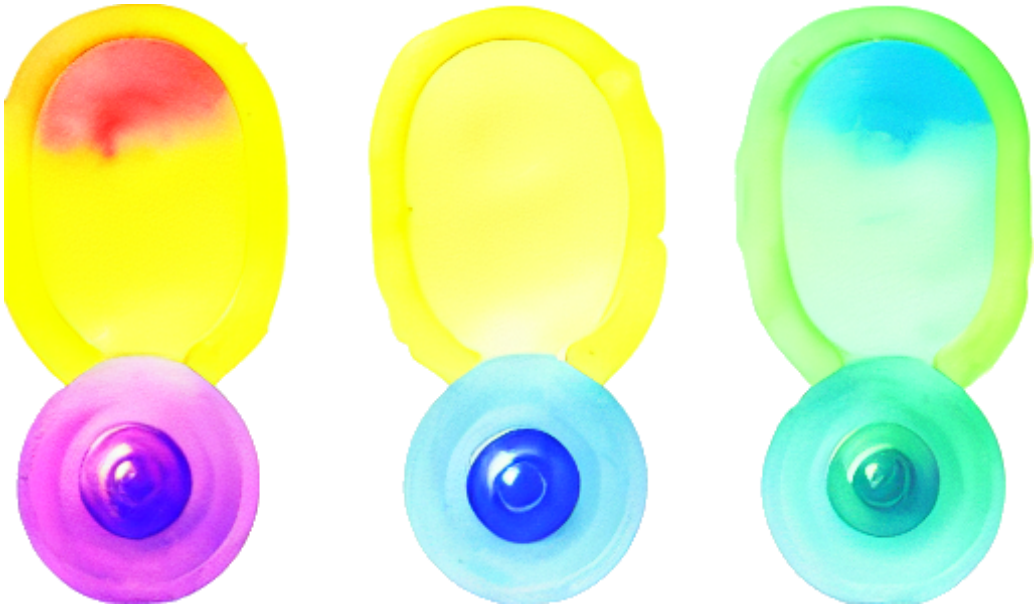
Chapter 5 is an experimental study that compares two different self-guided pitch matching methods for tinnitus. This chapter is focused on the research questions: 1) Can patients perform a tinnitus pitch matching task by themselves in a reliable and user-friendly way? 2) Do different methods give different results?

In **chapter 6**, a clinical trial was conducted to address the question on whether hearing aids can have an effect on tinnitus when they are tailored to the characteristics of each patient's tinnitus. Two settings with opposing approaches were tested in a clinical population by means of a double-blind and randomized-controlled trial.

Finally, an overall discussion is presented in **chapter 7**.

2

Investigating the stimulus parameters of Auditory Brainstem Responses in tinnitus



2.1. Introduction

Tinnitus, commonly known as phantom sound, is the perception of sound in the absence of an external acoustic source. The prevalence of tinnitus in European adults is about 15% on average, and this proportion increases with increasing age and worsening of hearing status (Biswas et al., 2022). Among the many possible classifications of tinnitus, one of them refers to its nature: objective, when the tinnitus source can be localized and heard or measured by the hearing specialist; and subjective, much more common, when the tinnitus is only perceived by the patient and the source is unknown (Baguley et al., 2013). To date, the assessment of subjective tinnitus is by means of self-report questionnaires (Hall et al., 2016). However, there is an increasing need for objective outcome measures, which can have important implications for clinical practice and the confirmation of symptoms in a medico-legal framework (Jackson et al., 2019).

Among different clinical tools, Auditory Brainstem Responses (ABRs) are a widely used method to assess hearing loss (specially in infants) and neurological disorders (Eggermont, 2019). By using acoustical signals as stimuli, this technique is able to represent the synchronized neural activation along the auditory pathway, which results in the so-called ABRs. The ABR consists of 7 waves that reflect the activation of different nuclei in the brainstem and thalamus , where waves I, III and V are reliably detected. (Burkard et al., 2007). Different nuclei contribute to a greater or lesser extent to the latency and amplitude of individual waves. There is a consensus that the main sources of the first five waves are: the distal and proximal portions of the auditory nerve for waves I and II, respectively, the cochlear superior olive for wave III, the lateral lemniscus for wave IV, and inferior colliculus for wave V (Han et al., 2021). Previous research suggested that the source of waves VI and VII might be the medial geniculate body (MGB), but this remains unclear (Akiyoshi et al., 2021). The way ABRs are used for diagnosis in clinical practice is usually by identifying waves I, III and V, and comparing their amplitudes and latencies to the normative values of a specific pathology, like hearing loss in newborns (Milloy et al., 2017).

The most common pathophysiological model of tinnitus suggests that tinnitus is triggered by auditory deprivation, which leads to changes in the auditory brainstem and the auditory cortex, such as increased spontaneous firing rates, increased neural synchrony and increased neural gain (Møller et al., 2010). Therefore, electrophysiological measures such as ABR have the potential to measure these maladjustments and provide a better understanding of tinnitus generation.

The tinnitus literature features many studies using ABR in humans to measure possible abnormalities associated with the pathology. A review by Milloy et al. (2017) analyzed 22 articles on the subject, finding several issues among the included studies that pointed out the difficulty of obtaining consistent measures in patients with tinnitus and drawing general conclusions. One of the highlighted problems to achieve this goal is the heterogeneity of the population's characteristics, such as presence or absence of hearing loss in controls, or unreported tinnitus etiology. Disentangling hearing loss and tinnitus is one of the current challenges in

tinnitus research due to the high comorbidity between both conditions (Ratnayake et al., 2009), which entails difficult interpretations when using the proposed diagnostic tool. Another issue that Milloy et al. discussed was the variety of techniques and assessments used among these studies. Differences in equipment, transducers, stimulus type, presentation levels or recording filters might underlie the heterogeneous outcomes, which also highlights the necessity for protocol standardization. Eggermont (2019) showed that changing the stimulus click rates affects the ABR signal. In particular, increasing the stimulus rate leads to longer latencies and the degradation of the entire averaged signal. This becomes a problem when trying to establish normative mean amplitudes and latencies of the ABR waves to assess tinnitus and/or hearing loss. Moreover, obtaining significant differences of amplitudes and latencies between test groups might only be attainable when using specific stimuli, as it has been shown by Hofmeier et al. (2018). In this study, the authors found differences in latencies and amplitudes of wave V between tinnitus subjects and hearing matched controls at certain stimulus' intensity levels only. These studies suggest that finding the correct stimulus might be crucial to determine whether ABRs are a suitable approach for tinnitus diagnostics and future research.

The purpose of this study was to address the problems presented above by measuring ABR with different stimuli in tinnitus patients and controls. Two hearing impaired groups were included, one with tinnitus and one without. These groups were homogeneous in the sense that all participants suffered from high-frequency sensorineural hearing loss. As a third group, we selected subjects with normal hearing, without tinnitus. For the three groups of participants we measured ABR, while systematically evaluating the effect of stimulus intensity and rate in the waves' latencies and amplitudes.

2.2. Methods

2.2.1. Participants

The study was conducted at the Ear, Nose and Throat (ENT) department of the University Medical Center Groningen (UMCG). A total of 41 participants were recruited. The first group consisted of 13 participants with normal hearing (NH group, thresholds ≤ 20 dB from 0.125 to 8 kHz). A second group of 13 participants was characterized by symmetric hearing loss and no reported tinnitus. For these participants (HL group), the average pure tone audiometry (PTA) at 2, 4 and 8 kHz was at least 30 dB in each ear. Symmetry in hearing loss was defined as a difference in threshold between both ears being no higher than 15 dB for each of those frequencies. A third group of 15 participants with chronic tinnitus and hearing loss was included (group TI+HL). These subjects had the same inclusion criteria for hearing loss than the HL group. Besides tinnitus and/or hearing loss, no other neurological or psychiatric disorders were reported for any of the subjects. All the participants in this study gave informed written consent before being enrolled. This study was approved by the ethics committee of the University

Medical Center Groningen (METc 2018/445).

2.2.2. Audiological assessment

All the participants underwent conventional audiometry during which hearing thresholds were measured at frequencies between 0.125 and 8 kHz. For this, a pair of TDH39 headphones (Telephonics) and an audiometer AC40 (Interacoustics) were used. Measurements were conducted in soundproof rooms. In addition to hearing thresholds, hyperacusis was assessed by means of the Hyperacusis Questionnaire (HQ; Khalfa et al., 2002). The group with tinnitus also filled in the Tinnitus Functional Index questionnaire (TFI; Meikle et al., 2012)

2.2.3. Auditory Brainstem Responses

ABRs were measured using the Interacoustics Eclipse and the 3A ABR EARtone insert earphones (3M) as transducers. Participants were comfortably sitting in a soundproof room during the measurements. Averaged ABRs were filtered using high and low pass filters of 0.1 and 2 kHz, respectively, with 100 μs click stimuli of alternating polarity and a contralateral masking of -40 dB relative to the stimulus ear. Stimuli were presented at 11.1, 29.1 and 47.1 clicks/s and at 50, 65 and 80 dB nHL, with a total of 9 recordings per ear and participant, each of them consisting of 2000 accepted responses. A fixed artifact rejection threshold of 40 μV was used.

Averaged ABR waves were exported and analyzed in Matlab, where peaks and troughs of waves I, III and V were marked following a manual procedure.

2.2.4. Analysis

Statistical analyses were carried on using R (version 4.0.2). In order to take into account all conditions of click rates and stimulus levels, between-group comparisons of waves' latencies and amplitudes were carried out by linear mixed models. During the fitting procedure, all predictors and interactions were tested on an iterative process of model comparison. For each model, fixed effects were tested through expected mean square approach, and random effects were set to the participants' identifiers. The proportion of explained variance was identified by conditional R2 (Nakagawa et al., 2017). For multiple comparisons, post-hoc Tukey tests were performed.

In addition to click rate and intensity, each model was tested using the following covariates: age, sex, hyperacusis, hearing and ear. Age and sex were reported by the participants, defining sex as a dichotomous variable. Hyperacusis was controlled by both including the HQ score as a numerical variable (with values between 0 and 42) and the positive/negative result of the questionnaire as a categorical variable (HQ score $\geq$ 22 for presence of hyperacusis, HQ score $<$ 22 for absence of it; Aazh and Moore, 2017; Koops et al., 2020). Hearing was characterized as the average PTA for 2, 4 and 8 kHz. Demographic data and questionnaires of the three groups were summarized and compared using the one-way ANOVA test.

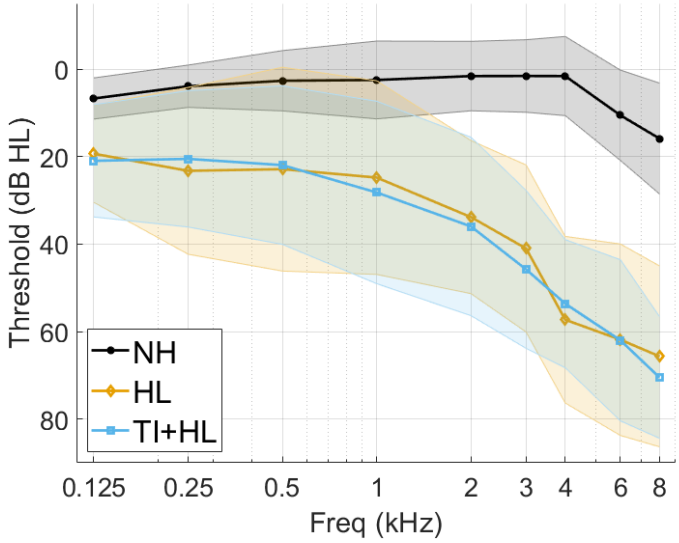


Figure 2.1 | Average pure tone audiograms of the three groups. Results are represented as mean values ± standard deviations as shaded contours. NH, normal hearing group (black line); HL, hearing loss group (yellow line); TI+HL, tinnitus and hearing loss group (blue line).

2.3. Results

2.3.1. Demographic characteristics

Demographic data, auditory information and the results of the questionnaire of the three groups are summarized in table 2.1. Between-group comparisons were made with ANOVA tests, in the case of continuous variables, and with Chi-Square test in case of categorical variables. Both HL and TI+HL groups were matched in age, sex and hearing profile, in contrast with the NH group, which consisted of younger subjects with no hearing loss in the measured frequencies. The average audiograms of the three groups are shown in figure 2.1. Regarding the hyperacusis questionnaire, all the groups showed significantly different scores, suggesting an increased sound sensitivity in the participants with tinnitus and hearing loss compared to the group with only hearing loss. For the participants with tinnitus, the results of the TFI indicated that tinnitus presented a moderate-to-big problem to them (Henry et al., 2016).

2.3.2. Waves' latencies and amplitudes

Latencies of the waves I, III and V in relation to intensities and click rates are shown in figure 2.2. Overall, latencies of the three waves decreased as the intensity of the stimulus increased. Waves I and III were too noisy to be registered

	NH	HL	TI+HL	P value	Adjusted $P_{HL - TI+HL}$
Number of subjects (n)	13	13	15		
Age (years)	36.4 ± 16.0	61.9 ± 10.9	61.2 ± 9.9	<0.001	1
Sex (n)				0.3	1
Male	6	7	10		
Female	7	6	5		
PTA (2, 4 and 8 kHz) (dB HL)					
Left ear	1.2 ± 7.7	38.7 ± 14.6	35.6 ± 11.6	<0.001	1
Right ear	4.0 ± 7.5	39.4 ± 14.8	35.4 ± 12.2	<0.001	1
HQ score	7.2 ± 3.3	13.8 ± 6.9	21.3 ± 8.1	<0.01	0.11
TFI score	–	–	51.6 ± 17.4		

Table 2.1 | Demographic characteristics of the participants. Mean values and standard deviation are presented. P values resulted from ANOVA for numerical variables and from the Chi-Square test for categorical variables. The adjusted P values resulted from a pairwise t-test with pooled standard deviation.

	numDF	denDF	F	p value
group	2	34.14	4.17	**<0.05**
intensity	1	411.05	31.39	**<0.001**
sex	1	34.31	4.81	**<0.05**

Table 2.2 | Fixed effects of wave I latency.

at intensities of 50 and 65 dB nHL. After checking for linearity, and in order to take into account the multiple conditions of intensity and click rates for the ABR comparisons, a linear mixed model per latency and wave was made. Significant differences in wave V latency at 11.1 and 29.1 clicks/s were found between the NH group and the two other groups, but not between the HL and the TI+HL groups. At the highest click rate tested, the difference is still present, but only between the groups NH and HL. s 2.2, 2.3 and 2.4 show the fixed effects of the best fitting models, obtained in a stepwise selection process, for the latencies of waves I, III and V, respectively.

For the above reasons, it was not possible to determine the amplitudes of waves I and III at 50 and 65 dB nHL. The averaged amplitudes of waves I, III and V at 80 dB nHL are shown in figure 2.3. For this intensity, the wave V amplitude of the NH group was significantly different to the two experimental groups for all the click rates tested, but no differences were found between the two of them. The fixed effects of the best fitting models of the amplitudes of waves I, III and V are shown in tables 2.5, 2.6 and 2.7, which varied among models. Both factors intensity and click rate contributed significantly to the variance of each linear model.

Multiple pairwise comparisons of all best fitting models for amplitudes and latencies are shown via post-hoc Tukey contrasts in table 5.

	numDF	denDF	F	p value
group	2	38.06	4.41	**<0.05**
intensity	1	431.37	55.88	**<0.001**

Table 2.3 | Fixed effects of wave III latency.

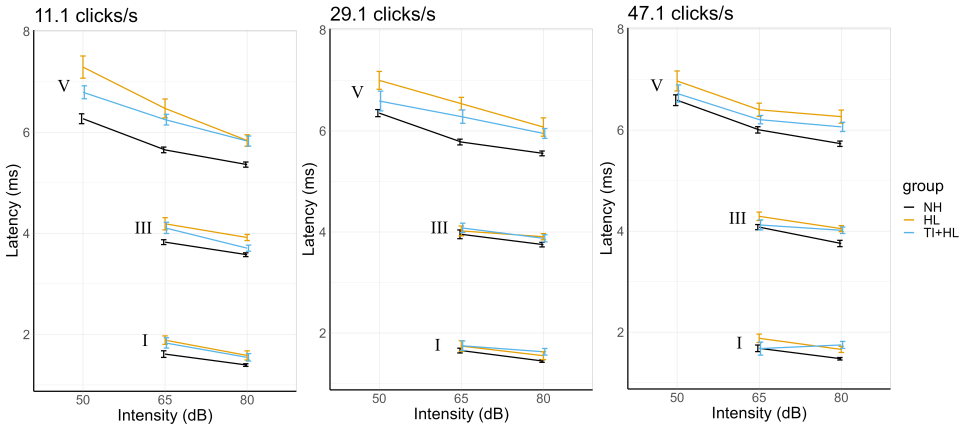


Figure 2.2 | Mean and standard error of waves I, III and V of the 3 groups, represented as a function of intensity. Each panel shows the obtained values for each tested click rate.

	numDF	denDF	F	p value
group	2	37.33	15.58	**<0.001**
click rate	2	637.45	3.40	**<0.05**
intensity	2	637.97	135.38	**<0.001**
ear	1	636.61	3.46	0.06
click rate * intensity	4	636.70	2.59	**<0.05**

Table 2.4 | Fixed effects of wave V latency.

	numDF	denDF	F	p value
group	2	37.99	13.30	**<0.001**
intensity	1	428.79	71.82	**<0.001**
click rate	1	429.11	8.24	**<0.001**

Table 2.5 | Fixed effects of wave I amplitude.

	numDF	denDF	F	p value
group	2	37.96	3.87	**<0.05**
intensity	1	427.81	86.90	**<0.001**
click rate	2	427.79	3.76	**<0.05**
ear	1	427.35	3.39	0.06

Table 2.6 | Fixed effects of wave III amplitude.

	numDF	denDF	F	p value
group	2	37.91	17.87	**<0.001**
intensity	2	642.14	146.92	**<0.001**
click rate	2	641.89	3.24	**<0.05**

Table 2.7 | Fixed effects of wave V amplitude.

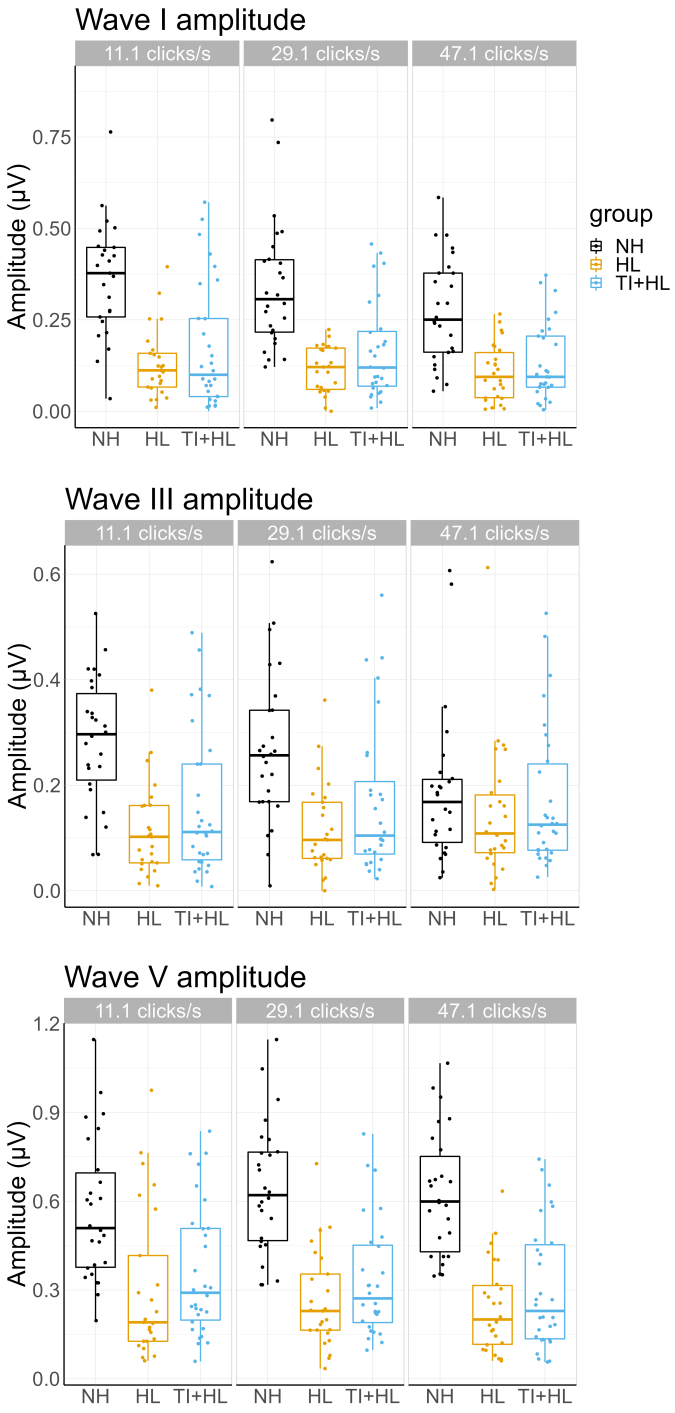


Figure 2.3 | Boxplots of the amplitudes of wave I (top), wave III (center) and wave V (bottom) at 80 dB nHL and the three click-rate conditions. Differences in amplitudes were found between the NH group and the hearing impaired groups. No significant difference was found between the wave amplitudes of the two hearing impaired groups groups.

2.4. Discussion

In this study, we measured the Auditory Brainstem Responses in three groups, consisting of patients with tinnitus and hearing loss, patients with hearing loss and no tinnitus, and healthy controls. We tested all combinations of three different click rates and three intensity levels, by recording, analyzing and comparing the resulting ABRs between the three groups. Between-group comparisons were made by including all the conditions in different linear mixed-effects models and measuring the latencies and amplitudes of waves I, III and V. No significant differences were found between the two hearing impaired groups. The hearing loss groups and the normal hearing group did show significant differences in wave latencies and amplitudes, where the largest differences were found between the control group and the group with hearing loss only.

Previous studies have shown that both intensity and click rate of the ABR's stimulus affect the amplitude and/or the latencies of the averaged responses (Eggermont, 2019; Hofmeier et al., 2018). Their findings, in addition to the heterogeneity of results reported by Milloy et al. (2017) provided the rationale for the present study. In this review —without addressing extensively the chosen stimuli of each of the papers reviewed— authors highlighted longer latencies for tinnitus compared to controls, with wave I being the most consistently affected one among the most well-controlled studies. In our study, despite not finding significant differences in latencies or amplitudes between the three groups, wave I did not provide any insight about potential trends in larger studies. As it has been shown previously in the literature (Turner et al., 2022), the low signal-to-noise ratio and the non-invasive nature of the ABR suggest that wave I has limited clinical utility for tinnitus diagnosis in humans.

In addition to hearing loss, hyperacusis is a well-known tinnitus comorbidity that has been previously assessed using ABR (Hofmeier et al., 2021; Refat et al., 2021). Both studies reported enhanced waves III and V amplitudes for high sound intensities for tinnitus with co-occurrence of hyperacusis. Hofmeier et al. (2021) theorized that tinnitus led to overall reduced and delayed responses as a consequence of a loss of fast auditory fibers, and the co-occurrence of hyperacusis could lead to a more widespread signal amplification through overactive thalamo-cortical activity. Refat et al. (2021) pointed out that the differences between tinnitus and the co-occurrence of hyperacusis might become more apparent the longer patients suffer from hyperacusis. In order to take these findings into account, we controlled for hyperacusis by means of the hyperacusis questionnaire (HQ, Khalfa et al., 2002) in our study. For this, we controlled using both the questionnaire scores as a numerical variable and the positive/negative result of the questionnaire as a categorical variable, using 22 as the threshold for hyperacusis/non-hyperacusis. Despite this double verification, and the fact that our 3 groups presented different averaged HQ scores (see table 2.1), hyperacusis resulted in a non-significant factor for either latency or amplitude in any of the best fitting models. However, we have no register of the onset of hyperacusis of the participants.

Recently, a larger study on ABR in tinnitus patients (Edvall et al., 2022) was conducted using two different devices (n=283 for Madsen EP200 Chartr and n=122 for Eclipse), using identical settings and conditions, and validated internally by 4 licensed audiologists. In their experiments, the chosen stimulus was 90 dB nHL at 9.1 clicks/s. The main finding of the authors was the difference in latency and amplitude of wave V between constant tinnitus, occasional tinnitus and controls. However, these findings differed between the ears tested and the devices used for measuring. Considering the high sample size of this study and the limited character of the results, it is feasible to consider that ABR might not be an effective assessment for tinnitus in clinical practice. In addition to this, the participants in this study showed better hearing thresholds in comparison to our both hearing impaired groups.

Although the differences in ABR waveforms between the normal hearing participants and the groups with hearing loss are notably different, no significant differences between the HL and the TI+HL groups were found in our study. Moreover, and regardless of significance, our HL group seems to present longer latencies of wave V on average, across conditions, compared to the tinnitus group, which differs from the findings in the above-mentioned studies. Previous studies have also reported no ABR differences in tinnitus when compared to matched hearing loss controls (Attias et al., 1993, 1996). It has been suggested that the effects found in some other studies of the literature might be the result of an accumulative effect of groups with different degrees of hearing loss and/or different size (Milloy et al., 2017). In our case, both groups are rather balanced in both size and hearing, so this possibility is not worth considering. Interestingly, a previous study on gray matter that included similar groups of participants showed more volume decline in the hearing loss group compared to the tinnitus group, in relation to the normal hearing controls (Koops et al., 2020).

It should be noted that most studies in the literature with positive results included participants with normal hearing thresholds (≤20 dB from 0.25 to 8 kHz) and the intensity of the stimulus used was, in most cases, higher than the ones we chose (above 90 dB nHL, where we maximally used 80 dB nHL), as it can be seen in the review by Milloy et al (2017). Only one study in this review included controls with hearing loss (Rosenhall and Axelsson, 1995), however, the authors did only find differences in latencies between the normal hearing groups. A meta-analysis (Jacxsens et al., 2022) reported longer wave latencies in tinnitus patients with normal hearing compared to controls. Nevertheless, the authors attributed these findings to non-reported high frequency sensorineural hearing loss or cochlear synaptopathy.

However, it is noteworthy that the statistical models used in similar studies often report significant contributions of different variables (such as hearing thresholds, age, sex, hyperacusis) but not how much of the total variance of the outcome is explained by these models, which might be far from a complete understanding of how tinnitus can be exhibited using ABR. Perhaps we need higher levels of abstraction to model the ABR waveforms instead of using linear relationships between such simple parameters. Possibly, further studies should focus on using

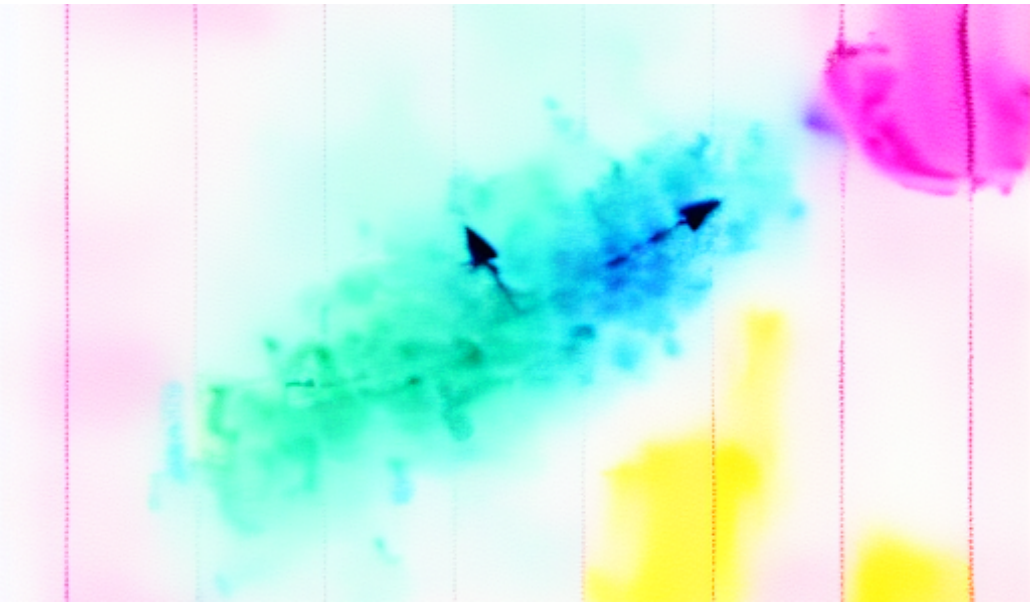
machine learning techniques with higher amounts of data. To our knowledge, there is at least a precedent in the literature for diagnosis of dysfunction in hearing using ABR and a bidirectional long short-term memory (BiLSTM) network (Chen et al., 2021), reporting remarkable accuracy levels. Despite the fact these artificial neural networks behave in a black box way, providing little to no understanding of how the inputs interact with each other (which would not contribute to shed light on the tinnitus mechanisms), the opportunity of obtaining a model that can diagnose tinnitus is worth exploring.

