Novel Approaches for Tinnitus Subphenotyping: Evidence Synthesis, Standardised Assessment, and Supervised and Unsupervised Machine Learning Applications

Eleni Genitsaridi

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

June 2021

Abstract

Clinical management of tinnitus is rather challenging and there is yet no cure for most tinnitus cases. It is speculated that tinnitus heterogeneity is hindering progress in scientific understanding and development of treatments. Phenotyping (i.e., assessment of observable characteristics) and subphenotyping (i.e., subgrouping based on differences in observable characteristics) are important for studying heterogeneous conditions like tinnitus. Identifying and defining clinically relevant tinnitus subphenotypes could help achieve transformational advances in the field. This dissertation reports the application of several advanced methodological approaches and has two main aims. The first aim is to contribute to an international standardisation of tinnitus assessment relevant for tinnitus phenotypic profiling and subphenotyping. The second aim is to further our understanding of tinnitus heterogeneity by investigating the presence of robust subphenotypes, consistent across multiple independent datasets.

Two chapters focus on the first aim. Chapter 2 reviews the literature, summarises current knowledge on tinnitus subphenotypes and identifies research gaps. It also summarises methods used so far and presents a novel framework of variable concepts that have been used for tinnitus subphenotyping. Chapter 3 describes the development of a self-report questionnaire intended to be used as a standard for tinnitus phenotyping. This questionnaire was developed through an international collaboration with tinnitus researchers from many centres. The questionnaire is already translated into 9 languages (Albanian, Dutch, French, German, Greek, Italian, Polish, Spanish, and Swedish) and is being used by multiple research teams as a tool for standardised tinnitus assessment.

The second aim is addressed in Chapters 4 and 5. Chapter 4 provides a detailed description of three tinnitus-specific datasets that were subsequently analysed in Chapter 5, and highlights commonalities and differences in the studied populations and the collected variables. Chapter 5 describes a novel data-driven approach for discovering tinnitus subphenotypes. This Chapter reports on a comprehensive unsupervised machine learning methodology applied to the three datasets. Findings indicate that this method was able to identify robust tinnitus subphenotypic patterns.

Finally, Chapter 6 relates the overall findings to the wider context of the published literature and presents suggestions and recommendations for future research. Age, sex, hearing ability, problems with sounds, symptoms of depression, and mandible problems

were highlighted as important variables for tinnitus subphenotyping and should be considered for assessment in future tinnitus studies. Overall, this work provides a basis for standardised tinnitus assessment in future studies and gives novel insights into the characteristics of tinnitus subphenotypes.

Declaration

I declare that this thesis and the work contained herein is my own work, except where indicated by referencing, and that no part of this thesis has been submitted elsewhere for any other degree or professional qualification.

Eleni Genitsaridi

Funding Statement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement number 722046

Acknowledgments

I would like to thank my supervisors Deborah Hall, Derek Hoare, and Theodore Kypraios for their guidance and support throughout my PhD studies. Also, many thanks to everyone at the British Tinnitus Association and the Karolinska Institute who welcomed me during my secondments; and especially to Christopher Cederroth for the scientific collaborations and for sharing the STOP dataset. Special thanks also to all my colleagues, mostly from the Hearing Sciences and the ESIT, for all the collaborations and the scientific and other interesting discussions. Socialising with people from so many different places of the world was a dream come true.

On a personal level, many thanks to my dear family and friends. To my parents Kosta and Stella, my sisters Yianna and Rena, and my life companion Yianni for their continuous support. To my dear friend Maryam for the special support in day-to-day PhD life. And to the rest of my beloved friends for being there. Also thanks to the cat Toula for playing with me during the breaks from writing. Sometimes I thought I would never finish as I would rather play with you forever. But then you would scratch me and send me back to work.

Finally, thanks to all humans who care for other beings, try for a better world, and stay positive. You motivate me.

List of Publications

Publications directly related to the PhD thesis

Genitsaridi, E., Hoare, D.J., Kypraios, T. and Hall, D.A., 2020. A Review and a Framework of Variables for Defining and Characterizing Tinnitus Subphenotypes. Brain Sciences, 10(12), p.938, doi: 10.3390/brainsci10120938.

Genitsaridi, E., Partyka, M., Gallus, S., Lopez-Escamez, J.A., Schecklmann, M., Mielczarek, M., Trpchevska, N., Santacruz, J.L., Schoisswohl, S., Riha, C., Lourenco, M., Biswas, R., Liyanage, N., Cederroth, C., Perez-Carpena, P., Devos, J., Fuller, T., Edvall, N.K., Hellberg, M., D'Antonio, A., Gerevini, S., Sereda, M., Rein, A., Kypraios, T., Hoare, D.J., Londero, A., Pryss, R., Schlee, W., Hall, D.A., 2019. Standardised profiling for tinnitus research: The European School for Interdisciplinary Tinnitus Research Screening Questionnaire (ESIT-SQ). Hearing Research, 377, pp.353-359, doi: 10.1016/j.heares.2019.02.017.

Additional publications achieved during the PhD studies

Genitsaridi, E., Kypraios, T., Edvall, N.K., Trpchevska, N., Canlon, B., Hoare, D.J., Cederroth, C. and Hall, D.A., The spatial percept of tinnitus is associated with hearing asymmetry: subgroup comparisons. Progress in Brain Research, in press (preprint available at: medrxiv.org/content/10.1101/2020.05.05.20073999v1).

Simoes, J.P., Daud, E., Shabbir, M., Amanat, S., Assouly, K., Biswas, R., Casolini, C., Dode, A., Enzler, F., Jacquemin, L., Joergensen, M., Kok, T., Liyange, N., Lourenco, M., Makani, P., Mehdi, M., Ramadhani, A., Riha, C., Santacruz, J.L., Schiller, A., Schoisswohl, S., Trpchevska, N., Genitsaridi, E., 2021. Multidisciplinary Tinnitus Research: Challenges and Future Directions from the Perspective of Early Stage Researchers. Frontiers in Aging Neuroscience, 13, p.179, doi: 10.3389/fnagi.2021.647285.

Edvall, N.K., Gunan, E., Genitsaridi, E., Lazar, A., Mehraei, G., Billing, M., Tullberg, M., Bulla, J., Whitton, J., Canlon, B. and Hall, D.A., Cederroth, C., 2019. Impact of temporomandibular joint complaints on tinnitus-related distress. Frontiers in Neuroscience, 13, p.879, doi: 10.3389/fnins.2019.00879.

Selected conference abstracts

Genitsaridi, E., Hoare, D.J., Kypraios, T., Hall D.A., 2019. An online survey to investigate the value of self-reported data for tinnitus subtyping. Basic Auditory Science Meeting, London, UK.

Genitsaridi, E., Kypraios, T., Hoare, D.J., Hall D.A., 2019. Discriminating tinnitus subgroups based on the audiometric profile. Tinnitus Research Initiative Conference, Taipei, Taiwan.

Genitsaridi, E., Hoare, D.J., Kypraios, T., Hall D.A., 2019. An online survey to investigate association between tinnitus-relevant characteristics. Tinnitus Research Initiative Conference, Taipei, Taiwan.

Genitsaridi, E., Kypraios, T., Hoare, D.J., Hall D.A., 2018. Relationship between audiometric and tinnitus characteristics: towards a variable set for subtyping tinnitus. British Tinnitus Association Annual Conference, Birmingham, UK.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Tinnitus

1.1.1 The health burden of tinnitus

Tinnitus, the perception of sound without any acoustic stimulus external to the head, is a very common and potentially debilitating condition. A systematic review showed that the prevalence of tinnitus ranged from 5.1 to 42.7% across different studies (McCormack et al., 2016). The highest figure (42.7%) was from a study investigating the prevalence of tinnitus in individuals older than 60 years (Gibrin et al., 2013). Prevalence of bothersome tinnitus also varied and was at least 1% of the population. One reason for these discrepancies, is that different studies use different definitions for tinnitus and bothersome tinnitus. Another reason is that there are many factors associated with tinnitus and, thus, prevalence varies depending on the characteristics of the groups sampled from the population. For example, prevalence of tinnitus increases with age peaking at 60-69 years old (Martinez et al., 2015; Shargorodsky et al., 2010). Also, tinnitus is associated with noise exposure and hearing loss (Norena, 2011b). Therefore, it is more frequent in populations exposed to these risk factors.

Considering the association between tinnitus and age, and the increasing life expectancy (Kontis et al., 2017; Mathers et al., 2015), we can speculate that the overall health burden from tinnitus will rise in the following years. Nevertheless, there are still many gaps in our understanding of tinnitus mechanisms and there is yet no cure for tinnitus in most cases. Despite the numerous available treatment options, very few are evidence based, and those approaches that are evidenced mainly involve specialised counselling (Hesse, 2016).

When trying to address any research question about tinnitus, there are major obstacles that need to be kept in mind including discrepancies in definitions among studies, difficulties in measuring tinnitus, and the complexity of tinnitus heterogeneity.

1.1.2 The problems of defining and measuring

1.1.2.1 The problem of defining

One issue that impedes tinnitus research is that, although tinnitus is a known condition since thousands of years, a standardised well-structured definition is still lacking (De Ridder et al., 2021 describe a recent attempt for an international consensus on tinnitus

definitions). One reason that makes reaching a commonly accepted definition difficult is that the same term is inconsistently used to refer to very heterogeneous cases. One of the main inconsistencies is whether somatosounds, meaning sounds generated in various parts of the head and neck with mechanisms related to blood flow, respiration, muscle contraction or movements of the temporomandibular joint (TMJ), are included (Henry et al., 2005). The presence of such sounds can usually, but not always, be confirmed with measurement instruments amplifying these sounds. For example, auscultation (the medical process of using a stethoscope to listen to sounds inside a body) can help doctors identify any body sounds that might lead to objective tinnitus. Some researchers consider such a condition as a tinnitus subtype and differentiate between objective (sound source within the body) and subjective tinnitus (no sound source) (Cima et al., 2019; Tunkel et al., 2014). Others refrain from including them in the tinnitus spectrum (Jastreboff, 1990). In this case, the definition of tinnitus usually includes a term emphasising that it is a 'phantom phenomenon' generated without any acoustic stimulus. Other terms that have been used for objective tinnitus include pseudo-tinnitus and somatosounds.

Another reason that a consensus definition of tinnitus is hindered is that most widely used definitions do not differentiate between tinnitus and other auditory hallucinations. Musical hallucinations and speech that does not make sense are generally considered as forms of tinnitus. The line is usually drawn when the phantom perception conveys meaning, which is suggestive of psychotic illness (Baguley et al., 2013). However, it has been argued that complex auditory hallucinations are not necessarily associated with psychiatric disorders (Jastreboff and Hazell, 2004). In Figure 1, I propose an algorithm towards defining subjective tinnitus.

A final remark about tinnitus definitions is that the most widely used definitions exclude tinnitus that is occasional and lasts less than 5 minutes (McCormack et al., 2016). This type of transient phantom noise perception is very common and, although its generating mechanisms are not known, it is considered normal rather than pathological (Dauman et al., 1991; Henry et al., 2005).

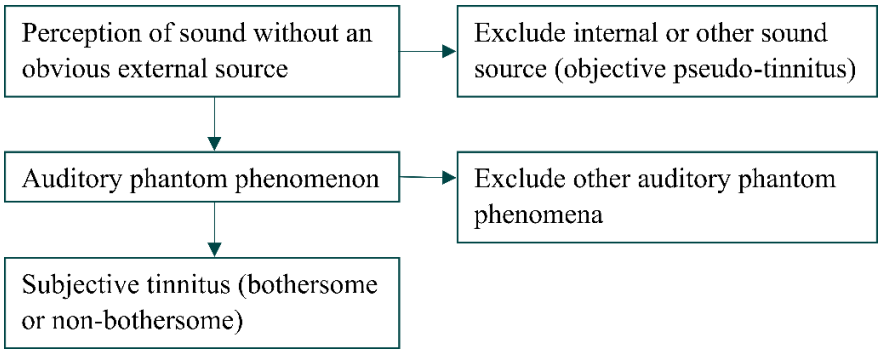

Figure 1. Diagnostic algorithm for subjective tinnitus.

1.1.2.2 A perception with many aspects

Another issue raising difficulties in investigating tinnitus mechanisms and treatment options is that tinnitus has many distinct aspects (or components) that are not clearly established. Any framework for the different aspects of tinnitus should at least separate the characteristics of the tinnitus perception from the impact it has on the individual. Shulman and Goldstein (2010) categorise tinnitus into the sensory component that refers to the 'subjective perception of the sensation of the aberrant auditory stimulus', the affect component that refers to 'the emotional response of the patient to the sensory component - tinnitus - as described by tinnitus patients', and the psychomotor component that refers to 'a patient's motor response accompanying the emotional and subjective perception of the sensation of the aberrant auditory stimulus' (pp. 76). De Ridder et al. (2014) explain that tinnitus is a perception rather than a sensation and note that 'a person's unified tinnitus percept is a complex percept encompassing multiple separable clinical cognitive and emotional aspects' (pp. 19). Considering these, the term tinnitus perception can be used to refer to the perceptual characteristics of tinnitus, the terms impact of tinnitus or reactions to tinnitus can be used to refer to changes in emotion or behaviour due to presence of a tinnitus, and the term tinnitus condition can be used to refer to all tinnitus-related aspects considering an individual with tinnitus. Recently, a publication authored by many tinnitus experts across the world proposed the use of the terms 'tinnitus' and 'tinnitus disorder' to differentiate between the presence of tinnitus 'without and with associated suffering' (De Ridder et al., 2021). However, such a binary classification should be used with caution considering that tinnitus severity can measured on a continuous scale.

1.1.2.3 The problem of measuring

Established instruments to measure and assess the various aspects of tinnitus are also lacking. Indicative of the difficulty of measuring tinnitus are the results of a systematic review that identified 35 domains and 78 instruments for measuring primary outcome in clinical trials for tinnitus (Hall et al., 2016). What is more, assessment of the subjective tinnitus perception depends on self-reports. Although objective markers (e.g., electrophysiological or neuroimaging) for the presence or impact of tinnitus are under investigation (Norena et al., 1999; Zimmerman et al., 2018), there are currently no established objective measures for tinnitus (Jackson et al., 2019).

1.1.3 Tinnitus heterogeneity: pathophysiologic mechanisms and clinical presentation

One of the main obstacles in tinnitus research and clinical management is that tinnitus is very heterogeneous in both underlying mechanisms and clinical presentation and there is no established method to define homogenous subtypes of tinnitus. Tinnitus perception is a symptom rather than a disease that can be associated with various underlying pathologies. As already discussed, tinnitus refers to at least two distinct pathophysiologic conditions, objective and subjective tinnitus. Objective tinnitus is rare, and the perceived sound can be generated by various sources (Table 1). Subjective tinnitus is far more common, but mechanisms for its generation have not been established yet. Aetiology of tinnitus can be examined at many levels (i.e., genome and environment, biological molecules, cells, and systems) and both peripheral and central neural mechanisms are likely to be involved in tinnitus generation (Eggermont and Roberts, 2004). Different mechanisms seem to be responsible for different aspects of the condition, e.g., tinnitus presence versus tinnitus severity (De Ridder et al., 2014; Hebert et al., 2012). However, it is still impossible to identify the exact pathophysiologic pathways that lead to subjective tinnitus. Therefore, tinnitus heterogeneity is usually examined in terms of phenotypic rather than pathophysiologic variability.

Tinnitus has been associated with various co-existing conditions, especially hearing-related conditions (including noise-induced hearing loss, presbycusis, Meniere's disease, and ototoxicity) but also various other conditions (including mood disorders, head or neck trauma, somatic disorders, psychological stress, and sleep disorders)

Table 1. Sources of objective tinnitus.

Vascular	Arterial or venous abnormalities (e.g., intra- and extra-cranial arteriovenous malformations, glomus tumour, carotid artery abnormalities, persistent stapedial artery, jugular bulb abnormalities, benign intracranial hypertension)
Musculoskeletal	Muscle contraction (e.g., palatal myoclonus, stapedial muscle spasticity), temporomandibular joint disorder, abnormally patent Eustachian tube
Other	Spontaneous otoacoustic emissions

References: Folmer et al. (2004), Jastreboff and Hazell (2004), Sismanis (2005), Lockwood et al. (2002)

(Baguley et al., 2013; Henry et al., 2005; Kim et al., 2018; Langguth et al., 2011b; Mahboubi et al., 2013; Nicolas-Puel et al., 2002; Norena, 2011b; Vernon and Press, 1994). The list of associated conditions is increasing (Yu et al., 2019). What is more, associations between tinnitus and other individual characteristics such as noise-exposure history, education, coping strategies, and personality traits have been shown in many studies (Budd and Pugh, 1996; Kim et al., 2015; Mahboubi et al., 2013; Martines et al., 2015; Shargorodsky et al., 2010; Wielopolski et al., 2017). The exact mechanisms for these associations (i.e., whether and how these factors affect the tinnitus perception and/or the impact of tinnitus) remain unclear. Nevertheless, it is common to assess presence of tinnitus-associated conditions and other individual characteristics for tinnitus phenotypic profiling (Evered and Lawrenson, 1981; Shulman, 1991; Tunkel et al., 2014; Zenner, 1998). Environment-related factors (e.g., noise exposure, social environment), genetic factors (e.g., family history), and factors related to body structure and functions (e.g., disorders of the ear, presence of depression, preceding head trauma, treatment history) are all variables that could be associated with the pathophysiology of tinnitus.

Tinnitus heterogeneity is also associated with differences in tinnitus characteristics such as loudness, pitch, timbre, localisation, and manifestation over time. Loudness can be assessed using visual or numerical rating scales, or with audiological loudness matching techniques (likeness to sample sounds at different intensities). Similarly, the quality of tinnitus can be specified by asking what tinnitus sounds like (e.g., hissing, ringing, pulsing, clicking, tonal, high pitch, low pitch) or by matching it to reference sounds. Psychoacoustic tinnitus pitch-matching techniques have evolved over the years. Recently proposed methods used standardised protocols taking into account that the

tinnitus bandwidth can be wider than a pure tone and have been shown to have sufficient stability across repeated measures (Moffat et al., 2009; Roberts et al., 2006). With regards to localisation, tinnitus can be perceived in one or both ears, inside the head or even in other locations in rare occasions. Moreover, manifestation of tinnitus over time can vary. Perception can be always present in a quiet room (constant tinnitus) or 'come and go' on its own (intermittent tinnitus) (Henry, 2016). Frequency of occurrence of intermittent tinnitus can also vary (e.g., daily, weekly, or monthly).

What is more, the impact of tinnitus and the reactions to tinnitus can take many forms. A systematic review identified 42 discrete complaints reported by people with tinnitus associated with physical and psychological health, quality of life and negative attributes of the tinnitus sound (Hall et al., 2018). Examples of such complaints include sleep difficulties, difficulties in hearing and other cognitive functions, annoyance, intrusiveness, worries, frustration, loss of sense of control, anxiety, and negative impact in social and work activities. To quantify the impact or severity of tinnitus is a rather complex task, and numerous scoring questionnaires have been developed that are usually structured into distinct concepts focusing on some of these complaints (Kuk et al., 1990; Meikle et al., 2012; Newman et al., 1996). Ideally, specific measures addressing each distinct impact domain should be developed.

In addition, it is important to assess the onset characteristics of tinnitus and its natural history since onset. Such assessments include age at onset, duration of tinnitus, onset related events (e.g., infection, head or neck trauma, noise exposure, psychological or physical stress, and other medical conditions or treatments), gradual versus abrupt onset, and whether the perceptual characteristics or impact have changed over time. These characteristics can be very different between individuals.

Finally, factors that modulate tinnitus are also important when examining tinnitus heterogeneity and trying to profile tinnitus. One example is the effects of treatment on tinnitus (Côté et al., 2019; Kloostra et al., 2019; Simoes et al., 2019). In addition, factors including somatic manoeuvres, changes in mental state (e.g., being stressed versus relaxed) and presence of external sounds have been shown to increase or reduce loudness or other aspects of tinnitus (Lugo et al., 2020; Schecklmann et al., 2014). Residual inhibition, the temporary reduction of tinnitus after a masking sound, is another variable that can differ across cases (Sanders et al., 2021). Differences in

modulating factors might imply different underlying mechanisms. For example, background sounds can have a masking effect in some patients, but for others external sounds aggravate tinnitus; these have been proposed to be distinct subtypes of tinnitus (Norena et al., 1999).