This study has some limitations. The low sample size makes it very difficult to distinguish between groups when using a measure with such a low signal-to-noise ratio. Before choosing the filters for the recordings, different cut-off frequencies were contemplated and tested. Finally, we chose high and low pass filters of 0.1 and 2 kHz, respectively, on the one hand because these parameters resulted in more distinguishable waveforms, but also in an attempt to replicate the settings of previous tinnitus studies and to advocate for standardization. However, prior to this study we conducted a pilot that showed the values of latencies and amplitudes slightly varied when using a high pass filter of 200 Hz, which is a value used before in several studies on tinnitus (Milloy et al., 2017).

Overall, we showed that distinguishing between hearing loss and normal hearing participants is an unambiguous task when using ABR. However, our study suggests that ABR might not have the potential to become a reliable diagnose tool for tinnitus. The heterogeneity of the individual responses and the indiscernible differences between the groups with hearing loss call for a change of direction in the context of tinnitus biomarkers.

3

Investigating the relation between minimum masking levels and hearing thresholds for tinnitus subtyping



Published as: Santacruz, J. L., De Kleine, E., & Van Dijk, P. (2021b). Investigating the relation between minimum masking levels and hearing thresholds for tinnitus subtyping. Progress in Brain Research, 263, 81-94.

Abstract

Heterogeneity of tinnitus imposes a challenge for its treatment. Identifying tinnitus subtypes might help to establish individualized diagnosis and therapies. The minimum masking level (MML) is a clinical tool defined as the minimum intensity of a masking sound required to cover tinnitus. Understanding the differences among masking patterns in patients could facilitate the task of subtyping tinnitus. Here, we studied the variability of hearing thresholds and MMLs among patients with tinnitus to identify tinnitus subgroups.

A population of 366 consecutive patients from a specialized tinnitus clinic were included in the analysis. Hearing thresholds and MMLs were determined for octave frequencies from 0.25 to 8 kHz, as well as for 3 and 6 kHz. Subjects were divided into two groups according to whether their tinnitus was maskable (M, 329 subjects) or non-maskable (NM, 37 subjects). Hearing thresholds and tinnitus loudness did not differ significantly between both groups. The dimensionality of the data was reduced by means of principal component analysis (PCA), and the largest resulting components were used for clustering the data. The cluster analysis resulted in five clusters with differences in tinnitus pitch, lateralization, hearing thresholds and MML, as well as on age and gender. Clusters differed in contours of hearing thresholds and MML, describing patterns of low or high thresholds in combination with low or high MML. The clustering solution presented a low silhouette value (0.45), implying that the clustering is weak and could be artificial. The analysis pointed out the diversity across tinnitus patients. Our results suggest that there might be a continuum of patients' characteristics rather than discrete subgroups.

3.1. Introduction

Tinnitus is a widespread phenomenon affecting up to 18% of the population in industrialized societies (Savage and Waddell, 2014). The etiology of tinnitus is frequently heterogeneous (Zagólski and Stręk, 2014) and its characteristics differs among patients. Possibly, several subtypes of tinnitus exist, each of them requiring a specific treatment. Subtyping different forms of tinnitus might be useful to find a homogeneous response to a treatment (Landgrebe et al., 2010). Several authors have suggested the importance of a personalized rehabilitation, attainable with an effective classification that is based on the physiological mechanism underlying the individual symptoms (Baguley et al., 2013; Ward et al., 2015).

Despite the literature about its underlying mechanisms has increased during the last years, the effectiveness of most of the existing treatment therapies has not followed the same trend. There is still a lack of high-quality evidence to support most of the available treatments for tinnitus. The limited available data makes it difficult to prove the benefit of sound therapies, but several studies using self-assessment tools like the Tinnitus Handicap Inventory (Newman et al., 1996) suggested that the effect might be better than a placebo when sound devices such as maskers, hearing aids or combination instruments are part of the therapy (Hobson et al., 2012). The exposure to acoustic stimulation like masking sounds can reduce or even temporarily suppress tinnitus (Feldmann, 1971; Hazell and Wood, 1981). Patients seem to best benefit from acoustic treatments when tinnitus is easily masked (Vernon, 1977), therefore psychoacoustical measures that assess more directly the tinnitus sensations such as masking might be helpful to improve the accuracy of the diagnosis and further treatment.

The minimum masking level (MML) is defined as the minimum intensity of a masking sound required to mask tinnitus, and it is a way to measure the intrusiveness of tinnitus and the acceptance of masking (Andersson, 2003; Vernon et al., 1990). There are only a few studies on MML, most of which tried to discriminate tinnitus subtypes. Feldmann (1971) described six types of masking pattern, defined according to the masking attainability and frequency contours of the maskers and hearing thresholds. Mitchell (1983) found similar classifications using Feldmann's congruence and convergence types of contours, characterized by thresholds running parallel to MML contours or crossing them, respectively. Unlike Feldman, no divergent pattern was reported in this study. Tyler and Conrad-Armes (1984b) obtained heterogeneous results in a similar study, suggesting the existence of more masking patterns than observed by Mitchell, and the lack of an apparent relation between the frequency contours of maskers and hearing thresholds. Zagólski and Stręk (2014) reported differences in mean MML values in relation to different tinnitus etiologies by using broadband noise. Particularly, the authors obtained lower MML values in acute acoustic trauma and congenital hearing loss cases, and higher MML values in patients after a stroke and with presbycusis. Etiology-dependent masking characteristics may be of value for patients whose etiology is doubtful. Despite the fact that MML is a widely used psychoacoustical technique in screening tinnitus patients, the relationship between hearing thresholds and

MML remains unclear.

In this paper, we address an analysis of patients' masking patterns from the tinnitus database of the University Medical Center Groningen (n = 366 patients). As a first step, we studied the relation between the hearing thresholds and the MMLs across frequencies. Next, we performed a principal component and cluster analysis to identify possible subgroups within the patients in the dataset.

The motivation to perform a cluster analysis stems from the large heterogeneity of tinnitus patients (Langguth et al., 2017; Tyler et al., 2008; Van den Berge et al., 2017). Patients differ with respect to many parameters, such as the degree of hearing loss, the underlying etiologies of the hearing loss itself, the factors that influence the tinnitus or their psychological response to the tinnitus. Often, these parameters are grouped in the literature according to several dimensions such as perception, causal risk factors, distress and treatment responses (Cederroth et al., 2019). The advantage of a cluster analysis is that it may recognize groups that are not evident upon superficial observation of the various variables. Clustering may serve the purpose of tailoring a treatment to the members of a particular cluster. Possibly, the members of a particular cluster are sensitive to a particular treatment. This treatment may be based on knowledge of the etiology of the tinnitus in that cluster, or on the symptoms of the tinnitus, such as the psychological response of a patient.

3.2. Methods

3.2.1. Participants

The study was conducted at the Otorhinolaryngology Department of the University Medical Center Groningen (UMCG). Since 2007, this clinic has been specialized for tinnitus patients through a multidisciplinary care group. After earlier consulting an audiologist and/or an otorhinolaryngologist, patients with bothersome and persistent tinnitus are referred to this tertiary care group for further consultation and psychological help. In the clinic, patients are evaluated by an audiologist, an ENT doctor, a medical social worker and a psychologist.

Data from 366 patients who visited the tinnitus clinic at the UMCG throughout a period of 3 years were collected. The dataset contains multiple variables like demographic information, audiometric diagnostics, results from quality of life and tinnitus questionnaires, psychoacoustical data and therapies followed by the patients.

3.2.2. Audiometry and MML

Hearing thresholds were obtained by means of pure tone audiometry for all patients, measuring octave frequencies from 0.25 to 8 kHz, as well as 3 and 6 kHz, using TDH39 headphones with an Interacoustics Affinity or Equinox audiometer. The MMLs were obtained using the same equipment. A narrowband noise (1/3-octave), with the same center frequencies as the tone audiometry, was used

as a stimulus. A monolateral up procedure in each ear was followed, presenting maskers for each frequency at an initial level of 15 dB below the hearing thresholds. The sound level was increased in steps of 5 dB until the patient could no longer perceive his/her tinnitus. The level at which the tinnitus was no longer audible, was recorded as MML. When a level of 90 dB HL was reached, the process was stopped regardless of whether the tinnitus was still audible or not.

Tinnitus pitch and loudness were obtained by means of the regular clinical protocol in which pure tones, narrowband noise or wideband noise are presented monaurally to the contralateral ear (in case of lateralized tinnitus), or to the better hearing ear (in case of bilateral tinnitus). The patient provides feedback to the clinician, who adapts the frequency and loudness of the stimulus until it matches the tinnitus percept best. Patients reporting similar tinnitus loudness in both ears or localizing their tinnitus "in the head", were considered as bilateral cases in this study. Those who described their tinnitus as partially or totally predominant in one ear were considered as lateralized cases.

3.2.3. PCA and cluster analysis

In order to reduce the dimensionality of the data, a principal component analysis (PCA) was carried out. Each subject contributed 8 audiometric thresholds and 8 MMLs, for each ear. Thus 2 ears * (8 thresholds + 8 MMLs) = 32 values entered the PCA. The components obtained from the PCA that described most of the data's covariance were included in the subsequent cluster analysis.

Since only continuous variables were included in the analysis, the selected clustering method was k-means. As this method requires fixing the number of clusters in advance, the choice was made by the elbow criterion algorithm to minimize bias. A silhouette coefficient was obtained as a measure of cohesion and separation between clusters.

3.3. Results

3.3.1. Hearing thresholds and MML

Masking was not possible in 37 patients. The tinnitus of a patient was considered non-maskable (NM) when the MML values obtained were 90 dB for a minimum of six frequencies in each ear out of the eight tested (from 250 to 8 kHz), and this was true for both ears. Table 3.1 shows the comparison between the NM and the M group, this latter group consisting of 329 subjects.

Boxplots of the hearing thresholds and MMLs of both the NM and the M group are shown in figure 3.1. The mean hearing threshold per subject was calculated as the pure tone average in both ears (PTA) of frequencies 2, 4 and 8 kHz, and they did not significantly differ between the two groups (see table 3.1). Most of the patients presented no hearing loss larger than 25 dB for frequencies up to 2 kHz. For higher frequencies, this gradually shifted toward mild and moderate hearing loss. MML values were, on average, constant across frequencies (for the

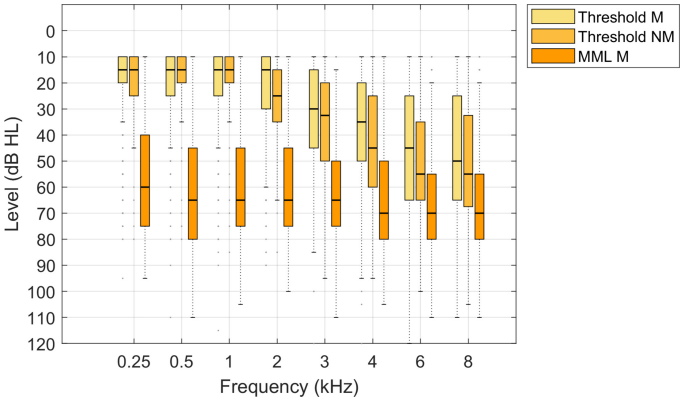


Figure 3.1 | Hearing thresholds and MML across frequencies for the M group (n = 329) and the NM group (n = 37). For each subject, data from both ears were averaged. Median values are represented with lines crossing the boxplots.

Empty Cell	**Cluster 1**	***Cluster 2***	***Cluster 3***	**Cluster 4**	**Cluster 5**	***p-Value***
Cluster size (n)	68	92	92	58	19	
Silhouette coefficient	0.55	0.54	0.39	0.37	0.26	
Age	56.0 ± 13.2	55.5 ± 14.1	64.5 ± 9.5	66.0 ± 9.2	63.0 ± 11.8	**<0.001**
Gender (% M/F)	43/57	63/37	73/27	67/33	68/32	**0.002**
Tinnitus pitch (kHz)	3.2 ± 2.6	4.1 ± 2.5	4.7 ± 2.3	3.2 ± 2.3	0.9 ± 2.1	**<0.001**
Tinnitus loudness (dB SL)	11.2 ± 13.6	*8.8 ± 10.1*	*6.3 ± 12.8*	6.3 ± 15.1	6.7 ± 10.3	0.315
Tinnitus lateralized (% yes/no)	65/35	63/37	59/41	70/30	89/11	**<0.001**
Hearing threshold (dB HL)	20.8 ± 11.2	23.3 ± 8	47.0 ± 11.9	48.7 ± 10.2	52.5 ± 20.4	**<0.001**
MML (dB HL)	30.9 ± 9.4	65.6 ± 10.1	80.9 ± 5.9	54.6 ± 8.7	68.1 ± 9.6	**<0.001**
Asymmetry of HL (dB HL)	3.3 ± 6.5	5.0 ± 8.4	6.6 ± 15.9	6.6 ± 10.2	31.6 ± 18.7	**<0.00**

One-way ANOVA

Table 3.1 | Characteristics of masking (M) and non-masking (NM) groups: median values with standard deviations. Values of tinnitus loudness and pitch represent the matching stimulus. The average of hearing thresholds was obtained by pure tone average (PTA) of frequencies 2, 4 and 8 kHz, for both ears. MML values are averaged across all frequencies in both ears.

M group), reaching values closer to the hearing thresholds as frequency increases. No significant differences between the two groups in tinnitus loudness nor pitch were found.

Lateralized cases had a prevalence of 65% in our database, and their characteristics are shown in table 3.2. These subjects presented higher thresholds in the ipsilateral ear for both M and NM groups. MML means did not differ significantly between the ipsilateral and contralateral ear within the M group, although the hearing was significantly worse in the ipsilateral ear.

3.3.2. PCA and cluster analysis

Using the hearing thresholds and MMLs of the 329 subjects in the M group, a principal component analysis was conducted to obtain the main vectors, the so-called higher eigenvalues that describe most of the variance of the data. In order

	NM (n = 21)		*p-Value*	*M (n = 216)*		*p-Value*
	Ipsilateral	Contralateral		Ipsilateral	Contralateral	
Hearing threshold (dB HL)	41.6 ± 23.8	28.3 ± 17.8	**0.049**	40 ± 20.7	26.6 ± 16.8	**<0.001**
MML (dB HL)	–	–		63.4 ± 18.8	66.8 ± 21.2	0.573

One-way ANOVA

Table 3.2 | Characteristics of patients with lateralized tinnitus for both the masking (M) and non-masking (NM) groups: median values with standard deviations. The average of hearing thresholds was obtained by pure tone average (PTA) of frequencies 2, 4 and 8 kHz. MML values are averaged across all frequencies in both ears. Bold p-values indicate a significance level < 0.05.

to assess the feasibility of the PCA, a Kaiser-Meyer-Olkin test and a Barlett's test were performed. The first one resulted in a sampling adequacy of 0.912, and the second test rejected the sphericity assumption ($p < 0.001$). Both results suggested that a component analysis might be appropriate (Bartlett, 1937; Kaiser, 1974). From all the eigenvalues obtained, the first four components were chosen based on the criteria of eigenvalues > 1, which are considered stable (Girden and Kabacoff, 2010). These four components explained the 83.8% of the variance of the entire dataset.

Based on the four components obtained, the elbow criterion was used to determine the optimal number of clusters (Syakur et al., 2018), which suggested a solution of five clusters. The chosen clustering method was k-means, and the average silhouette coefficient (SC) obtained was 0.458, interpreted by the literature as "weak structure and could be artificial" ($0.26 \leq SC \leq 0.50$) (Kaufman and Rousseeuw, 2009).

Boxplots of the hearing thresholds and the MMLs of the five clusters are shown in figure 3.2. Clusters were also compared according to variables that were not included in the PCA. Table 3.3 shows characteristics of the clusters based on significant differences between these variables.

Cluster 1 (figure 3.2A) includes a majority of middle-aged female patients with values of tinnitus pitch lower than the median. Patients in this cluster present a mild hearing loss and lower MML values indicating an easily maskable tinnitus.

Cluster 2 (figure 3.2B) is a large group of middle-aged patients with a mild hearing loss. MML values are approximately 70 dB across all frequencies, much higher than the median tinnitus loudness.

Cluster 3 (figure 3.2C) is another large group which includes a majority of older male patients with moderate hearing loss. Tinnitus pitch presents high values, and MML values are in the limit of the non-maskable condition.

Cluster 4 (figure 3.2D) represents another group of older patients with moderate hearing loss and MML median values in the range from 40 to 65 dB, close to the hearing thresholds.

Cluster 5 (figure 3.2E) mainly consists of patients with a large asymmetry in their hearing loss. This is the smallest cluster which also has the lowest silhouette value. Tinnitus presents relatively low frequency values.

All clusters presented similar tinnitus loudness values when these are expressed in dB SL, in line with previous studies reporting that most patients experience

	Cluster 1	*Cluster 2*	*Cluster 3*	Cluster 4	Cluster 5	*p-Value*
Cluster size (n)	68	92	92	58	19	
Silhouette coefficient	0.55	0.54	0.39	0.37	0.26	
Age	56.0 ± 13.2	55.5 ± 14.1	64.5 ± 9.5	66.0 ± 9.2	63.0 ± 11.8	**<0.001**
Gender (% M/F)	43/57	63/37	73/27	67/33	68/32	**0.002**
Tinnitus pitch (kHz)	3.2 ± 2.6	4.1 ± 2.5	4.7 ± 2.3	3.2 ± 2.3	0.9 ± 2.1	**<0.001**
Tinnitus loudness (dB SL)	11.2 ± 13.6	*8.8 ± 10.1*	*6.3 ± 12.8*	6.3 ± 15.1	6.7 ± 10.3	0.315
Tinnitus lateralized (% yes/no)	65/35	63/37	59/41	70/30	89/11	**<0.001**
Hearing threshold (dB HL)	20.8 ± 11.2	23.3 ± 8	47.0 ± 11.9	48.7 ± 10.2	52.5 ± 20.4	**<0.001**
MML (dB HL)	30.9 ± 9.4	65.6 ± 10.1	80.9 ± 5.9	54.6 ± 8.7	68.1 ± 9.6	**<0.001**
Asymmetry of HL (dB HL)	3.3 ± 6.5	5.0 ± 8.4	6.6 ± 15.9	6.6 ± 10.2	31.6 ± 18.7	**<0.00**

One-way ANOVA

Table 3.3 | Characteristics of the five clusters obtained: median values with standard deviations. Clusters are sorted by descending silhouette value. Differences between groups were calculated by using one-way ANOVA. The average hearing threshold was obtained by pure tone average (PTA) of frequencies 2, 4 and 8 kHz, for both ears. Asymmetry of HL was obtained by the absolute difference between the PTA of the same frequencies in both ears. Bold p-values indicate a significance level < 0.05.

tinnitus at a level of within 5–10 dB SL of their thresholds (Goodwin and Johnson, 1980).

3.4. Discussion

To our knowledge, there are only a few studies relating MML contours and tinnitus (Feldmann, 1971; Mitchell, 1983; Tyler and Conrad-Armes, 1984b; Zagólski and Stręk, 2014). The aim of this study was to identify specific subtypes of tinnitus through the analysis of MML contours and hearing thresholds. First, our patients' database was divided into a masking and a non-masking group, resulting in a proportion of 90/10 with no significant differences in hearing thresholds, tinnitus loudness or pitch between both groups. The majority of subjects in both groups presented lateralized tinnitus with poorer hearing thresholds in the ipsilateral ear. The diversity in findings between subjects motivated a cluster analysis for subtyping. The analysis resulted in five clusters underlining differences between the groups, but the separation and cohesion of the clusters was weak.

Minimum masking levels may seem counterintuitive in the psychoacoustical context. In principle, it can be expected that louder tinnitus would require higher levels of masking to relieve patients from its perception. The neurophysiological model of tinnitus (Jastreboff et al., 1994) suggests that the tinnitus percept might be originated in the auditory system at a subcortical level. According to this model, the level of intrusiveness can decrease as a result of habituation, therefore the neuronal networks are detuned from recognizing the "tinnitus signal" and lower levels of masking should be required to suppress it. In line with this model, it can be derived that MML can provide very useful clinical information when measured at different times, since it could be an indicator of changes in the auditory system during the course of a treatment. A decrease in MML would suggest that the tinnitus percept is more difficult to detect than before (Jastreboff et al., 1994).

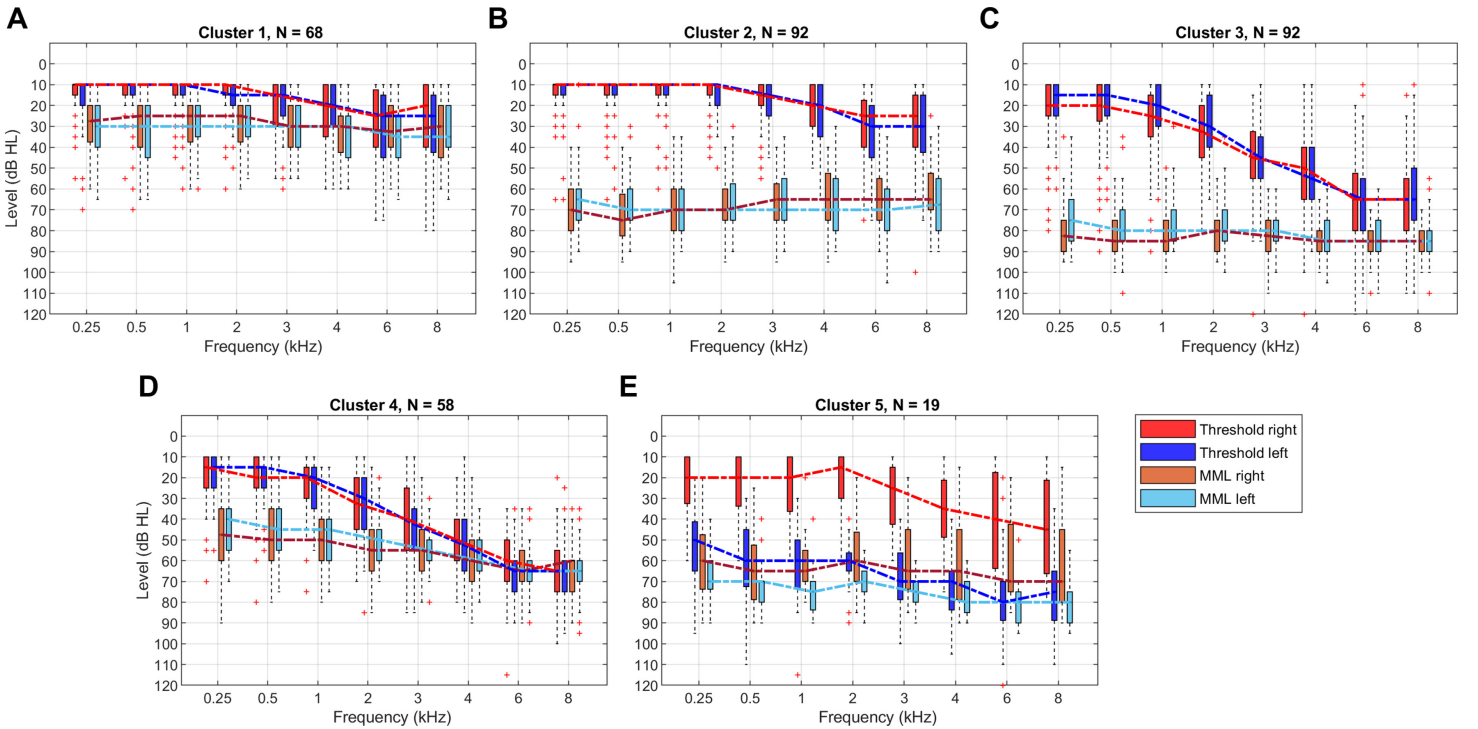


Figure 3.2 | Boxplots of hearing thresholds and MMLs for all the clusters found. (A, B) Clusters 1 and 2 represent subjects with relative good hearing thresholds. However, cluster 1 includes patients for whom tinnitus is easily maskable across frequencies, while patients in cluster 2 present higher MMLs. (C, D) Both clusters 3 and 4 include patients with moderate hearing loss, but differ in MMLs. Cluster 3 presents higher MMLs while cluster 4 represents the classical convergency type (Feldmann, 1971). (E) Cluster 5 represents patients with asymmetrical HL and high MMLs.

In our study, the non-masking group represented 10.1% of the population, which is consistent with the reports from the literature (Feldmann, 1971). Interestingly, these subjects presented similar thresholds, tinnitus loudness and pitch to those in the masking group. Previous studies reported no correlation between neither the loudness nor the pitch of the tinnitus percept and its maskability (Mitchell, 1983; Mitchell et al., 1993). A recent study reported lower intensity levels of maskers for frequencies close to the tinnitus pitch, but only for normal hearing subjects (Fournier et al., 2018). In our data, no clear relation between pitch, loudness and maskability was found. Moreover, masking contours showed a large variability for all degrees of hearing loss. Considering maskability as the attainability of covering the tinnitus sensation, our study suggests that it does not rely on hearing loss or on tinnitus loudness. However, in the cases in which tinnitus is maskable, the MML contours were relatively flat and the thresholds present different slopes, resulting in a decrease of masking levels relative to the thresholds with increasing frequency.

Previous studies using MML to discriminate between groups showed different masking patterns determined by the relation of hearing thresholds and MML contours (Feldmann, 1971; Mitchell, 1983). Despite the fact that similar patterns could be found in our database, we considered that this classification method is not systematic and is open to interpretation. Thus, we opted for an unsupervised method, such as clustering, whereby more flat frequency patterns were obtained. It seems that the key distinction between our clusters lies in the difference between hearing thresholds and MML, resulting in two types of masking pattern: easy and difficult to mask; and two types of hearing loss profile: mild and moderate hearing loss. These findings are consistent with the study from Tyler and Conrad-Armes (1984b), which suggested a lack of relation between the frequency contours of maskers and hearing thresholds.

One of the clusters found, was distinctively represented by female patients (cluster 1). This group was mainly characterized by mild hearing loss and low masking levels. Unlike Seydel et al. (2013), we did not find significantly lower tinnitus loudness among female subjects, but rather a good acceptance of masking. In addition to gender differences in tinnitus patients, one study also reported lower tinnitus pitch among female subjects (Oakes et al., 2013). This gender difference in tinnitus pitch was only partially consistent with our results, where cluster 1 described a pitch lower than the median of the entire dataset but similar to another cluster with higher male representation. The gender differences found, could reaffirm the predominance of mild hearing loss and among female patients (Puel et al., 2002; Seydel et al., 2013).

In our study, while most clusters were characterized by symmetric hearing loss, cluster 5 presented a considerable asymmetry. Despite being the smallest group and having the lowest silhouette value, it is to be noted that the vast majority of subjects in this cluster had unilateral tinnitus. Particularly worth mentioning is the concordance between the worst ear and the lateralization of tinnitus, both on the left side. Sixteen subjects out of 17 with left-side tinnitus in this group had higher hearing loss in the left ear. In our cluster 5, the high asymmetry of hearing

thresholds is remarkable (≈ 30 dB), whereas the same is not true for MML. A recent latent class analysis from Langguth et al. (2017) found the same link between tinnitus laterality and the side of unilateral hearing loss. To our knowledge, the largest study providing evidence of the relation between tinnitus laterality and hearing asymmetry is the one from Nuttall et al. (2004), which is based on the grand average audiometric thresholds of 1033 subjects. The study found averaged threshold differences between 5 and 20 dB higher in the poorer ear in case of unilateral tinnitus. Results of another study pointed to 15 dB as the minimum average threshold difference in order to predict tinnitus laterality (Tsai et al., 2012). Another study from Vielsmeier et al. (2015) restricted the difference in thresholds to the high frequency (HF) audiometry (> 8 kHz), highlighting that the mean HF threshold difference was higher for left-side tinnitus compared to right-side or bilateral cases. In our study, it is worth stressing that the silhouette value of this cluster highlighted its instability, indicating that this deviation could be the residual outcome of the cluster analysis instead of an accurate representation of the data.