A summary of the various aspects of tinnitus heterogeneity is presented in Figure 2. Most of these variables can change over time. Therefore, heterogeneity can refer not only to different characteristics between different individuals (inter-individual variability), but also to changes in characteristics of the same individual over time (intra-individual variability).

1.1.4 The importance of tinnitus subphenotyping

Considering the complexity of tinnitus, the lack of methods to identify the exact underlying pathology, and the high variability between cases, a framework for subphenotyping tinnitus would be very beneficial for both scientific advancement and clinical management. For clarity of terminology, the term subtype is used for subgroups of people with tinnitus that can be used to guide treatment decisions, whereas the term subphenotype is used for subgroups of tinnitus with distinct observable characteristics. In addition to these, the term endotype can be used for subgroups with established distinct underlying pathological or functional mechanisms (Anderson, 2008; Lötvall et al., 2011; Saria and Goldenberg, 2015). Table 2 presents an overview of these terms and definitions used across this thesis. Subtypes and endotypes can also be defined on the bases of having distinct genotypes (Dahl and Zaitlen, 2020).

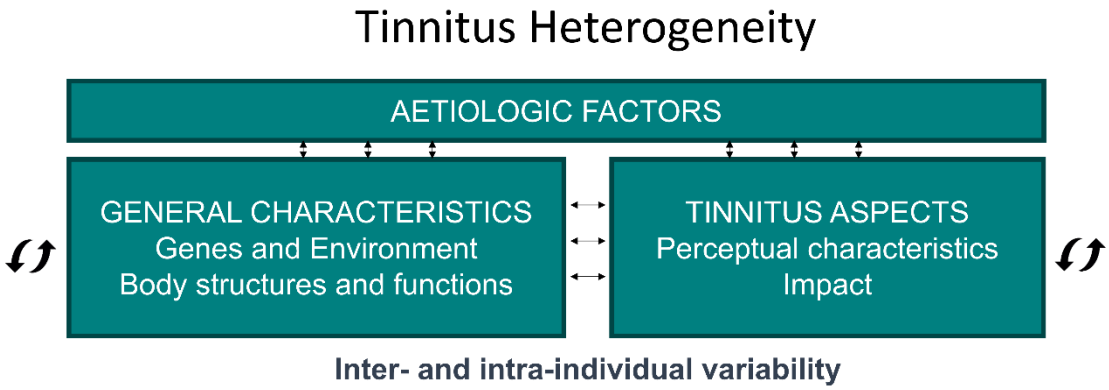

Figure 2. Aspects of tinnitus heterogeneity.

Table 2. Terminology used across this thesis.

Term	Definition
Subgroup	A part of a population (generic term)
Subphenotype	A part of a population with a distinct set of observable characteristics (based on Lötvall et al., 2011)
Subtype	A part of a population with a condition that can be used to guide treatment decisions (based on Saria and Goldenberg, 2015)
Endotype	A part of a population with a condition with distinct underlying mechanisms (based on Anderson, 2008; Lötvall et al., 2011)

It has been suggested that tinnitus heterogeneity is at least partially responsible for the lack of significant treatment outcomes in various clinical trials for tinnitus (Landgrebe et al., 2010). Tinnitus interventions might prove to be effective for specific subgroups, even if they have not been proven effective for the whole tinnitus population (D'Arcy et al., 2017). That means that identification of homogeneous subphenotypes could offer researchers a basis for designing and analysing results from clinical trials. Such knowledge could also be used for retrospective analysis of data from previous clinical trials which could have a high economic benefit considering the expenses associated with clinical trials.

In addition, knowledge of tinnitus subphenotypes could be very important for the design of studies seeking markers for tinnitus or seeking evidence for tinnitus mechanisms. Schmidt et al. (2017) suggested that tinnitus heterogeneity could be one of the factors responsible for inconsistencies across studies investigating resting state functional connectivity in tinnitus, measured using functional magnetic resonance imaging. In their study, Schmidt et al. (2017) grouped people with tinnitus based on tinnitus severity (measured with the Tinnitus Handicap Inventory [THI], Newman et al., 1996) and tinnitus duration (2 categories: long-term if more than 1 year or recent onset if less 1 year) and observed differences between subgroups in resting state functional connectivity. In another study investigating the heritability of tinnitus, researchers showed that heritability was much higher in people with bilateral tinnitus compared to unilateral tinnitus (Maas et al., 2017). These examples indicate that, even when researchers consider tinnitus heterogeneity for their analysis, the selection of variables to examine varies largely. Therefore, a framework for assessing and subphenotyping tinnitus could prove very beneficial for research design.

It is worth noting, that the requirements for a tinnitus assessment and subphenotyping framework differ depending on whether it would be used in clinical or research settings. In research settings, novel assessment tools are often available and assessment time can be adjusted to study needs. In clinical settings, on the other hand, the available time and measures for assessment are usually restricted compared to research settings. What is more a subphenotyping framework should have proven value in guiding assessment or treatment approaches. Nevertheless, such a framework aiming to facilitate research design would also benefit clinical practice. In the short term, it would provide a basis to identify subgroups of patients that may require specific care. For example, if a subgroup characterised by psychological comorbidities emerges, that could be an indication for treating patients belonging in this subgroup with psychology-based approaches. In the long term, a subphenotyping framework would improve clinical practice by advancing basic knowledge and treatment options.

1.1.5 Confounds hindering tinnitus subphenotyping

A review of the literature on tinnitus heterogeneity and subphenotyping identifies various confounds that can hinder tinnitus subphenotyping and subtyping. An overview of such confounds is presented in Table 3.

1.2 Machine learning applications to unravel clinical heterogeneity

1.2.1 Introduction

Statistical and computational advances have contributed a lot to medical research, helping generate new knowledge from large amounts of data. Particularly, machine leaning algorithms (i.e., computer algorithms that learn and improve automatically from the data) have been applied widely in multiple medical research fields. Two main categories are supervised and unsupervised learning algorithms (Theodoridis and Koutroumbas, 2009). In the first case, a target variable exists (known class labels), and the algorithm tries to learn how a set of predictor variables are associated with the target. In the second case, no target variable (a priori knowledge of class labels) is available. Instead, the algorithms try to find patterns based on the associations between all available data. Clustering algorithms, a common type of unsupervised learning algorithms, aim to categorise observations based on their similarity. Algorithms combining supervised and unsupervised learning (semi-supervised approaches) have also been developed. Indicative of the wide application of such techniques in medical

Table 3. Confounds hindering tinnitus subphenotyping and subtyping.

Confound	Impact
Complex aetiology (Evered and Lawrenson, 1981; Shulman, 1991): ▪ Many contributing aetiologies and moderating factors (e.g., personality, emotional resilience). ▪ Pathophysiologic mechanisms not always identifiable. ▪ Differences in individual characteristics can result in different tinnitus subphenotypes even if aetiology is the same.	➢ Difficult to subtype based on aetiology.
Many relevant variables (Dauman et al., 1991; Nodar, 1978; Shulman, 1991): ▪ Many different aspects of tinnitus (perceptual characteristics and impact to the individual) and other relevant individual characteristics. ▪ Some patients perceive more than one type of tinnitus that should be examined separately. ▪ Tinnitus-related characteristics can vary over time and depend on context (e.g., ambient sound, stress).	➢ Difficult to decide on a set of variables that are essential and sufficient to profile the tinnitus population. ➢ Results of any data analysis depend on the set of included variables.
Associations between variables (Hoekstra et al., 2014; Mahboubi et al., 2013): ▪ Some key variables are correlated with one another (e.g., age and degree of hearing loss).	➢ Results from studies that have not considered such correlations should be interpreted with caution.
Lack of standard definitions and measures (Hall et al., 2016; McCormack et al., 2016): ▪ No standardised definitions for tinnitus-related aspects (e.g., clinically relevant tinnitus, chronic tinnitus). ▪ Various measures and procedures used in different studies.	➢ Limited comparability across studies.
Non-representative sampling populations (Langguth et al., 2017b; Probst et al., 2017; Schecklmann et al., 2014; Vielsmeier et al., 2012): ▪ Many studies have recruited patients from specialised centres (e.g., tinnitus clinics), or with specific eligibility criteria (e.g., chronic tinnitus).	➢ Data not representing the whole tinnitus population.
Rare tinnitus subphenotypes (Levine and Oron, 2015): ▪ Some characteristic tinnitus subphenotypes (e.g., musical tinnitus) are only rarely encountered.	➢ Difficult to identify rare subphenotypes.

research are the results of a systematic review by Caballé et al. (2020) on the use of machine learning in diagnosing diseases, pattern discovery, and making predictions for medical research. They showed the wide application of multiple techniques (e.g., k-means clustering, support vector machines, and linear and logistic regression) across multiple medical research areas (e.g., cancer, metabolic diseases, and Parkinson's disease).

A lot of research focuses on the use of such methods towards understanding heterogeneity in medical conditions and personalising medicine (Feczko et al., 2019; Forte et al., 2019). A common aim is to identify subphenotypes or subtypes within a population having homogeneous characteristics. Unsupervised machine learning algorithms are very useful for this purpose. They have been applied widely across multiple medical fields as highlighted by multiple reviews such as for subphenotyping or subtyping Alzheimer disease (Ferreira et al., 2020), autism spectrum disorders (Wolfers et al., 2019), depression (Beijers et al., 2019), asthma (Horne et al., 2020), schizophrenia (Habtewold et al., 2020), and psychosis spectrum disorders in general (Green et al., 2019), to name but a few.

1.2.2 Clustering algorithms

Clustering can be very useful for many data analysis tasks. However, no perfect clustering algorithm exists yet. Thus, many different types of algorithms have been developed over the years (Xu and Tian, 2015). Three traditionally described broad categories are partitioning, hierarchical, and density-based methods (Halkidi et al., 2001). Partitioning-based methods divide observations in a prespecified number of clusters (subgroups), so that observations within subgroups are more similar than observations across subgroups. Popular algorithms in this category are k-means and partitioning around medoids clustering. In hierarchical (or connectivity-based) clustering, a hierarchical tree is built based on the similarity of all observations. This can start by having all observations in one group and then splitting until each observation is its own group (divisive approaches), or by having each observation in its own group and merging them until all are in one group (agglomerative approaches). This tree can be 'cut' in various positions, yielding different clustering results. Density-based clustering, on the other hand, decides on subgrouping based on regions of high density of observations. Another category, distribution-based (or model-based) clustering, offers a more statistical approach, trying to define clusters based on whether

observations are likely to belong to the same distribution (Fraley and Raftery, 2007). Other popular methods include deep learning approaches (unsupervised neural networks; Herrero et al., 2001), archetypal analysis (aiming to represent observations as combinations of extreme points; Eugster and Leisch, 2009), subspace clustering (discovering clusters that exist in multiple, potentially overlapping subspaces; Elhamifar and Vidal, 2013; Parsons et al., 2004), ensemble clustering (generating multiple clusterings from the same data and combining them into a final solution; Vega-Pons and Ruiz-Shulcloper, 2011), hybrid or semi-supervised approaches (combining supervised and unsupervised learning; Feczko et al., 2018), and multiple kernel clustering (combining multiple data sources; Mariette and Villa-Vialaneix, 2018; Yu et al., 2011).

Clustering algorithms have various parameters (arguments) for the user to specify. A common argument for most algorithms is the choice of a distance metric used to assess dissimilarity. Possible options include Euclidean distance, Pearson correlation distance, and Mahalanobis distance (Xu and Tian, 2015). Another example is the choice of linkage method for agglomerative hierarchical clustering. This refers to the method for assessing the similarity between clusters that will define which clusters will be grouped together. Possible options include single linkage (minimum distance between observations of two clusters), complete linkage (maximum distance), average linkage (average distance), and Ward linkage (aiming to minimise the variance within clusters) (Pina et al., 2020).

Application of different algorithms or the same algorithm with different arguments on the same data will likely produce different results (Rodriguez et al., 2019). The aim of the analysis can help decide on selecting algorithms and their arguments. Nevertheless, trying out different methods is a useful approach (Pierre-Jean et al., 2020). To enable application of many of the above mentioned methods multiple R packages have been developed (e.g., Hennig, 2020; Maechler et al., 2012).

1.2.3 Sampling and variable selection

The results of an unsupervised learning analysis will also be dependent on both the characteristics of the sampling population and the variables included in the analysis. The aim of the analysis will determine how to optimise these parameters.

Regarding sampling population characteristics, if the aim is for example to identify subphenotypes of people that are severely impacted by tinnitus, collecting data from specialised tinnitus clinics would be a logical approach. If, however, the question is whether there are distinct subphenotypes within the whole population experiencing tinnitus, then the general population should be sampled. The importance of the method for recruitment in the resulting tinnitus sample characteristics was investigate by Probst et al. (2017), looking at how samples recruited through outpatient clinics, self-help web platforms, or tinnitus mobile apps differed.

Equally important for the interpretation of the results is the selection of variables included in the analysis. Including or not relevant or irrelevant traits will be deterministic for the identified subgroups. To give a simple example, a clustering algorithm using hair length data from dogs and humans will not be able to differentiate dogs from humans. A variable reporting the presence or not of a tail would be much more suitable for this purpose. For variable selection, researchers can prespecify relevant variables using prior knowledge on the subject, and/or data driven techniques. Such techniques are more often studied for supervised machine learning application, but there are also numerous options for unsupervised learning (Alelyani et al., 2013). Depending on whether variable selection is conducted before and independently of the machine learning analysis or it is based on the quality of results of the algorithm, methods can be categorised into filter and wrapper, respectively. If there is a relevant target variable available, that can be used for a supervised learning analysis, the variables that are important for the classification based on this target can be identified. These variables can then be used for the unsupervised learning analysis, ensuring that the discovered subgroups are based on variables that are relevant for the task (Galili et al., 2014). For tinnitus research, one such example would be to identify variables that are important for differentiating people with or without tinnitus and use these to discover tinnitus subphenotypes in an unsupervised leaning approach.

1.2.4 Clustering Validation

Another important consideration is that most clustering algorithms will always provide a clustering solution, meaning that a partitioning of the data will be suggested even in cases where no such structure exists in the data. It is therefore essential to validate the goodness of any identified subgrouping. For unsupervised algorithms, the true labels are not known making validation more challenging. As a result, numerous validation

techniques have been suggested. Indicatively, the 'nbclust' R package has implemented 30 indices for clustering validation (Charrad et al., 2014).

Validation techniques can be classified into internal, external, and relative clustering validation (Theodoridis and Koutroumbas, 2009). Internal validation evaluates the clustering solution using the same data used to generate the clusters. External validation evaluates the clustering solution considering a prespecified structure. Relative validation is used to compare different clustering solutions. In a broader context, internal validation can refer to techniques based only on the information intrinsic to the data used for clustering, whereas external validation can refer to techniques based on previous knowledge about the data (Rendón et al., 2011). Just as different clustering algorithms might focus on different qualities to decide on subgrouping the data, internal validation measures are also quality specific. Compactness, connectedness, and separation are three main attributes defining the type of validation measure (Handl et al., 2005). Compactness or homogeneity can be represented by intra-cluster variance, connectedness by the presence of local densities and the distances of nearest neighbours, and separation by inter-cluster distance. Some measures combine these attributes, such as the silhouette width that simultaneously assesses the intra-cluster homogeneity and inter-cluster separation (Rousseeuw, 1987). In such cases, the benefit of having a single measure is contradicted by the information loss, since two clustering solutions might be better in different ways. Another important attribute for clustering validation is stability, meaning the consistency of the identified solutions in repeated applications of the method while resampling or perturbing the data (Handl et al., 2005). Assessing cluster stability is essential because meaningful clusters should be robust to non-essential changes in the data (Hennig, 2007).

If no true labels are available to validate results, external validation can be achieved by utilising other prior knowledge. In medical contexts, this could include assessment of subgroup characteristics with regards to important clinical attributes that were not used for clustering. In other words, clinical usefulness of identified subgroups can help validate results (Gamberger et al., 2017). If multiple datasets with overlapping information are available, replicability across datasets can also be used for validation (Planey, 2015).

1.3 Thesis aims and overview

Given the complexity of tinnitus heterogeneity and the importance of tinnitus subphenotyping, this dissertation aims to contribute to a better understanding of these domains. The term tinnitus is used with the generic meaning of the perception of sound in the absence of an external acoustic stimulus, whereas tinnitus phenotyping refers to the characterisation of the observable traits of an individual with tinnitus. Tinnitus subphenotyping refers to the identification of subgroups of people with tinnitus having a distinct set of observable characteristics. The impact of and reactions to tinnitus are investigated as attributes related to the tinnitus perception.

This thesis has two main aims. The first is to advance methods and contribute to the standardisation of tinnitus assessment relevant for tinnitus profiling and subphenotyping. The second is to contribute to our understanding of tinnitus heterogeneity by applying novel computational methods to discover robust tinnitus subphenotypes across multiple datasets. Chapters 2 and 3 focus on the first aim and Chapters 4 and 5 on the second aim.

Specifically, Chapter 2 describes a review of the literature summarising studies on tinnitus subphenotyping and identifying research gaps. This review led to the creation of a framework of variables that have been used for tinnitus subphenotyping. In addition, some variables that were shown to be important for this purpose across multiple studies were highlighted. Part of this work is published open access in Brain Sciences (Genitsaridi et al., 2020a).

Chapter 3 describes the development of a self-report questionnaire for standardised tinnitus phenotyping. This project involved the collaboration of an international team of tinnitus researchers from many centres, which I coordinated. The questionnaire is already translated into nine languages (Albanian, Dutch, French, German, Greek, Italian, Polish, Spanish, and Swedish) and is being used by multiple research teams as a tool for standardised tinnitus assessment. Part of this work is published open access in Hearing Research (Genitsaridi et al., 2019). The projects described in Chapters 2 and 3 happened in parallel chronologically, which is why the literature review influenced but was not directly used for the questionnaire development.

Chapter 4 provides a detailed description of three tinnitus-specific datasets, highlighting commonalities and differences in the studied populations and in the collected variables.

This allows an in depth understanding of the available data that was essential for informing the subsequent analysis. Chapter 5 describes that analysis and specifically a novel data-driven approach for discovering tinnitus subphenotypes. Various supervised and unsupervised machine learning approaches were applied to the three datasets. Analyses focused on robustness within and across datasets, and utilised information about both people with and without tinnitus. Through this approach robust tinnitus subphenotypes were identified and characterised.

Finally, Chapter 6 provides an overview of the main findings and relates these to the wider published literature. Recommendations about minimum standards for data collection and research directions for future tinnitus studies are also provided.

Chapter 2. Dimensions of tinnitus heterogeneity: Reviewing and synthesising evidence for tinnitus subphenotyping

2.1 Introduction

This chapter summarises evidence for variables that have been suggested in the literature as important for tinnitus subphenotyping and thus contribute to tinnitus heterogeneity. Since there is no single established methodology for defining subphenotypes or subtypes of medical conditions, the relevant published literature is very heterogeneous in terms of methods (Cianfrone et al., 2015; Saria and Goldenberg, 2015; van den Berge et al., 2017). It includes both articles suggesting profiling or subgrouping frameworks based only on the author's research or clinical expertise (experience-based approaches) and studies using statistical and computational methods to investigate tinnitus subphenotypes (evidence-based approaches). The methodology of these approaches is fundamentally different.

Experience-based frameworks for tinnitus subtyping or profiling are based on the expert's opinion about patterns of characteristics in subgroups of patients that can have some scientific or clinical interest (Nodar, 1996). These opinions can be based in previous research evidence. However, such information is usually not used in a systematic way in this type of work. On the other hand, evidence-based studies use statistical analyses of relevant data to characterise clinical subphenotypes. Given the advances in computational power in recent years, such methods can be very efficient in detecting patterns that clinicians might have not been able to observe (Krittanawong et al., 2017).