It is to be noted that the chosen clustering method may play a role in the separation and cohesion of the resulting clusters, since the performance of the algorithm is constrained by the number of clusters which are previously established (Pelleg et al., 2000). However, whereas this constraint may exert some influence in the result, the choice was made by an unsupervised method (the elbow criterion) to avoid bias. Also, the available audiometric data included in the analysis might have been decisive in the resulting groups, but this limitation is characteristic of every cluster analysis. For instance, hearing thresholds and MMLs were included in the PCA up to 8 kHz, while values above in frequency were not taken into consideration. In a study of Roberts et al. (2006) it was shown that 25% of the participants presented normal hearing up to 8 kHz (thresholds ≤ 25 dB HL), but they revealed hearing loss for higher frequencies. However, this large proportion might be due to their method of patient selection. In our study, 48 subjects (13.1%) had normal hearing up to 8 kHz, following the same criteria. In a study from Shim et al. (2009), 9.6% of subjects did not show a decrease in hearing ability in this frequency range, but when using an extended high frequency audiogram, the prevalence of hearing loss in this specific group was 67%. Although in our dataset hearing thresholds at 10, 12.5 and 16 kHz were available for most participants, this was not the case for MML. Therefore, these frequencies were excluded from the analysis in order to adhere to the corresponding frequencies between thresholds and MML. Another example of the limitation that the audiometric data entails is when it is treated as a continuous metric of hearing loss instead of ranges of threshold levels such as mild, moderate or severe hearing loss, which can better describe different possible pathological mechanisms (Langguth et al., 2017; Shekhawat et al., 2014). In our study, this problem was handled by the PCA, reducing the dimensionality of both MML and hearing threshold, which in turn can elude pre established ranges.

Here we explored a cluster analysis based on audiometric data. This did reveal a diverse pattern of patient characteristics. Although our analysis showed five

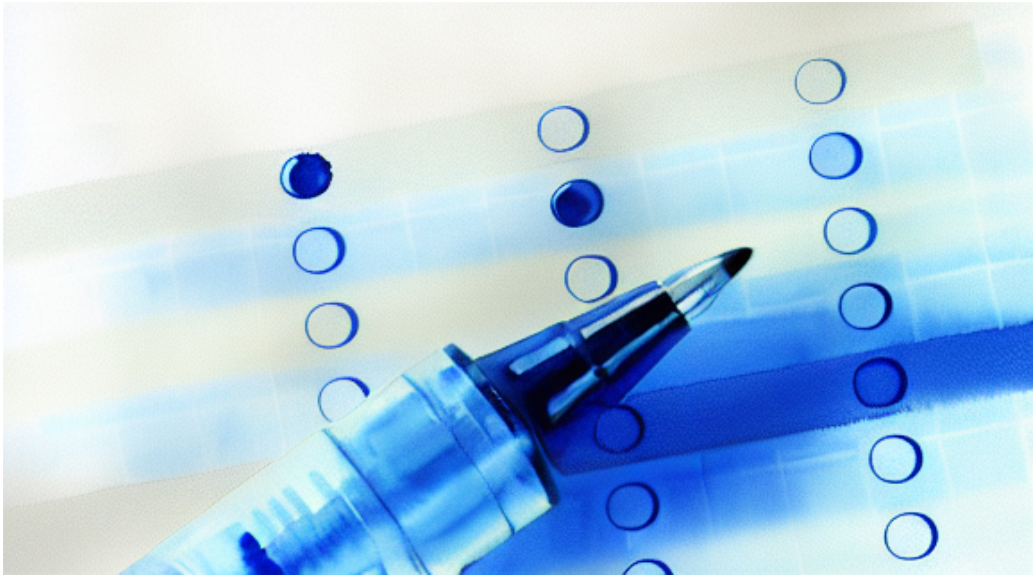
clusters, these are not very sharply divided from each other (as expressed in a low silhouette value). Our findings do not preclude the existence of tinnitus subtypes, but the use of audiometric data alone does not provide clear clusters. Future research could sharpen these clusters, if the analysis is combined with the outcome of treatments, such as hearing aid fitting or psychological interventions. On the other hand, there might not be different etiologies of tinnitus and clustering might be drawing our attention away from a common origin. Tinnitus percept and its associated distress could be an individual-dependent manifestation of a general source. Whatever the cause, tinnitus could unbalance the regular neuronal activity resulting in a rather arbitrary new state which might explain the different phenotypes.

3.5. Conclusion

To summarize, our findings suggest that the main groups found in this analysis are characterized by either easy or difficult maskability, and by better or worse hearing, which is expressed by the relation between frequency contours of MML and hearing thresholds. Although the analysis highlighted the diversity across tinnitus patients, there seems to be a continuum of patients' characteristics rather than discrete subgroups.

4

Validation of a Dutch version of the Tinnitus Functional Index in a tertiary referral tinnitus clinic



Published as: Santacruz, J. L., Arnold, R., Tuinstra, J., Stewart, R. E., & Van Dijk, P. (2021). Validation of a Dutch version of the Tinnitus Functional Index in a tertiary referral tinnitus clinic. Heliyon, 7(8).

Abstract

Tinnitus is a condition with a subjective nature that requires self-report questionnaires for its assessment. Aspects such as quality of life, sleep or intrusiveness have been addressed by multiple tinnitus questionnaires, but the high responsiveness to treatment effects of the Tinnitus Functional Index (TFI) makes this questionnaire part of the standard practice in tinnitus screening. To date, the TFI has been translated to more than 20 languages and used in more than 22 countries. In this study, the TFI was translated to Dutch and validated through a clinical population in the Netherlands.

After a back-translation procedure, the Dutch TFI was filled-out by 377 patients in the tinnitus outpatient clinic at the Ear, Nose and Throat (ENT) department of the University Medical Center Groningen, in the Netherlands. Reliability and construct validity of the questionnaire were assessed by correlations with one other tinnitus questionnaire (Tinnitus Handicap Inventory, THI) and with three psychological functioning questionnaires (Rand-36, Cantril's ladder and the Hospital Anxiety and Depression Scale (HADS)). The eight-factor structure of the Dutch TFI was tested by means of exploratory factor analysis using three different models (ICM-CFA, ESEM and ESEM-CFA).

The Dutch TFI showed a high internal consistency ($\alpha = 0.95$), and construct validity was proven by moderate-to high-convergent correlations with the THI ($r = 0.47–0.79$) and by moderate convergent ($r = 0.55–0.67$) and good-to moderate-divergent ($r = 0.12–0.47$) correlations with the psychological functioning questionnaires. The eight-factor structure of the TFI was confirmed for the Dutch version by the three models.

The Dutch version of the TFI is a reliable instrument for screening tinnitus impact in a clinical population, and its psychometric properties are comparable to the original TFI and other validated tinnitus questionnaires.

4.1. Introduction

Tinnitus ("ringing in the ears") is usually defined as the perception of a sound for which no external sound source exists. Most people experience episodes of tinnitus at times (ringing, buzzing or other sounds), either spontaneously or after being exposed to loud noise. In most cases, these sounds diminish or disappear after a certain period of time, from a few minutes to several days. If this perception persists for a period of 6 months or longer, the problem is considered chronic tinnitus (Mazurek et al., 2010).

Tinnitus is a common complaint, but its mechanisms are still poorly understood. Although different theories have been proposed, consensus has arisen with respect to a "central model" for the etiology of tinnitus, which is built on the assumption that tinnitus is the result of a change in spontaneous neural activity in the central auditory system (Eggermont and Roberts, 2004; Noreña, 2011). Most cases of tinnitus are associated with some degree of hearing loss (Shargorodsky et al., 2010). Disentangling the two of them is still a challenge today since hearing loss and tinnitus are closely related (Ratnayake et al., 2009): proportions from 70% to 80% of substantial hearing loss among tinnitus patients have been reported (Jastreboff, 2011). The prevalence of tinnitus in the adult population has been estimated to fall in the range of 10%–15% (De Ridder et al., 2014). Although there is no clear consensus in the literature on the association between sex and tinnitus (Gallus et al., 2015; Biswas and Hall, 2020), several studies have shown an increase in tinnitus prevalence and reported severity as a function of age (McCormack et al., 2014; Gallus et al., 2015; Bhatt et al., 2016). Despite clinical experience shows some examples of tinnitus in children, there is still a lack of a robust research on this issue (Rosing et al., 2016; Smith et al., 2019).

Although the consequences of tinnitus are diverse, for most patients these symptoms affect their quality of life (QoL) to a certain degree (Zeman et al., 2014). When patients severely suffer from tinnitus, several aspects of their daily functioning are also affected (Andersson and Westin, 2008). Consequences often reported by patients are sleep disturbance (Schecklmann et al., 2015), fatigue (Burke and Naylor, 2020), difficulties with hearing and with concentration (Mohamad et al., 2016), and a higher sensitivity to everyday sounds (hyperacusis Schecklmann et al., 2014). Relationships between tinnitus and psychological distress have been reported in several s tudies, highlighting that substantial percentages of the tinnitus patients had symptoms of depression or anxiety (Holmes and Padgham, 2009; Durai and Searchfield, 2016).

Since the consequences of tinnitus can be significant, research has aimed at finding effective treatments for tinnitus (Dobie, 1999; Savage and Waddell, 2014), such as pharmacological, electrophysiological or psychological approaches (Hall et al., 2016). Since a cure for tinnitus has not yet been found, the treatment of patients with tinnitus has shifted towards tinnitus management (Henry et al., 2005; Hoare et al., 2011). Tinnitus management aims at assisting patients in living with their condition as good as possible and to improve their quality of life. In order to assess the effect of tinnitus treatments on managing the complaints, there is a need

for standardised outcome measures. Numerous self-report questionnaires have been developed to assess the impact of tinnitus on patients' quality of life (Meikle et al., 2008; Kamalski et al., 2010; Hall et al., 2016), although these questionnaires were not specifically designed to study treatment outcomes (Kamalski et al., 2010). In order to study the effects of treatment options on the quality of life of the patients, it is necessary to use instruments that are responsive to treatment effects (Meikle et al., 2008). Therefore, Meikle et al. (2012) developed the Tinnitus Functional Index (TFI), to be able to assess both the impact of tinnitus and the treatment-related effects on the quality of life of the patients. In the developing process, an original prototype consisting of 175 items belonging to 9 different tinnitus questionnaires were evaluated by an expert panel and 13 different domains or subcategories were identified. After a refining process of clinical evaluations and restructurations, the final TFI resulted in 25 questions organized in 8 subscales of factors: intrusive, sense of control, cognitive, sleep, auditory, relaxation, quality of life and emotional.

The aim of the present study is to assess the psychometric properties of a Dutch version of the Tinnitus Functional Index and to test whether the same structure of 8 factors can be found, taking into consideration how these factors relate to each other. The original English version of the TFI has recently been validated within several cultures and for different languages (Oron et al., 2018; Kam et al., 2018; Peter et al., 2017; Hoff and Kähäri, 2017; Wrzosek et al., 2016; Fackrell et al., 2016, 2018; Rabau et al., 2014; Suzuki et al., 2019; Müller et al., 2016). It is worth noting that the TFI version of Rabau et al. (2014) is written in Dutch language from Belgium (also known as Flemish Dutch), different from the one proposed in our study. Here, the performance of the Dutch version of the Tinnitus Functional Index was studied in a clinical setting, as part of the assessments in a tinnitus outpatient clinic at the ENT department of a university hospital in the Netherlands.

4.2. Methods

4.2.1. Participants and procedure

As part of a standard diagnostic protocol, the data for this study were collected in a tertiary referral tinnitus clinic at the University Medical Center Groningen. All patients who visited this clinic filled in several questionnaires in order to gather information on their tinnitus characteristics as well as to screen for potential psychosocial problems. These data are used in the multidisciplinary assessment of the patients to determine the advice for further treatment. The Dutch version of the TFI was administered to a group of 377 consecutive tinnitus patients, who visited the specialised multidisciplinary outpatient clinic between September 2013 and September 2015.

Data were included in this study when patients were 18 years or older, and mastered the Dutch language sufficiently to fill in the questionnaires. Since the data were collected as part of the routine assessment in the tinnitus outpatient

clinic and are anonymously reported in this paper, no informed consent was asked of the participants. The study met the criteria for an exemption from institutional review board approval (METc2013/400).

4.2.2. Measurements

The Tinnitus Functional Index

The original TFI (Meikle et al., 2012) was translated by means of a back-translation procedure, following Guilliman et al. (1993) guidelines. First, the translation to Dutch was carried on by two independent translators with Dutch as native language. Our Dutch translation of the questionnaire was translated back into English by another translator with English as native language. Thus, the accuracy of the translation process could be checked. None of the translators involved in the process were medically skilled. The comparison of the original TFI with the translated version was carried out by bilingual experts in the field, and it did not reveal differences in the meaning of the individual items. The TFI consists of 25 items, which are divided into 8 subscales: intrusive (3 items), sense of control (3 items), cognitive (3 items), sleep (3 items), auditory (3 items), relaxation (3 items), quality of life (4 items), and emotional (3 items). All items are scored on a 10-point rating scale, with "0" and "10" indicating the lowest and highest impact on functioning, respectively. Items 1 and 3 are scored as percentages and have to be re-coded into a 10-point scale. Each subscale is scored individually: scores on the separate items are added up, divided by the number of items in the scale, and multiplied by 10. For the total TFI score, all items are added up, divided by 25 (the total number of items) and multiplied by 10. Figure 1 shows the Dutch version of the TFI. Total scores between 0-17 are interpreted as "not a problem", total scores between 18-31 as "small problem", total scores between 32-53 as "moderate problem", total scores between 54-72 as "big problem", and total scores between 73-100 as "very big problem".

Tinnitus Handicap Inventory

In the present study, scores on the TFI were compared to scores on the Dutch version of the Tinnitus Handicap Inventory (THI; Newman et al., 1996), a validated (Newman et al., 1998; Brussee, 2003) and widely used questionnaire developed to assess the severity of patients' tinnitus handicap. The THI consists of 25 items, scored on a 3-point self-rating scale (0 = "no", 2 = "sometimes", and 4 = "yes"). In addition to the total score, three different subscales are scored as well: functional (11 items), emotional (9 items), and catastrophic (5 items). Higher scores indicate a higher tinnitus impact.

Psychological functioning

Mental health or psychological functioning was measured by the Mental Health subscale of the Rand-36 Health Survey (Ware Jr and Sherbourne, 1992; Vander Zee et al., 1996). This subscale consists of 5 items, scored on a 6-point self-rating scale (0 = never to 6 = always) and assesses mood, including symptoms of depression

and tension. The total score on this subscale varies from 0 to 100, with higher scores indicating a better psychological functioning or mental health.

Overall wellbeing was measured on Cantril's ladder (Cantril, 1965), which is a scale ranging from 0 to 10. Patients answered the following question: 'Here is a picture of a ladder. Suppose the top of the ladder represents the best possible life for you and the bottom represents the worst possible life for you. Where on the ladder do you feel you personally stand at the present time?'

Symptoms of anxiety or depression were assessed by the Hospital Anxiety and Depression Scale (Zigmond and Snaith, 1983; Spinhoven et al., 1997). The HADS is a 14-item self-report screening instrument, developed to identify possible cases with anxiety or depression. The instrument consists of two 7-item scales, one of them with items assessing symptoms of anxiety, and the other one with items assessing symptoms of depression. The subscales vary from 0 to 21, with higher scores indicating a higher amount of anxiety or depression. The authors of the original questionnaire identified scores from 0 to 7 as "non-cases", scores from 8 through10 as "doubtful cases", and scores higher than 11 as "cases" with anxiety or depression (Zigmond and Snaith, 1983).

4.2.3. Data analysis

All descriptive analyses, reliability analyses, and construct validity analyses were performed with IBM SPSS Statistics 23. The factor structure of the Dutch version of the TFI was tested with M-Plus version 8.

Reliability and construct validity of the TFI

Reliability scores of the Dutch TFI were assessed by calculating the internal consistency coefficient Cronbach's alpha (α) for each subscale as well as for the total questionnaire (Cronbach, 1951). In general, Cronbach's alphas of ≥ .80 are considered good for diagnostic instruments, although Cronbach's alphas of > .90 are recommended in case of screening instruments (Nunnally and Bernstein, 1994).

Construct validity was evaluated by means of convergent and divergent correlations between TFI and measures of tinnitus handicap and psychological functioning. For it, Spearman's correlation coefficients between these measures were obtained. Correlation coefficients between .10 and .30 were considered small, correlations between .30 and .50 were considered moderate, and correlations higher than .50 were considered large (Cohen, 2013).

Factor structure of the TFI

In the original study (Meikle et al., 2012), the eight-factor structure of the TFI was derived from a principal component analysis (PCA, aimed to reduce the dimensionality of data) as an independent cluster model (ICM). An ICM (Marsh et al., 2009) is a factor structure in which each of the 25 items is loaded on only one of the eight factors. Three models were tested and compared to confirm the factor structure of the TFI.

First, a confirmatory factor analysis (CFA) was performed to check whether the ICM eight-factor structure of the original study could be confirmed (model ICM-CFA). In an ICM model, items load at their respective factor with no cross-loads on the other latent factors. A critical comment on the ICM model is that the zero factor loadings of items usually displays poor fit and leads to distorted factors with overestimated factor correlations (Marsh et al., 2009).

Second, in order to investigate whether cross-loading could be found in the ICM-CFA, an exploratory structural equation model (ESEM) (Asparouhov and Muthén, 2009) of eight factors was performed as an exploratory factor analysis (EFA).

And third, based on the ESEM model, we investigated whether an ESEM-CFA model could be obtained. An ESEM-CFA model means that non-significant loadings that are larger than zero of the ESEM solution then become zero loadings. This involves obtaining a model with cross-loadings, but the cross-loadings were retrieved from an ESEM model. Since the data were comprised of continuous variables, parameter estimation of the ESEM model was estimated by maximum likelihood (ML) with oblique factor rotation Geomin. A Geomin criterion of 0.01 with 30 random starts was used.

Finally, a goodness of fit test (GOF) (Schreiber et al., 2006) was used to compare the three models (ICM-CFA, ESEM and ESEM-CFA).

4.3. Results

4.3.1. Participants

Table 4.1 shows the demographic characteristics of the patients that were included in this study. In total, 377 patients participated in the study. More men (60.7%) than women (39.3%) were included, with a mean age of 54.8 years (range 19–88 years). Tinnitus duration was on average 7.1 years. The number of patients with an acute or gradual onset of tinnitus was almost equally divided. Most of the patients in this study experienced a continuous tinnitus (89.6%), whereas a smaller amount of the patients experienced tinnitus at intervals (10.4%). The majority of the patients reported hearing loss (68.7%). The demographic data described a wide range of characteristics in our clinical population.

4.3.2. Instruments

Table 4.2 gives an overview of all instruments used in the present study. The average TFI score fell into the ‘moderate problem’ category with a value of 48 ± 20.4, characteristic of a common tinnitus population as previous studies reported (Fackrell et al., 2018; Wrzosek et al., 2016; Peter et al., 2017; Jacquemin et al., 2019). In line with it, the THI presented also a ‘moderate handicap’ on average with a score of 44 ± 22.3. Psychological functioning tests such as Rand-36, Cantril’s ladder and HADS presented relatively normal average values as well.

Demographic characteristics	N = 377 (%)
Sex	
Male	229 (60.7)
Female	148 (39.3)
Age (years)	
Mean	54.8
SD	13.6
Range	19-88
Marital status	
With partner	301 (80.3)
Without partner	74 (19.7)
Missing	2
Educational level	
Low	87 (24.0)
Middle	166 (45.7)
High	110 (30.3)
Missing	14
Tinnitus duration (years)	
Mean	7.1
SD	8.1
Range	0-47
Onset of tinnitus	
Acute	174 (47.7)
Gradual	191 (52.3)
Missing	12
Presence of tinnitus	
Continuous	329 (89.6)
With intervals	38 (10.4)
Missing	10
Perceived hearing loss	
No	118 (31.3)
Yes	259 (68.7)

Table 4.1 | Demographic data and tinnitus characteristics of the subjects.

	N	Items	Possible Score Range	Observed Score Range	Mean	SD
Tinnitus Functional Index						
Intrusive	356	3	0-100	0-100	61.61	21.98
Sense of Control	359	3	0-100	3.33-100	65.04	20.44
Cognitive	362	3	0-100	0-100	43.43	25.38
Sleep	369	3	0-100	0-100	47.27	34.29
Auditory	361	3	0-100	0-100	43.52	30.5
Relaxation	366	3	0-100	0-100	47.48	27.96
Quality of Life	360	4	0-100	0-100	37.26	26.74
Emotional	363	3	0-100	0-100	39.61	27.66
Total	371	25	0-100	3.20-100	47.93	20.41
Tinnitus Handicap Inventory						
Functional	368	11	0-44	0-44	21.68	10.41
Emotional	368	9	0-36	0-36	14.08	9.01
Catastrophic	370	5	0-20	0-20	7.96	4.88
Total	374	25	0-100	0-98	43.84	22.33
RAND-36						
Mental health	372	5	0-100	0-100	63.89	19.27
Cantril's ladder	363	1	0-10	0-10	6.31	1.85
Hospital Anxiety & Depression Scale						
Anxiety	368	7	0-21	0-21	6.96	4.18
Depression	369	7	0-21	0-21	5.82	4.54

Table 4.2 | Questionnaires and subscales used. The maximum score for TFI, THI and RAND-36 is 100. The maximum scores for Cantril's ladder and HADS is 10 and 21, respectively.

	N	Items	Cronbach's Alpha
Tinnitus Functional Index			
Intrusive	356	3	0.82
Sense of Control	359	3	0.72
Cognitive	362	3	0.92
Sleep	369	3	0.96
Auditory	361	3	0.95
Relaxation	366	3	0.94
Quality of Life	360	4	0.89
Emotional	363	3	0.9
Total	309	25	0.95

Table 4.3 | Internal consistency scores of the Dutch version of the TFI.

4.3.3. Reliability

Table 4.3 summarizes the internal consistency scores of the subscales of the Dutch version of the TFI. Most of the subscales of the TFI, as well as the total scale, showed good internal consistency scores (Cronbach's alphas ranged from 0.82–0.96). Subscale "sense of control" showed a satisfactory internal consistency with a Cronbach's alpha of .72.

The internal consistency scores of the Dutch TFI were comparable to the scores of the original English version of the TFI, with only a lower internal consistency score for subscale "sense of control" of the Dutch version of the TFI (Meikle et al., 2012). The obtained values of internal consistency highlighted the reliability of each subscale.

	THI Functional	THI Emotional	THI Catastrophic	RAND-36 Mental Health	Cantril's Ladder	HADS Anxiety	HADS Depression
TFI							
Intrusive		+					
Sense of Control			+				
Cognitive	+			0	0	0	0
Sleep	+			0	0	0	0
Auditory	+			0	0	0	0
Relaxation	+					+	
Quality of Life	+	+		-	-		+
Emotional		+		-	-	+	+

+ = expected positive correlation
- = expected negative correlation
0 = no association expected

Table 4.4 | Expected convergent and divergent correlations between TFI subscales, THI subscales, and measures of psychological functioning.

4.3.4. Construct validity

Table 4.4 shows the expected convergent and divergent correlations between TFI subscales, THI subscales, and measures of psychological functioning. Convergent correlations were expected between the TFI subscales and the corresponding subscales of the THI. Also, subscales Quality of Life and Emotional were expected to be related to measures of psychological functioning. Divergent correlations were expected between TFI subscales Cognitive, Sleep and Auditory and measures of psychological functioning. These assumptions were made by the authors and based on their own clinical experience. In the case of the expected correlations between TFI and THI, both questionnaires contain similar questions.

Table 4.5 displays the actual convergent and divergent correlations that were found in the study population. With respect to convergent validity, all TFI subscales showed significant moderate-to strong-correlations (range 0.47–0.79) with the corresponding subscales of the THI and measures of psychological functioning. Subscales Intrusive and Auditory correlated less strongly with THI subscales Emotional and Functional, respectively (r = 0.47). All expectations regarding the direction of the convergent correlations were confirmed by the results.

With respect to divergent validity, significant, but small-to moderate-correlations were found for TFI subscales Cognitive, Sleep, and Auditory with measures of psychological functioning. Almost all of the correlation coefficients were smaller than 0.50 (range 0.12–0.47), which is indicative of a satisfactory divergent validity. Subscale Cognitive correlated strongly with overall wellbeing as measured by Cantril's ladder (r = 0.50), which indicates that some association exists between these constructs. All expectations with respect to the direction of the divergent correlations were confirmed by the results.

Overall, the construct validity showed smaller divergent correlations compared to convergent correlations for the subscales of the TFI. These correlations indicated a strong construct validity of the questionnaire for almost all subscales, which might infer that these factors are adequate for assessing the aspects of tinnitus that they are intended to measure.

	THI Functional	THI Emotional	THI Catastrophic	RAND-36 Mental Health	Cantril's Ladder	HADS Anxiety	HADS Depression
TFI							
Intrusive		**.47****					
Sense of Control			**.54****				
Cognitive	**.76****			*-.46***	*-.50***	*.47***	*.56***
Sleep	**.60****			*-.37***	*-.33***	*.39***	*.44***
Auditory	**.47****			*-.12**	*-.12**	*.19***	*.26***
Relaxation	**.64****					**.53****	
Quality of Life	**.79****	**.68****		**-.55****	**-.56****		**.67****
Emotional		**.78****		**-.66****	**-.57****	**.60****	**.64****

* = p <.05
** = p <.01
Bold values indicate the expected convergent correlations; Italic values indicate expected divergent correlations.

Table 4.5 | Convergent and divergent Spearman correlations obtained between TFI subscales, THI subscales, and measures of psychological functioning.

4.3.5. Confirmation of the 8-factor structure of the TFI

The 8-factor structure was tested by three different models (ICM-CFA, ESEM, and ESEM-CFA).

Tables 4.6, 4.7 and 4.8 show the standardized factor loadings (β) for all 25 TFI items and the 8 factors. The loadings of the ICM-CFA are shown in Table 6A, where only the items of each factor are considered and the empty cells represent zero loadings. All values indicate good associations with their designated factor since they are above the recommended cut-off ≥ 0.40 (Wülferth, 2013). Table 6B contains the loadings of the ESEM model. Values in bold correspond to the significant loadings ($p \leq 0.05$), which occurs for items that are either associated with their factor or not. For this model, several items showed significant cross-loadings with other factors (i.e., item 20 and factor Emotional). However, none of these cross-loadings scored above the cut-off value of 0.40. The loadings of the model ESEM-CFA are shown in Table 6C, which includes only the significant loadings obtained in the ESEM-CFA model, zero loadings appear blank. As in the previous model, none of the cross-loadings scored above 0.40.

Tables 4.9, 4.10 and 4.11 contain the correlations between factors of the 3 models. Values presented in bold are below or above the recommended criteria (<0.30 to >0.85) (Hair et al., 2010). For all models, the Auditory factor showed the weakest correlations with the rest of the factors.

Table 4.12 shows the results of the goodness of fit test (GOF). Values of root mean square error of approximation (RMSEA) for the three models are below 0.08, indicating good fitting (MacCallum et al., 1996). Despite RMSEA values should normally be below 0.05, the limit of 0.08 is reasonable when the standardized root mean square residual (SRMR) is lower than 0.06 (Hu and Bentler, 1999), which was true for the three models. However, the models ESEM and ESEM-CFA showed better values of GOF. Although AIC and BIC values were smaller for the ESEM-CFA model, these parameters did not differ to a great extent between the three models.

	Intrusiveness	Sense of Control	Cognitive	Sleep	Auditory	Relaxation	Quality of life	Emotional
TF1	0.778							
TF2	0.826							
TF3	0.77							
TF4		0.415						
TF5		0.849						
TF6		0.778						
TF7			0.903					
TF8			0.918					
TF9			0.839					
TF10				0.901				
TF11				0.985				
TF12				0.925				
TF13					0.918			
TF14					0.997			
TF15					0.897			
TF16						0.896		
TF17						0.957		
TF18						0.891		
TF19							0.868	
TF20							0.85	
TF21							0.812	
TF22							0.768	
TF23								0.781
TF24								0.875
TF25								0.949

Table 4.6 | Standardized loadings (β) of ICM-CFA model: Eight factors based on 25 items of the TFI. All values are above the recommended cut-off ≥ 0.40.