In this chapter, I describe a systematic review of studies that have used evidence-based methods to investigate tinnitus subphenotypes. Information regarding variables used across studies and their importance in tinnitus subphenotyping were extracted. Based on these, a framework of variables for tinnitus subphenotyping was created and the most prominent variables were highlighted. In addition, a comparison of selected articles describing experience-based frameworks for tinnitus profiling or subgrouping is presented.

2.2 Methods

To identify relevant evidence-based studies, PubMed was searched using the following search string: (Tinnitus(Title)) AND (profil*(Title/Abstract) OR subtyp*(Title/Abstract) OR subgroup*(Title/Abstract) OR class*(Title/Abstract) OR subphenotyp*(Title/Abstract) OR phenotyp*(Title/Abstract) OR type*(Title) OR group*(Title) OR cluster*(Title/Abstract) OR unsupervised(Title/Abstract)). This search syntax was successful in identifying six pre-defined target articles and so it was decided that a further search of additional electronic databases was unnecessary. To select articles describing experience-based frameworks, results of this search and other literature sources (e.g., references list of relevant articles from previous literature searches I had conducted) were consulted.

Analyses from evidence-based studies were classified into one of three categories:

i) Hypothesis-driven approaches refer to analyses where researchers predefined tinnitus subgroups based on one or more variables, hypothesising that these can define phenotypically distinct subgroups. Resulting subgroups were then compared against other characteristics.
ii) Data-driven approaches refer to analyses that used unsupervised machine learning for defining tinnitus subgroups.
iii) Treatment response approaches refer to analyses that based the definition of tinnitus subgroups on the patient's response to a tinnitus treatment, such as by subgrouping into responders versus non-responders.

For each eligible analysis, data were extracted about included variables (how these were measured, whether they were used for defining or for comparing subgroups, and whether the variable differed significantly across subgroups or was important for classification).

Each extracted variable was labelled with:

- A main domain name, being either tinnitus-specific (traits that can be measured only in people with tinnitus) or not tinnitus-specific (generic traits that can be measured in both people with and without tinnitus).
- A subdomain name (e.g., demographics and socio-economic characteristics, or tinnitus perceptual characteristics).

- A variable-concept name, which was common for variables measuring the same concept, even if the method for assessing it was different. For example, the variable concept 'hearing ability' could be measured with either audiological tests or self-reported questions.

Subdomain names and variable-concept names were assigned and updated in an iterative way while extracting data about included variables by regularly assessing the resulting list of labels. To reduce personal bias, the initial list developed in the end of the data extraction process was assessed by my supervisory team and changes were made as deemed appropriate.

R version 4.0.2 was used to analyse the extracted data (R Core Team, 2020). A framework of variables was created and visualised into a tree diagram using the 'collapsibleTree' function (Khan, 2018). Other R packages used included 'ggplot2' (Wickham, 2016), 'viridis' (Garnier, 2018), and 'flextable' (Gohel, 2020). The following frequency statistics were calculated for each variable:

- Number of studies that used it for subgrouping.
- Number of studies that, although did not use it for subgrouping, showed that it differed significantly among tinnitus subgroups or that it was important for classification models.
- Number of studies that assessed this variable but did not use it for subgrouping and did not find it to be important for subphenotyping.

If either condition 1 or 2 was true, the variable was referred to as 'important for subphenotyping'. A spreadsheet with all extracted data is available online as a supplement to Genitsaridi et al. (2020a).

2.3 Results

2.3.1 Overview of studies

2.3.1.1 Theoretical frameworks

Fourteen theoretical studies proposing frameworks for tinnitus patient profiling or subgrouping were identified through various sources and summarised (Appendix 2.1). The publication year of these 14 studies ranged from 1950 to 2016 (Briner, 1995; Cianfrone et al., 2015; Dauman and Tyler, 1992; Evered and Lawrenson, 1981; Goodhill, 1950; Henry, 2016; Jastreboff and Hazell, 2004; Langguth et al., 2011a;

Levine and Oron, 2015; Nodar, 1978; Nodar, 1996; Shulman, 1992; Tunkel et al., 2014; Zenner, 1998). This body of work did not report any novel data analysis but was drawn from authors' previous experience and knowledge. In general, the authors refer to broad variable concepts, without suggesting specific measures or metrics.

Although the exact frameworks varied from study to study some common aspects emerged. First, they all suggested assessing multiple traits to characterise tinnitus. Initially subtyping of tinnitus into objective or subjective was commonly suggested. Additionally, it was often proposed to consider factors related to the aetiology, triggering or modulation of tinnitus (e.g., comorbidities and onset related events), perceptual characteristics of tinnitus (e.g., localisation and patterns of presence), and factors related to the impact and reactions to tinnitus (e.g., coping and distress) for tinnitus profiling. Particularly, assessing hearing function (e.g., audiological assessment and presences of other hearing related problems such hyperacusis) was often proposed.

2.3.1.2 Evidence-based tinnitus subphenotyping

Sixty-four studies using statistical or computational methods for tinnitus subtyping were identified by the systematic review (Genitsaridi et al., 2020a). From these, 55, eight, and two studies used hypothesis-driven, data-driven, and, treatment-response-based approaches, respectively (one study described two types of approaches). The sample sizes ranged from 30 to 2838 participants.

Across the 55 hypothesis-driven studies (summarised in Appendix 2.2), multiple variables were used for subgrouping such as the hearing loss profile, presence of problems tolerating sounds, presence of somatic disorders, presence and type of headache, and tinnitus manifestation (constant or intermittent) (Al-Swiahb and Park, 2016; Andersson et al., 1999; Attias et al., 1995; Bartels et al., 2008; Bartels et al., 2010; Boecking et al., 2020; Burkart et al., 2019; Carpenter-Thompson et al., 2015; Cederroth et al., 2020; De Ridder et al., 2011; Edvall et al., 2019; Erlandsson et al., 1991; Han et al., 2020; Heijneman et al., 2013; Hiller and Goebel, 2007; Holgers et al., 2005; Jeon et al., 2016; Kim et al., 2016; Kirsch et al., 1989; Kojima et al., 2017; Koops et al., 2019; Kreuzer et al., 2012; Langguth et al., 2017a; Lindblad et al., 2011; Lugo et al., 2020; Michiels et al., 2015; Michiels et al., 2019; Milner et al., 2020; Mores et al., 2019; Moring et al., 2016; Niemann et al., 2020a; Niemann et al., 2020b; Niemann et al., 2020c; Oostendorp et al., 2016; Peters et al., 2020; Ralli et al., 2016; Ralli et al., 2018;

Sand et al., 2012; Schecklmann et al., 2014; Schecklmann et al., 2012; Schmidt et al., 2017; Schmidt et al., 2018; Simões et al., 2019; Song et al., 2013; Song et al., 2015; Suzuki et al., 2018; Sztuka et al., 2010; Van der Wal et al., 2020; Vielsmeier et al., 2011; Vielsmeier et al., 2015; Vielsmeier et al., 2012; Wang et al., 2018; Ward et al., 2015; Watabe et al., 2020; Yüksel et al., 2018). Most studies used only univariate tests for statistically significant differences to compare the subgroups against other variables. The variables used to compare subgroups differed between studies.

The eight studies using data-driven approaches to subphenotype the tinnitus population are summarised in Appendix 2.3 (Andersson and McKenna, 1998; Erlandsson et al., 1991; Langguth et al., 2017b; Newman et al., 1997; Rizzardo et al., 1998; Schecklmann et al., 2012; Tyler et al., 2008; van den Berge et al., 2017). Across these studies, the identified subgroups were highly heterogeneous, which is expected given the largely different study designs. For example, the sampling population differed across studies, with recruitment including patients visiting audiology clinics, patients visiting specialised tinnitus clinics, or participants from tinnitus clinical trials. The set of included variables was also very different across studies. All but one study depended solely on researcher's decision for the selection of included variables. The exception used Principal Component Analysis (PCA) for variable selection (van den Berge et al., 2017). Despite heterogeneity in the methodology and results, some variables were found to discriminate subgroups across more than one study (i.e., findings were replicable). These variables included tinnitus laterality, effect of external sounds, sound tolerance, and audiometric characteristics. Independent replication increases confidence in these findings.

The two studies using treatment response approaches are summarised in Appendix 2.4 (Côté et al., 2019; Kloostra et al., 2019). These examined the effect of physiotherapy and cochlear implantation on tinnitus, respectively. Response to treatment was quantified either using the THI before and after treatment or using prespecified questions asking about any change in tinnitus after treatment. Table 4 gives one illustrative example of the study summaries presented in appendices 2.2-2.4 for each study type.

Table 4. Examples of study summaries presented in appendices 2.2, 2.3, and 2.4.

Example of a hypothesis-driven study summary (Appendix 2.2)						
Reference	Sample size	Sample characteristics	Subgroup definitions	Statistical methods	Significant level / Importance definition	
Lugo et al. (2020)	2539	People with tinnitus from a population-based cohort	2 subgroups based on presence of headache: 1.Yes, 2. No.	chi-squared test, Wilcoxon's test, Cliff's δ	$p \leq 0.05$, Benjamini and Hochberg method for p-value correction	
Example of a data-driven study summary (Appendix 2.3)						
Reference	Sample size	Sample characteristics	Variables for clustering	Algorithm (method for selecting number of subgroups)	Subgroup descriptions	Validation
Tyler et al. (2008)	153	Tinnitus patients enrolled in clinical trials	51 variables: demographics, symptomatology, case history variables, questionnaire scores, psychoacoustic tinnitus measures. Authors did not explicitly mention all variables.	2-step cluster analysis (tested 4-6 cluster solutions, choice of 4 because it resulted in about equal subgroup sizes)	4 subgroups: 1. Loud, persistent, and distressing tinnitus, suffering from loudness hyperacusis 2. Varying tinnitus pitch and loudness, worse in noise 3. Copers, tinnitus not influenced by touch 4. Copers, tinnitus worse in quiet and better in noise	None
Example of a treatment-response study summary (Appendix 2.4)						
Reference	Sample size	Sample characteristics	Treatments	Outcomes and subgrouping		
Kloostra et al. (2019)	61	Patients with tinnitus who received a cochlear implant from a University medical centre	Cochlear implantation	Two subgroups based on whether cochlear implantation had or had not a positive effect on tinnitus. Details: In response to the question 'Were there any changes concerning your tinnitus after the cochlear implantation?', responses 'Yes, the tinnitus I experienced before implantation disappeared after implantation.' and 'Yes, my already existing tinnitus got better after implantation' were considered positive response. Responses 'No, my already existing tinnitus remained the same after implantation' and 'Yes, my already existing tinnitus got worse after implantation' were considered no positive response.		

2.3.2 Framework of variable concepts for tinnitus subphenotyping

From each included study, information about variables were extracted in separate rows on a standardised data extraction file. Data for 2559 variables were extracted. These were categorised into 15 subdomains (8 non-tinnitus-specific and 7 tinnitus-specific) and further into 94 variable concepts to create a framework of variables that have been used for tinnitus subphenotyping as shown in Figure 3. This figure highlights the high dimensionality of tinnitus phenotypic heterogeneity. Variable concepts included variables measuring the same concept even if the exact measure used for assessing these was different.

Figure 4 shows bar plots for all variables that were shown to be important for subphenotyping in at least two studies. Overall, the top five variables, if ordered by number of studies that showed these to be important for subphenotyping, were tinnitus overall severity, hearing ability, depression symptomatology, age, and sex. As described in section 2.2., a variable was considered important for subphenotyping if it was used for subgrouping, it significantly differed among subgroups, or it was important for classification models. The green portion of each bar in Figure 4 represents studies that used each variable for defining subgroups. If variables were ordered only based on how many times they were used to define subgroups, the top eight variables were overall severity, hearing ability, problems with external sounds, depression symptomatology, effect of somatic manoeuvres on tinnitus, tinnitus localisation, problems with the mandible, and various somatic symptoms (five variables shared 4th position). The purple portion of each bar in Figure 4 represented studies in which the variable differed significantly between subgroups or was important for classification models. If ordering was based on this, then the top six variables were age, overall severity, depression symptomatology, hearing ability, sex, and impact of tinnitus on emotion and mental health (two variables shared 4th position). Considering these alternative approaches for variable ranking, 11 variables could be highlighted as important for tinnitus subphenotyping. These are summarised in Table 5.

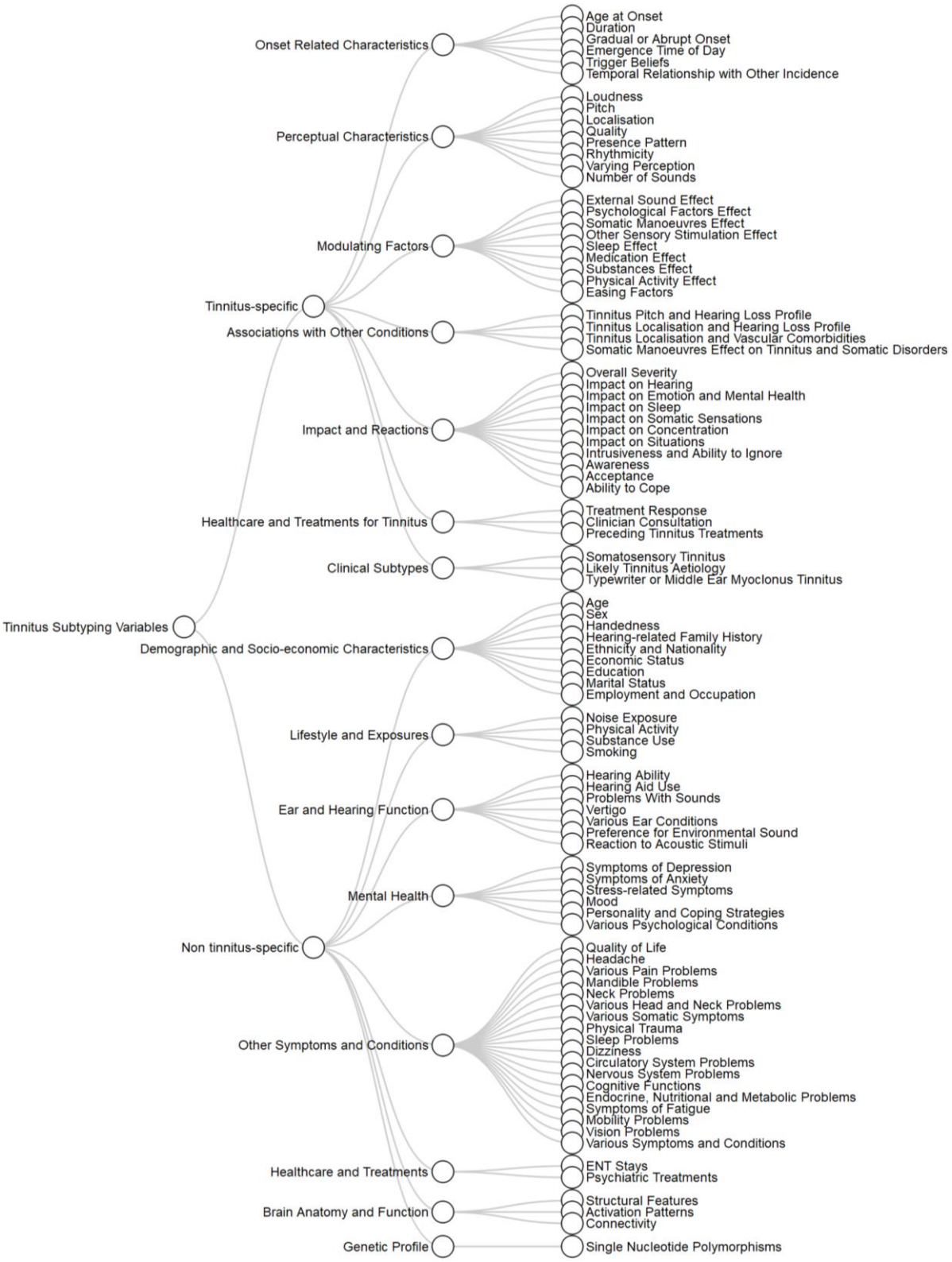

Figure 3. Framework of variable concepts for tinnitus subphenotyping.

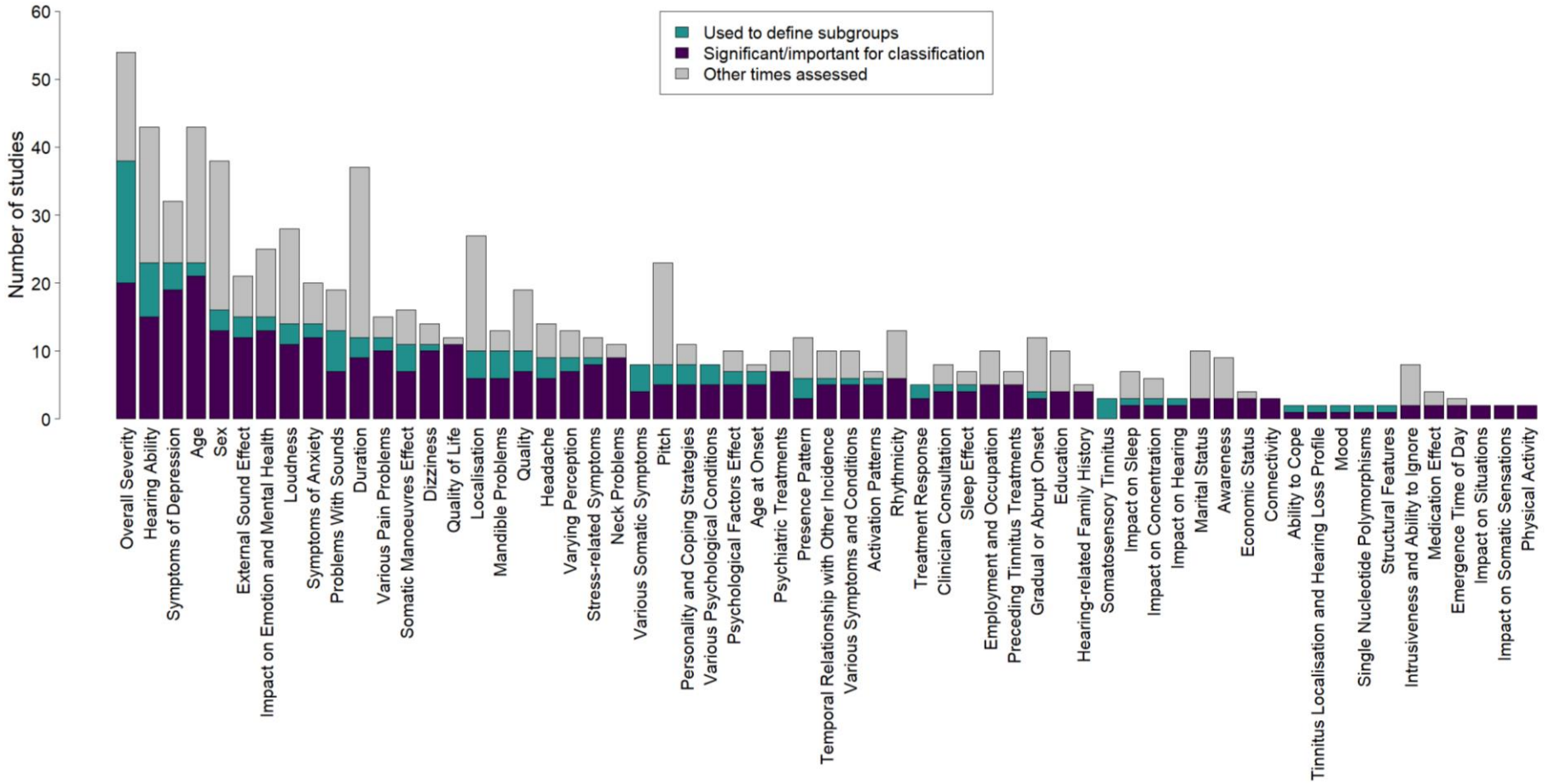

Figure 4. Variables that have been shown to be important for subphenotyping in at least two studies. The green portion of each bar represents studies that used each variable for defining subgroups, the purple portion represents studies in which the variable differed significantly between subgroups or was important for classification models, and the grey portion represents studies that used the variable in their analysis but did not show it to be important for subphenotyping.