	Intrusiveness	Sense of Control	Cognitive	Sleep	Auditory	Relaxation	Quality of life	Emotional
TF1	**0.924**	-0.049	-0.068	-0.011	0.001	0.006	0.073	-0.011
TF2	**0.627**	0.134	0.051	0.01	**0.121**	0.065	-0.038	0.021
TF3	**0.448**	0.159	0.12	0.041	-0.034	-0.037	-0.015	0.268
TF4	-0.004	**0.473**	0.013	-0.09	0.089	0.08	0.044	-0.1
TF5	0.014	**0.593**	-0.003	0.098	0.01	-0.028	-0.03	0.38
TF6	0.078	**0.63**	0.042	0.011	-0.02	0.045	0.093	0.06
TF7	0.072	0.079	**0.736**	0.076	0.035	-0.004	0.064	-0.025
TF8	-0.002	-0.128	**0.992**	0.009	0.037	0.019	-0.014	0.013
TF9	-0.015	0.079	**0.607**	0.013	-0.041	0.085	**0.162**	0.074
TF10	0.012	0.044	0.069	**0.86**	-0.034	-0.015	-0.039	0.041
TF11	0.004	-0.022	-0.006	**0.962**	**0.035**	0.014	0.024	0.019
TF12	0.005	0.003	-0.006	**0.891**	0.016	0.063	0.041	-0.045
TF13	0.028	-0.02	0.008	-0.004	**0.893**	0.027	-0.013	0.046
TF14	-0.029	0.014	-0.037	0.017	**1.014**	0.017	-0.005	0.041
TF15	0.033	0.011	0.083	0.006	**0.815**	-0.029	**0.109**	-0.092
TF16	-0.003	0.018	0.07	0.019	**0.071**	**0.808**	0.01	-0.009
TF17	0.005	-0.039	0.055	0.024	0.004	**0.911**	0.01	0.019
TF18	0.037	0.052	-0.026	0.012	-0.02	**0.813**	0.018	0.063
TF19	0.054	0.019	0.004	0.004	0.035	0.029	**0.872**	-0.04
TF20	0.024	0	-0.019	0.041	**-0.086**	**0.142**	**0.52**	**0.353**
TF21	-0.028	-0.026	0.066	0.024	0.056	-0.12	**0.795**	0.101
TF22	0.001	0.079	0.295	-0.051	0.021	0.079	**0.452**	0.037
TF23	-0.039	0.052	-0.124	0.041	0.039	0.072	0.088	**0.714**
TF24	0.034	0.045	0.059	-0.068	0.053	0.023	-0.009	**0.818**
TF25	0.031	-0.096	0.077	-0.006	-0.007	0.01	0.067	**0.895**

Values in bold correspond to the significant loadings (p ≤ 0.05).

Table 4.7 | Standardized loadings (β) of ESEM model: Eight ESEM factors based on 25 items of the TFI.

	Intrusiveness	Sense of Control	Cognitive	Sleep	Auditory	Relaxation	Quality of life	Emotional
TF1	1.266	-0.33	-0.219					
TF2	0.806							
TF3	0.599							0.24
TF4		0.555						-0.153
TF5		0.646						0.26
TF6		0.819						
TF7			0.897					
TF8		-0.297	1.142					
TF9			0.685				0.204	
TF10				0.9				
TF11				0.985				
TF12				0.925				
TF13					0.915			
TF14			-0.075		1.04			
TF15					0.896			
TF16						0.895		
TF17						0.957		
TF18						0.89		
TF19							1.001	-0.135
TF20					-0.136	0.12	0.592	0.307
TF21						-0.193	0.979	
TF22			0.32				0.519	
TF23								0.779
TF24								0.872
TF25								0.955

Table 4.8 | Standardized loadings (β) of ESEM-CFA model: Eight ESEM factors based on 25 items of the TFI. Only the significant loadings ($p \leq 0.05$) of the ESEM-CFA model are shown, zero loadings appear blank.

Factor	1	2	3	4	5	6	7	8
(1) Intrusiveness	1	0.766	0.624	0.495	0.388	0.59	0.588	0.595
(2) Sense of Control		1	0.646	0.505	**0.293**	0.632	0.614	0.676
(3) Cognitive			1	0.562	0.505	0.7	0.756	0.625
(4) Sleep				1	**0.239**	0.586	0.501	0.474
(5) Auditory					1	0.366	0.494	**0.269**
(6) Relaxation						1	0.749	0.696
(7) Quality of life							1	0.78
(8) Emotional								1

Values in bold are below or above the recommended criteria (0.85). 1 = Intrusiveness; 2 = Sense of control; 3 = Cognition; 4 = Sleep; 5 = Auditory; 6 = Relaxation; 7 = Quality of life; 8 = Emotional.

Table 4.9 | ICM-CFA model: Correlations between factors.

Factor	1	2	3	4	5	6	7	8
(1) Intrusiveness	1	0.547	0.456	0.393	0.312	0.446	0.381	0.433
(2) Sense of Control		1	0.488	0.342	**0.215**	0.444	0.32	0.412
(3) Cognitive			1	0.483	0.467	0.605	0.607	0.523
(4) Sleep				1	**0.177**	0.527	0.384	0.434
(5) Auditory					1	0.308	0.47	**0.176**
(6) Relaxation						1	0.63	0.633
(7) Quality of life							1	0.607
(8) Emotional								1

Values in bold are below or above the recommended criteria (0.85). 1 = Intrusiveness; 2 = Sense of control; 3 = Cognition; 4 = Sleep; 5 = Auditory; 6 = Relaxation; 7 = Quality of life; 8 = Emotional.

Table 4.10 | ESEM model: Correlations between factors.

Factor	1	2	3	4	5	6	7	8
(1) Intrusiveness	1	0.795	0.691	0.497	0.421	0.6	0.573	0.565
(2) Sense of Control		1	0.688	0.473	**0.334**	0.596	0.529	0.564
(3) Cognitive			1	0.558	0.542	0.697	0.699	0.621
(4) Sleep				1	**0.268**	0.586	0.48	0.48
(5) Auditory					1	0.399	0.544	**0.292**
(6) Relaxation						1	0.727	0.691
(7) Quality of life							1	0.734
(8) Emotional								1

Values in bold are below or above the recommended criteria (0.85). 1 = Intrusiveness; 2 = Sense of control; 3 = Cognition; 4 = Sleep; 5 = Auditory; 6 = Relaxation; 7 = Quality of life; 8 = Emotional.

Table 4.11 | ESEM-CFA model: Correlations between factors. Correlations between factors.

	ICM-CFA	ESEM	ESEM-CFA
AIC	38197	37951	37913
BIC	38602	38824	38373
RMSEA (90%CI)	0.071 (0.065 - 0.077)	0.046 (0.036 - 0.055)	0.044 (0.036 - 0.051)
SRMR	0.047	0.011	0.027
CFI	0.946	0.988	0.98
TLI	0.943	0.972	0.975

AIC = Akaike information criterion; BIC = Bayesian information criterion; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardised Root Mean Square Residual; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index.

Table 4.12 | Goodness of fit (GOF) statistics for the models ICM-CFA, ESEM, and ESEM-CFA.

4.4. Discussion

The aim of the present study was to assess the psychometric properties of the Dutch translation of the Tinnitus Functional Index. The original English TFI was translated into Dutch and tested on a population of 377 tinnitus patients. Reliability of the questionnaire was tested by means of internal consistency, and construct validity was estimated through convergent and divergent correlations with the THI and 3 psychological functioning questionnaires. A factor analysis was performed to confirm the 8-factor structure of the TFI by using 3 different models. Overall, the Dutch version of the TFI has shown good qualities with respect to the internal consistency and convergent validity, comparable to the values of the original TFI but also to those obtained for the validation of the Dutch version of the THI (Brussee, 2003).

In line with the study of Meikle (Meikle et al., 2012) and previous validations of the TFI, the Auditory subscale showed the lowest correlation values with the rest of the factors (Wrzosek et al., 2016; Fackrell et al., 2018). A possible explanation of this effect is the comorbidity between tinnitus and hearing loss and the challenge of disentangling the two of them, which is the rationale for the creation of the Tinnitus and Hearing Survey (THS) (Henry et al., 2015b). However, the THS addresses tinnitus, hearing and sound tolerance with 4 items per factor. Therefore, the THS takes into account these covariates but is less responsive for assessing

tinnitus impact separately. The authors of the original version of the TFI suggested studying the impact of removing the Auditory factor from the questionnaire. This analysis was carried out later by Fackrell et al. (2018), who tested a modified TFI-22 version, which performed better in their UK clinical population. Nevertheless, the authors suggested not removing the Auditory factor from the TFI but using a different scoring system. Taking this into account, we consider that the Auditory factor of the Dutch TFI provides complementary information due to the association between tinnitus and hearing loss and, therefore, it is a useful supplement to the questionnaire. Further studies could investigate the impact of a modified scoring system that increases the Auditory correlation values with the rest of the factors without undermining the TFI performance.

It is noteworthy how the TFI scores were interpreted in the original study of Meikle: mild (scores below 25), moderate (scores between 25 and 50) and severe (scores above 50) problem. As it has been pointed out by Gos et al. (2021), the averaged TFI score is often close to the severe limit, and with a quite high dispersion. This applies in particular to our data (M = 47.93; SD = 20.41), which raises the question of whether the original cut-off for diagnosing severe tinnitus is too low. The study of Gos et al. (2021) suggested that this boundary should be set at 65 points, in order to limit the most severe rating to a smaller sample. In line with Gos et al. findings, 37% of the patients in our study obtained total score above 50. The study by Fackrell et al. (2016) included both clinical and non-clinical populations, obtaining a lower proportion of participants with global TFI scores above 50 (30%). The difference in severe cases might be explained by the tinnitus symptoms of a patient population who seek medical help, compared to a general population who might report milder tinnitus on average. In our dataset, a proportion of 17.2% of patients scored above 65, which is a rather small group and might not represent the distress reported by the patients who visited our clinic. Due to the similarities between the global scores of the THI and the TFI, a potential solution to this problem is to increase the number of categories as in the THI, instead of raising the limit of the group with severe tinnitus.

Construct validity of the Dutch TFI showed strong correlations with the THI for almost all factors. One of the exceptions was the convergent validity between the TFI-factor Intrusive and the THI-factor Emotional. Previous studies highlighted the importance of evaluating tinnitus intrusiveness for studying treatment outcomes (Hoare et al., 2011; Hall et al., 2018). However, tinnitus intrusiveness seems to be a complex construct that can be interpreted in different ways, as it can be deduced from comparing different tinnitus questionnaires (Jacquemin et al., 2019). In the case of the Dutch TFI, the three items belonging to this factor are focused on annoyance, awareness and loudness of the tinnitus percept. Two of these items (awareness and loudness) do not necessarily correlate with the items included in the Emotional factor of the THI, which mostly covers anxiety, depression and psychological impact. The low correlation obtained for this particular comparison between the two questionnaires might be due to this effect, since only one of the items evaluating intrusiveness is clearly connected to the THI-Emotional. A similar effect occurs when comparing the TFI-Auditory to the THI-Functional, for which

a weak correlation was obtained as well. The Functional factor covers aspects such as concentration, sleep, intrusiveness and fatigue. Only 2 out of 11 items of this THI factor are surely related to the TFI-Auditory, and these are "Does the loudness of your tinnitus make it difficult for you to hear people?" and "Does your tinnitus interfere with your ability to enjoy your social activities (such as going out to dinner, to the movies)?". The wide-ranging design of the THI-Functional is presumably the reason for the low convergent validity obtained. Moreover, a strong correlation is expected when comparing two subscales with the same name from different questionnaires, however, they might measure different underlying aspects (Jacquemin et al., 2019). Nevertheless, it should be noted that previous translations of the THI have shown that the subscales are unreliable, and a THI-total scale might be a valid measure of general tinnitus related distress (Zachariae et al., 2000). Further validations of the TFI might benefit most by analyzing construct validities of the global scores.

One aspect of our study that should be considered is the confirmation of the 8-factor structure by means of 3 different models of factor analysis. Most of the available TFI translations used a CFA model based on independent clusters (ICM). This method assumes no crossloadings between factors which leads to poor fit and overestimated factor correlations (Marsh et al., 2009). In addition to this model, the 8-factor structure of the Dutch TFI was confirmed by 2 more models (ESEM and ESEM-CFA) that take into account possible crossloadings between factors and, consequently, further ensuring the fit. This overestimation can be seen when comparing Tables 7B and 8A all correlations are higher in the first table. We think that the models ESEM and ESEM-CFA are more adequate for a factor analysis in a study like this one, given the complexity and the subjective nature of a tinnitus questionnaire.

Another aspect worth to note with regard to previous TFI translations, is that the Dutch TFI was validated through a broad and diverse clinical population of 377 patients whose characteristics corroborate the values of reliability and construct validity that have been obtained in this study. Both sample size and techniques of factor analysis used in this study make the validation process more robust. It should be noted that the Dutch language used in Rabau et al. (2014) refers to Flemish, which is mostly spoken in Belgium. One of the main motivations of this study was to obtain a new Dutch version that could be fully understood by a clinical population in The Netherlands.

Some items of the models ESEM and ESEM-CFA loaded on to their designated factor but also on to others, resulting in the so-called crossloadings. Although the significance of a factor loading depends on the sample size (Stevens, 2012), it's common practice in exploratory factor analysis to ignore loadings below 0.3 (Field et al., 2012). Using the recommendation of Guadagnoli and Velicer (1988), only scores greater than 0.4 are considered stable. In our study, none of the crossloading scores in any of the models exceeded this threshold, resulting in only stable items with loadings on to their designated factor. Despite the crossloadings of both models can be ignored, ESEM-CFA showed better correlations between factors and better GOF values when compared to ESEM. Therefore, we suggest

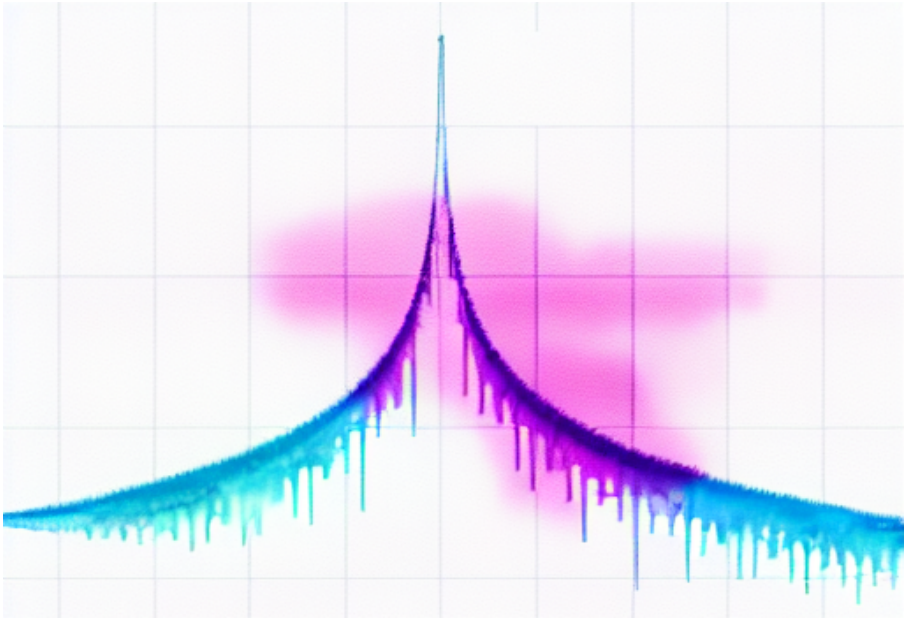
that the ESEM-CFA is the most optimal model out of the three.

Although the Dutch TFI showed a good reliability as a screening tool, responsiveness to treatment for different follow-up groups was not evaluated in this study. The main goals of the original TFI were evaluating both the impact of tinnitus and the treatment-related effects on the patients. Further analyses should focus on evaluating treatment efficacy by measuring the changes before and after treatment for the total score and for each subscale.

Overall, the results of this study show that most of the subscales of the Dutch version of the TFI have a good internal consistency. The reliability scores are considered good for use as a diagnostic instrument as well as a screening instrument (Nunnally and Bernstein, 1994). Furthermore, these results are comparable with the reliability scores of the original TFI (Meikle et al., 2012). Only the subscale "sense of control" showed a low internal consistency, which indicates that its use for screening should be done carefully, although the scale is acceptable for using it as a research instrument.

5

Comparison between two self-guided tinnitus pitch matching methods



Published as: Santacruz, J. L., de Kleine, E., & van Dijk, P. (2023). Comparison between two self-guided tinnitus pitch matching methods. Frontiers in Aging Neuroscience, 15, 1095178.

Abstract

Tinnitus pitch matching is a procedure by which the frequency of an external sound is manipulated in such a way that its pitch matches the one of the tinnitus. The correct measure of the tinnitus pitch plays an important role in the effectiveness of any sound-based therapies. To date, this assessment is difficult due to the subjective nature of tinnitus. Some of the existing pitch matching methods present a challenge for both patients and clinicians, and require multiple adjustments of frequency and loudness, which becomes increasingly difficult in case of coexisting hearing loss. In this paper, we present the comparison in terms of reliability between two self-guided pitch matching methods: the method of adjustment (MOA) and the multiple-choice method (MCM).

Twenty participants with chronic tinnitus and hearing loss underwent the two assessments in two different sessions, 1 week apart. Measures of intraclass correlation (ICC) and difference in octaves (OD) within-method and within-session were obtained.

Both methods presented good reliability, and the obtained values of ICC and OD suggested that both methods might measure a different aspect of tinnitus.

Our results suggest that a multiple-choice method (MCM) for tinnitus pitch matching is as reliable in a clinical population as more conventional methods.

5.1. Introduction

Tinnitus is often defined as the perception of sound without an external source. Several studies have reported the tinnitus prevalence in the population, which ranges from 5.1% to 42.7% (McCormack et al., 2016). One of the main challenges of health care when addressing tinnitus is the large heterogeneity of its symptoms and etiologies (Langguth, 2011), making it improbable that a specific therapy would be suitable for every patient (Hall et al., 2019). Some authors have highlighted the importance of personalized treatments, which are prescribed according to the physiological mechanisms that underlie each individual's symptoms. The most frequent comorbidity of tinnitus is hearing loss which, in the case of the Dutch population, has an association with an odds ratio of 8.5 (Schubert et al., 2021).

There is an increasing interest in sound-based therapies for tinnitus treatment (e.g., Henry et al., 2008; Hobson et al., 2012; McNeill et al., 2012; Shekhawat et al., 2013; Searchfield et al., 2017). The most common sound-based tinnitus therapy by far are hearing aids, and it has been estimated that up to 90% of the tinnitus population may benefit from their use (Henry et al., 2015a). Hearing aids increase the volume of external sounds, improving the communication of users. They may help to reduce other tinnitus symptoms like stress or anxiety, but also mask or provide distraction from tinnitus (Sereda et al., 2015). Nevertheless, patients differ with respect to many audiological characteristics, such as the degree of hearing loss, the tinnitus pitch and loudness, the factors that influence their tinnitus or their psychological response to the tinnitus percept (Schaette et al., 2010; Cederroth et al., 2019).

The potential dependency of the tinnitus pitch and the effectiveness of a sound-based therapy has motivated the development of different pitch-based treatments. Examples of these are the vagus nerve stimulation combined with a sound stimulus (De Ridder et al., 2015), tailor-made notch noise training (Stracke et al., 2010), notch filter amplification (Marcrum et al., 2021), harmonic sound therapy (Mahboubi et al., 2012), phase-shift sound therapy (Heijneman et al., 2012), or different discrimination/attention tasks focused on re-adjusting the attention to the tinnitus percept (Hoare et al., 2010; Wise et al., 2015).

Sound-based therapies are often fine-tuned to the pitch of the tinnitus (Hoare et al., 2014). A procedure well-known in the tinnitus field, is tinnitus matching, where the frequency of an external sound is manipulated such that its pitch matches that of the tinnitus (Henry and Meikle, 2000). Although pitch matching is part of the standard audiological assessment of a tinnitus clinic, its reliability is often questioned due to its self-reported nature and the large variability between consecutive sessions (Hoare et al., 2014), which can even vary over 2 octaves (Henry et al., 2004). It remains unclear whether these variations are the result of the patients' difficulties when performing the tests or whether they reflect a change of the percept between sessions (Penner and Bilger, 1992). Even though clinicians have to rely on patients' feedback when performing a pitch matching test, the procedure does not entirely resemble a "black box." Many authors have investigated the relationship between the audiogram and the tinnitus pitch and,

more specifically, several instances can be found in the literature where authors theorize on the link between audiogram edge and pitch (Schaette and Kempter, 2009; Moore et al., 2010; Jain et al., 2021). However, there seems to be a broader consensus on the relationship between the pitch and the whole frequency region of hearing loss (Norena et al., 2002; Roberts et al., 2006; Schecklmann et al., 2012; Jain et al., 2021).

The literature reports plenty of different approaches to carry out the pitch matching, and their performance have been extensively compared with each other (Tyler and Conrad-Armes, 1983; Penner, 1995; Henry et al., 2004; Neff et al., 2019). Some of these methods consist of several steps of choices where the distance in frequency between the presented tones is narrowed step by step, just as in the case of the two-alternative forced-choice method (2AFC; Penner and Bilger, 1992) or the forced-choice double staircase (FCDS; Henry et al., 2013). Other methods, such as the likeness rating (LR; Norena et al., 2002), aim to broaden the tinnitus characterization from a single frequency to a wider spectrum by means of comparisons between the subject's percept and several pure tones of different frequencies. Unlike these approaches, which are usually based on the interaction between audiologist and patient through a series of questions and adjustments, the method of adjustment (MOA) allowed subjects to self-guide the test by using a computer interface or a noise generator and dial (Tyler and Conrad-Armes, 1983; Henry et al., 2004). The MOA involves the constant presentation of a stimulus (typically a pure tone or a narrow-band noise) whose frequency and loudness can be controlled by the subject. The finer adaptability of this method might provide a more accurate representation of the subject's tinnitus. However, the MOA can be difficult to perform for some patients due to a steep slope of their audiogram, which leads to numerous adjustments of the loudness dial (Penner and Bilger, 1992). It is worth mentioning that most pitch matching methods require extra time for the adjustment of the stimulus loudness, despite the fact that pitch-based therapies (as their name suggests) do not usually need loudness data to be implemented.

Due to the above-mentioned reasons, we decided to develop a different pitch matching method and to compare its performance to the MOA. In this paper, we report the reliability of a self-guided multiple-choice method (MCM) for tinnitus pitch matching, and we compare the results to the MOA between sessions. With the MCM, we aim for an easy-to-conduct method, with higher reliability and a user-friendly interface to simplify the procedure.

5.2. Methods

5.2.1. Participants

A total of 20 adult patients of the Otorhinolaryngology Department of the University Medical Center Groningen (UMCG) were recruited to participate in this study between September of 2020 and April of 2021. All of the 20 participants had chronic tinnitus (suffering tinnitus for at least 3 months; Vesterager, 1997) and presented a symmetric hearing loss (≤ 15 dB difference between both ears at 2, 4

and 8 kHz) with an averaged pure-tone audiometry (PTA at the same frequencies) of at least 30 dB. Excluding tinnitus and hearing loss, participants had no history of either neurological or psychiatric disorders. All participants gave written informed consent to join the study, which was approved by the ethics committee of the University Medical Center Groningen (METc 2018/445).

5.2.2. Questionnaires

After giving written informed consent and prior to being invited to the clinic, participants received by mail a series of questionnaires that were sent back to us with a return envelope. These questionnaires were the Tinnitus Functional Index (TFI; Meikle et al., 2012), the Hyperacusis Questionnaire (HQ; Khalfa et al., 2002) and the European School for Interdisciplinary Tinnitus Research Screening Questionnaire (ESIT-SQ; Genitsaridi et al., 2019). The latter was used to gather demographic data and additional tinnitus characteristics.

5.2.3. Method of adjustment

When the process starts, the question “Hoe klinkt uw tinnitus?” (“what does your tinnitus sound like?”) appears on the screen, followed by two clickable answers: “Pieptoon” (“Beep”) for pure tone and “Ruis” (“Noise”) for narrow-band noise with a bandwidth of 1/3 of an octave. After choosing one of the two, the stimulus is presented initially at 1 kHz and 60 dB SPL, while the interface shows the sentence “Verplaats de balk totdat het geluid het meest op uw tinnitus lijkt,” meaning “Move the bar until the sound most resembles your tinnitus” (figure 5.1). The subject then can adjust the central frequency and the loudness of the stimulus by using two sliders. The stimulus, which is continuously presented during the entire test, can also be changed between pure tone and noise at this stage. The subject can finalize this stage by pressing the button “Kies” (“Choose”), by means of which the frequency of the stimulus is stored. Next, an octave confusion test is performed. For this, the selected frequency is tested against two other stimuli that are centered at an octave below and an octave above, with the three of them presented at the same intensity level. Here, the participant has to choose one out of the three options, which is stored as the final frequency and loudness of the pitch matching process. Onset and offset times of the stimuli were 100 ms. The frequency slider (range from 0.05 to 16 kHz) allows minimum changes in linear steps of 18.5 Hz, and the loudness slider (range from 10 to 95 dB SPL) can be adjusted in steps of 0.81 dB.

Figure 5.1 | MOA's interface.

5.2.4. Multiple-choice method

Like MOA, this method starts by asking the participant to choose between noise and pure tone. After the subject chooses one of the two options, the interface shows 22 different clickable buttons that can be activated one by one (see figure 5.2). Each of them then presents a stimulus of 1 s duration and centered at the following frequencies: 0.1, 0.3, 0.5, 0.6, 0.7, 0.8, 1, 1.1, 1.2, 1.35, 1.5, 2, 2.5, 3, 3.5, 4, 5, 6, 7, 9, 10, and 12 kHz. Stimuli are presented at a comfortable level and adjusted to the participant's audiogram, according to the following procedure: the level of each stimulus is adjusted by adding 60 dB of baseline presentation level to the dBs of the nearest frequency available of the audiogram, with a maximum level of 95 dB SPL. Bandwidth of the noise, onset and offset times are identical to MOA.

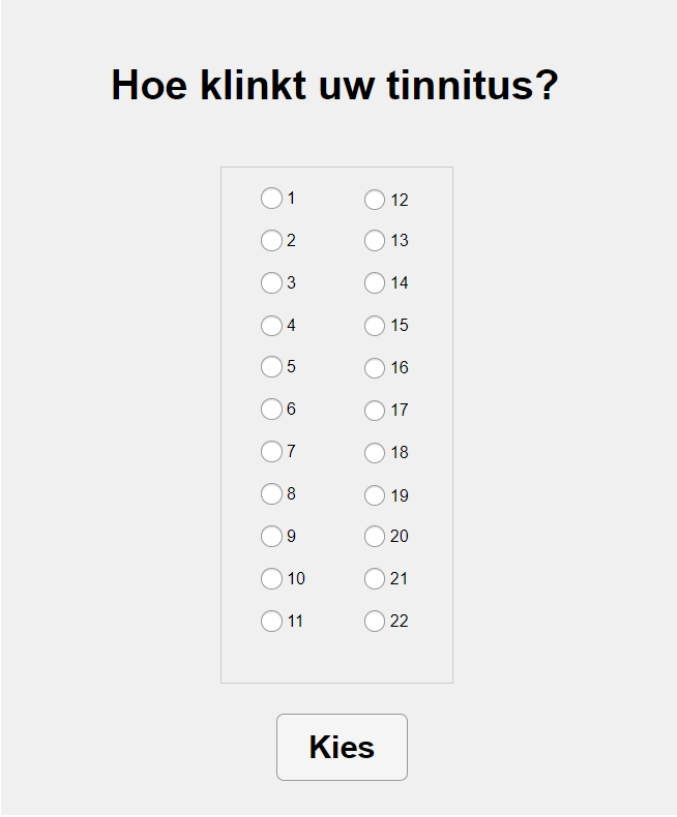


Figure 5.2 | MCM's interface.

5.2.5. Procedure

Figure 5.3 shows the timeline of the experiment. The participants were invited to come to the clinic for two sessions, 1 week apart. Hearing thresholds were measured during the first session with a conventional audiometry at frequencies between 0.125 and 8 kHz in octave steps, as well as 3 and 6 kHz. For this, an audiometer AC40 (Interacoustics) and a pair of TDH39 headphones (Telephonics) were used. All measurements were carried out in sound proof rooms. For the pitch matching procedures, a MOTU UltraLite audio interface and a pair of Sennheiser HD660S headphones were used. All sounds were delivered monoaurally. In case of unilateral tinnitus, sounds were presented in the contralateral ear. For the bilateral cases, sounds were presented in the best hearing ear.

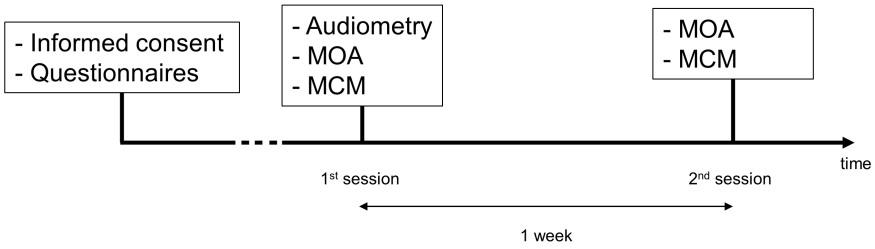


Figure 5.3 | Timeline of the experiment. The measurements are shown in chronological order for each session.

5.2.6. Analysis

Data was analyzed in R version 4.0.2. Sample size was determined based on a power analysis with an expected reliability of 0.9, a minimum acceptable reliability of 0.65 and a significance level of 0.05 (α). Reliability of the two matching methods was estimated by means of several coefficients and measures. Intraclass correlation (ICC) was used to quantify the reliability of each method within and between sessions. ICC was estimated using the package "irr" of R (version 0.84.1). Data were tested for normality by the Shapiro–Wilk Test. Mean frequencies and standard deviations over all participants for both methods between and within sessions were calculated. Moreover, the within-method and within-session differences in octaves were also estimated. Mean loudness and standard deviation of MOA was obtained.