Table 5. Eleven variables highlighted as important for tinnitus subphenotyping. Order of variables is as they appear in Figure 4.

Overall severity
Hearing ability
Symptoms of depression
Age
Sex
Impact of tinnitus on emotion and mental health
Problems with sounds
Somatic manoeuvres effect
Localisation
Mandible problems
Various somatic symptoms

Tinnitus overall severity was found to be the most assessed variable and the one most often shown to be important for subphenotyping. This variable was assessed using various tinnitus-specific questionnaires that measure multiple aspects of tinnitus symptom severity. Examples include the Fear of Tinnitus Questionnaire (FTQ) and the Tinnitus Catastrophising Scale (TCS) (Cima et al., 2011), the Tinnitus Questionnaire (TQ) (Hallam et al., 1988), the Tinnitus Handicap Questionnaire (THQ) (Kuk et al., 1990), the Tinnitus Functional Index (TFI) (Meikle et al., 2012), the THI (Newman et al., 1996), and the Tinnitus Reaction Questionnaire (TRQ) (Wilson et al., 1991). In addition, tinnitus overall severity was sometimes measured using a specific question asking about the overall impact of tinnitus (e.g., How much of a problem is your tinnitus?; Langguth et al., 2017b). Hearing ability was also assessed in various ways; most commonly using pure tone audiometry but also with other audiological measures (e.g., Distortion Product Otoacoustic Emissions [DPOAE] and speech in noise tests) and self-reported questions. Symptoms of depression were assessed with various self-reported questionnaires such as the Hospital Anxiety Depression Scale (HADS) (Zigmond and Snaith, 1983) and the Beck Depression Inventory (BDI) (Beck et al., 1961). Details for the different measures used to assess variable concepts can be found in Supplementary Table 3 in Genitsaridi et al. (2020a).

2.4 Discussion

2.4.1 Overview of results

This study provided a comprehensive review of the literature aiming to investigate tinnitus subphenotypes using evidence-based methodologies, including 64 relevant

articles. For comparison, selected theoretical studies suggesting tinnitus subtyping frameworks based on experience were also reviewed.

The main novel contribution of this study is the creation of a framework of variables that have been used for tinnitus subphenotyping, which can be used to guide the design of future studies. Despite the many tinnitus subphenotyping variables identified, quantitative analysis of the extracted data from evidence-based studies were used to highlight eleven variables as the most promising based on current literature. These were overall tinnitus severity, hearing ability, symptoms of depression, age, sex, having problems with sounds, effect of somatic manoeuvres on tinnitus, tinnitus localisation, having mandible problems, having various somatic symptoms, and the impact of tinnitus on emotion and mental health. Future studies investigating tinnitus subphenotypes should consider these variables in their protocols.

2.4.2 Methodological aspects

Studies proposing theoretical frameworks for tinnitus subtyping or profiling can be useful by providing researchers with a basis to choose which variables to assess when conducting tinnitus studies. Nevertheless, the proposed variables are not validated with research data. Evidence-based studies come to fill this gap. These were grouped into three broad categories based on their methodology in this review. The first and most common category was hypothesis-driven studies, where researchers chose a variable or a set of variables to define subgroups, that were then compared against several other clinically relevant characteristics. Fifty-five such studies were included in this review. The choice of variables used for subgrouping depended on the knowledge and experience of the researchers. The second category was data-driven studies using unsupervised learning algorithms to identify tinnitus subphenotypes. Eight such studies were included in this review. Across these, five validated their clusters by comparing subgroups against characteristics that were not used by the clustering algorithm (Andersson and McKenna, 1998; Erlandsson et al., 1991; Langguth et al., 2017b; Newman et al., 1997; Rizzardo et al., 1998). None of the eight studies explored stability of the identified clusters within dataset (e.g., using bootstrapping) or across independent datasets. The third category was treatment response studies, defining subgroups based on responses to specific tinnitus treatments and subsequently assessing their distinct phenotypic characteristics. Only two such studies were included in this review. However, it should be noted that the search methodology was not optimised to identify

subphenotypes linked to specific treatment responses. For example, treatment response is often measured with numerical variables, whereas this review included only studies investigating clearly defined subgroups.

The advantage of hypothesis-driven studies is that if the resulting subgroups have distinct characteristics the researcher's original hypothesis is validated. Nevertheless, these studies use only one or a few variables to define subgroups, while it is far more likely that a diverse set of variables would be required to discriminate clinically significant tinnitus subtypes. This has been shown to be true in other conditions such as diabetes and asthma (Ahlqvist et al., 2018; Hamilton and Lehman, 2020).

Data-driven studies on the other hand, can consider multiple variables simultaneously and can thus be very useful in advancing understanding for tinnitus heterogeneity. Results, however, are highly dependent on the details of the methodology used, including choice of the algorithm, method for selecting number of clusters, and selection of included variables (Horne et al., 2020; Pina et al., 2020). Therefore, the resulting subgroups might not be clinically relevant. This underlines the importance of externally validating the results of such data-driven approaches (Marquand et al., 2016), which is a step that is frequently overlooked. External validation requires data that where not used for the discovery of subgroups. These could be either additional variables on the same participants used to characterise identified subgroups (the most common method applied in tinnitus and other fields), or additional participants whose data can be used to investigate if the originally discovered subphenotypes are replicable (Beijers et al., 2019; Horne et al., 2020). The second method has not been applied in tinnitus research so far but is used here in this thesis work (see Chapter 5). Another way of external validation would be to investigate stability of identified subgroups over time. This approach is much more challenging as it requires longitudinal data. This was not possible within the constraints of a PhD, but some interesting applications have started to emerge in other (non-tinnitus) fields (Boudier et al., 2013; Dwyer et al., 2020; Lee and Van Der Schaar, 2020; Lim et al., 2020; Stronge et al., 2019).

Regarding treatment-response studies, one challenge with this approach is that subgroups and their distinct characteristics are treatment-specific. Considering the huge number of treatment options for tinnitus, designing a prospective study to explore the effect of the same variables on the outcome of a representative set of treatments for

tinnitus is very challenging. Standardisation in tinnitus research with regards to the set of variables that are collected at baseline would allow to identify potential treatment predictors across various studies investigating different tinnitus interventions in a prospective set up. An effort to promote standardised tinnitus assessment using a self-report case history questionnaire is described in the next chapter.

2.4.3 Variable aspects

Regardless of the study design, the choice of variables that are included in an analysis in part determines the results obtained, since a variable can only be proven important if it is measured. Results from this review showed that researchers chose many different variables in their studies investigating tinnitus subphenotypes. Even across studies examining data within the standardised Tinnitus Research Initiative (TRI) database (Kreuzer et al., 2012; Langguth et al., 2017a; Schecklmann et al., 2014; Schecklmann et al., 2012; Vielsmeier et al., 2011; Vielsmeier et al., 2015; Vielsmeier et al., 2012), there were many differences in the set of included variables (data not presented). This observation further emphasises the need for standardisation of tinnitus assessment protocols that would enable accurate comparisons of findings among independent studies.

In addition, some variables were under investigated across the included studies. For example, genetic and neuroimaging markers were not often assessed in the reviewed studies, which is also the case for other disciplines such as depression (Beijers et al., 2019). Although such data are more challenging to collect in terms of time, cost, and required equipment, efforts to include them in future studies should be reinforced considering their objective nature and their potential significance in pinpointing to disease endotypes (Luo et al., 2020).

Looking at the variables included in theoretical studies compared to evidence-based studies, in both approaches it was commonly suggested to assess multiple phenotypic traits to profile people with tinnitus. In theoretical frameworks, differentiating objective and subjective tinnitus was a common suggestion. In evidence-based approaches, however, this important trait was not always considered (e.g., Niemann et al., 2020a). Often, nevertheless, having either subjective or objective tinnitus was an inclusion criterion (e.g., Vielsmeier et al., 2015). Objectivity of tinnitus is ideally assessed by a researcher or clinician, usually by auscultation. However, for studies relying completely

on self-reported information this would not be possible. An alternative self-report solution could be by using specifically formulated questions, as will be described in the next chapter.

2.4.4 Limitations

A limitation of this review was that it only included studies with tinnitus subphenotyping being a main objective (primary or secondary). There are far more studies reporting an exploratory analysis relevant for tinnitus subphenotyping. However, these analyses are expected to have less rigorous methodologies. In addition, many studies investigating tinnitus heterogeneity do so using continuous scales rather than distinct subgroups. Considering tinnitus heterogeneity in a dimensional rather than categorical perspective might be beneficial for some traits, as is being discussed for mental disorders in general (Widiger and Samuel, 2005), but this remains to be proven (van den Berge et al., 2017). Inclusion of such studies would lead to a more comprehensive review. However, it would require much greater effort and a different study design. A future review focusing on studies investigating tinnitus heterogeneity in a dimensional perspective could yield interesting findings, especially in comparison to the findings of this review.

Another limitation was that we did not introduce any critical evaluation of the study design for the included studies. Such criteria would improve the overall confidence in the evidence synthesis and a 'risk of bias' assessment is common to many systematic reviews of interventions (Moseley et al., 2009). The reason such critical evaluation was not included in this review is that there is no established methodology. Instead, risk of bias criteria are more common to clinical interventions and include assessments such as allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data and selective reporting. Another reason is that the aim of this study was to chart information rather than synthesise the evidence. Typically, studies that chart (or 'scope') the evidence do not critically appraise the sources of evidence.

2.4.5 Conclusions

In this project, a framework of variables that have been used in previous studies investigating tinnitus subphenotypes was created. This framework can serve as a reference for future studies. A consensus across the tinnitus scientific community

regarding which of these variables should be assessed in every future study would be extremely beneficial for future research, allowing comparisons of results across studies. Importantly, considering studies included in this review used a variety of methods to measure the same variable concept, consensus needs to be reached also with regards to how these variables should be measured. In the next chapter, Chapter 3, the development of a tinnitus profiling questionnaire aiming to contribute to the standardisation of the collection of self-reported, tinnitus-relevant information is presented. In addition, this review showed clearly that there are numerous variables associated with tinnitus heterogeneity. Identifying the subsets of these variables that are essential and sufficient to characterise tinnitus subtypes or subphenotypes is a challenging task that would require simultaneous analysis of multiple domains. The aim of Chapters 4 and 5 is to identify important variables for tinnitus subphenotyping using such analyses.

2.5 Chapter contributions

All work presented in this chapter is my own contribution.

Chapter 3. Development of a tool for standardised self-report tinnitus profiling: the ESIT-SQ

3.1 Introduction

The initial motivation for this project was to enable coherent data collection across the different projects forming the European School for Interdisciplinary Tinnitus Research (ESIT). The ESIT is a European Union funded Marie Skłodowska-Curie Innovative Training Network including 15 tinnitus-related PhD projects based in 12 European cities (Granada, Groningen, Lodz, Maastricht, Marseille, Milan, Nottingham, Regensburg, Salzburg, Stockholm, Ulm, and Zurich) (Schlee et al., 2018). One of ESIT's main aims was to provide novel insights into tinnitus heterogeneity and to create a tinnitus profiling framework. One objective towards this aim was the creation of a large multinational tinnitus-specific database. To enable collection of common information across the different ESIT projects, a standardised self-report questionnaire, available in many European languages, was proposed in the original grant application and the deliverable was assigned to my project.

More broadly, a questionnaire suitable both for screening for the presence of tinnitus and for profiling different tinnitus characteristics would fill an important research gap. Although some case history questionnaires for profiling people with tinnitus have been published (Langguth et al., 2007; Schechter and Henry, 2002; Stouffer and Tyler, 1990), these are not structured in a way to be applicable to the population without tinnitus. In addition, these questionnaires do not always capture sufficient self-report information for tinnitus profiling such as medical history and lifestyle-related information.

Therefore, the aim of this chapter was to develop a set of questions that would enable standardised self-reported data collection relevant to tinnitus assessment, considering the multiple dimensions of tinnitus heterogeneity. One further criterion for the instrument was that it should be able to be administered to people with and without tinnitus. These questions would be organised into a questionnaire, namely the ESIT Screening Questionnaire (ESIT-SQ). This would allow its use both as a screening tool for the presence of tinnitus, and as a profiling tool for the characterisation of unique individual subphenotypes.

3.2 Methods

3.2.1 ESIT-SQ development

Development of the ESIT-SQ was a collaborative project across the ESIT consortium, under my coordination. The first step of the process was the creation of an initial list of questions and response options, as potential items of the questionnaire. In total this initial list comprised 41 items. Various sources were consulted to create this list including:

i) A widely used tinnitus case history questionnaire (Tinnitus Sample Case History Questionnaire - TSCHQ, Langguth et al., 2007). Development of the TSCHQ was also based on an international consensus for tinnitus assessment, thus, most of its items were considered for the initial list (either unchanged or adapted), including questions about tinnitus onset, perceptual characteristics, and tinnitus modulating factors.

ii) An unpublished tinnitus case history questionnaire in French previously developed by an otorhinolaryngologist (Alain Londero, personal communication). This questionnaire is used at the Georges Pompidou tinnitus hyperacusis clinic (Paris) for self-completion before the first medical appointment and contains a comprehensive set of items with diagnostic utility for tinnitus. It was first translated into English by an ESIT colleague in Salzburg (Marta Partyka). From this translated version, some of the questions were modified for further usage. Many questions, mainly on medical and tinnitus history, were adapted and included in the initial list of questions for the ESIT-SQ.

iii) Questions from an ongoing European epidemiological survey for tinnitus which was also a project in the ESIT consortium (Biswas et al., 2019). Four questions relevant to tinnitus were adapted to fit the flow of the new questionnaire (Table 6).

Table 6. Comparison of the four questions adapted from the European epidemiological survey for tinnitus (Biswas et al., 2019). Modifications shown in blue text.

ESIT-SQ	European epidemiological survey for tinnitus (Biswas et al., 2019)
Presence of tinnitus	
Tinnitus refers to the perception of noise in your head or ears (such as ringing or buzzing) in the absence of any corresponding source of sound external to your head. Over the past year, have you had tinnitus in your head or in one or both ears that lasts for more than five minutes at a time?	Over the past year, have you had noises (such as ringing or buzzing) in your head or in one or both ears that lasts for more than five minutes at a time?
Response options: Yes, most or all of time; Yes, a lot of the time; Yes, some of the time; No, not in the past year; No, never; Do not know	Response options: Yes, most or all of time; Yes, a lot of the time; Yes, some of the time; No, not in the past year; No, never; Do not know/ Prefer not to answer
Tinnitus severity	
Over the past year, how much does your tinnitus worry, annoy or upset you when it is at its worst?	Over the past year, how much do these noises in your head or ears worry, annoy or upset you when they are at their worst?
Response options: Severely; Moderately; Slightly; Not at all; Do not know	Response options: Severely; Moderately; Slightly; Not at all; Do not know/ Prefer not to answer
Use of healthcare resources for tinnitus	
Over the past year, have you seen your family doctor, or seen a healthcare professional at a clinic or hospital about your tinnitus?	Over the past year, have you seen your family doctor, or seen a healthcare professional at a clinic or hospital about problems with noises in your head or ears?
Response options: Yes, 5 or more visits; Yes, from 2 to 4 visits; Yes, just one visit; Not at all; Do not know	Response options: Yes, 5 or more visits; Yes, from 2 to 4 visits; Yes, just one visit; Not at all; Do not know/ Prefer not to answer
Presence of hearing difficulty	
Do you currently have any other difficulty with your hearing, such as listening to speech in a noisy situation?	Do you currently have any other difficulty with your hearing, such as listening to speech in a noisy situation?
Response options: Yes, cannot hear at all; Yes, severe difficulty; Yes, moderate difficulty; Yes, slight difficulty; No difficulty; Do not know	Response options: Yes, cannot hear at all; Yes, severe difficulty; Yes, moderate difficulty; Yes, slight difficulty; No difficulty; Do not know/ Prefer not to answer

Additional questions not available in the above-mentioned questionnaires were added in the initial list, including two questions on self-reported anthropometric characteristics, one question on the level of education, and three questions on smoking and alcohol habits, adapted from a European survey on tobacco smoking (Gallus et al., 2020). These were selected with the help of an ESIT colleague in Milan (Silvano Gallus) who is an epidemiologist. The abovementioned European epidemiological

survey for tinnitus (Biswas et al., 2019) was supplementary to the European survey on tobacco smoking (Gallus et al., 2020), with the inclusion of the four tinnitus relevant questions shown in Table 6, which were presented to participants after the questions of the tobacco survey. What is more, a question about the familial history of tinnitus and hearing loss was developed by the two ESIT colleagues with expertise in tinnitus genetics, one located in Stockholm (Christopher Cederroth) and one in Granada (Jose Antonio Lopez-Escamez). All sources used to develop items of the ESIT-SQ can be found in Appendix 3.1.

Members of the ESIT consortium, including both early career researchers and supervisors, were invited to review this list of 41 items and suggest changes, additions, or exclusions. Thirteen colleagues responded and provided a total of 71 comments. Based on their feedback, items were modified and sent back to them for a second round of consultation. As an example of feedback and the resulting modifications, the original question asking about factors reducing tinnitus did not include the effect of using hearing aids. This response option was therefore added after a comment from one of the reviewers. Examples of items added to the initial list include questions about handedness, employment, and economic status. The initial list also included six follow-up questions asking about the relationship between the onset of tinnitus and the onset of other conditions. These follow-up questions were combined into a single multi-part item.

Only one of the reviewers provided further comments leading to additional minor changes. For example, the first response option to the question ‘How often do you have tinnitus?’ was modified from ‘Almost daily’ to ‘Daily or almost daily’. After these minor changes, the list of questions and response options was finalised.

Seventeen items applicable to people with and without tinnitus were categorized into ‘Part A’ and 22 items applicable only to people experiencing tinnitus were categorized into ‘Part B’. In addition, a list of 17 optional questions for Part A was created. All optional question were suggestions for additional items coming from consultation. These questions were not considered essential to the main questionnaire, which needed to be kept relatively short, but were standardised to be available for use for individual projects. The final list comprised 56 items in total.

3.2.2 ESIT-SQ translations for ESIT projects

Final items of the ESIT-SQ, and instructions for navigating across the questionnaire (e.g., initial instructions for participants) were translated into six European languages: Dutch, German, Italian, Polish, Spanish, and Swedish. Languages were chosen based on requests from the various ESIT project leaders. To ensure that all translations followed the same methodology, a translation plan was designed based on available resources and a guideline for translating and adapting hearing-related questionnaires (Hall et al., 2018). This included the following main steps:

i) Preparation step including recruitment of three translators and informing them about the concepts of the questionnaire and the requirements of the translation. Translators had to be highly proficient in both the source and target languages (with at least one native speaker). A tinnitus expert and a healthcare practitioner, that could be one of the three translators, were part of the translation team. This corresponds to step 1 in Hall et al. (2018).

ii) Production of two independent forward translations, consolidated into a harmonized version by a third independent translation (in consultation with the first two translators when deemed necessary). This corresponds to step 2 in Hall et al. (2018).

iii) Final reviewing, proofreading, and formatting of the translated version. This corresponds to step 6 in Hall et al. (2018).

This plan was provided to ESIT colleagues who undertook the role of the translation coordinators (Roshni Biswas for the Italian translation, Jose Lopez-Santacruz for the Spanish translation, Marta Partyka for the Polish translation, Matheus Pereira da Cruz Gomez Lourenco for the Dutch translation, Stefan Schoisswohl for the German translation, and Natalia Trpchevska for the Swedish translation). Translation coordinators were either located in the country speaking the target language or were native speakers of the target language. Their role was to coordinate the translation process as described above and create the final translated paper version of the questionnaire. In addition to the translation plan, an excel template for recording translated items was created and provided to translation coordinators. This included concept descriptions for each item that were created to ensure concept and semantic equivalence with the original English version. An example of the format of the excel template is shown in Figure 5.

Text to be translated	Concept definition	Translation 1	Translation 2	Final	Notes
A5. What is the highest education level you have achieved?	A5 question. The highest form of educational institution is to be selected from the following 5 response options.				
1. No school 2. Primary (elementary school) 3. Lower secondary (middle school) 4. Upper secondary (high school) 5. University or higher degree	A5 response options. 1. No formal schooling 2. The first phase of compulsory education is primary school. It generally starts between the ages of 5 and 7 years and ends between 10 and 14 3. 'Lower secondary' school is defined as the phase that immediately follows primary level education, in those countries where the two levels are distinct from each other 4. The most common age of transfer from lower secondary to the next stage of secondary education (referred to as 'upper secondary') is between 14 and 16. Upper secondary education may fall outside the boundaries of compulsory education and generally caters for students up to the ages of 18 or 19 5. 'University' describes an optional education, usually leading to a degree qualification.				
A6. What is the average number of alcoholic drinks that you consume per week? __ drinks per week (1 drink= 125 ml glass of wine, 330 ml of beer, 40 ml of spirits)	A6 question and response options. The average amount of alcohol consumption should be recorded using a single number of drinks per week. An explanation is given about the volume that corresponds to 1 drink for the different types of alcoholic drinks.				

Figure 5. Example of the template used for the ESIT-SQ translations, following Hall et al. (2018).

3.2.3 Implementation of an online version

For the ESIT project, an online database for uniform data collection and data storing across ESIT was created, namely the ESIT database (ESIT, 2018). After finalising the ESIT-SQ and its translations, assistance was provided to the developers of the ESIT database to implement this questionnaire on the database. This involved consultation on automatically generating questions that are contingent on preceding responses (adaptive testing) and on the content of messages for the participants while navigating across the questionnaire, and extensive testing for identifying bugs and suggesting improvements.

3.2.4 Guide for further translations

To ensure that future translation of the ESIT-SQ would also follow a common methodology, a document to be shared with anyone interested in translating the ESIT-SQ was created. The suggested process for translations is similar to that followed for the six abovementioned translations (i.e., two forward translations consolidated into a single version by a third translator), including a field-testing step. Field testing would involve the completion of the questionnaire by a small number of people, with documentation of difficulties in understanding any items or response options. Any necessary changes would be discussed and agreed in a final review involving all translations. The detailed translation guide can be found in Appendix 3.2.

3.3 Results

3.3.1 English version and translations

The paper and pencil English version of the ESIT-SQ, including the optional questions, can be found in Appendix 3.3. Part A comprises 17 core and 17 optional questions that can be answered by every adult irrespective of whether they experience tinnitus or not. From the 17 core questions, there are seven questions about demographics, anthropometric and lifestyle characteristics, and education, one question about family history of tinnitus or hearing loss, and nine questions about the presence of hearing-related and other symptoms and medical conditions. The optional questions ask about further demographic and lifestyle characteristics. The last question in part A screens for the presence of tinnitus lasting more than five minutes during the past year. Part B is relevant only to participants who respond ‘yes’ to the tinnitus screening question and comprises 22 tinnitus-related questions. These include eight questions about tinnitus

perceptual characteristics, one question about the objectivity of tinnitus, one question about the overall impact of tinnitus, six questions about tinnitus onset related characteristics, four questions about factors modulating tinnitus and associations between tinnitus and other co-existing conditions, and two questions about healthcare and management for tinnitus.

The paper and pencil version of the core set of questions of the English ESIT-SQ and its translations into Dutch, German, Italian, Polish, Spanish, and Swedish are published and available open access (Genitsaridi et al., 2019). The electronic version of the ESIT-SQ is implemented within the ESIT database and can be used for online surveys. It is available for use by anyone willing to store the collected data into the ESIT database.

3.3.2 Usage across the ESIT projects

To get an overview of the usage of ESIT-SQ across ESIT projects, a short survey was sent to all ESIT early-stage researchers on 10 June 2020 via email. According to this survey, up to June 2020, the ESIT-SQ had been used for data collection for four completed projects within ESIT. Researchers from the Karolinska Institute used the Swedish version of the ESIT-SQ to enrich an existing tinnitus-specific database, including data from thousands of participants of the Swedish Tinnitus Outreach Project (STOP) (Swedish Tinnitus Outreach Project, 2015). Part of these data were shared with me, to be used in various analyses exploring phenotypic tinnitus heterogeneity (see Chapter 4 for more details). In addition, I used the English version of ESIT-SQ on an online survey and collected data from 200 participants with tinnitus (see Chapter 4 for more details). Researchers from the Ulm University, translated the ESIT-SQ into Albanian, following the guide for further translations. Using this newly developed Albanian version of the ESIT-SQ, they collected data from 150 participants from a survey aiming to describe phenotypic characteristics of patients with tinnitus in Albania. Finally, researchers from the Maastricht University used the Dutch version of the ESIT-SQ for a study with six participants looking at how Ecological Momentary Assessments (EMA) can influence tinnitus perception. The ESIT-SQ was used in at least one more ESIT project, from the University of Groningen, and recruitment was ongoing in June 2020.

3.4 Discussion

3.4.1 Overview

In Summary, this chapter presents the development in English and translation into Dutch, German, Italian, Polish, Spanish, and Swedish of a novel case history questionnaire for standardised tinnitus assessment, namely the ESIT-SQ. Although it was developed to enable standardisation in tinnitus research, it is applicable in both research and clinical settings. It allows both screening for and profiling tinnitus and its comprehensive list of tinnitus-relevant questions and availability in many languages make it a valuable tool for standardised tinnitus assessment internationally.

The discussion addresses the novel contribution that the ESIT-SQ makes to the tinnitus community, and the advantages of an electronic version. Strengths and limitations are also considered, ending with remarks on the wider usage beyond ESIT.

3.4.2 Standardised tinnitus screening and profiling

Compared to other tinnitus case history questionnaires (Langguth et al., 2007; Schechter and Henry, 2002; Stouffer and Tyler, 1990), the ESIT-SQ is unique in that it can be answered by people with and without tinnitus. The last question in Part A, asks about presence of tinnitus during the past year. The ESIT-SQ is therefore a suitable tool for screening for tinnitus. In addition, Part A questions capture information about a variety of potential risk factors for tinnitus, including demographics, lifestyle, medical history, and other co-existing conditions. Therefore, the ESIT-SQ can also be used to explore tinnitus risk factors and more broadly characteristics that differentiate tinnitus from non-tinnitus populations.

As a tool for tinnitus profiling, the ESIT-SQ can be used to collect standardised information about various tinnitus characteristics including onset related characteristics, perceptual characteristics, and temporal associations between tinnitus and co-existing conditions; particularly temporal associations with other conditions about which other tinnitus case history questionnaires include very few questions. The ESIT-SQ extensively questions for the presence of tinnitus-relevant co-existing conditions in Part A, and subsequently asks for their time of onset related to the onset of tinnitus in Part B. Variables such as the association between the onset of tinnitus and other hearing or somatic conditions have been recommended as important for tinnitus profiling in previous studies (Hallam et al., 1984; Michiels et al., 2018).

3.4.3 Advantages of an electronic version

The availability of an electronic version of the ESIT-SQ takes advantage of the benefits and popularity of online survey methodologies. Benefits include the ability to reach unique and global populations, the ability to reach large sample sizes in less time, and the reduced cost compared to traditional paper surveys (Evans and Mathur, 2005; Wright, 2005). In addition, online surveys have the advantage of convenience for the participant who can complete the survey in their own time (Evans and Mathur, 2005). Linked to this, adaptive testing, which allows for automatic adaptation of questions based on preceding responses, makes the completion of questionnaires more user friendly. For example, in the ESIT-SQ paper version, question B8 requires participants to list the previously reported co-existing conditions and document next to each the temporal relationship between their onset and tinnitus onset. In the electronic version, on the other hand, the previously reported conditions are listed automatically making the response to this question less effortful and (hopefully) less prone to errors. What is more, an electronic version allows for automatic data validation. For example, questions with pre-specified numerical ranges (e.g., age of all participants must be 18 years or more) can be programmed to not allow values outside these ranges (e.g., children below the age of 18). In addition, when saving current progress on completing the survey, participants can be informed about the number and location of unanswered questions, meaning lower rates of missing information (Ryan et al., 2002). Finally, electronic questionnaires and online surveys have the advantage of ease of data entry (Evans and Mathur, 2005). In the case of the electronic version of the ESIT-SQ, responses are automatically saved to the ESIT database, thus, no transcribing of the responses is required. This not only saves a lot of time, but also avoids the risk of mistakes during transcribing.

3.4.4 Limitations and considerations for a revised version

One of the limitations of collecting tinnitus-relevant data using the ESIT-SQ, is that some conditions, such as medical comorbidities and symptoms with diagnostic utility such as TMJ disorder or pain, would better be assessed by a clinician to ensure data quality. Nevertheless, this would require a lot of time and resources that are not available in many tinnitus studies. In such cases, having a standardised way to collect these data, even if based on self-report, is very beneficial.

Another limitation is that in order to keep the questionnaire relatively short characteristics that cannot be captured accurately by one or a few questions are not assessed. These include history of noise exposure, personality traits, and coping strategies, for which specifically developed questionnaires exist (Budd and Pugh, 1996; Durai and Searchfield, 2016; Guest et al., 2018). Also, although one question about the impact of tinnitus has been included, it is suggested that detailed questionnaires capturing the complexity of problems that can be associated with tinnitus are used alongside the ESIT-SQ. Many such questionnaires are available, such as the TFI (Meikle et al., 2012).

Finally, since the development of the ESIT-SQ was based on clinical expert opinion and the available scientific evidence at the time of its development, a need for regular reviewing and updating is evident. For example, new evidence might highlight variables not included in the ESIT-SQ as important for tinnitus profiling. In addition, some of the included variables might be proven less important and could be considered for exclusion from future versions. To aid future efforts for updating the ESIT-SQ or creating another questionnaire for self-reported tinnitus assessment, a list of suggestions for changes of the current ESIT-SQ version has been created summarising feedback and comments received during this project (Table 7). These suggestions are based on comments during the various translations of the ESIT-SQ, experience from the use of the ESIT-SQ in various studies (including comments from independent colleagues asked to review the questionnaire before the study, and comments from participants), and novel findings in the tinnitus literature. One example of these suggestions is with regards to the tinnitus defining question. This question was adapted from the European epidemiological survey for tinnitus. An improvement of this question would be the change of the phrase ‘perception of noise in your head or ears’ to ‘involuntary perception of sound in your head or ears’. The term sound is more generic, and it is suitable for tinnitus cases without noise-like tinnitus. Also, this way the negative attributed of the word ‘noise’ would be excluded from the definition. In addition, by including the word ‘involuntary’, confusion with the perception of auditory imagery would be avoided.

Table 7. Suggestions for changes for a revised version of the ESIT-SQ.

Section	Suggested change
Introduction of part A	Change: 'Everyone can complete part A, even if you've never had tinnitus' to '**You** can complete part A, even if you've never had tinnitus'.
Introduction of part A	Add the tinnitus definition. For example, change: 'If you have experienced tinnitus during the past year, you will be asked some more tinnitus-related questions in part B.' to 'Part B has some additional questions only for people that have experienced tinnitus (the involuntary perception of sound in the head or ears in the absence of any corresponding source of sound external to the head) during the past year.'. Alternatively, A17 question can become the first question.
A2 response option	Change from 'Intersex' to '**Other'**, for easier cross-cultural adaptation.
O5 question	Change from: 'How is your economic status relative to the average of the country where you leave in?' to '**What** is your economic status relative to the average of the country where you **live** in?'.
O5 response options	Change from: 'Much higher than the average, 'Quite higher than the average', 'On the average', 'Quite lower than the average', 'Much lower than the average' to '**A lot** higher than the average', '**Higher** than the average', 'On the average', '**Lower** than the average', '**A lot** lower than the average'.
O6 question	Change from: 'Which of the following describes best your current situation?' to 'Which of the following **best describes** your current situation?'.
O6 response options	Change from: 'Parental leave (since two months or longer)' to 'Parental leave (**for more than two months**)' to match other response option.
O7 question	Change from: 'Have you ever worked at night (i.e., between 24:00-5:00)?' to 'Have you ever worked at night (i.e., between **00:00-5:00am**)?'.
O17 responses	Change from 'Less than 6 hours per day' to 'Less than 6 hours', and from '9 or more hours per day' to '9 or more hours'.
A6 question.	Add a note for those that never drink alcohol: '**If you never drink alcohol please type 0 (zero)**'.
A7 response option	Change from 'Never smoker' to 'Never **smoked'**; and consider adding a response option '**Passive smoker**'.
A8 question	Split into two separate questions, one for tinnitus and one for hearing loss family history. Also, remove 'first degree'.
A9 question	Change from: 'Do you suffer from vertigo (sensation of spinning or tilting)?' to 'Do you **have** vertigo (sensation of spinning or tilting)?'.
A10 question	Move after the tinnitus question (A17) since some people report tinnitus in 'Other ear disorders. Please specify'.
A10 response options	▪ Change from 'Presbycusis (aging of ears)' to 'Presbycusis (**age-related hearing loss**)'. ▪ Change from 'Acoustic neuroma (auditory nerve tumour)' to 'Acoustic neuroma (vestibular schwannoma).

Section	Suggested change
A11 question	Change from 'Have you ever undergone any of the following procedures?' to 'Have you ever undergone any of the following **medical interventions**?'.
A11 response options	▪ Consider removing 'Electroconvulsive therapy'. ▪ Change 'None of these' to 'None'.
A13 response options	Reverse order to be in increasing severity like in A12.
A14 response option	Split 'Sound generator' into: 'Wearable sound generator' and 'Non-wearable sound generator'
A16 response options	Add: ▪ '**Chronic kidney dysfunction**' under 'Other' (Shih et al., 2017). ▪ '**Problems with smell**' under 'Otorhinolaryngological' (Park et al., 2018). ▪ '**Schizophrenia**' under 'Psychiatric or psychological' to be able to control for the possibility of confusion of tinnitus with auditory hallucinations ▪ Consider vitamin deficiencies.
A17 question	Change from: 'Tinnitus refers to the perception of noise in your head or ears (such as ringing or buzzing) in the absence of any corresponding source of sound external to your head.' to 'Tinnitus refers to the **involuntary** perception of **sound** in your head or ears (such as ringing or buzzing) in the absence of any corresponding source of sound external to your head.'.
Part B introduction	Remove 'Thank you for completing part A'.
B3 question	Change from 'How long ago did your tinnitus appear?' to 'How long ago did your tinnitus appear **for the first time**?'.
B5 question	Change from 'How long ago did your tinnitus start bothering you?' to 'How long ago did your tinnitus start bothering you **for the first time**?'.
B3 and B5 response options	Add space to specify onset in days for those with early onset tinnitus.
B10 question and response option	Change 'medicines' to 'medication'.
B10 response options	Change 'Quinine (muscle cramps, malaria)' to 'Quinine'.
B11 and B22 questions	Change 'Do you think any of the conditions mentioned before or any other conditions' to 'Do you **think any conditions (from those mentioned before or any other)**'.
Question B17	Change from: 'Has a clinician ever heard your tinnitus?' to 'Most people with tinnitus have a subtype of tinnitus, often called subjective tinnitus, where there is no sound source generating the perceived sound. Others, however, have another subtype, often called objective tinnitus, where the perceived sound is generated by the person's body. In the second case, the sound can sometimes be heard by an external observer, such as a clinician during auscultation. In your case, has a clinician ever heard your tinnitus?', to avoid confusing participants.
B18 and B19 response options	Change: ▪ 'Pressing your head, neck, or area around the ear' to 'Pressing your head, neck, or area around the **ears**'.

Section	Suggested change
	▪ 'Good sleep quality' and 'Poor sleep quality' to 'Good **night** sleep quality' and 'Poor **night** sleep quality'. ▪ 'Driving' to 'Driving **a car**'. ▪ 'Medications' to 'Medications. Please specify__'.
B20 response option	Change from 'Not at all' to '**No,** not at all'.
All questions and response options with medical terms.	Give descriptions for medical terms such as temporomandibular joint, pain syndrome, and Globus Hysterics.
All questions	Add a 'Do not know/Prefer not to say' option.
Phrases thanking for 'participating in this survey'	Change to thanking for 'answering this questionnaire' to be applicable in clinical settings.

3.4.5 Usage within the ESIT and wider impact

Translation of the ESIT-SQ into various languages ensured that the questionnaire would be available for use for various research projects within ESIT. Projects that involved prospective collection of data from humans after the questionnaire was ready were able to incorporate it into their protocols. This led to the collection of standardised data from four independent research projects so far, contributing to the creation of a new international tinnitus database, the ESIT database (ESIT, 2018). Further, it is expected that research groups that have participated in the ESIT project will include the ESIT-SQ in future projects when relevant.

The ESIT-SQ is expected to contribute to standardisation in tinnitus research even outside the ESIT project. Being published open access, the paper version is accessible to anyone that wants to use it. Access to the electronic version is open to everyone that is willing to contribute their collected data to the ESIT database. ESIT members from the University of Regensburg and Ulm University are responsible for queries regarding accessing the electronic version of the ESIT-SQ and intellectual property rights of the acquired data.

Besides the existing translations of the ESIT-SQ in six European languages, research groups external to the ESIT project have expressed interest in translating the ESIT-SQ in other languages to be used for their research and/or clinical needs. An Albanian translation of the ESIT-SQ was the first additional translation following the six initial translations (Genitsaridi et al., 2019). In addition, so far, the ESIT-SQ has been

translated into French and Greek. The availability of the ESIT-SQ in numerous languages makes it a strong candidate for a standardised tool for self-reported tinnitus assessment. This could lead to an overall positive effect in standardisation in tinnitus research and the quest for a better understanding of tinnitus heterogeneity.

3.5 Chapter contributions

I led and coordinated the development of the ESIT-SQ by developing the original set of questions, contacting researchers to act as reviewers, making changes according to reviewers' feedback, and developing the final formatted paper version. I also led the overall translation process by developing the translation protocol, assigning the roles of translation coordinators, and assisting translation coordinators throughout the translation process. The individual translations were conducted by translation coordinators and their translation teams. The online version of the ESIT-SQ was implemented by the ESIT database developers, who I assisted in various steps of the process by providing consultation regarding functionalities and testing of the questionnaire.

Chapter 4. Independent datasets to explore tinnitus heterogeneity

4.1 Introduction

Detailed investigation of phenotypic tinnitus heterogeneity requires collection of appropriate information from large representative populations. Meta-analysis of independently collected datasets would allow more robust analysis in three main ways. First, it would allow combining datasets to achieve larger samples in order to investigate research questions related to uncommon traits. Such an approach is described in Genitsaridi et al. (2020b), where two independent datasets were combined to investigate differences in subgroups of tinnitus with distinct types of tinnitus localisation. Second, it would allow finding answers that might be related to specific populations, and thus emerge from analysing one dataset but not another. This approach is used in Chapter 5, where the same clustering methodology was applied to independent datasets to search for unique subphenotypes that might be present in one dataset but not the other. Third, it would allow validation of findings based on their reproducibility across datasets. This is utilised again in Chapter 5, where subphenotypes identified in one dataset were defined in another, and characteristics of subphenotypes among datasets were compared in search of common patterns.

In this project, three independent tinnitus-specific datasets were accessed for the investigation of tinnitus subphenotypes, i.e., subgroups with similar phenotypic characteristics. The first two, namely the Swedish Tinnitus Outreach Project (STOP) dataset (Sweden) and the National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre (BRC) dataset (UK) were retrospective datasets. The STOP dataset was a subset from a population-based tinnitus-specific database that has been created with a main aim to explore tinnitus heterogeneity (Swedish Tinnitus Outreach Project, 2015). The BRC dataset was a collection of published data from three previous tinnitus clinical studies conducted at the University of Nottingham (Davies et al., 2014; Hoare et al., 2012; Hoare et al., 2014). The third dataset (ESIT dataset) was a prospective dataset, that I collected through an online survey using the ESIT-SQ. In this context, retrospective dataset refers to the fact that the dataset was collected in the past independent of this project, whereas prospective dataset refers to the fact that the dataset was collected specifically for the needs of this project.