5.3. Results

Demographic characteristics of the participants are presented in table 5.1. Hearing thresholds were assessed by estimating a Pure Tone Average (PTA of 2, 4 and 8 kHz) and did not differ significantly between ears. Averaged values of the TFI and HQ scores are shown.

Figure 5.4 shows the individual pitch-matching results during both sessions and using both MOA and MCM. Normality of the data could not be assumed for MCM. The reliability measures and pitch matching averages of both methods are represented in table 5.2. When comparing the two methods, the intraclass correlation coefficient was higher for MCM compared to MOA. However, the overlap between the two confidence limits of both ICCs indicated that there was no significant difference between both methods. There was no significant difference in the averaged tinnitus pitch between both methods. The mean octave difference (OD) between the two sessions was calculated for both methods, no significant difference was found. The within-method and within-session individual's ODs are shown in figure 5.5.

Demographic data	
Number of subjects (n)	20
Age (years)	62.2 ± 8.5
Sex — n (%)	
Male	15 (75)
Female	5 (25)
Average hearing threshold in both ears (dB HL)	51.3 ± 11.6
PTA (2, 4 and 8 kHz)	
Left ear	52.0 ± 11.5
Right ear	50.7 ± 12.6
TFI score (0-100)	51.4 ± 16.4
HQ score (0-42)	21.5 ± 7.7

Table 5.1 | Demographic characteristics of the participants.Mean values and standard deviation are presented, unless stated otherwise.

Comparison	**ICC (95% CI)**	**Freq (kHz)**	**OD**	**Loudness (dB SPL)**
Between sessions, within-method				
Method Of Adjustment (MOA)	0.77 (0.49 - 0.90)	4.4 ± 2.4	0.53 ± 0.60	78 ± 12
Multiple-Choice Method (MCM)	0.92 (0.81 - 0.97)	4.0 ± 2.8	0.39 ± 0.48	-
Within-session, between methods				
Session 1	0.43 (0.02 - 0.73)	4.4 ± 2.7	0.80 ± 0.97	
Session 2	0.62 (0.25 - 0.83)	4.1 ± 2.5	0.62 ± 0.55	

Table 5.2 | Averaged pitch matching results and reliability measures.ICC = intraclass correlation; OD = difference in octaves. Mean values and standard deviations are presented, unless stated otherwise.

5.4. Discussion

In this study, we compared two self-guided methods to measure the tinnitus pitch in 20 participants with chronic tinnitus. The participants used the both methods MOA and MCM to measure their tinnitus pitch in two sessions, 1 week apart. The comparison was made by means of reliability, mean frequencies and octave difference between and within sessions.

Both methods presented very good reliability. However, due to the relatively large confidence intervals of the ICC, it is not possible to determine which one of the methods is more reliable. Nonetheless, MCM presented an ICC ≥ 0.9, which is considered the required standard of a tool used for clinical decision making for individual patient data (Kottner et al., 2011). MOA presented an ICC ≥ 0.7, indicating good agreement between measures for group data (Nunnally and Bernstein, 1994). In terms of octave difference, no significant difference was found between both methods as a result of the spread of the data.

A previous study compared three different pitch matching methods including the MOA, for which they obtained analogous results of reliability (Neff et al., 2019). In this study, the authors mentioned the potential bias of the participants in the decision-making due to the initial presentation of the stimulus of this method (1 kHz in our case), which explains differences in pitch between methods.

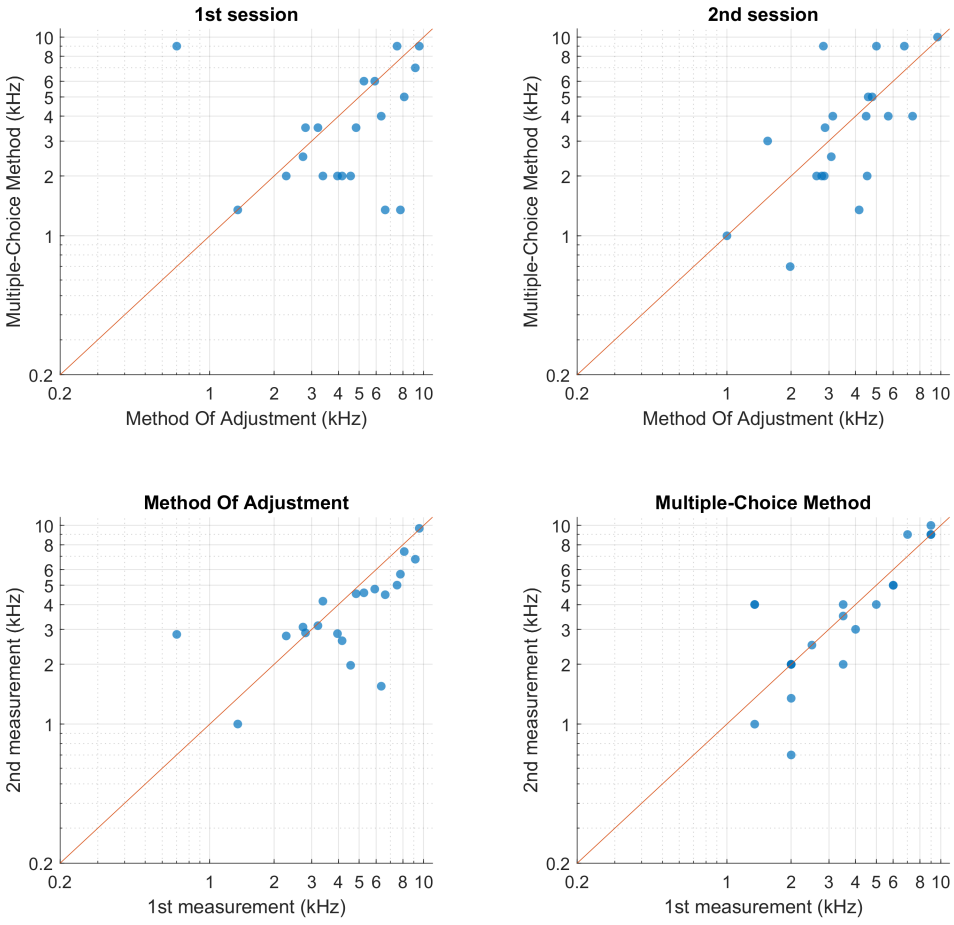


Figure 5.4 | Participant's pitch-matching results for both methods and both sessions, each data point represents one participant. (Upper-left corner) Comparison between bothmethods within the 1st session. (Upper-right corner) Comparison between both methods within the second session. (Bottom-left corner) Comparison within MOA between both sessions.(Upper-right corner) Comparison within MCM between both sessions.

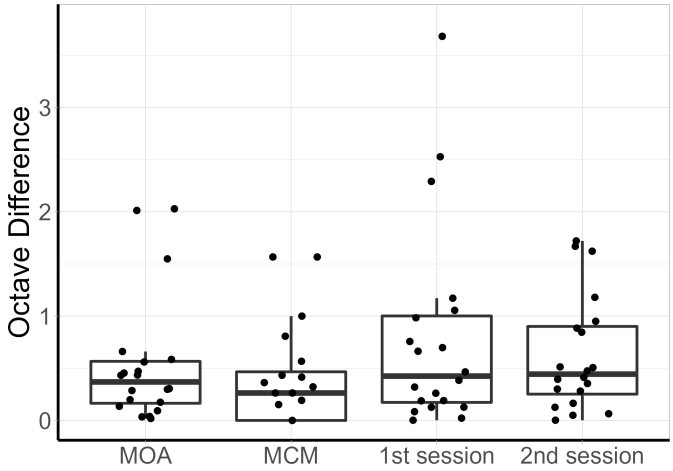


Figure 5.5 | Box plots of the difference in octaves, from left to right: MOA within-method, MCM within-method, 1st session between methods, 2nd session between methods. For each boxplot, the data points represent individual participants.

However, the authors did not report the standard deviation of the frequency selection. In our case, despite no significant difference in frequency was found between methods, we also suspect that the starting frequency can play a role in the procedure. This potential bias is avoided in the MCM, which is not initialized with any stimulus. However, some subjects have the tendency of starting the matching from the first option, which corresponds to the lowest frequency available. Future implementations could prevent this by removing the numbers from the buttons and keeping the same sequence of frequencies.

Pitch-dependent sound-based therapies such as the tailor-made notch noise training (Stracke et al., 2010), the notch filter amplification (Marcrum et al., 2021) or the harmonic sound therapy (Mahboubi et al., 2012) are based on narrow-band approaches which commonly use a bandwidth of half or a third of an octave. Consequently, frequency resolution might not be the most important characteristic of a pitch matching procedure. Instead, a self-guided method that allows the subject to choose the closest available option without having to constantly adapt the volume of the stimulus (as in the case of MOA), might be a practical solution for a clinical environment. Another advantage that the MCM presents is the automatic adaptation of the loudness of the stimuli to the hearing profile of the subject. In the case of MOA, patients with high-frequency tinnitus often have trouble adjusting the loudness of the stimulus due to the abrupt decrease of their audiogram, which could be solved by using loudness correction. The MCM addresses this issue so the subject can focus only on the frequency of the sound. Future implementations of the method could adjust the intensity in a more cautious way for high frequencies, following a half-gain rule as in hearing aids fitting (Lybarger, 1963). Moreover, the MCM can be implemented on mobile devices such as smartphones or tablets,

which have the potential to be used for hearing diagnosis after the corresponding validation (Wunderlich et al., 2015; Hauptmann et al., 2016; Swanepoel et al., 2019).

Unfortunately, the test duration was not recorded. This limitation prevents us from claiming that one of the methods has significantly lower duration than the other one. Nevertheless, by observing the participants during the experiment, we noted shorter durations during pitch matching with the MCM than with the MOA. Additionally, it's worth mentioning the fact that the order of test procedure was not randomized, which could potentially result in a learning effect when performing the second test. Another limitation that both methods had during the experiment is the constraint of 95 dB HL as the maximum level of presentation of the stimulus as a safety measure. A subject whose tinnitus' loudness is above that level is likely to choose the closest audible frequency during the matching procedure. For presentations within the extended high frequency range in the MCM, a similar problem can be seen: since the stimuli were adjusted to the audiogram, and this was measured up to 8 kHz, presentations for extended high frequencies will not be perceived equally loud by participants with high frequency hearing loss. An extended high frequency audiometry could mitigate this issue. Previous comparisons between pitch matching methods used repeated measurements in one single session, which might not be a sufficient time interval to reveal changes in cases of fluctuating tinnitus (Neff et al., 2019). Instead of several measurements in one session, we opted for measuring in 2 different sessions, 1 week apart. The fact that the obtained within-methods ICC values and OD values are higher and lower, respectively, than the between-method ICC and OD, suggests that each method is consistently measuring a different aspect of tinnitus. However, this aspect or feature differs between both methods, hence the higher between-methods OD and lower between-methods ICC. Based on these results, we cannot conclude whether the differences between the two sessions are a result of changes in the tinnitus or an overall difficulty that subjects may have to match an external stimulus to their tinnitus. In addition to this, it is noteworthy the difference in step sizes between both methods, which can affect the reliability results.

To conclude, our results suggest that a multiple-choice method (MCM) for tinnitus pitch matching is as reliable in a clinical population as more conventional methods such as the method of adjustment (MOA). This self-guided approach can be easily implemented on mobile devices. Due to the limited number of response options and the only requirement of having to include the subject's hearing threshold in advance, the MCM has the potential to speed-up the matching process.

6

Hearing Aid Amplification Schemes Adjusted to Tinnitus Pitch: a Randomized Controlled Trial

Abstract

Hearing aids can be used as a treatment for tinnitus. There are indications that this treatment is most effective when the tinnitus pitch falls in the frequency range of amplification of the hearing aid. Then, the hearing aid provides masking of the tinnitus. Alternatively, it has been suggested that a gap in the amplification around the tinnitus pitch would engage lateral inhibition and thereby reduce the tinnitus. To test these ideas, we conducted a randomized controlled trial. Patients were fitted with hearing aids using 3 different amplification schemes: (1) standard amplification according to the NAL-NL2 prescription procedure, (2) boosted amplification at the tinnitus frequency to enhance tinnitus masking, and (3) notch-filtered amplification at the tinnitus frequency to engage lateral inhibition and suppress tinnitus. The goal was to compare the boosted and notched amplification schemes to standard amplification. The primary outcome measure was tinnitus handicap as measured by the Tinnitus Functional Index (TFI). The trial was designed as a double-blind Latin square balanced crossover study. Eighteen tinnitus patients with moderate hearing loss were included. All of them were experienced hearing aids users. After two weeks of initial adaptation to the standard setting, each setting was tried for four weeks. There was a small reduction of the TFI score after the adaptation process, possibly due to a placebo effect. The TFI score did not differ significantly from the standard setting after using the notched or the boosted settings. However, notched amplification was significantly better than boosted amplification. Regardless of the TFI outcomes, most participants had an individual preference for a particular setting. In conclusion, notch-filtered and boosted amplification did not provide better tinnitus suppression than standard amplification, although notched amplification performed better than boosted amplification. The individual preferences highlighted the importance of tailor-made approaches to hearing aid amplification in clinical practice. Further studies should explore the differences among patient's tinnitus and their preference for a hearing aid setting.

6.1. Introduction

Tinnitus is described as 'ringing' or 'hissing' in the ears or the head in the absence of any external sound. About 15% of European adults experience tinnitus; however, as hearing status deteriorates with age, this proportion is larger in older adults (Biswas et al., 2022). The most common comorbidity of tinnitus is hearing loss, which has an odds ratio of 8.5 for developing tinnitus in the case of the Dutch population (Schubert et al., 2021).

Neurophysiological studies suggest that tinnitus can be based on maladaptive adaptation to hearing loss (Eggermont and Roberts, 2004; Knipper et al., 2013**?**; Koops et al., 2020). Hence, some of the treatments aim to reverse the changes produced by such maladaptive adaptation. Hearing aids and other sound therapies for tinnitus are examples of this. These treatments may revert the abnormal brain activity that could originate from acoustic deprivation, by recalibrating central gain (Schaette and Kempter, 2006), by preventing maladaptive neuroplasticity (Noreña, 2011), or by restoring connectivity in the auditory pathway (Boyen et al., 2014; Hofmeier et al., 2018).

A particular application of sound therapy is the Tailor-Made Notched Music Training (TMNMT), which aims at reducing the spontaneous activity in neurons by enhancing lateral inhibition from the frequencies above and below the tinnitus frequency (Pantev et al., 2012; Pape et al., 2014). These authors suggested that initialization, development and manifestation of tinnitus is due to the loss of lateral inhibition affecting the functioning of specific populations of auditory neurons. By creating a spectral gap at the frequency that matches the tinnitus pitch (referred to as the tinnitus frequency) lateral inhibition could be engaged and potentially ameliorate tinnitus. An application of this idea is the notched amplification in some hearing aids (Marcrum et al., 2021), which combines both the amplification of regular hearing aids and a notch filter centered at the user's tinnitus frequency.

Hearing aids mainly increase the volume of external sounds, which improves the communication of users and, consequently, can help to reduce other tinnitus symptoms like stress or anxiety. But they also provide distraction from tinnitus and, more importantly, masking (Sereda et al., 2015). Masking might be the main cause for the reduction of tinnitus symptoms. The tinnitus literature suggests that the tinnitus frequency might have an influence on the result of sound-based therapies. For example, previous studies suggested that the perceived pitch usually corresponds to frequencies where hearing is impaired (König et al., 2006); Norena et al., 2002; Roberts et al., 2008). It has been shown that masking is more likely to be achieved when the tinnitus pitch falls into the frequency range of the hearing aids (Joergensen et al., 2022; McNeill et al., 2012; Schaette et al., 2010). In this context, it's reasonable to think of an opposite approach to the notch therapy discussed above: by boosting the amplification of the frequencies close to the tinnitus percept, masking might be enhanced and, as a consequence, bothersome tinnitus diminished.

The existing literature highlights the lack of high quality evidence to support the clinical efficacy and effectiveness of hearing aids for tinnitus (Hoare et al.,

2014). Shekhawat et al. (2013) pointed out the scarce evidence for the effects of hearing aids in tinnitus management, underlining the low quality of the studies in the form of non-randomized controlled trials (non-RCTs). A similar conclusion has been drawn in a recent scoping review on hearing aids and tinnitus (Jacquemin et al., 2022) where, despite the positive results of most of the included studies, the authors emphasized the variability of the quality among these studies. Well-designed randomized controlled trials are necessary in tinnitus research to provide the highest grade of evidence quality for treatment efficacy, as it is described in clinical guidelines (Tunkel et al., 2014).

Due to all the aforementioned reasons, the aim of this study was to conduct a randomized controlled trial (RCT) to test the efficacy of different amplification schemes in hearing aids, adjusted to the characteristics of each patient's tinnitus frequency. Two settings were tested: notch amplification and boost amplification, and both were compared to a third setting, the standard amplification, serving as a control. For this, 18 patients with tinnitus used hearing aids fitted with the three different programs during a total period of 14 weeks, switching the program every 4 weeks of use, in a randomized and double blind procedure. Hearing aid fitting was ensured by real ear measurements (REM), and assessment of tinnitus and hearing aid benefit was carried out by means of questionnaires.

6.2. Methods

6.2.1. Participants

The study was conducted at the ENT department of the University Medical Center Groningen (UMCG), in the Netherlands. From 19 subjects who were enrolled, 18 completed the study. Two ways of recruiting participants were used. Most subjects were invited by an audiologist to participate after checking in their medical record that they met the inclusion criteria. A few subjects joined after replying to the study advertisement at the UMCG website. All the participants had a moderate to moderate-severe degree of hearing loss, with the following characteristics: average pure tone audiometry (PTA) at 1, 2 and 4 kHz of at least 35 dB HL; no more than 50 dB HL at 1 kHz and symmetric hearing loss, defined by no more than 15 dB HL of difference between both ears at each of the above frequencies. In addition, all participants had chronic tinnitus (suffering from tinnitus for at least 6 months). The tinnitus of the participants had to be categorized as tonal, meaning that they could identify their tinnitus frequency during a pitch matching task (see below). Furthermore, all participants were experienced hearing aids users, using their devices for at least 6 months before joining the study. All participants gave written informed consent before joining the study, which was approved by the ethics committee of the UMCG (METc 2021/128).

6.2.2. Questionnaires

After giving informed consent, participants received a series of questionnaires by mail that were sent back to us with a return envelope. These questionnaires were the Tinnitus Functional Index (TFI; Meikle et al., 2012), the Abbreviated Profile of Hearing Aid Benefit (APHAB; Cox and Alexander, 1995) and the Hyperacusis Questionnaire (HQ; Khalfa et al., 2002). This initial round of questionnaires served as creating a baseline. In order to monitor the participant's progress during the trial, they had to fill these three questionnaires repeatedly throughout the entire study (see section 2.5). The TFI was used to assess the effectiveness of the different treatments and their impact on the patients' tinnitus. The purpose of the APHAB was to evaluate whether the different hearing aids settings had an impact on the patient's hearing, independently of their tinnitus. Finally, the HQ was used to measure whether the settings had an impact on hyperacusis.

6.2.3. Hearing aids and amplification schemes

Participants used the receiver-in-canal (RIC, size M) devices Signia Pure 312 Nx 7 (Sivantos, Erlangen, Germany) with open click molds (vented) during the entire study. These hearing aids are specified (by the manufacturer) to be used in subjects with moderate to severe hearing loss. The devices had 20 frequency bands ranging from 0 to 12 kHz that could be adjusted individually through the fitting software Connexx (Sivantos, Erlangen, Germany). Participants could only increase or decrease the general gain of their devices if necessary using their devices' buttons or an app on their phones, but no changes in the frequency response was available to them. All the participants committed to not use their own devices until completion of the study. After enrollment, a document containing the instructions of use of the devices was given to them.

Hearing aids were fitted according to three different amplification techniques:

- **Standard amplification.** In this setting, all available frequencies are amplified according to the standard clinical fitting using the formula NAL-NL2 (Keidser et al. [2012]). This approach is consistent with the recommendation for patients with tinnitus and hearing loss in the ENT clinic of the UMCG.
- **Notch amplification.** This setting uses the previous fitting formula and, additionally, applies a notch filter centered at the participant's tinnitus frequency. The rationale of this approach is based on lateral inhibition, by which excitation of neurons with characteristic frequency close to the tinnitus frequency might suppress the tinnitus. The notch filter has a depth of 60 dB and a bandwidth of 0.5 octaves. The filter can be placed at each of 31 logarithmically distributed frequencies in the range of 0.250 to 8 kHz.
- **Boosted amplification.** The same fitting formula is applied. Additionally, the closest frequency band to the participant's tinnitus frequency is amplified with 5 dB extra.

All three settings were stored in each participant's profile in the software, ready to be loaded in the devices before each of the appointments.

6.2.4. Audiological assessment and hearing aid fitting

During the participant's first visit to the clinic, pure tone audiometry from 0.250 to 8 kHz was performed. If they met the inclusion criteria, tinnitus pitch matching was carried out. For this, a multiple-choice method was used (Santacruz et al., 2023). This self-guided method allows the participant to choose between 22 available frequencies (ranging from 0.1 to 12 kHz), the one that matches their tinnitus pitch as closely as possible. The center frequency or this matching stimulus is referred to as the tinnitus frequency. Stimuli were presented in the contralateral ear in case of unilateral tinnitus, and in the best hearing ear in case of bilateral tinnitus. The patient could choose between two main categories of stimuli: pure tones and noise. In the case of the latter, narrow-band noise with a bandwidth of 1/3 of an octave was used. Due to the limitation of 8 kHz as the maximum frequency at which the notch can be set in the study devices, participants with a higher tinnitus frequency were excluded.

With the obtained hearing thresholds and tinnitus frequency, a pair of hearing aids were fitted with the 3 settings according to section 2.3. In order to ensure the correct fitting of the different settings, both objective and subjective measures are an important common practice in the clinic (Sereda et al., 2015). For this reason, some small adjustments of the NAL-NL2 formula were made until each participant was comfortable with the gain setting. All three settings were verified by means of real ear measurements (REM) using an Affinity 2.0 (Interacoustics, Denmark) and in-ear probes. After measuring the open ear response with white noise at 65 dB SPL, without removing the probes, hearing aids were placed in the participants' ears. Then, the insertion gain of the three settings were measured by using the free-field International Speech Test Signal (ISTS; Holube et al., 2010) at 55, 65 and 75 dB SPL. Figure 6.1 shows the insertion gain measure of the three settings from a participant whose tinnitus was matched at 3.5 kHz. The curves shown were obtained using 65 dB SPL as the loudness of the ISTS stimulus. To confirm that none of the settings affected the speech understanding of the participants, free-field speech audiometry at 60 and 70 dB SPL was measured.

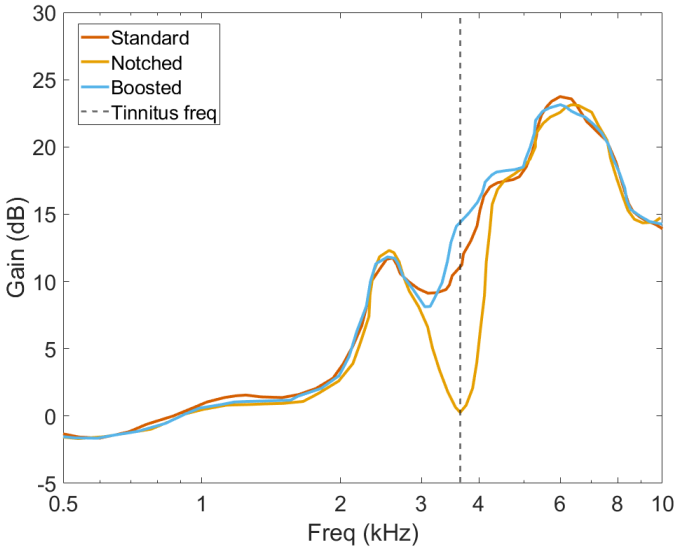


Figure 6.1 | REM of the three settings of a participant with tinnitus at 3.5 kHz, which is shown as a dotted line. The curves represent the insertion gain responses of the different settings in the ear canal at 65 dB SPL using the ISTS stimulus.

6.2.5. Study design and blinding

In order to maximize the power of the statistical tests considering the sample size, the study was designed as a crossover trial. An additional advantage of this design is the capacity of measuring the response to a specific treatment in comparison to the same participant's response to the other treatments (Sibbald and Roberts, 1998). In this sense, each participant becomes their own control subject, which is particularly interesting given the heterogeneity of treatment response in tinnitus. In long clinical trials such this, washout periods can increase the dropout rates (Amirah Fatin et al., 2016). To avoid this and also potential carryover effects, characteristic for crossover studies, we opted for a *balanced latin square* design (Edwards, 1951). Table 6.1 shows the treatment order of each group of three participants. The design contains all the possible combinations of treatment order in a balanced way. The sample size for this study was calculated following the indications for TFI repeated measures of Fackrell et al. (2022), stablishing a Standard Deviation of the difference of 7.2 points in the TFI, and $\alpha = \beta = 0.05$.

The study was performed in a randomized controlled and double-blind way. During the entire trial, none of the participants, nor the researchers involved in the study had any knowledge on which setting was programmed in the hearing aids given to each participant at each appointment. A researcher not involved in the study designed a table with the participants' identification number and their treatment orders, as in table 6.1. This document was printed and kept in an envelope. Prior to each participant's appointment, 3 pairs of identical hearing aids

were programmed according to their individual fitting, each of them with one of the settings (either standard, notched or boosted). When a participant came to the clinic for either starting the trial or switching between programs, the envelope with the appointments order, the patient's id and the three pairs of devices in separate boxes were given to a person not involved in the study. The boxes were labeled to reflect the three settings. The hearing aids themselves were not labeled. The person, in a separate room, would take all the devices out of their boxes and only give back to the researcher the corresponding pair of devices to be handed to the participant. The other two pairs were reset. Thus, when giving the selected pair to the participant, neither the researcher nor the participant would be aware of the setting in the device. The study remained blinded until all participants completed the trial.

Subjects	**Treatment order**		
3	Standard	Notched	Boosted
3	Notched	Boosted	Standard
3	Boosted	Standard	Notched
3	Standard	Boosted	Notched
3	Notched	Standard	Boosted
3	Boosted	Notched	Standard

Table 6.1 | Balanced latin square design, crossover study. Each row represents a group of 3 participants. Each group used the settings in the stated order.

6.2.6. Procedure

Participants were enrolled for 14 weeks as it is shown in Figure 2. During this period, participants visited the clinic 5 or 6 times until completion of the study. The figure shows all the procedures that took place during each of the visits. In the two first weeks, the participants got the new devices with a regular NAL-NL2 fitting formula (standard amplification) for adaptation. After 2 weeks, the participants would start the randomized trial itself. If the subjects were not satisfied with the gain setting, they could come for an extra appointment prior to the start of the trial (optional appointment in figure 6.2), where small adjustments of individual frequency bands could be made until loudness was pleasant.

Once the trial started (after 2 weeks of adaptation), participants would visit the clinic once every 4 weeks for getting a new pair of identical looking devices, fitted with the corresponding setting. During these visits, the procedure was the same: filling in the three questionnaires, informal discussion on their hearing and tinnitus status, handing in the devices with the previous setting, getting the new ones and undergoing a speech audiometry. During the informal discussions, participants were asked whether they preferred the current setting compared to the previous one. Information about usage time of the previous devices as logged by the devices would be stored, while ensuring the blinding. At the end of the trial, each participant ranked the settings in order of preference, ranging from worst to best in relation to their tinnitus status.

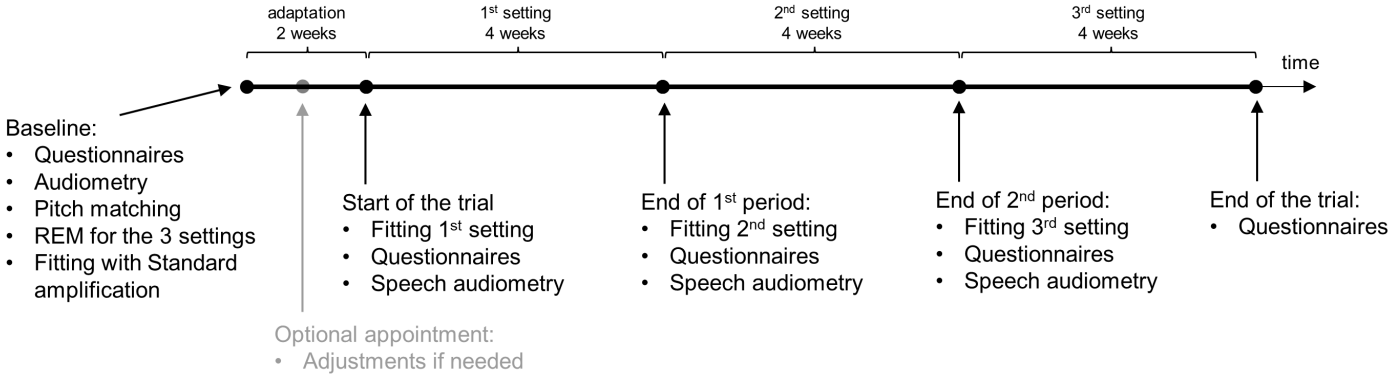


Figure 6.2 | Study timeline. Each dot in the line represents an appointment. Performed tasks are listed underneath each appointment.

At the end of the trial, participants could keep the devices if they wished or go back to using their own hearing aids.

6.2.7. Statistical analysis

Statistical analyses of the data were performed using R (version 4.0.2). In order to avoid the bias of a potential placebo effect produced by joining a tinnitus study and using new hearing aids, the main analysis was carried out to measure the changes of the TFI score between the Standard setting and the other two settings. Between-treatment comparisons were calculated using linear mixed effects models. For this, all predictors and interactions between these were tested on an iterative process of model comparisons. Fixed effects were tested by using the expected mean square approach, which involves comparing the variance associated with the different predictors to the residual variance through the F-statistic, and then determining their statistical significance. As to the random effects, we set up a model with random intercept for the participant level. Conditional and marginal R2 were used to measure the proportion of explained variance (Nakagawa et al., 2017). Post-hoc Tukey tests were used for multiple comparisons. The assumptions of linear mixed effects models were verified (linearity, normality of residuals, homogeneity of variance and no outliers).

The main model was used for measuring tinnitus status and had the TFI score as the response variable. The following covariates (predictors) were tested for this model: hearing thresholds, speech audiograms (averaged free-field score at 60 and 70 dB SPL), tinnitus frequency and type (tonal or noise-like), age, hyperacusis, average hours of use of the devices and order of setting. Hearing thresholds were characterized as the average PTA for 2, 4 and 8 kHz. A similar analysis was conducted for both the HQ score and the APHAB score.

A post-hoc analysis using linear mixed models was conducted to measure impact on TFI when having the preferred setting as the main predictor, instead of each individual setting. The rest of the predictors tested remained the same as in the previous analysis. For this, the same procedure and variables were used.

6.3. Results

6.3.1. Demographics

Demographic variables and the results of the questionnaires are shown in table 6.2. Participants were on average around 60 years old, two thirds of them were male. On average, the TFI scores at baseline indicated that tinnitus was a moderate problem for the participants (Henry et al., 2016). For hyperacusis, 6 of the participants scored above 22 in the HQ questionnaire when starting the trial, which is considered to indicate hyperacusis (Aazh and Moore, 2017). The APHAB questionnaire scores showed an average benefit above 25, which confirmed the benefit of the hearing aids fitting (Cox, 1997). One participant dropped out of the study during the adaptation period since he/she found the standard hearing

aid setting (implementing NAL-NL2) unpleasant. A new participant joined the study as a replacement. Nine out of the 18 participants made use of the optional appointment to adjust the gain settings during the adaptation period.

6.3.2. TFI during the adaptation period

The results of the TFI for the baseline, adaptation and standard setting can be seen with detail in figure 6.3. It shows a large spread in TFI scores and in some cases also large changes. In addition, table 6.3 shows the averaged reductions of the TFI scores between these stages of the trial. The differences of the TFI score during this period is summarized in table 6.4, which shows the results of Welch's t-test comparisons between baseline, adaptation and the standard setting for the three questionnaires, with the corresponding Bonferroni corrections (Bland and Altman, 1995). Note that the standard setting was not necessarily used right after the adaptation period for each participant, since the order of settings differs between participants. During the adaptation period of two weeks, every participant was using the standard setting. Although the mean score decreased, the t-test revealed no significant differences between baseline, adaptation and the standard setting for the TFI results. The rest of the questionnaires also did not differ significantly between these periods. Two of the questionnaire forms at baseline went missing, which is the reason for the two missing data points at baseline in figure 6.3.

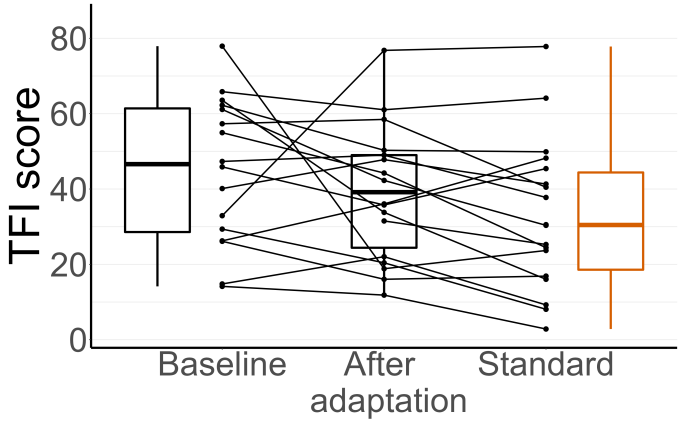


Figure 6.3 | TFI scores of participants at baseline, adaptation stage and after using the standard setting. After hearing aids were fitted at baseline, a two-week adaptation period started. Baseline and adaptation were sequential stages in the timeline of the trial, while standard setting occurred at different stages for each participant. Each participant is represented by dots that are connected. Median values are represented inside of the boxplots.

6.3.3. TFI's results after using each setting

All the assumptions for linear effects mixed models were met. Levene's test of homogeneity of variance was not significant (p value = 0.495). Residuals were

Demographic data	
Number of subjects (n)	18
Sex – n	
male	12
female	6
Age (years)	60.7 ± 12.7
Speech audiometry with hearing aids (% correct)	
at 60 dB SPL	87.7 ± 7.9
at 70 dB SPL	95.5 ± 5.5
PTA at 2, 4 and 8 kHz (dB HL)	57.3 ± 9
Questionnaires	
TFI score	**46.1 ± 13.2**
Intrusiveness	44.9 ± 19.4
Sense of control	52.7 ± 24.5
Cognition	38.3 ± 25.3
Sleep	40.4 ± 31.4
Auditory	56.2 ± 27.2
Relaxation	58.3 ± 28.3
Quality of life	40.7 ± 21.5
Emotional distress	28.3 ± 22.3
HQ score	**19.3 ± 6.9**
APHAB score*	**27.1 ± 12.7**
Ease of communication	30.3. ± 16.1
Reverberation	23.3. ± 14.6
Background noise	27.1. ± 14.7
Aversiveness	-11.9. ± 16.6
Tinnitus	
Frequency (kHz)	5.08 ± 3.0
Type – n (pure tone / noise-like)	6/12
Hours of use per day	
Standard setting	10.5 ± 4.5
Notched setting	8.5 ± 5.5
Boosted setting	9.8 ± 4.9

Table 6.2 | Demographic characteristics and results of the questionnaires at the beginning of the trial. Each questionnaire is accompanied by their own subscales. Mean values and standard deviation are presented. *The APHAB score reflects the benefit of the hearing aids that participants used before entering the trial.

Comparison	**Mean reduction of the TFI score (points)**
Baseline – Adaptation	6.9 ± 2.0
Baseline – Standard	12.0 ± 0.4
Adaptation – Standard	5.1 ± 2.4

Table 6.3 | Mean reductions of the TFI score between stages of baseline and adaptation. Standard setting is included as a reference for the settings tested.

Parameter	Comparison	df	p value	Adjusted p value
TFI score	Baseline – Adaptation	30.4	0.28	0.84
	Baseline – Standard	31.6	0.08	0.24
	Adaptation – Standard	33.4	0.41	1
HQ score	Baseline – Adaptation	31.5	0.89	1
	Baseline – Standard	32	0.83	1
	Adaptation – Standard	33.7	0.92	1
APHAB score	Baseline – Adaptation	44.6	0.52	1
	Baseline – Standard	25.4	0.31	1
	Adaptation – Standard	26.5	0.11	1

Table 6.4 | Welch's t-test comparisons of the questionnaire's scores at baseline, during the adaptation period and after using the standard setting. Adjusted p-values for Bonferroni correction are shown.

normally distributed and no outliers were observed. Figure 6.4 shows the pairwise comparisons between TFI scores of the three settings. The results of the fixed effects of the mixed model for the TFI score are shown in table 6.5. This model included standard as a reference, and considered the effects of all the predictors mentioned in section 6.2.7. No significant differences in the change of the TFI were found for the settings notched or boosted, therefore, the test failed in the rejection of the null hypothesis. While tinnitus frequency contributed significantly to the changes of the TFI, this was not the case for the type of tinnitus (tonal or noise-like) or the interaction between tinnitus frequency and the type of tinnitus. The tinnitus frequency contributed negatively to the TFI increase, meaning that participants with lower tinnitus pitch achieved poorer results. The HQ score also had a significant effect on the result of the TFI. In this case, participants with lower HQ scores performed better than those with higher HQ scores. For multiple comparisons, post-hoc Tukey contrasts on the settings were checked, and the results are shown in table 6.6. The notched setting reduced (non-significantly) the TFI compared to the standard setting, and the boosted setting increased it (again, non-significantly). A significant difference between settings notched and boosted was found, with the notched setting giving more TFI reduction. However, this difference didn't reach the clinical significance level of 14 points for the TFI (Fackrell et al., 2022).

Comparison	Estimate	Std. error	Adjusted p value
Notched – Standard	-3.2	3	0.516
Boosted – Standard	3.7	3	0.414
Boosted – Notched	7	3	**0.047**

Table 6.6 | Post-hoc Tukey contrasts of the settings for the TFI model. Estimate=mean change of the TFI score; Std. error=standard error of the TFI score change; p value=significance level.

predictor	numDF	Estimate	Std. error	p value
setting - Notched	2	-3.2	9.8	0.28
setting - Boosted	2	3.7	3	0.21
tinnitus type - noise-like	1	0.4	8	0.92
tinnitus frequency	1	-21	9.9	**<0.05**
tinnitus frequency * tinnitus type	1	20.8	10.5	0.06
HQ score	1	1.2	0.3	**<0.001**
			R2 conditional	**R2 marginal**
Model performance			0.78	0.39

Table 6.5 | Results from a mixed effects analysis of the effect of various predictors on the TFI. Only fixed effects of the predictors are shown. numDF=number of degrees of freedom; Estimate=mean change of the TFI score; Std. error=standard error of the TFI score change; p value=significance level. Goodness-of-fit is represented by: R2 conditional=proportion of variance in the TFI score that is explained by both the fixed and random effects; R2 marginal= proportion of variance in the TFI score that is explained by only the fixed effects.

6.4. Preferred setting

There was no consistent preference of setting at the end of the trial. Thirteen participants expressed a preference for one the three settings, where five participants did not have a preference. Figure 6.6 shows the hearing thresholds of the participants when these are grouped by preferred setting. Although some differences in thresholds in the high frequency range can be seen, no significant difference of PTAs between the four groups was found.

Figure 6.5 shows the results of the questionnaires at baseline arranged by preferred setting at the end of the trial. The participants who had no preferred setting scored lower on average in the APHAB questionnaire at the beginning of the study (a lower score meaning that they experienced less benefit of their hearing aids). Participants who preferred the settings standard or notched had a higher TFI score when starting the trial. Given the low number of subjects in each group, no further statistical analysis was conducted.

A Chi-squared test between the preferred settings and the settings that obtained the lowest TFI score for each participant resulted in a Chi-square statistic of 11.0 (p-value = 0.09). Nine out of the 18 participants preferred the setting that also reduced their TFI score the most. Figure 6.7 shows the TFI scores for the different preferences in order of those participants who did have a preference, being 1 the most preferred setting and 3 the least preferred setting. A pairwise comparison

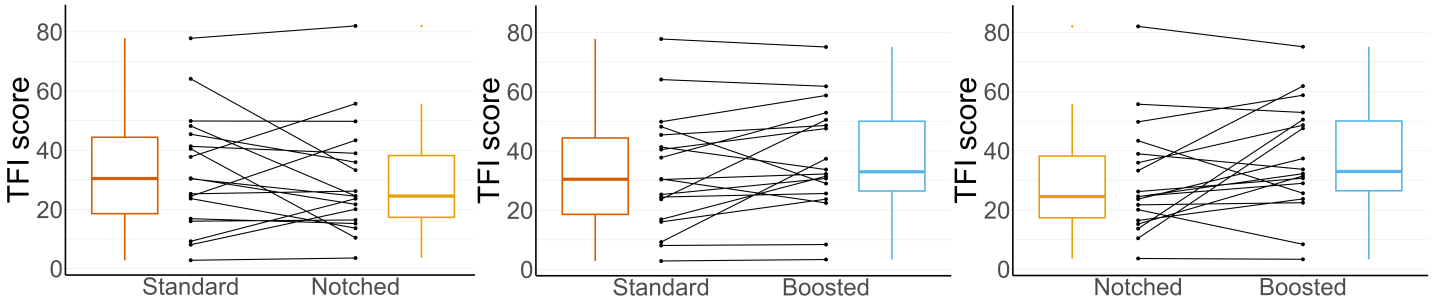


Figure 6.4 | Line series graphs of the pair-wise TFI comparisons between the three settings. Each participant is represented as two dots connected with a line between the settings. Boxplots with median values are represented at both sides of each graph. Only the difference between notched and boosted (right panel) is significant.

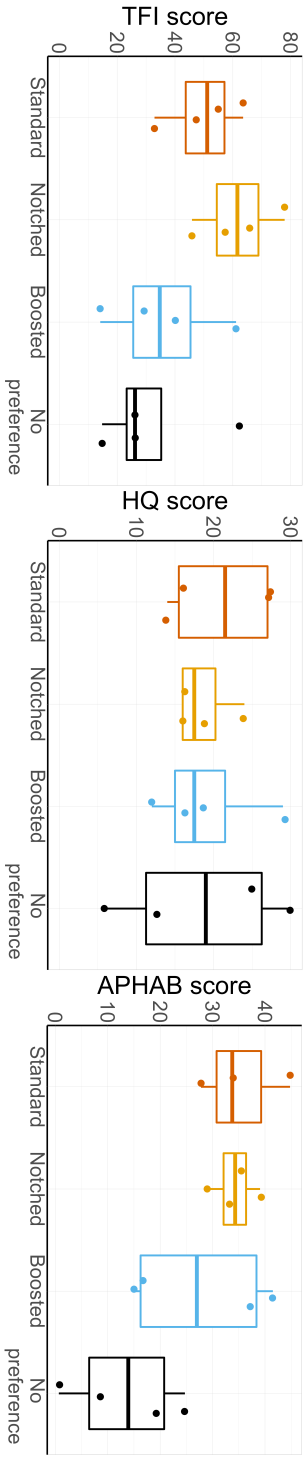


Figure 6.5 | Results of the questionnaires at baseline when arranging the participants by preferred setting.

resulted in no significant difference in TFI score between the preferences.

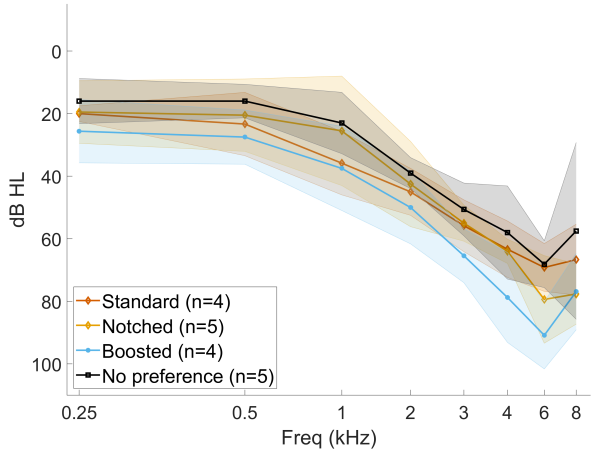


Figure 6.6 | Hearing thresholds of participants when arranging them by preferred setting. Each group size is expressed by n. Results are represented as mean values ± standard deviations as shaded contours.

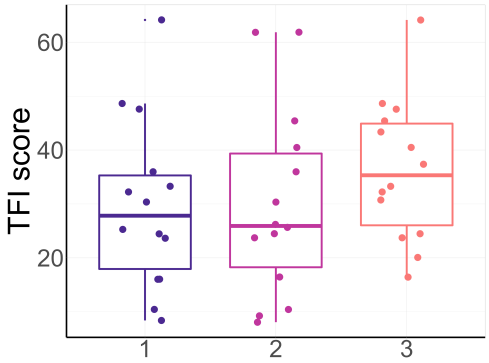


Figure 6.7 | TFI scores of all subjects after a certain setting; ordered by preferred setting. The numbers at the x axis refer to the order of preference, ranging from 1 (most preferred setting) to 3 (least preferred setting).

6.5. Discussion

In this study, we investigated the efficacy of two amplification strategies in tinnitus treatment and compared them to a regular amplification scheme. The settings of the tested strategies were adjusted to each patient's tinnitus frequency. The amplification schemes were fitted using the standard NAL-NL2 formula as a reference. For the control setting (standard amplification), no modifications were

made. For the notched setting, a notch filter was applied at the center of the participant's tinnitus frequency. For the boosted setting, a boost of 5 dB was applied at the tinnitus frequency. Overall, the TFI score was reduced after 2 weeks of adaptation, although not significantly, and remained reduced after using any of the three settings. The comparison between settings revealed no significant reduction of the TFI score in comparison to the standard setting. Despite the TFI score after notched-amplification was significantly better than that after Boost amplification, this difference did not reach the required clinical significance for the TFI. The HQ and APHAB questionnaires did not reveal any significant changes in hyperacusis or hearing aid benefit after using any of the tested settings.

The initial improvement in TFI scores during the initial 2-week adaptation period (figure 6.3) may be due to a placebo effect. The placebo effect has been explored in the tinnitus literature (Hoare et al., 2014) and, more generally, in the context of hearing aid studies (Dawes et al., 2013). Dawes et al. (2013) highlighted the importance of double-blind studies to reduce the placebo effect in hearing aids studies, which this study succeeded in. Nevertheless, we cannot exclude the possibility that the effects in the adaptation period were due to higher quality of the devices, compared to the previous ones, which might have affected the way the participants perceived their tinnitus during the adaptation process.

Sound-based therapies for tinnitus have shown some promising results. These therapies apply either a sound that includes the tinnitus frequency (McNeill et al., 2012; Schaette et al., 2010), or they place a notch filter at the tinnitus frequency thus presenting sounds that surround the tinnitus in frequency domain (Pantev et al., 2012, 2004; Pape et al., 2014). In the first approach, masking of the tinnitus is the potential mechanism of tinnitus suppression. This approach is similar to the standard and boosted amplification schemes in the current hearing aid trial. In the second approach, lateral inhibition is the proposed suppression mechanism, where sounds with frequencies surrounding the tinnitus frequency inhibit neural activity associated with the tinnitus. When hearing aids are used as sound therapy, the sounds presented to the ear are an amplified version of the sound in the surround. Its effect on tinnitus is rather mixed. Some patients achieve masking of their tinnitus (McNeill et al., 2012), but others only obtain it minimally (Marcrum et al., 2021). The result of Macrum et al. was similar to ours, in that the differences between standard and notch-amplification were only minimal. However, in our study, the notched amplification performed better than the boosted amplification. Since clinically significant levels were not reached for this comparison (Fackrell et al., 2022), future studies should not disregard the boost amplification as a treatment for tinnitus, moreover considering that a certain portion of patients (22%) did prefer this amplification strategy (figure 6.5).

For clinical application of the various hearing aid settings, it is important to identify which features determine the preference for one setting over another one in individual patients. In our study, preference for one setting or no preference at all was equally distributed across the participants. The speech audiometry performed at baseline did not reveal any differences in speech understanding between the three settings, therefore we do not consider speech understanding

as a potential factor in the decision making. Despite the setting preferences being equally distributed and the fact that the groups based on preferred settings were too small to come to a conclusion, the between-groups differences in the questionnaires' scores at baseline suggest that a further study could explore more in-depth the correlation between setting preference and the questionnaires' scores before treatment.

During the informal discussions in the appointments, patients would describe their reasons to opt for one specific setting over another one, mainly influenced by how their tinnitus percept evolved during the last few weeks but also by their stress or anxiety levels during the same period, which are well-known correlated factors of tinnitus (Mazurek et al., 2019). Differences in daily life routines may have had an impact on the decision making. For instance, participants with active lifestyles, with more exposure to noisy environments such as public places, or simply by listening to the radio or the tv at home might get a bigger benefit when using any of the settings. In contrast, participants whose lifestyle is more quiet and their activities require less listening effort might not be able to distinguish between settings. Some participants used the devices in a more augmented way, by listening to podcasts or audiobooks through the Bluetooth connection, which shaped the sound according to the gain settings. Ultimately, hearing aids, and especially those with tinnitus settings, need to be exposed to the circumstances that are relevant in the life of a patient, in order to determine a preference for a specific setting.

Assessing treatment efficacy in the context of tinnitus is not straightforward. Due to the subjective nature of tinnitus, self-report questionnaires are used to measure its severity. Numerous questionnaires have been developed to assess the impact of tinnitus on patients' quality of life (Hall et al., 2016; Meikle et al., 2008) although these questionnaires were not specifically designed to study treatment outcomes (Kamalski et al., 2010). In order to study the effects of treatment options on the quality of life of the patients, it is necessary to use instruments that are responsive to treatment effects (Meikle et al., 2008). In this regard, Meikle et al. (2012) developed the Tinnitus Functional Index (TFI), facilitating the assessment of both the impact of tinnitus and the treatment-related effects on the patients.

Some limitations must be acknowledged. Tinnitus subjectiveness imposes the use of questionnaires for assessing tinnitus relief. The TFI in particular measures the impact of tinnitus during the week prior to the appointment, which leaves a longer period of time without evaluation in between appointments. This might present a bias in the assessment, as it was pointed out by some participants during the appointments. Sometimes, participants would describe a couple of bad initial weeks after a new fitting, and a substantial tinnitus relief during the third or fourth week, right before the assessment. This study covered the comparison between two different settings and a control setting over a total period of 14 weeks. The amount of comparisons might have had an effect on the significance of the results. Future studies on specific hearing aids strategies might benefit from simplifying the study design by prolonging or maintaining the duration of it and reducing the number of tested settings to one.

On a different note, each setting was tested for a period of one month. Studies involving hearing aids typically aim for a period of at least three months to achieve significant improvements of tinnitus relief, although in most studies naive hearing aid users are studied (Parazzini et al., 2011). Lastly, the two tested settings in this study (notched and boosted amplification) required a tinnitus pitch matching procedure for the correct fitting. Despite using a reliable pitch matching method, it is still possible that the notch filter and the boost filter were applied in the wrong frequencies for some participants. This is a feasible scenario due to the subjectivity and complexity of a pitch matching test.

Nevertheless, some strengths of this study can also be acknowledged. Most of the studies on hearing aids in tinnitus relief are conducted with participants who never used hearing aids before. In our study, we included experienced hearing aids users (with at least 6 months of use) in order to dim the initial improvement characteristic of getting a new treatment. The study design (a balanced and double-blind crossover study) allowed to minimize the risk of confounding, increasing the power of statistical tests performed with a relatively low sample size (Wellek and Blettner, 2012). We consider that these design decisions strengthened our results and we encourage researchers to design tinnitus studies in a similar way. Future trials with hearing devices might benefit from the blinding procedure laid out in this study.

The tinnitus literature often states the necessity of RCTs in the context of tinnitus treatment (Tunkel et al., 2014). There is a consensus in all the European guidelines with regard to the use of hearing aids in tinnitus treatment when patients suffer from hearing loss. Positive results are present in most studies, however, the quality of the evidence on the effectiveness of hearing aids for tinnitus relief is insufficient to this day (Jacquemin et al., 2022). The current study was designed to shed more light on the efficacy of hearing aids for tinnitus treatment and, in particular, to assess the effectiveness of devices adapted to the patient's tinnitus frequency. Our study showed that a notched or a boosted amplification setting did not significantly improve the patient's tinnitus compared to a standard setting. Individual preferences were different between participants, showing that a tailor-made approach to hearing aid amplification is important in clinical practice. Further studies should explore the differences among patient's tinnitus and their preferences for hearing aid settings. Moreover, hearing aids trials might benefit from the blinding procedure used in this study.

7

General discussion

The subjective nature and heterogeneity of tinnitus entail a challenge in its diagnosis, evaluation and treatment. Despite the growing body of literature on the topic, there is insufficient available evidence on the tinnitus mechanisms and the effectiveness of treatments in subjects. The high correlation with hearing loss makes it particularly difficult differentiating one from the other. This thesis investigated aspects of tinnitus in the contexts of diagnosis, assessment and sound-based therapies. The different chapters in this thesis were conceived as steps towards the goal to conduct a clinical trial as described in chapter 6. Chapters 2 to 5 provided potentially useful tools when conducting the trial, and substantiated the importance of individual approaches in tinnitus management. The current chapter discusses the findings and the interpretations of the obtained results in these studies.

7.1. Auditory Brainstem Response (ABR) as an objective measure for tinnitus.

Finding an objective measure for tinnitus does not only arise from the necessity of assessing the impact of tinnitus and better guiding treatment decisions, but it is also motivated by obtaining a reliable tool for diagnosis. Better tinnitus diagnosis could avoid unnecessary tests and procedures and allow for more personalized and targeted treatment approaches. Moreover, detecting and addressing tinnitus at an earlier, more treatable stage, could potentially prevent the further development of the condition. This could reduce the tinnitus-related health care costs, which has been estimated at €1544 per patient and year in the Netherlands (Maes et al., 2013) and much higher numbers in the US and their tinnitus-related compensation to veterans (McFerran et al., 2019).

In chapter 2 we investigated the potential of ABR as an objective measure for tinnitus. Previous studies on ABR and tinnitus obtained mixed results, possibly due to differences in experiment designs. For instance, the stimuli used varied among these studies, in particular, ABR click rates and intensity levels. The literature also suggested that varying these parameters could lead to different results in wave amplitudes and/or latencies, which encouraged us to perform a more methodical analysis. Therefore, we decided to systematically vary click rates and intensity levels to observe potential differences in wave amplitudes and latencies between patients with and without tinnitus. In view of this measure, we hoped to strengthen the assessment of the treatment response during the clinical trial, in combination with the self-reported questionnaires.

Despite controlling for hearing loss, hyperacusis, age and other variables that can influence the results, no significant difference in latencies or amplitudes was found between the patients with hearing loss and those with both hearing loss and tinnitus. The intrinsically low signal-to-noise ratio of this measure, in addition to hardware and stimuli dependencies, suggests that it is unlikely that we will obtain a set of normative latencies and amplitude values to diagnose tinnitus on an individual basis, as we do with other pathologies.

Most studies with positive results included normal hearing participants with and without tinnitus, but these findings could possibly be attributed to non-reported high frequency hearing loss or cochlear synaptopathy in the case of subjects with tinnitus (Jacxsens et al., 2022). Ultimately, the available evidence indicates that ABR cannot be used as a biomarker for tinnitus.

7.2. Minimum Masking Levels and their relationship to hearing thresholds.

One of the reasons to study the relationship between Minimum Masking Levels and hearing thresholds in a clinical population was to obtain etiology-dependent masking characteristics to, potentially, prescribe treatments according to them. Another reason was to rely on this measurement to assess the treatment effects during the clinical trial. As it has been suggested before, a decrease in MML after using a specific hearing aid setting would mean that tinnitus is easier to be masked, which suggests that the tinnitus is more difficult to be perceived by the patient (Jastreboff, 1990).

Although the results of the cluster analysis revealed a clear diversity in patients characteristics, the separation between clusters was weak, suggesting that tinnitus subtypes might not exist. At least, they cannot be identified when examining audiometric data. It is noteworthy that, despite a clear imbalance between the non-masking and the masking groups, both presented similar thresholds, tinnitus loudness and pitch. Our study suggested that maskability is not related to these variables.

Masking thresholds could differentiate between tinnitus patients. One could expect that hearing aids are more effective for tinnitus in subjects with lower MML. However, since no correlation was found between maskability, hearing thresholds, tinnitus loudness and pitch, we decided to not include this measurement in the hearing aid trial. Moreover, the test duration would be too long, potentially causing dropouts.

Our results, in line with those from Van de Berge et al. (2017), suggested that there seems to be a continuum of patients' characteristics rather than discrete subgroups. These findings support the importance of personalizing therapies for tinnitus, instead of prescribing a specific one based on etiology.

It is worth mentioning that there might not be different tinnitus etiologies. Instead, tinnitus percept could be a manifestation of a broader source, exhibited differently in each individual. Using the analogy of Alzeimer's disease, previous studies were focused on its different phenotypes and the accumulation of abnormal protein deposits in the brain as the central feature of the different clinical presentations. However, most recent studies suggest that chronic neuroinflammation is the central mechanism of Alzheimer's onset and its progression (Kinney et al., 2018). Similarly, inflammation has been linked to both noise-induced hearing loss (Frye et al., 2019) and the pathogenesis of tinnitus (Mennink et al., 2022). Inflammation or a different underlying mechanism could be the origin of every

form of tinnitus that we observe.