In this chapter, I describe the three datasets and report on preliminary data explorations to familiarise with these. In Chapter 5, I describe the subphenotype investigations themselves. A description of data collection and pre-processing methods for each dataset is provided. Emphasis is given in describing variables that are common among datasets and comparing characteristics of each dataset based on these.

4.2 Statistical methods

All analyses were conducted in R version 4.0.2. Categorical variables were summarised with counts for each value (category) of the variable. Numerical variables were summarised using medians, 1st, and 3rd quartiles. Differences in the distribution of categorical variables among datasets were assessed using Fisher's exact tests. In the case of numerical variables, Kruskal-Wallis rank sum tests were used for comparisons of the three datasets, and two-sample Wilcoxon tests (Mann-Whitney tests) for comparisons between any two datasets. R functions used for these tests were implemented in the core package 'stats'. Due to the multiple comparisons, a significance alpha level of 0.001 was used to minimise risk of over-interpreting the findings. Using a statistical procedure for p-value correction for multiple comparisons was considered too stringent for the exploratory analyses conducted across this dissertation (White et al., 2019). Therefore, the common approach of setting a reasonably strict alpha level at $p<0.001$ was chosen (Langguth et al., 2017a; Schecklmann et al., 2014; Schmidt et al., 2018). These statistical methods are used across this thesis, unless otherwise stated.

4.3 Retrospective datasets: the STOP and the BRC datasets

4.3.1 Data collection

In the hearing theme of the NIHR Nottingham BRC, a database was created using data collected from previous studies conducted in the department. For the purposes of this project, 205 tinnitus datasets were selected from three previous studies with complete audiological assessment including extended high frequency audiometric thresholds (Davies et al., 2014, Hoare et al., 2012, Hoare et al., 2014). Each of these studies had received ethical approval from a National Research Ethics Committee (Nottingham or Derby, UK). These studies had a common protocol for audiological assessment including manual pure tone audiometry (spanning frequencies from 0.125 kHz to 14 kHz) using a Siemens Unity 2 system and Sennheiser HDA 200 headphones in sound-

proofed conditions, and psychoacoustic tinnitus assessment following the Tinnitus Tester method (Roberts et al., 2006). The latter included tinnitus pitch matching (likeness Borg scale at 0.5, 1, 2, 3, 4, 5, 6, 7, 8, and 10 kHz), tinnitus loudness matching (decibel [dB] sound pressure level [SPL] at 0.5, 1, 2, 3, 4, 5, 6, 7, 8, and 10 kHz), and assessment of comfortable hearing levels at 0.5 and 5 kHz. Besides the audiological information, this dataset also included self-reported information collected using the TSCHQ (Langguth et al., 2007), and scores from questionnaires assessing tinnitus severity (THQ) (Kuk et al., 1990), depression (BDI) (Beck et al., 1988), anxiety (Beck Anxiety Inventory, BAI) (Steer and Beck, 1997), and hyperacusis (Hyperacusis Questionnaire, HQ) (Khalfa et al., 2002).

The STOP project has led to the creation of a large tinnitus-specific database, including data from thousands of participants with and without tinnitus (Swedish Tinnitus Outreach Project, 2015). From this database, a subset of 657 datasets (395 with tinnitus, 262 without tinnitus) were extracted for this project. This subset had complete audiometric data as well as a response to the ESIT-SQ question (A17) about the presence of tinnitus during the past year. An affiliation with the Karolinska Institute made the remote access of this dataset through a virtual private network (VPN) possible. The STOP project received ethical approval by the local ethics committee '*Regionala etikprövningsnämnden*' in Stockholm (2015/2129-31/1). Audiological assessment included fixed frequency Bekesy audiometry using an Astera 2 audiometer (Otometrics) and Sennheiser HDA 200 headphones. As in the case of the BRC dataset, frequencies from 0.125 kHz to 14 kHz were assessed in sound-proofed conditions. In addition, the STOP dataset included information from tympanometry, speech in noise, DPOAEs, and Uncomfortable Loudness Levels (ULL) assessments. Psychoacoustic tinnitus assessment included:

i) Pitch matching using a two alternative forced choice procedure ranging from 0.18 to 16 kHz with $1/24^{th}$ octave resolution, and measurement of the hearing threshold (dB hearing level, HL) at the pitch matched frequency.
ii) Loudness matching (dB HL) at the pitch-matched frequency using a two forced choice procedure with 1 dB steps.
iii) Minimal masking level (MML) in dB sensation level (SL; number of dB above threshold of the masking noise) starting at Masking Noise Threshold (determined using narrow band noise) and increasing level in 1 dB steps

every second; and masking effect characterisation as complete masking, exacerbation, none, or partial.

iv) Residual inhibition type (using narrow band noise 10 dB above MML for 1 minute) characterised as complete inhibition, partial inhibition, no inhibition, or tinnitus increase rather than inhibition.

The STOP dataset also included detailed self-reported information collected using the ESIT-SQ (Swedish version), and scores from questionnaires assessing tinnitus severity (THI, FTQ, TCS) (Cima et al., 2011; Newman et al., 1996), depression (HADS for Depression, HADS-D), anxiety (HADS for Anxiety, HADS-A) (Levenstein et al., 1993), stress (Perceived Stress Questionnaire, PSQ) (Levenstein et al., 1993), quality of life (World Health Organization's [WHO] Quality of Life – physical, psychological, social, and environmental subscales) (Whoqol Group, 1998), and hyperacusis (HQ) (Khalfa et al., 2002).

4.3.2 Pre-processing

Data pre-processing on the two datasets included recoding values of categorical variables, making these consistent across datasets whenever possible. For example, tinnitus localisation for the BRC dataset was captured using the TSCHQ question 9 'Where do you perceive your tinnitus?' with response options 'Right ear', 'Left ear', 'Both ears, worse in right', 'Both ears, worse in left', 'Both ears, equally', 'Inside the head', and 'Elsewhere'. In addition to this variable, for the ESIT dataset the B15 ESIT-SQ question was available with response options 'Right ear', 'Left ear', 'Both ears, worse in right', 'Both ears, worse in left', 'Both ears, equally', 'Inside the head', 'Other. Please specify', and 'Do not know'. A common variable for localisation was created by recoding values 'Both ears, equally' and 'Inside the head' into one 'No lateralisation' value, and recoding values 'Elsewhere', 'Other. Please specify', and 'Do not know' as missing values. Appendix 4.1 provides details for all included variables, including original values and adaptations.

In addition, ranges for all numerical variables were examined to confirm they were within expected limits. Also, specific cases of missing information were addressed. Since audiometers have a maximum output of stimulus intensity, audiometric data for cases with more severe hearing loss than this limit were given a standardised value of 110 dB HL.

What is more, interesting features were created from existing variables. For example, age at tinnitus onset was calculated by subtracting tinnitus duration in years from participant's age. A single value for mean hearing loss is another such example, calculated as the mean hearing loss from both ears at 0.125, 0.25, 0.5, 1, 2, 4, and 8 kHz. Appendix 4.1 includes details for these new features, in addition to basic variables.

4.4 Prospective dataset: the ESIT dataset

4.4.1 Data collection

For prospective data collection, an online survey was conducted using the ESIT-SQ, including the optional questions. This study received ethical approval by the University of Nottingham Faculty of Medicine and Health sciences Research Ethics Committee (reference number 170-1812).

An online system for informed consent, completion of the survey, and data storage on the ESIT database was first developed. Various online-based tools were combined to ensure automatisation of the process and secure storage of anonymised responses on the ESIT database. Specifically, the first step was to create a study homepage (using Microsoft Sway) including all important participant information, eligibility screening and a consent form (created using Microsoft Forms). This platform allowed secure storage of responses to a University of Nottingham cloud storage server. Upon completion of the consent form, a unique identifier was created linking each participant to their stored data. In addition, an email was sent to participants including a unique survey link and guidance on how to proceed with the survey. Automatisation of steps was implemented using Microsoft Flow. An overview of the flow process and the email sent to participants is shown in Figure 6. For the implementation of the automatic storage of responses through the unique survey link, I worked in collaboration with the developers of the ESIT database.

The online survey was open from 19 June 2019 until 10 July 2019. Participants were recruited by sending an email invitation and link to the study homepage (on 19 June 2019) to 1101 members of a database of the hearing theme of the NIHR Nottingham BRC. In total, 290 consent forms were submitted and 222 participants completed the ESIT-SQ questionnaire. There were 22 participants without tinnitus and only the 200 who reported tinnitus during the past year were considered for further analysis.

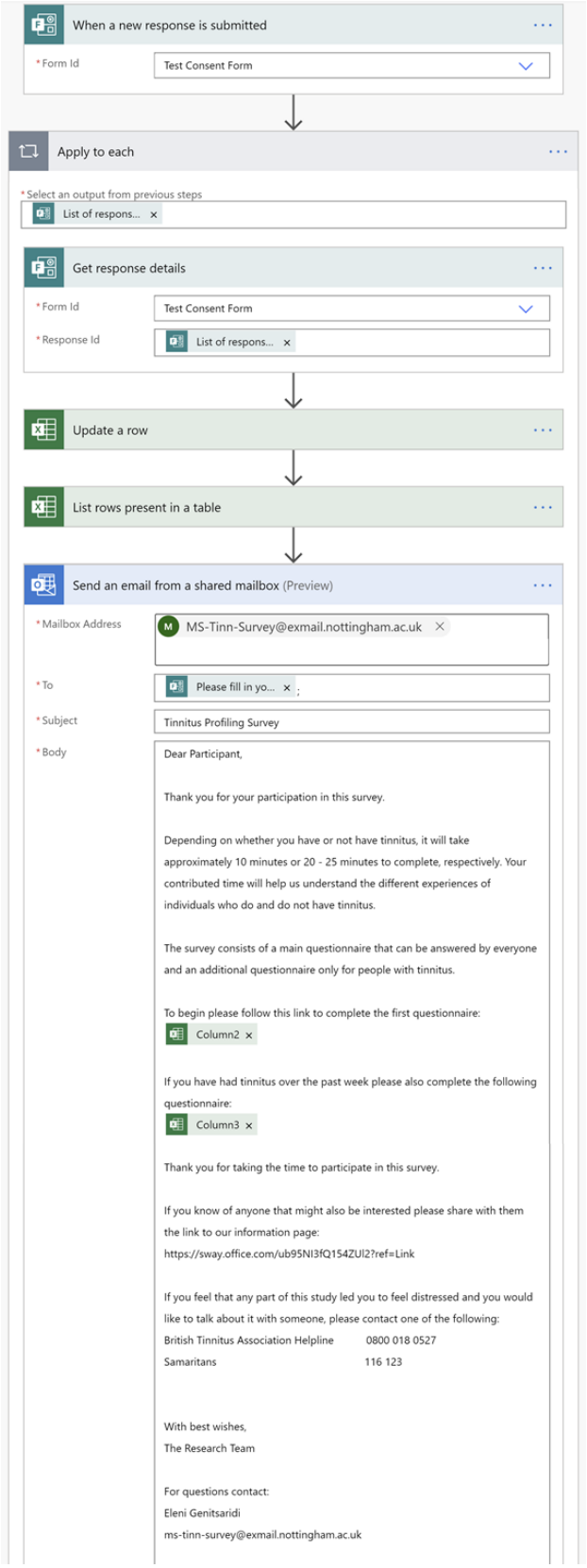

Figure 6. Microsoft flow process initiated upon completion of the consent form, that created a personalised email with a unique link for the completion of the survey and data storage on the ESIT database.

It is worth noting that there were many missing values for sex in the ESIT dataset (33%). That was due to a technical issue related to saving the data from the online survey to the ESIT database. I informed the ESIT database developers about this issue, and they confirmed that it would be fixed for any future projects.

4.4.2 Pre-processing

Pre-processing steps were as described for the STOP and BRC datasets in section 4.3.2, including recoding of variables, checking ranges of numerical variables, and creating new features. All included variables with details for how they were manipulated or created can be found in Appendix 4.1.

Some ESIT-SQ questions had an 'Other, please specify' free-text response option. Such values were recoded as 'Other' or missing to create clean variables for the analyses described in the following chapters. However, an overview of the provided free-text responses from such questions is reported in section 4.5.6.

4.5 Descriptions and comparisons of the datasets

4.5.1 Overview of available variables

Across the three datasets, there was a substantial amount of common information and this commonality is summarised in Figure 7. The STOP dataset contained the most detailed information, including detailed phenotyping and audiological assessment. The BRC dataset included basic audiological information, with many tests in common with the STOP dataset. The ESIT and STOP datasets both included detailed phenotyping as captured with the ESIT-SQ. Comprehensive information for each of 397 assessed variables can be found in Appendix 4.1. This information describes in which of the three datasets each variable was available.

Even though some variables assessed the same characteristic, it was sometimes measured in slightly different ways across the different datasets. If differences were substantial, these variables were not considered directly comparable across datasets because it would be impossible to know whether any observed effect was because of differences in the way the information was collected. For example, the vertigo variable for the BRC dataset corresponded to the TSCHQ question 31 'Do you suffer from vertigo or dizziness?' with response options 'No' or 'Yes'. For the STOP and ESIT

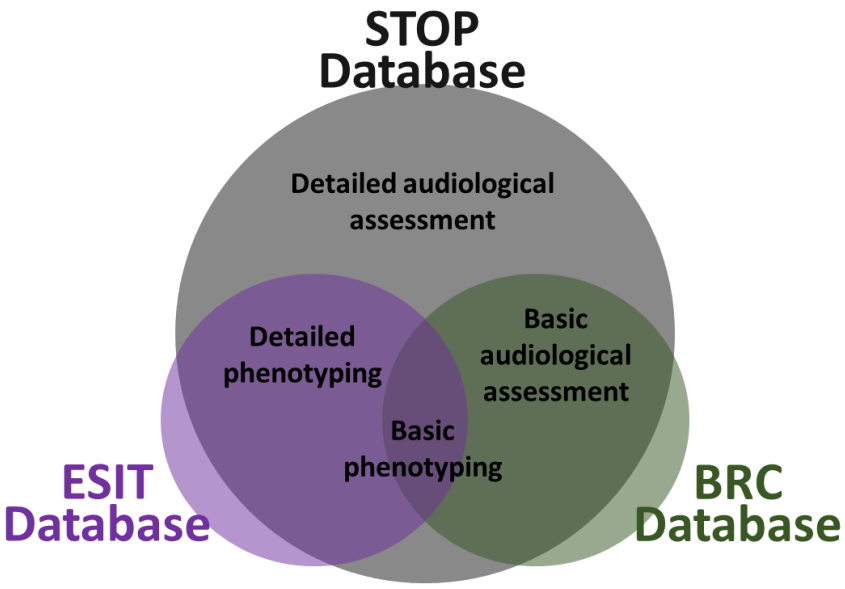

Figure 7. Overlap of available information across the three datasets.

datasets, this variable corresponded to the ESIT-SQ question A9 'Do you suffer from vertigo (sensation of spinning or tilting)?'. The original response options were 'Never', 'Yes, less than one episode per year', and 'Yes, at least one episode per year'. The latter two options were combined into 'Yes' to create a binary variable for vertigo as for the BRC dataset. However, the two variables (from the BRC and from the STOP or ESIT datasets) captured different information since the question from the BRC asked for the presence of either vertigo or dizziness (a much more general symptom).

Table 8 presents a list of the variables that were common to all three datasets and could therefore support across-study evaluations and comparisons. The descriptive statistics for those variables that were common to the STOP and the ESIT datasets can be found in Appendix 4.2. The descriptive statistics for those variables that were common to the STOP and the BRC datasets can be found in Appendix 4.3. Descriptive statistics for some additional variables unique to the STOP, ESIT, and BRC datasets are presented in Appendices 4.4, 4.5, and 4.6, respectively. The STOP dataset included more detailed audiological assessment (including tympanometry, DPOAEs, speech in noise hearing, and uncomfortable loudness levels) and some additional information from self-reported questionnaires; the ESIT dataset included some additional information from optional ESIT-SQ questions; the BRC dataset including some additional audiological variables (hearing thresholds at additional frequencies, and frequency-specific tinnitus matching) and some unique self-reported information.

Table 8. Common variables across the three datasets: descriptions and comparisons.

	All	BRC	ESIT	STOP	P value
All	800	205	200	395	-
Age (y)	57 (47.25, 66), n=798	61 (52.75, 66), n=204	62 (52, 70), n=200	52 (40.25, 64), n=394	<0.001
Sex (female/male)*	342/385	77/122	69/65	196/198	0.022
Handedness (both/left/right)	23/58/712	6/13/179	13/17/170	4/28/363	0.007
Family history of tinnitus (or hearing loss for ESIT) (no/yes)	517/271	138/55	103/97	276/119	NA
Hearing aid use (no/yes)	588/162	117/39	114/85	357/38	<0.001
Headaches (no/yes)*	563/213	122/77	133/49	308/87	<0.001
Vertigo (no/yes)*	381/413	121/78	94/106	166/229	NA
TMJ disorder (no/yes)*	742/44	167/31	184/9	391/4	NA
Tinnitus duration (y)	10 (5, 20), n=588	8 (3, 20), n=199	10 (5, 20), n=168	15 (7, 21), n=221	<0.001
Age at tinnitus onset (y)	44 (27, 56), n=587	50 (35.5, 59.25), n=199	47.92 (32, 57), n=168	32.5 (18.75, 46), n=220	<0.001
Head trauma at tinnitus onset (no/yes)	761/25	185/13	186/7	390/5	0.004
Infection at tinnitus onset (no/yes)	728/58	173/25	176/17	379/16	NA
Change in hearing at tinnitus onset (no/yes)*	714/72	178/20	157/36	379/16	<0.001
Sound exposure at tinnitus onset (no/yes)*	522/264	154/44	138/55	230/165	<0.001
Stress at tinnitus onset (no/yes)*	661/125	176/22	168/25	317/78	0.012
Tinnitus spatial perception (left ear/both ears, more left/no lateralisation [both ears equally or in the head]/both ears, more right/right ear)*	102/140/329/118/51	37/42/29/38/24	30/38/77/32/11	35/60/223/48/16	<0.001
Presence during the day (constant/intermittent)	644/141	181/16	178/15	285/110	NA
Rhythmic tinnitus (no/yes, other/yes, with heartbeat)*	671/61/46	163/18/17	150/21/14	358/22/15	NA
Tinnitus increased by stress (no/yes)	449/326	94/101	114/71	241/154	NA

	All	BRC	ESIT	STOP	P value
Previous treatments or healthcare visits for tinnitus (no/yes)	585/199	139/59	111/82	335/58	NA

*Variable assessed using the TSCHQ question for BRC dataset and the ESIT-SQ question for the STOP and ESIT datasets.
If variables were assessed in substantially different ways, no statistical comparison across datasets was conducted; denoted with NA (not applicable).

4.5.2 Variance and missing data

Some values of binary variables assessed using the ESIT-SQ (such presence of specific comorbidities) were not present in any of the participants. These variables had zero variance (every sample had the same value) and were thus excluded from further analysis. Any other variables with zero variance in one dataset and very low variance in another dataset (majority response higher than 95%), were also not considered in further analysis. Variables excluded for these reasons were use of sound generator, history of electroconvulsive therapy, multiple sclerosis, cochlear implantation, cerebrovascular conditions, dementia, heart attack, lupus, hyperinsulinemia, syphilis, and globus hystericus, tinnitus being reduced by coffee, steroid use at tinnitus onset, and psychiatric management for tinnitus.

Percentage of missing data was also assessed separately per dataset. Overall, missing data for the STOP (tinnitus participants), BRC, and ESIT datasets were 3, 4, and 6%, respectively. Appendix 4.8 shows variables with more than 5% missing data for each of the three tinnitus datasets. The methods for dealing with missing information are described in Chapter 5 (section 5.2.2).

4.5.3 Participant locations

All participants in the STOP and BRC dataset were Sweden and UK residents, respectively. Participants from the ESIT dataset were mostly UK residents (191/200), but eight (4%) were from overseas; five from the USA, one from Argentina, one from France, and one from Germany. One participant did not report country of residence.

4.5.4 Observed differences in participant characteristics between datasets

To allow better interpretation of the findings from analyses in the following chapter, it is important to have a detailed overview of the characteristics of participants across the three datasets. These are presented in detail in this section. One difference worth

highlighting was that both the ESIT and the BRC datasets were characterised by greater hearing and tinnitus burden than the STOP dataset.

Looking at variables available in all three datasets (Table 8), it was found that the median age in the STOP dataset (52 years) was younger than that in the ESIT (62 years) and BRC datasets (61 years) ($p<0.001$). Sex was reasonably balanced in the STOP and ESIT datasets, although the sex data from the ESIT dataset was not reliable due to the level of missingness. There were more males than females in the BRC dataset, although the difference in sex ratio among datasets did not reach statistical significance ($p=0.022$). Hearing aid use was more common in the ESIT dataset. This is expected since the BRC participant database, that was used to recruit participants for the online survey, includes many people with hearing impairment. The datasets also differed significantly in presence of headaches, tinnitus duration, age at tinnitus onset, tinnitus onset related to change in hearing or sound exposure, and spatial perception of tinnitus ($p<0.001$). Although we did not statistically compare datasets against the temporal manifestation of tinnitus (being constant or intermittent) due to differences in the assessment methodology, the STOP dataset clearly included a much higher proportion of people with intermittent tinnitus. This is expected given the population-based methodology for recruitment. Related to this, the STOP dataset also had a lower proportion of people that have had healthcare visits for their tinnitus, indicating a cohort with overall lower tinnitus burden.

Looking at variables available only in the STOP and the ESIT datasets (Appendix 4.2), it was shown that tinnitus participants from the STOP dataset were taller, with lower body mass index (BMI), had more often attended higher education, and less often reported any ear condition, sudden hearing loss or other hearing loss ($p<0.001$) than participants from the ESIT dataset. Differences in height and weight were expected considering population standards between Sweden and the UK (de Almeida et al., 1999; Grasgruber et al., 2014). Participants from the two datasets also differed in self-reported hearing difficulties, use of a combination device (hearing aid and sound generator), use of any hearing device, having undergone ear surgery or any medical procedure, presence of ear pain, face pain, or any pain syndromes, and presence of anxiety, rheumatoid arthritis, nasal septum deviation, or Lyme disease ($p\leq0.001$). Considering the high incidence of Lyme disease in Sweden (Vandekerckhove et al., 2019), the higher percentage of Lyme disease in the STOP dataset compared to the ESIT dataset

served as a confirmation for the validity of the collected information. In addition, in the STOP dataset there was a higher percentage of people reporting non-daily tinnitus and the age at bothersome tinnitus onset (question asked only to those with bothersome tinnitus) was smaller ($p<0.001$). The two datasets also differed in the relationship between tinnitus onset and onset of other conditions including hearing difficulties, sudden hearing loss, other type of hearing loss (not related to acoustic trauma, barotrauma, presbycusis, or sudden hearing loss), dental surgery, ear or face pain, rheumatoid arthritis, Lyme disease, and gastroesophageal reflux ($p\leq0.001$). Also, significant differences were observed in the degree of being affected by tinnitus (worried, annoyed, or upset), the number of different sounds perceived, the variability of tinnitus loudness over a day, tinnitus being reduced by good sleep or by using hearing aids, tinnitus being increased by driving, and having had audiological management, self-management, or any management for tinnitus ($p<0.001$).

Comparing variables available in the STOP and the BRC datasets (Appendix 4.3), hearing thresholds were significantly higher in the BRC dataset ($p<0.001$). Subjects from the STOP dataset more often reported tinnitus loudness varying from day to day and tinnitus worsened by loud noise and scored lower on a numerical rating scale for tinnitus annoyance ($p<0.001$).

4.5.5 Data from participants without tinnitus

The STOP dataset included people without tinnitus (n=262) and so this gave an opportunity to assess how well-matched subgroups with and without tinnitus were on other key variables (see Appendix 4.7). Compared to the non-tinnitus subgroup, the tinnitus subgroup had a higher percentage of females and tended to report a family history of tinnitus ($p\leq0.001$). There were also statistically significant differences in many variables related to hearing function, showing an overall greater hearing impairment in the tinnitus sample. In addition, people with tinnitus more often reported vertigo, and had higher scores of perceived stress ($p<0.001$).

4.5.6 Free-text information

The ESIT dataset provided some free-text information (ESIT-SQ questions A10, A11, A15, A16, B9, B10, B11, B13, B15, B18, B19, B21 and B22 with a free text option ‘Other - please specify’). Although these responses were recoded and standardised (as

described in section 4.4.2), the actual responses were summarised to gain a better understanding of the heterogeneity of tinnitus subphenotypes (Appendix 4.9).

Briefly, in response to question A10, other ear conditions reported by participants included hyperacusis (n=7), balance disorders (n=8), and tinnitus (n=10). 'Other' procedures (question A11) reported included bladder and urethra interventions (n=4), external ear intervention including wax removal (n=4), orthopaedic procedures (n=5), and tonsillectomy (n=4). Additional pain syndromes (question A15) included other musculoskeletal pain syndromes (n=20) and migraine (n=5). Other specified conditions diagnosed by a clinician (question A16) included heart problems such as angina and arterial fibrillation (n=4), arthritis (n=8), and osteoporosis (n=5). In part B, free-text responses regarding conditions related to tinnitus onset included specific noise exposures (n=6) and 'do not know' or 'cannot remember' due to long duration (n=12). Participants reported taking various 'other' medication at tinnitus onset (question B10), most commonly statins (n=6) and levothyroxine (n=5). Interestingly, statins are being discussed as potential otoprotective agents (Prayuenyong et al., 2020). Conditions thought to be related with tinnitus onset (question B11) included noise exposure (n=38), hearing loss (n=18), and anxiety or depression (n=15). Descriptions of tinnitus quality (question B13) included noise-like (n=12), hissing (n=12), pulsating (n=7), buzzing (n=8), ringing (n=6), aeroplane sound (n=6), and whistling (n=5). Free-text responses for tinnitus localisation (question B15) included mainly combinations of different locations. Regarding tinnitus rhythmicity (question B16) most common responses were constant (n=6), varying types (n=4), and combinations of types (n=4). Reported conditions that were perceived to reduce tinnitus (question B18) included listening to music, radio or television or background music or sounds (n=14), situational distraction including focusing on other things, being busy and socialising (n=15), and physical activities (n=5). As for conditions that were perceived to increase tinnitus (question B19), the most common response was loud sounds or noise (n=11). Free-text responses for types of tinnitus management (question B21) included descriptions of hearing-related options such as sound exposure and auditory distractions (n=6) and use of hearing aids or sound generators (n=4). Finally, reported conditions related with periods of increased tinnitus included mental health problems such as anxiety and stress (n=33), sound exposure (n=20), sleep problems (n=5), and silence (n=4).

4.6 Discussion

4.6.1 Online survey

One of the three datasets (the ESIT dataset) was prospectively collected through an online survey. Online surveys are powerful tools allowing convenient collection of large data (see section 3.4.3). They can be widely distributed to people with internet access speaking the survey's language, being largely unaffected by country barriers. Their main disadvantage is that collected data are self-reported, lacking validation from a researcher or clinician.

Many software systems have been developed to facilitate online surveys and these are easy to set up and user-friendly. In our case, a more complicated pipeline had to be developed to automatically save anonymized data into the ESIT database. The system required relatively high technological literacy, which led to the need of providing guidance through emails for many participants (approximately 20% of those completing the questionnaire). This prohibited advertising the study more widely due to time constraints. Improvements in the system allowing less involvement of the researchers would be essential to take advantage of the automatisation that online surveys offer. The experience gained throughout the various stages of our survey helped improve this system for the benefit of future studies. One example of such an improvement is that, based on our feedback, the survey navigation system was updated with messages notifying participants when the survey was completed, and with extra features that help navigate to questions that were missed or unanswered. It is anticipated that such improvements will have a positive impact in future ESIT-SQ completion rates, quality of collected information, and overall user satisfaction.

4.6.2 Importance of the independent datasets

Analysing independent datasets with common information is very important, allowing to investigate in depth relevant research question. This approach has been undertaken in other medical fields. For example, Ahlqvist et al. (2018) applied cluster analysis on different datasets to discover diabetes subphenotypes and could identify stable and replicable subgroups. Such an approach has not yet been reported for tinnitus research and is one of the aims of this dissertation.

In this chapter I provided a detailed description of the available datasets used in the planning of subsequent analyses and in the interpretation of results. Evaluations and

comparisons demonstrated a substantial amount of common information across datasets. Some variables were present in all three datasets, while even more were common in two (the more comprehensive STOP dataset and either the BRC or ESIT dataset). In further analyses, it is important to keep in mind that the three datasets represent different sampling populations, as was highlighted by the statistical comparisons conducted in this chapter. The STOP dataset comes from a population-based cohort, whereas the BRC and ESIT datasets are more representative of people with more severe tinnitus or hearing-related burden. Therefore, the observed lower overall tinnitus burden in the STOP dataset was expected. In addition, the STOP dataset included people living in Sweden, whereas the other two included mainly UK residents. This contributed to some of the observed differences among datasets such as with regards to height, weight, and frequency of Lyme disease. Despite these differences, the three datasets provided a unique opportunity for in-depth exploration of tinnitus heterogeneity. In addition, data collected using the ESIT-SQ for the STOP and ESIT datasets have provided valuable knowledge that can be used to improve future versions of this questionnaire. These insights have informed suggestions for improvements reported in Table 7.

4.7 Chapter contributions

The design and conduction of the online survey is my own work. I developed all steps of the automatic online process up to and including providing the link to the online questionnaire. Everything that followed clicking the link to the questionnaire (i.e. the online implementation of the ESIT-SQ and the automatic storage of the data on the ESIT database) was implemented by the ESIT database developers. The BRC and STOP datasets had been previously collected by other researchers from the University of Nottingham and the Karolinska Institute, respectively. All analyses presented in this chapter is my own work.

Chapter 5. Data-driven discovery of tinnitus subphenotypes

5.1 Introduction

The challenge of the large number of variables contributing to tinnitus heterogeneity has been discussed in previous chapters. In this chapter, I describe a comprehensive application of unsupervised machine learning approaches for the identification of robust tinnitus subphenotypes. These techniques allow simultaneous analysis of multiple variables towards the identification of subgroups with common phenotypic patterns. The main objective of this chapter was to investigate whether application of such algorithms on independent tinnitus-specific datasets could lead to the discovery of robust phenotypic patterns within and across datasets. To my knowledge, there is no previously published study aiming to identify tinnitus subphenotypes by applying unsupervised machine learning techniques and using data from more than one research centre.

Some of the challenges I needed to overcome were related to the selection of included variables, unsupervised learning algorithms and their arguments, and methods for the validation of results. To overcome these challenges and achieve robustness, I applied a novel semi-automatic methodology, assessing various clustering algorithms and validation approaches. The devised validation protocols included measuring stability, compactness, and separation of identified clusters (internal validation) and their ability to differentiate variables or participants' data which were not used for cluster identification (external validation). Such robust application of unsupervised learning techniques had not been previously applied for tinnitus research. Another novelty of this work, considering also other medical fields, was the assessment of specific subgroups within clustering solutions, rather than the overall assessment of the clustering solutions. The results of this project led to the identification and characterisation of robust tinnitus subphenotypes, that could help define more homogeneous subgroups for future tinnitus research.

5.2 Methods

5.2.1 Software and packages

All analyses were conducted in R version 4.0.2 (R Core Team, 2020). Packages used included 'caret' (Kuhn, 2015), 'FactoMineR' (Lê et al., 2008), 'factoextra' (Kassambara and Mundt, 2017), 'Boruta' (Kursa and Rudnicki, 2010), 'glmnet'

(Friedman et al., 2010), 'missForest' (Stekhoven, 2015), 'Proc' (Robin et al., 2011), archetypes (Eugster and Leisch, 2009), 'fpc' (Hennig, 2020), 'cluster' (Maechler et al., 2012), 'clValid' (Brock et al., 2011), 'mclust' (Scrucca et al., 2016), 'gridExtra' (Auguie et al., 2017), 'ggpubr' (Kassambara, 2018), and 'rlist' (Ren, 2016).

5.2.2 Dealing with missing values

Missing values were imputed with the 'missForest' function (Stekhoven, 2015) whenever these were not accepted in any analysis across this study. This function fits random forest classifiers for each variable with missing values using all observed values. These are then used to predict missing values. In some analyses, variables and/or participants with certain percentages of missing values were excluded before imputation (see sections 5.2.5.1, 5.2.5.4 and 5.3.1.3).

5.2.3 Workflow overview

In order to discover subgroups of people with tinnitus with common phenotypic characteristics, I applied unsupervised learning algorithms on the three datasets. The overall methodology was motivated by ensemble machine learning techniques, that develop and combine multiple models to achieve improved results. The aim was to try out a number of different approaches and algorithms expected to produce different results, and select a small number of identified subgroups that optimised prespecified criteria.

An overview of the workflow is presented in Figure 8. Four main approaches for variable selection and validation of results were undertaken. Each approach was applied in one or two of the available datasets leading to a total of seven distinct clustering analyses. For each of these the following four main steps were undertaken:

Step 1. Data pre-processing This step included participant selection, variable transformations, and/or dimension reduction processes.

Step 2. Clustering solutions Eight different unsupervised learning algorithms were applied to discover potential data subgroupings (clustering solutions). These were used with default settings. Results from 2-10 cluster solutions were examined for all algorithms except one that did not require specifying the number of cluster but did require specifying other arguments. For each analysis there was a total of 191

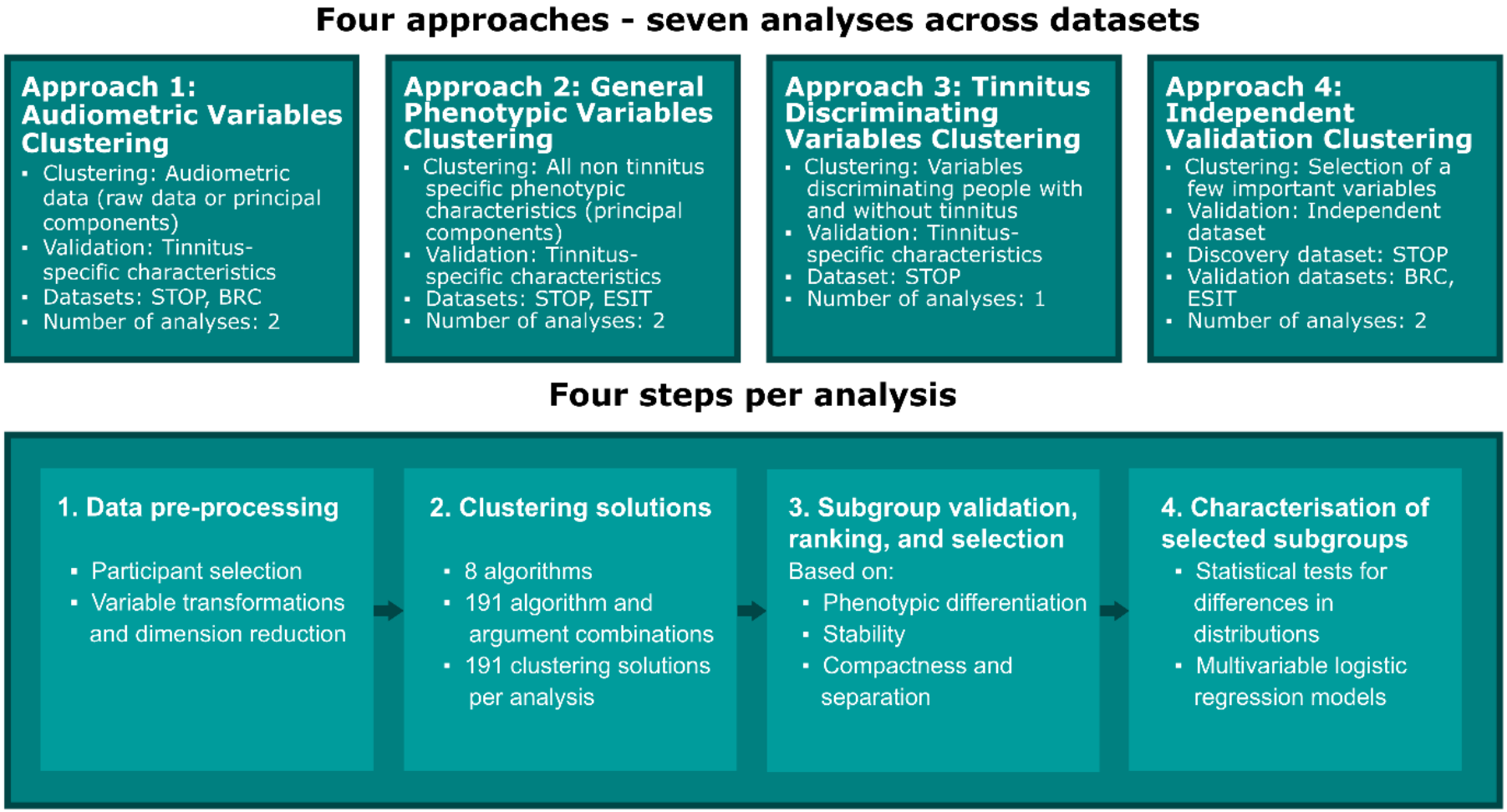

Figure 8. Workflow diagram. The upper boxes describe the four approaches for variable selection and clustering validation. The lower boxes describe the individual steps undertaken for each clustering analysis.

parameter combinations (application of a specific algorithm with specific arguments) leading to an equal number of discovered clustering solutions.

Step 3. Validation, ranking, and selection of subgroups For each of the seven analyses, all identified subgroups from each clustering solution were assessed separately and ranked based on prespecified criteria (phenotypic differentiation of subgroups, stability, compactness, and separation of discovered subgroups). The most highly ranked subgroups meeting prespecified criteria were considered for further evaluation from each clustering analysis.

Step 4. Characterisation of selected subgroups Selected subgroups were then characterised by comparing all available variables using statistical tests for differences in distributions and multivariable LASSO logistic regression models.

Methodological details for the four approaches and the clustering analysis steps are described in sections 5.2.4 and 5.2.5. For clarity of terminology used across this chapter, *clustering approach* refers to each of the four methodologies applied for variable selection and validation of results, *clustering analysis* refers to the application of these approaches to specific datasets, *parameter combination* refers to a specific clustering algorithm with specific arguments, and *clustering solution* to the result of the application of a specific parameter combination. The terms cluster and subgroup are used interchangeably.

5.2.4 Four approaches for variable selection and cluster validation

As previously discussed, variable selection and validation are important elements of a clustering process. Since there is no gold standard method for these, in this study four different approaches were used. For the first three approaches, a subset of non-tinnitus-specific variables was used to identify (discover) subgroups and a distinct subset of tinnitus-specific variables for their validation. For the fourth approach, a few important variables that were common in at least two datasets were selected for clustering discovery in the STOP dataset. Identified patterns were then validated in either the BRC or ESIT dataset.

Clustering approach 1: Audiometric Variables Clustering Audiometric data were selected for cluster discovery given their well-established importance for tinnitus subphenotyping (Kim et al., 2016; Langguth et al., 2017b; Vielsmeier et al., 2015).

Tinnitus-specific variables were used to externally validate the resulting clusters. This approach was applied on the STOP and the BRC datasets that had available audiometric data. Audiometric thresholds are expected to be highly correlated with one another, especially thresholds from neighbouring frequencies, despite the possibility of presence of audiometric notches (a frequency having higher threshold than its neighbouring frequencies). Therefore, these two analyses were conducted using either all original threshold values or a few principal components (for method details see section 5.2.5.1). Details about variables used for discovery and validation can be found in Appendices 5.1 and 5.2.