7.3. Assessing the impact of tinnitus and treatment effects through the Tinnitus Functional Index

For our clinical trial (detailed in chapter 6), the Dutch TFI was the chosen tool for assessing tinnitus impact and the responsiveness to treatment of each patient throughout the study. Following the recommendations of the COMiT'ID study (Hall et al., 2018), which laid the foundations for clinical trials of sound-, psychology- and pharmacology-based interventions for tinnitus, we decided to use the TFI as the optimal tool for assessing the recommended domains of "ability to ignore," "concentration," "quality of sleep," and "sense of control", among others. Given the lack of objective measures for tinnitus, assessment is usually carried out by self-reported questionnaires, of undoubtedly subjective nature.

In chapter 4, we translated and validated the Dutch version of the Tinnitus Functional Index, originally developed by Meikle et. al (2012). The TFI has been translated to more than 20 languages and it constitutes a standard instrument in both research and clinical settings, due to its psychometric properties and its responsiveness to treatment effects.

It should be noted that the Dutch translation of the TFI was validated in a diverse clinical population of 377 patients, which substantiates the results of reliability that we obtained. Before the publication of our study (Santacruz et al., 2021a), the only Dutch translation of the TFI (Rabau et al., 2014) in the literature was in the Flemish language, which is mainly spoken in Belgium. For clinical trials, it is crucial that a translated questionnaire achieves similar validity measures to those from the original version. Besides the importance of developing a reliable tool to be used in a clinical setting, validated translations allow scientists to conduct a multicentre study in different countries minimizing the risk of bias. On a more national scale, this study granted a tool that can be used in tinnitus clinics in the Netherlands.

Subjectivity is inherent in self-reported questionnaires, as it is the nature of tinnitus. From a scientific perspective, it is important to not disregard these tools that are focused on the patient's sensations instead of any specific biomarker, especially considering that the main focus in tinnitus research and care should be on the patients rather than the condition itself.

7.4. Tinnitus pitch matching can be an effortless self-guided method

Tinnitus pitch matching is a method by which a sound is presented to an individual with the goal of determining the pitch of their tinnitus percept. Usually, this process involves a hearing care professional who, by means of questions to the patient, makes changes in the stimulus until it closely matches the pitch of their tinnitus. In the fifth chapter, we developed two distinct approaches to assess the

tinnitus pitch. These methods were controlled by the participants using a software interface. The method of adjustment (MOA) consisted of two sliders to control the frequency and the intensity of the stimulus presented to the participants through headphones. The multiple-choice method (MCM) included several predefined sound samples that the participant could choose among.

The two methods developed in chapter 5 presented good reliability, which was measured by means of intra-class correlation (ICC). Despite not being significantly different from each other, the multiple-choice method (MCM) presented an ICC ≥ 0.9, which is the requirement for being used for clinical decision making in individual patient's data (Kottner et al., 2011). The main benefit of these two methods is that they give the patient the control of the process without compromising the accuracy of the results. Unlike the classical pitch-matching procedure, in which the clinician asks the patient for feedback while changing the stimulus, these methods don't require the clinician's intervention, similarly to an automated audiometry (Shojaeemend and Ayatollahi, 2018). Both pitch matching methods were developed in Matlab (version 2020b) and can be easily exported as desktop applications.

What can be inferred from the fact that both methods were not significantly different from each other? Participants underwent both pitch matching methods twice, and we assessed between-methods and within-methods comparisons. In addition to the ICC, we also used an octave difference measure (OD) for this. The higher within-methods ICC values and lower within-method OD values, as compared to the between-method ICC and OD, suggested that each method was consistently measuring a distinct aspect of tinnitus. Despite no significant differences in reliability between both methods were found, these results suggested that both could capture different information about tinnitus and how participants engage with these tools.

It should be noted the potential effect of the stimulus loudness during a pitch matching procedure. The method of adjustment (MOA) included a slider to control the loudness of the stimulus, while MCM this was automatically adjusted to the hearing thresholds of the participants. Apart from differences in frequency step sizes, this additional control parameter could be the reason why both methods obtained higher within-method ICC than between methods ICC. Previous studies have recommended to not use high-intensity stimuli during pitch matching to avoid the residual inhibition effect (McFadden et al., 1982). Besides residual inhibition, the stimulus intensity can affect the pitch perception during the matching process (e.g., a pure tone can produce aural harmonics as a form of distortion in the inner ear). To avoid this, an octave confusion test is recommended. Differently to previous reports, where this additional test resulted in a majority of participants changing their decision to a lower octave (Tyler and Conrad-Armes, 1983), no participant changed their decision in our study when using the MOA. We did not conduct an octave confusion test in the case of MCM, which could have had an influence on the results. However, no significant differences in average frequency were observed between the two methods. Perhaps even within a range of stimulus intensity that is not producing distortion artifacts, small variations in intensity

could affect the result of the matching process. To our knowledge, there is no study that systematically investigated the influence of the stimulus loudness in the reliability of a pitch matching procedure.

In the context of this thesis, this chapter provided an instrument for the application of both the notch and the boost filters of the hearing aids during the clinical trial. Then, the 18 participants underwent the MCM by themselves, and the result of the matching was directly applied to their devices. I believe that including the direct assessment of the patients in the process, in addition to the periodic interviews, encouraged them and gave them the possibility of being involved in the trial.

Despite the fact that the MCM is reliable and user-friendly, it is noteworthy to mention that the difference in octaves (OD) obtained during this study was above half an octave in some cases. The notch filter applied in the hearing aids had a bandwidth of half an octave, implying that it is possible to miss the tinnitus pitch during the matching. Needless to say, the same goes for any other existing pitch matching procedure. Like any tinnitus assessment, the subjectiveness of the process cannot prevent the risk of bias.

7.5. Adapting hearing aids to the characteristics of each patient's tinnitus

The previous chapters led to this clinical trial (detailed in chapter 6), in which different settings for hearing aids were tested by means of a double-blinded randomized controlled trial (RCT). A standard setting served as a control, and two additional settings were adjusted to each participant's tinnitus frequency. The two tested settings were a notch amplification, where a notch filter is placed at the frequency of the tinnitus pitch; and a boost amplification, where the tinnitus frequency is amplified over the standard level. Pitch matching was carried out by using the MCM method developed in chapter 5, and tinnitus impact and responsiveness to treatment were mainly assessed by the TFI, validated and translated to Dutch in chapter 4. In addition, two other questionnaires measured hyperacusis and hearing aid benefit (HQ and APHAB).

In the tinnitus literature, the necessity of well-designed RCTs to prove the efficacy of hearing aids for tinnitus management has been systematically highlighted (Tunkel et al., 2014). We believe that our clinical trial was designed accordingly and, despite doing it, no differences between approaches were obtained. For the three settings, we observed an initial reduction of the tinnitus handicap, probably related to the enthusiasm of participating in a clinical trial. Also, this effect might have been produced by the difference in quality of the devices used in the trial, since most participants preferred these hearing aids to their own ones.

Several participants pointed out that, when using the devices of our trial, they could hear birds for the first time in a long while. This seemingly meaningless remark underlines the consequences of using not only the right hearing aids but also undergoing the correct fitting and visiting a hearing care professional

regularly for the adequate adjustments as hearing declines. An updated hearing aid fitting, independently of a specific pitch-dependent setting, can help with masking the tinnitus percept more efficiently.

In connection with the above mentioned issue, the preferences for hearing aids settings were very different among the participants. However, the majority of the participants had a preference for a specific setting. Although these preferences did not always correspond to lower TFI results, they did highlight that hearing aid fitting requires a tailor-made approach. Notwithstanding that future trials could find out that one particular setting is more effective than another one, individual preferences must be taken into account in the prescription process.

In perspective, the fact that we did not find significant differences between the tested settings and the standard approach could be derived from including too many variables in the analysis. By testing only one setting (e.g., the notch therapy), this comparison could have resulted differently. In particular, testing a single setting would have not only increased the statistical power of the analysis, but also allowed longer treatment periods. In addition to this, registering life events of the participants during the trial (such as stressful events) could have helped to explain the variance in the outcomes.

7.6. Future directions

Future studies should continue the search for an objective measure in tinnitus. Regarding Auditory Evoked Potentials (AEP, an umbrella term that comprises early, middle and late auditory responses), recent studies have incorporated machine learning approaches in the analysis of these signals with promising results (Manta et al., 2023). This line of research could benefit from the latest advances in Deep Neural Networks and their high performance in time-series data processing. These networks, however, need a large amount of data to raise their accuracy to acceptable levels, therefore large tinnitus-specific databases are necessary (Simoes et al., 2021). We have shown that conducting a double-blind trial with hearing aids for tinnitus is feasible. This blinding procedure is particularly important for future trials to compare different amplification techniques. Future device trials might benefit from using our blinding method to minimize the risk of bias.

To date, it is still unclear whether the observed variations in repeated measures of pitch matching sessions are due to changes in tinnitus or the difficulty of the test itself. Potential pitch changes could be detected via longitudinal studies. User-friendly or semi-automatic approaches that minimize the interaction between patient and health care professional could help with the data collection process, especially in research settings, where using consistent procedures is crucial.

Despite the literature suggesting that the tinnitus pitch might influence the efficacy of sound-based therapies, this still remains unclear since tinnitus profiles are highly individual. Future research on hearing aids and tinnitus should focus on the study of individualized therapies. Identifying auditory profiles, developing new amplification strategies and further studying noise therapies (such as in combination aids) may help with increasing the available evidence of sound treatments for

tinnitus.

7.7. Clinical implications

Our findings in chapter 2 indicate that ABR is not a suitable measure for tinnitus diagnosis. By means of larger available datasets and adequate machine learning approaches it might be possible to shed some light on the brainstem tinnitus mechanisms using this technique, but our results suggest that it is not feasible to detect tinnitus in individuals.

The translation and validation of the Dutch TFI provided a useful tool for tinnitus clinics in Dutch speaking countries. In fact, the tinnitus guidelines of the KNO-vereniging (ENT association in The Netherlands) recommend using a questionnaire. These clinics can make use of our validated TFI, as it is being currently used at the ENT polyclinic of the UMCG. Using the Dutch TFI could assist clinicians to measure both tinnitus impact and responsiveness to treatment of patients with tinnitus, and help with the assessment in future research settings.

The reliability of the two self-guided pitch matching methods developed in chapter 5 showed that these techniques can be used in a clinical environment. Their ease of use and simplicity create an opportunity for online implementations that can facilitate remote consultation in hybrid models of tinnitus management.

The results of the clinical trial highlighted the diversity of preferences for specific settings among the patients. And, despite these preferences being different for each patient, the majority of them preferred one setting over the rest. This observation underlined the importance of tailoring treatments to each participant's needs and adopting patient-driven strategies in tinnitus management. For this, it's important that clinicians engage patients in an "active listening" conversation. Patient-centered communication has been reported to be a missing factor in audiological consultation (Grenness et al., 2015). This is also important in the context of tinnitus research. Investigators' goal is often finding objective measures and/or proving the efficacy of a specific treatment. However, it's especially important to not neglect the impact of effective patient care, with greater reason in tinnitus management, given the subjective nature of the condition and the lack of evidence for most available therapies.

7.8. Conclusions

In this thesis, I investigated different aspects of tinnitus from a clinical point of view, which included measurement and assessment of tinnitus impact and treatment. Overall, the ABR study showed that this technique does not have the potential to become a reliable diagnosis tool for tinnitus and to disentangle it from hearing loss, highlighting the heterogeneity of individual responses. In line with this observation, and despite masking contours having the potential to assess the acceptability of tinnitus masking, the cluster analysis on MML and hearing thresholds suggested that there seems to be a continuum of patient's characteristics rather than distinct tinnitus subgroups. This indicates that identifying tinnitus

profiles for prescribing therapies by means of MML and hearing threshold is not a viable solution. Additionally, this thesis provided a validated questionnaire for assessing tinnitus impact and responsiveness to treatment for Dutch speaking clinics, and a self-guided pitch matching method that showed to be reliable in a clinical population. The RCT showed that the notch and the boost amplification are not significantly better than a standard amplification in hearing aids for tinnitus treatment. However, the individual preferences of the trial participants highlighted the importance of tailor-made approaches in sound-based therapies for tinnitus.

Appendices

TFI questionnaire

TINNITUS FUNCTIONAL INDEX

Lees elke vraag zorgvuldig door. Beantwoord de vragen door één van de getallen bij een vraag als volgt te omcirkelen: (10%) of (1.)

In de afgelopen week…

1. Hoeveel procent van de tijd dat u wakker was, was u zich ***bewust*** van uw tinnitus?

Nooit	▶	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	◀	Altijd

2. Hoe ***sterk*** of ***luid*** was uw tinnitus?

Helemaal niet sterk of luid	▶	0	1	2	3	4	5	6	7	8	9	10	◀	Extreem sterk of luid

3. Hoeveel procent van de tijd dat u wakker was, was u ***geïrriteerd*** door uw tinnitus?

Geen moment	▶	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	◀	De hele tijd

4. Had u het gevoel dat u ***controle*** had over uw tinnitus?

Altijd	▶	0	1	2	3	4	5	6	7	8	9	10	◀	Nooit

5. Hoe gemakkelijk was het voor u om met uw tinnitus ***om te gaan***?

Heel gemakkelijk	▶	0	1	2	3	4	5	6	7	8	9	10	◀	Onmogelijk

6. Hoe gemakkelijk was het voor u om uw tinnitus te ***negeren***?

Heel gemakkelijk	▶	0	1	2	3	4	5	6	7	8	9	10	◀	Onmogelijk

Hoezeer verstoorde uw tinnitus in de afgelopen week...

7. ...uw ***concentratievermogen***?

Niet verstoord	▶	0	1	2	3	4	5	6	7	8	9	10	◀	Volledig verstoord

8. …uw vermogen om ***helder te denken?***

Niet verstoord	▶	0	1	2	3	4	5	6	7	8	9	10	◀	Volledig verstoord

9. ...uw vermogen om de ***aandacht te richten*** op andere dingen dan uw tinnitus?

Niet verstoord	▶	0	1	2	3	4	5	6	7	8	9	10	◀	Volledig verstoord

In de afgelopen week...

10. Hoe vaak maakte uw tinnitus het moeilijk om ***in slaap te vallen of in slaap te blijven***?

Nooit	▶	0	1	2	3	4	5	6	7	8	9	10	◀	Altijd

11. Hoe vaak maakte uw tinnitus het u moeilijk om ***de hoeveelheid slaap te krijgen*** **die u nodig had**?

Nooit	▶	0	1	2	3	4	5	6	7	8	9	10	◀	Altijd

12. Hoe vaak weerhield uw tinnitus u ervan zo ***diep of zo rustig te slapen*** als u graag had gewild?

Nooit	▶	0	1	2	3	4	5	6	7	8	9	10	◀	Altijd

Hoezeer verstoorde uw tinnitus gedurende de afgelopen week...

13. …uw vermogen om ***goed te horen***?

Niet verstoord	▶	0	1	2	3	4	5	6	7	8	9	10	◀	Volledig verstoord

14. …uw vermogen om ***mensen te verstaan*** die aan het praten zijn?

Niet verstoord ▶ 0 1 2 3 4 5 6 7 8 9 10 ◀ Volledig verstoord

15. …uw vermogen om ***gesprekken te volgen*** in een groep of tijdens vergaderingen?

Niet verstoord ▶ 0 1 2 3 4 5 6 7 8 9 10 ◀ Volledig verstoord

16. ...uw ***rustige ontspannende activiteiten*****?**

Niet verstoord ▶ 0 1 2 3 4 5 6 7 8 9 10 ◀ Volledig verstoord

17. …uw vermogen om u te ***ontspannen***?

Niet verstoord ▶ 0 1 2 3 4 5 6 7 8 9 10 ◀ Volledig verstoord

18. ...uw vermogen om van ***rust*** **te genieten?**

Niet verstoord ▶ 0 1 2 3 4 5 6 7 8 9 10 ◀ Volledig verstoord

19. ...uw plezier in ***sociale activiteiten*****?**

Niet verstoord ▶ 0 1 2 3 4 5 6 7 8 9 10 ◀ Volledig verstoord

20. ...uw ***levensvreugde*****?**

Niet verstoord ▶ 0 1 2 3 4 5 6 7 8 9 10 ◀ Volledig verstoord

21. ...uw ***contact*** met familie, vrienden en anderen?

Niet verstoord ▶ 0 1 2 3 4 5 6 7 8 9 10 ◀ Volledig verstoord

22. Hoe vaak had u door uw tinnitus moeite met het uitvoeren van uw ***werk of andere taken,*** zoals huishoudelijke werkzaamheden, werk, school of de zorg voor kinderen of anderen?

Nooit ▶ 0 1 2 3 4 5 6 7 8 9 10 ◀ Altijd

In de <u>afgelopen week</u>...

23. Hoe ***angstig of bezorgd*** heeft u zich gevoeld door uw tinnitus?

Helemaal niet ▶ 0 1 2 3 4 5 6 7 8 9 10 ◀ Extreem

24. Hoe ***geïrriteerd of boos*** heeft u zich gevoeld door uw tinnitus?

Helemaal niet ▶ 0 1 2 3 4 5 6 7 8 9 10 ◀ Extreem

25. Hoe ***somber*** was u door uw tinnitus?

Helemaal niet ▶ 0 1 2 3 4 5 6 7 8 9 10 ◀ Extreem

References

Aazh, H. and Moore, B. C. (2017). Factors related to uncomfortable loudness levels for patients seen in a tinnitus and hyperacusis clinic. *International Journal of Audiology*, 56(10):793–800.

Akiyoshi, R., Shindo, M., and Kaga, K. (2021). Subcortical deafness as a subtype of auditory agnosia after injury of bilateral auditory radiations caused by two cerebrovascular accidents–normal auditory brainstem responses with i–vii waves and abolished consciousness of hearing–. *Acta Oto-Laryngologica*, 141(4):374–380.

Amirah Fatin, I., Khatijah, L., Zilany, M., Zuheir, A., Ong, S., and Tan, M. (2016). The effect of hearing augmentation on cognitive assessment scores: a pilot crossover randomized controlled trial. In *International Conference for Innovation in Biomedical Engineering and Life Sciences: ICIBEL2015, 6-8 December 2015, Putrajaya, Malaysia 1*, pages 24–29. Springer.

Andersson, G. (2003). Tinnitus loudness matchings in relation to annoyance and grading of severity. *Auris Nasus Larynx*, 30(2):129–133.

Andersson, G. and Westin, V. (2008). Understanding tinnitus distress: introducing the concepts of moderators and mediators. *International Journal of Audiology*, 47(sup2):S106–S111.

Asparouhov, T. and Muthén, B. (2009). Exploratory structural equation modeling. *Structural equation modeling: a multidisciplinary journal*, 16(3):397–438.

Attias, J., Pratt, H., Reshef, I., Bresloff, I., Horowitz, G., Polyakov, A., and Shemesh, Z. (1996). Detailed analysis of auditory brainstem responses in patients with noise-induced tinnitus. *Audiology*, 35(5):259–270.

Attias, J., Urbach, D., Gold, S., and Shemesh, Z. (1993). Auditory event related potentials in chronic tinnitus patients with noise induced hearing loss. *Hearing Research*, 71(1-2):106–113.

Baguley, D., McFerran, D., and Hall, D. (2013). Tinnitus. *The Lancet*, 382(9904):1600–1607.

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences*, 160(901):268–282.

Beukes, E. W., Manchaiah, V., Allen, P. M., Andersson, G., and Baguley, D. M. (2021). Exploring tinnitus heterogeneity. *Progress in brain research*, 260:79–99.

Bhatt, J. M., Lin, H. W., and Bhattacharyya, N. (2016). Prevalence, severity, exposures, and treatment patterns of tinnitus in the united states. *JAMA Otolaryngology–Head & Neck Surgery*, 142(10):959–965.

Biswas, R. and Hall, D. A. (2020). Prevalence, incidence, and risk factors for tinnitus. In *The behavioral neuroscience of tinnitus*, pages 3–28. Springer.

Biswas, R., Lugo, A., Akeroyd, M. A., Schlee, W., Gallus, S., and Hall, D. (2022). Tinnitus prevalence in europe: a multi-country cross-sectional population study. *The Lancet Regional Health–Europe*, 12.

Bland, J. M. and Altman, D. G. (1995). Multiple significance tests: the bonferroni method. *Bmj*, 310(6973):170.

Boyen, K., de Kleine, E., van Dijk, P., and Langers, D. R. (2014). Tinnitus-related dissociation between cortical and subcortical neural activity in humans with mild to moderate sensorineural hearing loss. *Hearing research*, 312:48–59.

Brussee, L. (2003). Tinnitus een permanente stoorzender. *De invloed van tinnitus op het dagelijks functioneren en de kwaliteit van leven. Groningen: Wetenschapswinkel Geneeskunde en Volksgezondheid, Rijksuniversiteit Groningen*.

Burkard, R. F., Eggermont, J. J., and Don, M. (2007). *Auditory evoked potentials: basic principles and clinical application*. Lippincott Williams & Wilkins.

Burke, L. A. and Naylor, G. (2020). Daily-life fatigue in mild to moderate hearing impairment: An ecological momentary assessment study. *Ear and hearing*, 41(6):1518.

Cantril, H. (1965). Discovering people's aspirations: The method used. the pattern of human concerns. pages 21–29.

Cederroth, C. R., Gallus, S., Hall, D. A., Kleinjung, T., Langguth, B., Maruotti, A., Meyer, M., Norena, A., Probst, T., Pryss, R., et al. (2019). Towards an understanding of tinnitus heterogeneity. *Frontiers in aging neuroscience*, 11:53.

Chen, C., Zhan, L., Pan, X., Wang, Z., Guo, X., Qin, H., Xiong, F., Shi, W., Shi, M., Ji, F., et al. (2021). Automatic recognition of auditory brainstem response characteristic waveform based on bidirectional long short-term memory. *Frontiers in Medicine*, 7:613708.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic press.

Cox, R. M. (1997). Administration and application of the aphab. *The Hearing Journal*, 50(4):32–35.

Cox, R. M. and Alexander, G. C. (1995). The abbreviated profile of hearing aid benefit. *Ear and hearing*, 16(2):176–186.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, 16(3):297–334.

Dawes, P., Hopkins, R., and Munro, K. J. (2013). Placebo effects in hearing-aid trials are reliable. *International Journal of Audiology*, 52(7):472–477.

De Ridder, D., Kilgard, M., Engineer, N., and Vanneste, S. (2015). Placebo-controlled vagus nerve stimulation paired with tones in a patient with refractory tinnitus: a case report. *Otology & Neurotology*, 36(4):575–580.

De Ridder, D., Vanneste, S., Weisz, N., Londero, A., Schlee, W., Elgoyhen, A. B., and Langguth, B. (2014). An integrative model of auditory phantom perception: tinnitus as a unified percept of interacting separable subnetworks. *Neuroscience & Biobehavioral Reviews*, 44:16–32.

Dobie, R. A. (1999). A review of randomized clinical trials in tinnitus. *The Laryngoscope*, 109(8):1202–1211.

Durai, M. and Searchfield, G. (2016). Anxiety and depression, personality traits relevant to tinnitus: a scoping review. *International journal of audiology*, 55(11):605–615.

Edvall, N. K., Mehraei, G., Claeson, M., Lazar, A., Bulla, J., Leineweber, C., Uhlén, I., Canlon, B., Cederroth, C. R., et al. (2022). Alterations in auditory brain stem response distinguish occasional and constant tinnitus. *The Journal of Clinical Investigation*, 132(5).

Edwards, A. L. (1951). Balanced latin-square designs in psychological research. *The American journal of psychology*, 64(4):598–603.

Eggermont, J. J. (2019). Auditory brainstem response. *Handbook of clinical neurology*, 160:451–464.

Eggermont, J. J. and Roberts, L. E. (2004). The neuroscience of tinnitus. *Trends in neurosciences*, 27(11):676–682.

Fackrell, K., Hall, D. A., Barry, J., and Hoare, D. J. (2022). Integrating distribution-based and anchor-based techniques to identify minimal important change for the tinnitus functional index (tfi) questionnaire. *Brain Sciences*, 12(6):726.

Fackrell, K., Hall, D. A., Barry, J. G., and Hoare, D. J. (2016). Psychometric properties of the tinnitus functional index (tfi): Assessment in a uk research volunteer population. *Hearing research*, 335:220–235.

Fackrell, K., Hall, D. A., Barry, J. G., and Hoare, D. J. (2018). Performance of the tinnitus functional index as a diagnostic instrument in a uk clinical population. *Hearing research*, 358:74–85.

Feldmann, H. (1971). Homolateral and contralateral masking of tinnitus by noise-bands and by pure tones. *Audiology*, 10(3):138–144.

Field, Z., Miles, J., and Field, A. (2012). Discovering statistics using r. *Discovering Statistics Using R*, pages 1–992.

Fournier, P., Cuvillier, A.-F., Gallego, S., Paolino, F., Paolino, M., Quemar, A., Londero, A., and Norena, A. (2018). A new method for assessing masking and residual inhibition of tinnitus. *Trends in hearing*, 22:2331216518769996.

Frye, M. D., Ryan, A. F., and Kurabi, A. (2019). Inflammation associated with noise-induced hearing loss. *The Journal of the Acoustical Society of America*, 146(5):4020–4032.

Gallus, S., Lugo, A., Garavello, W., Bosetti, C., Santoro, E., Colombo, P., Perin, P., La Vecchia, C., and Langguth, B. (2015). Prevalence and determinants of tinnitus in the italian adult population. *Neuroepidemiology*, 45(1):12–19.

Genitsaridi, E., Partyka, M., Gallus, S., Lopez-Escamez, J. A., Schecklmann, M., Mielczarek, M., Trpchevska, N., Santacruz, J. L., Schoisswohl, S., Riha, C., et al. (2019). Standardised profiling for tinnitus research: the european school for interdisciplinary tinnitus research screening questionnaire (esit-sq). *Hearing research*, 377:353–359.

Girden, E. R. and Kabacoff, R. (2010). *Evaluating research articles from start to finish*. Sage.

Gos, E., Rajchel, J. J., Dziendziel, B., Kutyba, J., Bienkowska, K., Swierniak, W., Gocel, M., Raj-Koziak, D., Skarzynski, P. H., and Skarzynski, H. (2021). How to interpret tinnitus functional index scores: a proposal for a grading system based on a large sample of tinnitus patients. *Ear and Hearing*, 42(3):654–661.

Grenness, C., Hickson, L., Laplante-Lévesque, A., Meyer, C., and Davidson, B. (2015). The nature of communication throughout diagnosis and management planning in initial audiologic rehabilitation consultations. *Journal of the American Academy of Audiology*, 26(01):036–050.

Guadagnoli, E. and Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological bulletin*, 103(2):265.

Guillemin, F., Bombardier, C., and Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of clinical epidemiology*, 46(12):1417–1432.

Hall, D. A., Haider, H., Szczepek, A. J., Lau, P., Rabau, S., Jones-Diette, J., Londero, A., Edvall, N. K., Cederroth, C. R., Mielczarek, M., et al. (2016). Systematic review of outcome domains and instruments used in clinical trials of tinnitus treatments in adults. *Trials*, 17(1):1–19.

Hall, D. A., Hibbert, A., Smith, H., Haider, H. F., Londero, A., Mazurek, B., Fackrell, K., and in Tinnitus (COMiT) initiative, C. O. M. (2019). One size does not fit all: developing common standards for outcomes in early-phase clinical trials of sound-, psychology-, and pharmacology-based interventions for chronic subjective tinnitus in adults. *Trends in hearing*, 23:2331216518824827.

Hall, D. A., Smith, H., Hibbert, A., Colley, V., Haider, H. F., Horobin, A., Londero, A., Mazurek, B., Thacker, B., Fackrell, K., et al. (2018). The comit'id study: developing core outcome domains sets for clinical trials of sound-, psychology-, and pharmacology-based interventions for chronic subjective tinnitus in adults. *Trends in hearing*, 22:2331216518814384.

Han, M. S., Jeong, Y. J., Im, G. J., Song, J.-J., Chae, S.-W., Rah, Y. C., and Choi, J. (2021). Auditory brainstem response test results in normal hearing adolescents with subjective tinnitus. *International Journal of Pediatric Otorhinolaryngology*, 146:110775.

Hauptmann, C., Wegener, A., Poppe, H., Williams, M., Popelka, G., and Tass, P. A. (2016). Validation of a mobile device for acoustic coordinated reset neuromodulation tinnitus therapy. *Journal of the American Academy of Audiology*, 27(09):720–731.

Hazell, J. and Wood, S. (1981). Tinnitus masking-a significant contribution to tinnitus management. *British journal of audiology*, 15(4):223–230.

Heijneman, K. M., de Kleine, E., and van Dijk, P. (2012). A randomized double-blind crossover study of phase-shift sound therapy for tinnitus. *Otolaryngology–Head and Neck Surgery*, 147(2):308–315.

Henry, J., Reavis, K., Griest, S., Thielman, E., Theodoroff, S., Grush, L., and Carlson, K. (2020). Tinnitus: an epidemiologic perspective. *Otolaryngologic Clinics of North America*, 53(4):481–499.

Henry, J. A., Dennis, K. C., and Schechter, M. A. (2005). General review of tinnitus.

Henry, J. A., Flick, C. L., Gilbert, A., Ellingson, R. M., and Fausti, S. A. (2004). Comparison of manual and computer-automated procedures for tinnitus pitch-matching. *Journal of Rehabilitation Research & Development*, 41(2).

Henry, J. A., Frederick, M., Sell, S., Griest, S., and Abrams, H. (2015a). Validation of a novel combination hearing aid and tinnitus therapy device. *Ear and Hearing*, 36(1):42–52.

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

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

Henry, J. A., McMillan, G. P., Thielman, E. J., Galvez, G., Zaugg, T. L., Porsov, E., and Silaski, G. (2013). Evaluating psychoacoustic measures for establishing presence of tinnitus. *Journal of Rehabilitation Research & Development*, 50(4).

Henry, J. A. and Meikle, M. B. (2000). Psychoacoustic measures of tinnitus. *Journal of the American Academy of Audiology*, 11(03):138–155.

Henry, J. A., Zaugg, T. L., Myers, P. J., and Schechter, M. A. (2008). Using therapeutic sound with progressive audiologic tinnitus management. *Trends in Amplification*, 12(3):188–209.

Henton, A. and Tzounopoulos, T. (2021). What's the buzz? the neuroscience and the treatment of tinnitus. *Physiological Reviews*, 101(4):1609–1632.

Hoare, D., Searchfield, G., El Refaie, A., and Henry, J. (2014). Sound therapy for tinnitus management: practicable options. *Journal of the American Academy of Audiology*, 25(01):062–075.

Hoare, D. J., Kowalkowski, V. L., Kang, S., and Hall, D. A. (2011). Systematic review and meta-analyses of randomized controlled trials examining tinnitus management. *The Laryngoscope*, 121(7):1555–1564.

Hoare, D. J., Stacey, P. C., and Hall, D. A. (2010). The efficacy of auditory perceptual training for tinnitus: a systematic review. *Annals of Behavioral Medicine*, 40(3):313–324.

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

Hoff, M. and Kähäri, K. (2017). A swedish cross-cultural adaptation and validation of the tinnitus functional index. *International journal of audiology*, 56(4):277–285.

Hofmeier, B., Wertz, J., Refat, F., Hinrichs, P., Saemisch, J., Singer, W., Rüttiger, L., Klose, U., Knipper, M., and Wolpert, S. (2021). Functional biomarkers that distinguish between tinnitus with and without hyperacusis. *Clinical and translational medicine*, 11(5).

Hofmeier, B., Wolpert, S., Aldamer, E., Walter, M., Thiericke, J., Braun, C., Zelle, D., Rüttiger, L., Klose, U., and Knipper, M. (2018). Reduced sound-evoked and resting-state bold fmri connectivity in tinnitus. *NeuroImage: Clinical*, 20:637–649.

Holmes, S. and Padgham, N. D. (2009). More than ringing in the ears: a review of tinnitus and its psychosocial impact. *Journal of Clinical Nursing*, 18(21):2927–2937.

Hu, L.-t. and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1):1–55.

Jackson, R., Vijendren, A., and Phillips, J. (2019). Objective measures of tinnitus: a systematic review. *Otology & Neurotology*, 40(2):154–163.

Jacquemin, L., Gilles, A., and Shekhawat, G. S. (2022). Hearing more to hear less: a scoping review of hearing aids for tinnitus relief. *International Journal of Audiology*, 61(11):887–895.

Jacquemin, L., Mertens, G., Van de Heyning, P., Vanderveken, O. M., Topsakal, V., De Hertogh, W., Michiels, S., Van Rompaey, V., and Gilles, A. (2019). Sensitivity to change and convergent validity of the tinnitus functional index (tfi) and the tinnitus questionnaire (tq): Clinical and research perspectives. *Hearing Research*, 382:107796.

Jacxsens, L., De Pauw, J., Cardon, E., Van der Wal, A., Jacquemin, L., Gilles, A., Michiels, S., Van Rompaey, V., Lammers, M. J., and De Hertogh, W. (2022). Brainstem evoked auditory potentials in tinnitus: A best-evidence synthesis and meta-analysis. *Frontiers in Neurology*, 13:941876.

Jain, S., Cherian, R., Nataraja, N. P., and Narne, V. K. (2021). The relationship between tinnitus pitch, audiogram edge frequency, and auditory stream segregation abilities in individuals with tinnitus. *American Journal of Audiology*, 30(3):524–534.

Jastreboff, P. J. (1990). Phantom auditory perception (tinnitus): mechanisms of generation and perception. *Neuroscience research*, 8(4):221–254.

Jastreboff, P. J. (2011). Tinnitus retraining therapy. *Textbook of tinnitus*, pages 575–596.

Jastreboff, P. J., Hazell, J. W., and Graham, R. L. (1994). Neurophysiological model of tinnitus: dependence of the minimal masking level on treatment outcome. *Hearing Research*, 80(2):216–232.

Joergensen, M. L., Hyvärinen, P., Caporali, S., and Dau, T. (2022). Broadband amplification as tinnitus treatment. *Brain Sciences*, 12(6):719.

Kaiser, M. (1974). Kaiser-meyer-olkin measure for identity correlation matrix. *Journal of the Royal Statistical Society*, 52(1):296–298.

Kam, A. C. S., Leung, E. K. S., Chan, P. Y. B., and Tong, M. C. F. (2018). Cross-cultural adaptation and psychometric properties of the chinese tinnitus functional index. *International journal of audiology*, 57(2):91–97.

Kamalski, D., Hoekstra, C., Zanten, B., Grolman, W., and Rovers, M. (2010). Measuring disease-specific health-related quality of life to evaluate treatment outcomes in tinnitus patients: a systematic review. *Otolaryngology—Head and Neck Surgery*, 143(2):181–185.

Keidser, G., Dillon, H., Carter, L., and O'Brien, A. (2012). Nal-nl2 empirical adjustments. *Trends in amplification*, 16(4):211–223.

Khalfa, S., Dubal, S., Veuillet, E., Perez-Diaz, F., Jouvent, R., and Collet, L. (2002). Psychometric normalization of a hyperacusis questionnaire. *Orl*, 64(6):436–442.

Kinney, J. W., Bemiller, S. M., Murtishaw, A. S., Leisgang, A. M., Salazar, A. M., and Lamb, B. T. (2018). Inflammation as a central mechanism in alzheimer's disease. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 4:575–590.

König, O., Schaette, R., Kempter, R., and Gross, M. (2006). Course of hearing loss and occurrence of tinnitus. *Hearing research*, 221(1-2):59–64.

Koops, E. A., de Kleine, E., and van Dijk, P. (2020). Gray matter declines with age and hearing loss, but is partially maintained in tinnitus. *Scientific Reports*, 10(1):21801.

Kottner, J., Gajewski, B. J., and Streiner, D. L. (2011). Guidelines for reporting reliability and agreement studies (grras).

Landgrebe, M., Zeman, F., Koller, M., Eberl, Y., Mohr, M., Reiter, J., Staudinger, S., Hajak, G., and Langguth, B. (2010). The tinnitus research initiative (tri) database: a new approach for delineation of tinnitus subtypes and generation of predictors for treatment outcome. *BMC medical informatics and decision making*, 10:1–7.

Langguth, B. (2011). A review of tinnitus symptoms beyond 'ringing in the ears': a call to action. *Current medical research and opinion*, 27(8):1635–1643.

Langguth, B., Landgrebe, M., Schlee, W., Schecklmann, M., Vielsmeier, V., Steffens, T., Staudinger, S., Frick, H., and Frick, U. (2017). Different patterns of hearing loss among tinnitus patients: a latent class analysis of a large sample. *Frontiers in Neurology*, 8:46.

Lybarger, S. (1963). Simplified fitting system for hearing aids. *Canonsburg, Pennsylvania: Radioear Corp*.

MacCallum, R. C., Browne, M. W., and Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological methods*, 1(2):130.

Maes, I. H., Cima, R. F., Vlaeyen, J. W., Anteunis, L. J., and Joore, M. A. (2013). Tinnitus: a cost study. *Ear and hearing*, 34(4):508–514.

Mahboubi, H., Ziai, K., and Djalilian, H. (2012). Customized web-based sound therapy for tinnitus. *The International Tinnitus Journal*, 17(1):26–30.

Manta, O., Sarafidis, M., Schlee, W., Mazurek, B., Matsopoulos, G. K., and Koutsouris, D. D. (2023). Development of machine-learning models for tinnitus-related distress classification using wavelet-transformed auditory evoked potential signals and clinical data. *Journal of Clinical Medicine*, 12(11):3843.

Marcrum, S., Picou, E., Steffens, T., Hannemann, R., Vielsmeier, V., Schecklmann, M., Langguth, B., and Schlee, W. (2021). Conventional versus notch filter amplification for the treatment of tinnitus in adults with mild-to-moderate hearing loss. *Progress in Brain Research*, 260:235–252.

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J., and Trautwein, U. (2009). Exploratory structural equation modeling, integrating cfa and efa: Application to students' evaluations of university teaching. *Structural equation modeling: A multidisciplinary journal*, 16(3):439–476.

Mazurek, B., Boecking, B., and Brueggemann, P. (2019). Association between stress and tinnitus—new aspects. *Otology & Neurotology*, 40(4):e467–e473.

Mazurek, B., Olze, H., Haupt, H., and Szczepek, A. J. (2010). The more the worse: the grade of noise-induced hearing loss associates with the severity of tinnitus. *International journal of environmental research and public health*, 7(8):3071–3079.

McCormack, A., Edmondson-Jones, M., Fortnum, H., Dawes, P., Middleton, H., Munro, K. J., and Moore, D. R. (2014). The prevalence of tinnitus and the relationship with neuroticism in a middle-aged uk population. *Journal of Psychosomatic Research*, 76(1):56–60.

McCormack, A., Edmondson-Jones, M., Somerset, S., and Hall, D. (2016). A systematic review of the reporting of tinnitus prevalence and severity. *Hearing research*, 337:70–79.

McFadden, D., Council, N. R., et al. (1982). *Tinnitus: Facts, theories, and treatments*. National Academy Press.

McFerran, D. J., Stockdale, D., Holme, R., Large, C. H., and Baguley, D. M. (2019). Why is there no cure for tinnitus? *Frontiers in neuroscience*, 13:802.

McNeill, C., Távora-Vieira, D., Alnafjan, F., Searchfield, G. D., and Welch, D. (2012). Tinnitus pitch, masking, and the effectiveness of hearing aids for tinnitus therapy. *International journal of audiology*, 51(12):914–919.

Meikle, M., Henry, J., Griest, S., Stewart, B., Abrams, H., McArdle, R., Myers, P., Newman, C., Sandridge, S., Turk, D., and Folmer, R. (2012). The tinnitus functional index: development of a new clinical measure for chronic, intrusive tinnitus. *Ear and Hearing*, 33(2):153–176.

Meikle, M., Stewart, B., Griest, S., and Henry, J. (2008). Tinnitus outcomes assessment. *Trends in Amplification*, 12(3):223–235.

Mennink, L. M., Aalbers, M. W., van Dijk, P., and van Dijk, J. M. C. (2022). The role of inflammation in tinnitus: A systematic review and meta-analysis. *Journal of Clinical Medicine*, 11(4):1000.

Milloy, V., Fournier, P., Benoit, D., Noreña, A., and Koravand, A. (2017). Auditory brainstem responses in tinnitus: a review of who, how, and what? *Frontiers in Aging Neuroscience*, 9:237.

Mitchell, C. (1983). The masking of tinnitus with pure tones. *Audiology*, 22(1):73–87.

Mitchell, C. R., Vernon, J. A., and Creedon, T. A. (1993). Measuring tinnitus parameters: loudness, pitch, and maskability. *Journal of the American Academy of Audiology*, 4(3):139–151.

Mohamad, N., Hoare, D. J., and Hall, D. A. (2016). The consequences of tinnitus and tinnitus severity on cognition: a review of the behavioural evidence. *Hearing research*, 332:199–209.

Møller, A. R., Langguth, B., DeRidder, D., and Kleinjung, T. (2010). *Textbook of tinnitus*. Springer Science & Business Media.

Moore, B. C. et al. (2010). The relationship between tinnitus pitch and the edge frequency of the audiogram in individuals with hearing impairment and tonal tinnitus. *Hearing research*, 261(1-2):51–56.

Müller, K., Edvall, N. K., Idrizbegovic, E., Huhn, R., Cima, R., Persson, V., Leineweber, C., Westerlund, H., Langguth, B., Schlee, W., et al. (2016). Validation of online versions of tinnitus questionnaires translated into swedish. *Frontiers in aging neuroscience*, 8:272.

Nakagawa, S., Johnson, P. C., and Schielzeth, H. (2017). The coefficient of determination r 2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, 14(134):20170213.

Neff, P., Langguth, B., Schecklmann, M., Hannemann, R., and Schlee, W. (2019). Comparing three established methods for tinnitus pitch matching with respect to reliability, matching duration, and subjective satisfaction. *Trends in Hearing*, 23:2331216519887247.

Newman, C. W., Jacobson, G. P., and Spitzer, J. B. (1996). Development of the tinnitus handicap inventory. *Archives of Otolaryngology–Head & Neck Surgery*, 122(2):143–148.

Newman, C. W., Sandridge, S. A., and Jacobson, G. P. (1998). Psychometric adequacy of the tinnitus handicap inventory (thi) for evaluating treatment outcome. *Journal of the American Academy of Audiology*, 9(2).

Norena, A., Micheyl, C., Chéry-Croze, S., and Collet, L. (2002). Psychoacoustic characterization of the tinnitus spectrum: implications for the underlying mechanisms of tinnitus. *Audiology and Neurotology*, 7(6):358–369.

Noreña, A. J. (2011). An integrative model of tinnitus based on a central gain controlling neural sensitivity. *Neuroscience & Biobehavioral Reviews*, 35(5):1089–1109.

Nunnally, J. C. and Bernstein, I. H. (1994). Psychometric theory new york. *NY: McGraw-Hill*.

Nuttall, A., Meikle, M., and Trune, D. (2004). Peripheral processes involved in tinnitus. *Tinnitus: Theory and management*, pages 52–68.

Oakes, M. P., Anwar, M. N., and Panchev, C. (2013). Data mining for gender differences in tinnitus.

Oron, Y., Shemesh, S., Tamir, S., Goldfarb, A., Marom, T., Gluck, O., and Shushan, S. (2018). A hebrew adaptation of the tinnitus functional index. *Clinical Otolaryngology*, 43(2):662–665.

Pantev, C. (2004). Training-induced plasticity of the human auditory cortex. In *BRAIN AND COGNITION*, volume 54, pages 129–129. ACADEMIC PRESS INC ELSEVIER SCIENCE 525 B ST, STE 1900, SAN DIEGO, CA 92101 . . . .

Pantev, C., Okamoto, H., and Teismann, H. (2012). Music-induced cortical plasticity and lateral inhibition in the human auditory cortex as foundations for tonal tinnitus treatment. *Frontiers in Systems Neuroscience*, 6:1–19.

Pape, J., Paraskevopoulos, E., Bruchmann, M., Wollbrink, A., Rudack, C., and Pantev, C. (2014). Playing and listening to tailor-made notched music: Cortical plasticity induced by unimodal and multimodal training in tinnitus patients. *Neural Plasticity*.

Parazzini, M., Del Bo, L., Jastreboff, M., Tognola, G., and Ravazzani, P. (2011). Open ear hearing aids in tinnitus therapy: An efficacy comparison with sound generators. *International journal of audiology*, 50(8):548–553.

Pelleg, D., Moore, A. W., et al. (2000). X-means: Extending k-means with efficient estimation of the number of clusters. In *Icml*, volume 1, pages 727–734.

Penner, M. (1995). Tinnitus synthesis: fluctuant and stable matches to the pitch of tinnitus. *Int. Tinnitus J*, 1(2):79–83.

Penner, M. and Bilger, R. (1992). Consistent within-session measures of tinnitus. *Journal of Speech, Language, and Hearing Research*, 35(3):694–700.

Peter, N., Kleinjung, T., Jeker, R., Meyer, M., Klaghofer, R., and Weidt, S. (2017). Tinnitus functional index: validation of the german version for switzerland. *Health and quality of life outcomes*, 15:1–9.

Puel, J.-L., Mondain, M., and UzieP, A. (2002). Characteristics of tinnitus and etiology of associated hearing loss: a study of 123 patients. *International Tinnitus Journal*, 8(1):37–44.

Rabau, S., Wouters, K., and Van de Heyning, P. (2014). Validation and translation of the dutch tinnitus functional index. *B-ent*, 10(4):251–258.

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

Refat, F., Wertz, J., Hinrichs, P., Klose, U., Samy, H., Abdelkader, R. M., Saemisch, J., Hofmeier, B., Singer, W., Rüttiger, L., et al. (2021). Co-occurrence of hyperacusis accelerates with tinnitus burden over time and requires medical care. *Frontiers in Neurology*, 12:627522.

Roberts, L. E., Moffat, G., Baumann, M., Ward, L. M., and Bosnyak, D. J. (2008). Residual inhibition functions overlap tinnitus spectra and the region of auditory threshold shift. *Journal of the Association for Research in Otolaryngology*, 9:417–435.

Roberts, L. E., Moffat, G., and Bosnyak, D. J. (2006). Residual inhibition functions in relation to tinnitus spectra and auditory threshold shift. *Acta Oto-Laryngologica*, 126(sup556):27–33.

Rosenhall, U. and Axelsson, A. (1995). Auditory brainstem response latencies in patients with tinnitus. *Scandinavian audiology*, 24(2):97–100.

Rosing, S. N., Schmidt, J. H., Wedderkopp, N., and Baguley, D. M. (2016). Prevalence of tinnitus and hyperacusis in children and adolescents: a systematic review. *BMJ open*, 6(6).

Santacruz, J. L., Arnold, R., Tuinstra, J., Stewart, R. E., and Van Dijk, P. (2021a). Validation of a dutch version of the tinnitus functional index in a tertiary referral tinnitus clinic. *Heliyon*, 7(8).

Santacruz, J. L., De Kleine, E., and Van Dijk, P. (2021b). Investigating the relation between minimum masking levels and hearing thresholds for tinnitus subtyping. *Progress in Brain Research*, 263:81–94.

Santacruz, J. L., de Kleine, E., and van Dijk, P. (2023). Comparison between two self-guided tinnitus pitch matching methods. *Frontiers in Aging Neuroscience*, 15:1095178.

Savage, J. and Waddell, A. (2014). Tinnitus. *BMJ clinical evidence*, 2014.

Schaette, R. and Kempter, R. (2006). Development of tinnitus-related neuronal hyperactivity through homeostatic plasticity after hearing loss: a computational model. *European Journal of Neuroscience*, 23(11):3124–3138.

Schaette, R. and Kempter, R. (2009). Predicting tinnitus pitch from patients' audiograms with a computational model for the development of neuronal hyperactivity. *Journal of neurophysiology*, 101(6):3042–3052.

Schaette, R., König, O., Hornig, D., Gross, M., and Kempter, R. (2010). Acoustic stimulation treatments against tinnitus could be most effective when tinnitus pitch is within the stimulated frequency range. *Hearing Research*, 269(1-2):95–101.

Schecklmann, M., Landgrebe, M., Langguth, B., and Group, T. D. S. (2014). Phenotypic characteristics of hyperacusis in tinnitus. *PloS one*, 9(1):e86944.

Schecklmann, M., Pregler, M., Kreuzer, P. M., Poeppl, T. B., Lehner, A., Crönlein, T., Wetter, T. C., Frank, E., Landgrebe, M., Langguth, B., et al. (2015). Psychophysiological associations between chronic tinnitus and sleep: a cross validation of tinnitus and insomnia questionnaires. *BioMed research international*, 2015.

Schecklmann, M., Vielsmeier, V., Steffens, T., Landgrebe, M., Langguth, B., and Kleinjung, T. (2012). Relationship between audiometric slope and tinnitus pitch in tinnitus patients: insights into the mechanisms of tinnitus generation. *PloS one*, 7(4):e34878.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., and King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of educational research*, 99(6):323–338.

Schubert, N., Rosmalen, J., van Dijk, P., and Pyott, S. (2021). A retrospective cross-sectional study on tinnitus prevalence and disease associations in the dutch population-based cohort lifelines. *Hearing Research*, 411:108355.

Searchfield, G. D., Durai, M., and Linford, T. (2017). A state-of-the-art review: personalization of tinnitus sound therapy. *Frontiers in Psychology*, 8:1599.

Sereda, M., Hoare, D. J., Nicholson, R., Smith, S., and Hall, D. A. (2015). Consensus on hearing aid candidature and fitting for mild hearing loss, with and without tinnitus: Delphi review. *Ear and hearing*, 36(4):417.

Seydel, C., Haupt, H., Olze, H., Szczepek, A. J., and Mazurek, B. (2013). Gender and chronic tinnitus: differences in tinnitus-related distress depend on age and duration of tinnitus. *Ear and hearing*, 34(5):661–672.

Shargorodsky, J., Curhan, G. C., and Farwell, W. R. (2010). Prevalence and characteristics of tinnitus among us adults. *The American journal of medicine*, 123(8):711–718.

Shekhawat, G., Searchfield, G., and Stinear, C. (2013). Role of hearing aids in tinnitus intervention: a scoping review. *Journal of the American Academy of Audiology*, 24:747–762.

Shekhawat, G. S., Searchfield, G. D., and Stinear, C. M. (2014). The relationship between tinnitus pitch and hearing sensitivity. *European Archives of Oto-Rhino-Laryngology*, 271:41–48.

Shim, H. J., Kim, S. K., Park, C. H., Lee, S. H., Yoon, S. W., Ki, A. R., Chung, D. H., and Yeo, S. G. (2009). Hearing abilities at ultra-high frequency in patients with tinnitus. *Clinical and experimental otorhinolaryngology*, 2(4):169–174.

Shojaeemend, H. and Ayatollahi, H. (2018). Automated audiometry: a review of the implementation and evaluation methods. *Healthcare informatics research*, 24(4):263–275.

Sibbald, B. and Roberts, C. (1998). Understanding controlled trials crossover trials. *Bmj*, 316(7146):1719–1720.

Simoes, J. P., Daoud, E., Shabbir, M., Amanat, S., Assouly, K., Biswas, R., Casolani, C., Dode, A., Enzler, F., Jacquemin, L., et al. (2021). Multidisciplinary tinnitus research: challenges and future directions from the perspective of early stage researchers. *Frontiers in Aging Neuroscience*, 13:647285.

Smith, H., Fackrell, K., Kennedy, V., Barry, J., Partridge, L., and Hoare, D. J. (2019). A scoping review to catalogue tinnitus problems in children. *International journal of pediatric otorhinolaryngology*, 122:141–151.

Spinhoven, P., Ormel, J., Sloekers, P., Kempen, G., Speckens, A. E., and van Hemert, A. M. (1997). A validation study of the hospital anxiety and depression scale (hads) in different groups of dutch subjects. *Psychological medicine*, 27(2):363–370.

Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences*. Routledge.

Stracke, H., Okamoto, H., and Pantev, C. (2010). Customized notched music training reduces tinnitus loudness. *Communicative & Integrative Biology*, 3(3):274–277.

Suzuki, N., Oishi, N., and Ogawa, K. (2019). Validation of the japanese version of the tinnitus functional index (tfi). *International Journal of Audiology*, 58(3):167–173.

Swanepoel, D. W., De Sousa, K. C., Smits, C., and Moore, D. R. (2019). Mobile applications to detect hearing impairment: opportunities and challenges. *Bulletin of the World Health Organization*, 97(10):717.

Tsai, B. S., Sweetow, R. W., and Cheung, S. W. (2012). Audiometric asymmetry and tinnitus laterality. *The Laryngoscope*, 122(5):1148–1153.

Tunkel, D. E., Bauer, C. A., Sun, G. H., Rosenfeld, R. M., Chandrasekhar, S. S., Cunningham Jr, E. R., Archer, S. M., Blakley, B. W., Carter, J. M., Granieri, E. C., et al. (2014). Clinical practice guideline: tinnitus. *Otolaryngology–Head and Neck Surgery*, 151(2_suppl):S1–S40.

Turner, K., Moshtaghi, O., Saez, N., Richardson, M., Djalilian, H., Zeng, F.-G., and Lin, H. (2022). Auditory brainstem response wave i amplitude has limited clinical utility in diagnosing tinnitus in humans. *Brain sciences*, 12(2):142.

Tyler, R., Coelho, C., Tao, P., Ji, H., Noble, W., Gehringer, A., and Gogel, S. (2008). Identifying tinnitus subgroups with cluster analysis. *Age (years)*, 1(2):3.

Tyler, R. and Conrad-Armes, D. (1984a). Masking of tinnitus compared to masking of pure tones. *Journal of Speech, Language, and Hearing Research*, 27(1):106–111.

Tyler, R. S. and Conrad-Armes, D. (1983). Tinnitus pitch: a comparison of three measurement methods. *British journal of audiology*, 17(2):101–107.

Tyler, R. S. and Conrad-Armes, D. (1984b). Masking of tinnitus compared to masking of pure tones. *Journal of Speech, Language, and Hearing Research*, 27(1):106–111.

Van den Berge, M., Free, R., Arnold, R., De Kleine, E., Hofman, R., Van Dijk, J., and Van Dijk, P. (2017). Cluster analysis to identify possible subgroups in tinnitus patients. *Frontiers in Neurology*, 8:115.

Vander Zee, K. I., Sanderman, R., Heyink, J. W., and de Haes, H. (1996). Psychometric qualities of the rand 36-item health survey 1.0: a multidimensional measure of general health status. *International journal of behavioral medicine*, 3:104–122.

Vernon, J. (1977). Attempts to relieve tinnitus. *Ear and Hearing*, 2(4):124–131.

Vernon, J., Griest, S., and Press, L. (1990). Attributes of tinnitus and the acceptance of masking. *American journal of otolaryngology*, 11(1):44–50.

Vesterager, V. (1997). Fortnightly review: tinnitus—investigation and management. *Bmj*, 314(7082):728.

Vielsmeier, V., Lehner, A., Strutz, J., Steffens, T., Kreuzer, P. M., Schecklmann, M., Landgrebe, M., Langguth, B., Kleinjung, T., et al. (2015). The relevance of the high frequency audiometry in tinnitus patients with normal hearing in conventional pure-tone audiometry. *BioMed research international*, 2015.

Ward, J., Vella, C., Hoare, D. J., and Hall, D. A. (2015). Subtyping somatic tinnitus: a cross-sectional uk cohort study of demographic, clinical and audiological characteristics. *PLoS One*, 10(5):e0126254.

Ware Jr, J. E. and Sherbourne, C. D. (1992). The mos 36-item short-form health survey (sf-36): I. conceptual framework and item selection. *Medical care*, pages 473–483.

Wellek, S. and Blettner, M. (2012). On the proper use of the crossover design in clinical trials: part 18 of a series on evaluation of scientific publications. *Deutsches Ärzteblatt International*, 109(15):276.

Wise, K., Kobayashi, K., and Searchfield, G. (2015). Feasibility study of a game integrating assessment and therapy of tinnitus. *Journal of neuroscience methods*, 249:1–7.

Wrzosek, M., Szymiec, E., Klemens, W., Kotyło, P., Schlee, W., Modrzyńska, M., Lang-Małecka, A., Preis, A., and Bulla, J. (2016). Polish translation and validation of the tinnitus handicap inventory and the tinnitus functional index. *Frontiers in Psychology*, 7:1871.

Wunderlich, R., Stein, A., Engell, A., Lau, P., Waasem, L., Shaykevich, A., Rudack, C., and Pantev, C. (2015). Evaluation of ipod-based automated tinnitus pitch matching. *Journal of the American Academy of Audiology*, 26(02):205–212.

Zachariae, R., Mirz, F., Johansen, L. V., Andersen, S. E., Bjerring, P., and Pedersen, C. B. (2000). Reliability and validity of a danish adaptation of the tinnitus handicap inventory. *Scandinavian Audiology*, 29(1):37–43.

Zagólski, O. and Stręk, P. (2014). Tinnitus pitch and minimum masking levels in different etiologies. *International Journal of Audiology*, 53(7):482–489.

Zeman, F., Koller, M., Langguth, B., Landgrebe, M., and Group, T. R. I. D. S. (2014). Which tinnitus-related aspects are relevant for quality of life and depression: results from a large international multicentre sample. *Health and quality of life outcomes*, 12:1–10.

Zigmond, A. S. and Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta psychiatrica scandinavica*, 67(6):361–370.