Clustering approach 2: General Phenotypic Variables Clustering All available non-tinnitus-specific variables were used for cluster discovery, whereas tinnitus-specific variables were used for external cluster validation. This approach was applied on the STOP and ESIT datasets since these had many non-tinnitus-specific variables. Due to the many selected variables, a few principal components were included in the analysis (for method details see pre-processing step in section 5.2.5.1). Details about variables used for discovery and validation can be found in Appendices 5.3 and 5.4.

Clustering approach 3: Tinnitus Discriminating Variables Clustering Variables that were important for discriminating tinnitus from non-tinnitus participants were used for cluster discovery (for method details see pre-processing step described in section 5.2.5.1). Non-tinnitus-specific variables were used for external cluster validation. This approach was applied only on the STOP dataset, where information from participants without tinnitus were also available. Details about variables used for discovery and validation can be found in Appendix 5.5.

Clustering approach 4: Independent Validation Clustering A small set of important variables (based on prior knowledge) was selected for clustering. These included both general and tinnitus-specific characteristics. A few different combinations of variables were tried out initially in a non-systematic way and some variables that did not differ significantly among identified subgroups were excluded from the final analysis. The STOP dataset was used for cluster discovery (discovery dataset). All identified clusters were then externally validated based on their replicability in a different dataset (validation dataset). Specifically, observations in the validation dataset were first labelled according to the subgroups discovered in the STOP dataset using a random

forest classifier. This classifier had been trained on the discovery dataset using the same variables that had been used for clustering in the discovery dataset. Then, for both the discovery and the validation datasets, each identified subgroup was compared to all other observations using all variables that were common across the two datasets. Two such analyses were conducted. In both cases, the STOP dataset was used for cluster discovery. Validation was conducted on either the BRC or the ESIT datasets. Details about variables used for discovery and validation can be found in Appendices 5.6 and 5.7. An overview of this approach is presented in Figure 9.

In total, seven analyses were carried out: two from approach 1 (Audiometric Variables Clustering), two from approach 2 (General Phenotypic Variables Clustering), one from approach 3 (Tinnitus Discriminating Variables Clustering), and two from approach 4 (Independent Validation Clustering) (Figure 8).

5.2.5 Analysis steps

5.2.5.1 Step 1. Data pre-processing

Initial pre-processing steps for all datasets are reported in sections 4.3.2 and 4.4.2. In addition to these, for all analyses presented in this chapter, participants reporting tinnitus heard by a clinician (ESIT-SQ question B17) or pulsatile tinnitus synchronous with the heartbeat (TSCHQ question 8 and ESIT-SQ question B16) were excluded, to focus on cases with subjective tinnitus. The resulting sample sizes for the STOP, BRC, and ESIT datasets were 332, 188, and 167, respectively. In addition, for each of the subsequent analyses, a subset of variables was selected excluding those variables assessing the same characteristics but using different questionnaires. For example, the STOP dataset included two variables for sex, one from the TSCHQ and one from the ESIT-SQ, but only one was used in each analysis. In addition, some variables with high percentage (more than 30%) of missing values were excluded.

Three data-driven approaches for data transformation and/or dimension reduction were applied across the various clustering approaches: (i) PCA, (ii) Factor Analysis of Mixed Data, and (ii) Random Forest approach combined with the Boruta algorithm. These were selected according to the needs of the different clustering approaches.

(i) PCA This technique was used for dimensionality reduction of correlated data (Kassambara, 2017). The aim of the method is to summarise multiple intercorrelated quantitative variables using fewer uncorrelated new variables. These new variables are

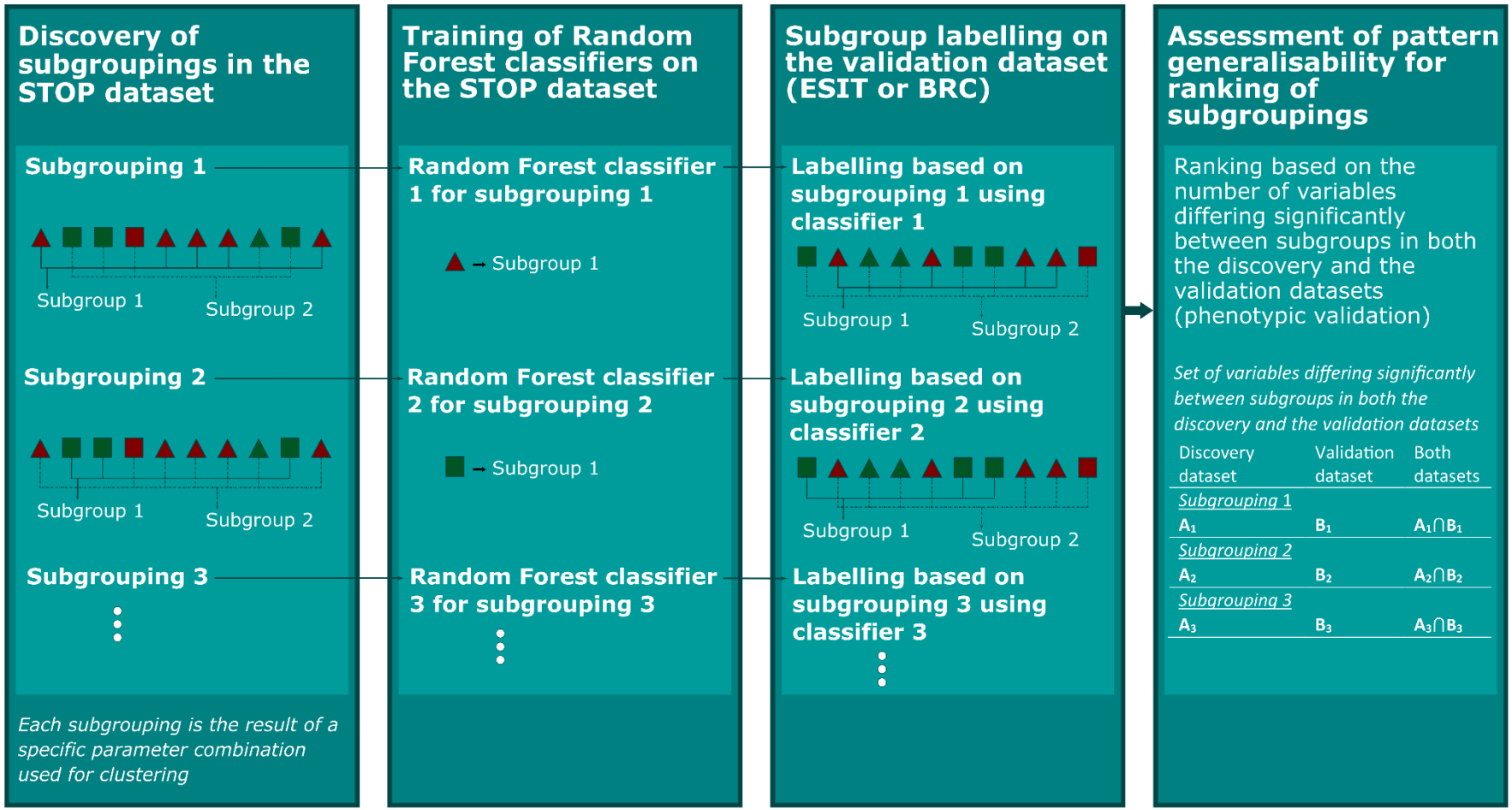

Figure 9. Schematic overview of clustering approach 4. Squares and triangles represent different participants. Shape (square or triangle) and colour (red or green) represent variables available in both the discovery and the validation datasets that were used for discovery of subgroupings.

called principal components and correspond to linear combinations of the original variables. The method aims to identify components along directions that maximise data variation. The first principal direction would be the one with the largest variance. The next would have an orthogonal direction to the previous, maximising again the variance. In this study, PCA was applied on audiometric thresholds from the BRC and STOP datasets (tinnitus participants) separately. Thresholds from each ear and from 0.125, 0.25, 0.5, 1, 2, 3, 4, 6, 8, 10, 12.5, and 14 kHz were analysed (24 variables). Although thresholds were measured in the same scale, higher frequencies tend to have higher thresholds and higher variance. Thus, variables were scaled to have unit variance before PCA. The identified components were used for clustering in clustering approach 1 (Audiometric Variables Clustering) and for validation in clustering approach 4 (Independent Validation Clustering), as described in section 5.3.

(ii) Factor Analysis of Mixed Data (FAMD) This technique is similar to PCA and is applicable to datasets containing both qualitative and quantitative variables (mixed variables). It was used to transform mixed data into representative numerical components (Kassambara, 2017). This transformation was applied any time quantitative variables were used for clustering. That was the case for all analyses except from the Audiometric Clustering Approach. Including all components in the clustering means that all available information is maintained and is a very useful way to deal with mixed data for clustering. In addition to its application for variable transformation, FAMD was also used for dimension reduction in clustering approach 2 (General Phenotypic Variables Clustering). This will be described in more detail in section 5.3.

(iii) Boruta algorithm For clustering approach 3 (Tinnitus Discriminating Variables Clustering), variable selection was based on variables' performance on discriminating tinnitus from non-tinnitus participants. For this purpose, a random Forest approach, the Boruta algorithm (Kursa and Rudnicki, 2010), was applied to identify variables that were important for differentiating tinnitus from non-tinnitus participants from the STOP dataset. This technique selects variables by comparing the importance of original variables for classification to that of randomly generated variables.

5.2.5.2 Step 2. Clustering solutions

For all four approaches, eight unsupervised learning algorithms were applied, spanning a range of clustering techniques. For seven of these algorithms, the numbers of clusters

(or archetypes in one case) should be specified in advance. The algorithms therefore were run for two to 10 clusters, resulting in 72 (8*9) parameter combinations. The 8th algorithm included an inherent method for selecting the number of clusters but required defining two other arguments (Eps and MinPts; see below for details). Eight and 16 values were tried out for these two arguments, resulting in a total of 128 parameter combinations. Therefore, for each analysis, 191 (72+128) parameter combinations were explored. For clustering approach 1 (Audiometric Variables Clustering), this was repeated twice, once using the raw audiometric thresholds as inputs and again using the first four principal components as input. This resulted in 382 parameter combinations. Overall, 1719 (5*191+2*382) clustering solutions were generated. In all cases Euclidean distance was the chosen metric for calculating dissimilarities. Most of the algorithms focus on finding representative subgroups to categorise all observations. More details are provided below:

Algorithm 1: K-means clustering The function 'kmeansCBI' from the 'fpc' package (using the base R function 'kmeans') was used with default settings (Hennig, 2020). K-means is a very popular clustering algorithms in clinical research (Green et al., 2019; Habtewold et al., 2020). It is a partitioning-based algorithm, that tries to categorise observations into k subgroups in a way that minimises the sum of squares from observations to the centre of their assigned subgroup. The initial cluster centres are randomly selected, which can lead to different results. However, unstable results are minimised by initialising the algorithm multiple times.

Algorithm 2: Partitioning Around Medoids (PAM) clustering The function 'claraCBI' from the 'fpc' package (using the function 'pam' from the 'cluster' package) was used with default settings (Hennig, 2020; Maechler et al., 2012). This is another partitioning-based algorithm, very similar to k-means. The difference is that in PAM, subgroups are represented by one of the observations (medoid), rather than from subgroup centres that might not correspond to an observed point (Pina et al., 2020).

Algorithms 3 and 4: Agglomerative and divisive hierarchical clustering The function 'hcut' from the 'factoextra' package was used in both cases (Kassambara and Mundt, 2017). For agglomerative hierarchical clustering, the function 'hcut' called the base R function 'hclust' with the 'ward.D2' method. For divisive hierarchical clustering, function 'hcut' called the 'diana' function from the 'cluster' package with

default settings (Maechler et al., 2012). Hierarchical clustering is different from partitioning-based clustering in that it builds a hierarchy of the observations based on their similarity (Pina et al., 2020). These techniques have been used extensively in clinical research (Green et al., 2019; Habtewold et al., 2020). Agglomerative and divisive hierarchical clustering are the two main categories of hierarchical clustering techniques. In the first case, the algorithms start by considering each observation as its own clusters, subsequently trying to merge observations based on their similarity. In this case, the criterion for linkage of observations with their most similar cluster needs to be specified. The Ward's criterion, used in this study, aims to minimise variance (average of the sum of squares) within clusters (Murtagh and Legendre, 2011). Divisive hierarchical clustering, on the other hand, starts with all observations belonging to one cluster, that is then split recursively until each observation is its own cluster.

Algorithm 5: Hierarchical k-means clustering The function 'hkmeans' from the 'factoextra' package was used (Kassambara and Mundt, 2017). This method starts with hierarchical clustering. For a pre-specified number of clusters, this generates a specific partitioning of the data. The clusters identified by this approach are then used to define cluster centres to initialise the k-means algorithm (Rueda and Krishnan, 2018).

Algorithm 6: Self-Organizing Tree Algorithm (SOTA) The function 'sota' from the 'clValid' package with default settings was used (Brock et al., 2011). This algorithm is based on an unsupervised neural network (self-organising map, SOM) growing with a binary tree topology (Herrero et al., 2001). It is a divisive method that can be stopped when the chosen number of clusters is reached. This method combines the benefits of hierarchical clustering and SOMs, such as having a hierarchic topology and being robust to noise.

Algorithm 7: Archetypal analysis The function 'archetypes' from the package 'archetypes' with the default settings was used (Eugster and Leisch, 2009). The problem this method tries to solve is 'to find a few, not necessarily observed, points (archetypes) in a set of multivariate observations such that all the data can be well represented as convex combinations of the archetypes' (Eugster and Leisch, 2009). This makes this algorithm distinct from other approaches, that focus on finding representative subgroups to categorise all observations. Like all previously described methods that required prespecifying number of clusters, this method also requires prespecifying

number of archetypes. The method provides coefficients for each archetype for each observation, representing how much each archetype contributes to each observation. In the analyses reported here, after identification of archetypes, observations were grouped according to the archetype with the highest coefficient for this observation.

Algorithm 8: Density-based clustering The function 'dbscan' from the package 'fpc' was used (Hennig, 2020). This algorithm is different than all the above-mentioned, as it aims to discover clusters based on the density of the observations, allowing clusters with arbitrary shapes (Ester et al., 1996). The number of clusters is inferred by the algorithm based on the data. To find these clusters, the algorithm requires that the neighbourhood (of specified radius) of points within a cluster must contain a minimum number of other points. Observations that do not belong to any cluster (noise observations) are grouped together in a 'noise' cluster. For this algorithm, an argument specifying the radius of the neighbourhood (Eps) and an argument specifying the minimum number of points required within each core point's neighbourhood (MinPts) need to be specified. In this analysis 16 values for Eps were assessed (1.5 to 3, with step of 0.1). For MinPts eight values were assessed (3 to 10, with step of 1). Therefore, a total of 128 (8*16) parameter combinations were assessed for this algorithm.

5.2.5.3 Step 3. Subgroup validation, ranking, and selection

It is possible that some but not all subgroups discovered by a clustering solution have a robust structure (high compactness, connectedness, and/or separation) or are clinically meaningful. Therefore, in this project, validity was assessed separately for each identified cluster, rather than for the whole clustering solution. For this purpose, observation from each identified cluster were compared to all remaining observations, which were treated as a separate subgroup. This novel approach ensured that any important subgroup would not be hidden by the average measures of the whole clustering solution.

Three main methods of clustering validation were used for ranking the importance of discovered subgroupings (see Table 9 for an overview).

Validation method 1: Phenotypic validation For clustering approaches 1-3 that were conducted on a single dataset, this method involved statistically comparing distribution of tinnitus-specific variables (that were not used for clustering) between each discovered subgroup and the subgroup including all other observations. The number of

Table 9. Overview of criteria for ranking of subgroupings. For each clustering analysis, the most highly ranked subgrouping having Jaccard stability ≥ 0.75 was selected for further assessment.

1. First level ranking: phenotypic validation • For approaches 1-3: number of variables differing between subgroups • For approach 4: number of variables differing between subgroups in both the discovery and the validation dataset (see Figure 9)
2. Second level ranking: Jaccard stability
3. Third level ranking: Silhouette index

statistically significant differences was used as a measure for phenotypic validation. For clustering approach 4, the number of characteristics that differed significantly in both the discovery and the validation dataset (representing generalisability of the identified patterns) was used for phenotypic validation. All statistical comparisons were conducted as described in section 4.2. These measures provided external clustering validation, since the information used for validation was not used for cluster discovery. For all analyses, identified subgroups were first ranked based on these phenotypic validation measures (i.e., either the number of variables differing that were not used for cluster discovery, or the number of variables differing in both the discovery and the validation datasets).

Validation method 2: Stability of identified clusters This was assessed using the Jaccard coefficient, as suggested by Hennig (2007). This index can be used to quantify similarity between clusters generated after resampling the data, and thus assess stability of clusters resulting from a specific algorithm. Given two sets of observations (C and D), the Jaccard coefficient is the proportion of observations belonging to both sets of all the observations belonging to at least one set:

$$\gamma(C, D) = \frac{|C \cap D|}{|C \cup D|}.$$

After applying the same clustering algorithm on 100 bootstrap replications of the data, the Jaccard stability for an identified cluster from the initial clustering solution is defined as the mean of the 100 Jaccard coefficients when comparing this cluster to the most similar from each of the 100 clustering solutions from the bootstrap replications. Jaccard stability ranges from 0 to 1. Values smaller than 0.5 can be interpreted as the cluster being dissolved, and values larger than 0.75 as the cluster being recovered across

the bootstrap replications (Hennig, 2004). The 'clusterboot' function from the 'fpc' package was used to compute Jaccard stability, with boot method set to 'boot' and multiboot set to 'FALSE'. This index was used for a second level ranking of discovered subgroups.

Validation method 3: Silhouette index This index combines information about the compactness and separation of clusters (Rousseeuw, 1987). The 'silhouette' function from the 'cluster' package was used for this purpose. For each observation *i*, the silhouette *s(i)* is defined as:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i),\ b(i))},$$

where *a(i)* is the average distance of *i* to all other observations in the same cluster, and *b(i)* is the average distance of *i* to all other observations in a cluster different than *i* (choosing the cluster that gives the smallest *b(i)*). For a cluster with only one observation, *a(i)* is set to zero. *B(i)* represents the dissimilarity between *i* and the nearest cluster other than the one *i* belongs to. Cluster-wise average silhouettes can be computed by averaging individual silhouettes within a cluster. To guide the otherwise rather subjective interpretation of the average silhouette, Kaufman and Rousseeuw (1990) proposed the following categorisation:

- 0.71-1.00: Strong structure
- 0.51-0.70: Reasonable structure
- 0.26-0.50: Weak structure that could be artificial
- ≤ 0.25 : No substantial structure.

The reported silhouette indices in this study were always computed considering two subgroups: the one being evaluated and the subgroup containing all other observations. Since this index depends on the number of clusters, assessing each subgroup separately also allowed comparison of results from different parameter combinations. This index was used for a third level ranking of discovered subgroups.

The most highly ranked subgroups from each analysis were considered for further evaluation, if their Jaccard stability was higher than 0.75. Selected subgroups were compared between themselves, to identify any two subgrouping that were too similar. For subgroupings coming from the same dataset, the Jaccard coefficient was used again to assess overlap of identified subgroups.

5.2.5.4 Step 4. Characterisation of selected subgroups

Characteristics of the two subgroups from each selected subgrouping were compared using statistical tests for differences in distributions (methods as described in section 4.2). In addition, multivariable logistic regression models were built to assess 1) which variables were associated with the identified subgroups when considering all variables simultaneously, and 2) the performance of predicting membership to a subgroup by other phenotypic characteristics. The response (dependent) variable for these models was the subgroup membership variable. The predictor (independent) variables used for modelling were the available phenotypic variables (see details in Appendix 5.8). Audiological information (and tinnitus psychoacoustic measurements per frequency for the BRC dataset) were not included in these models, except of a single mean value from audiometric thresholds, due to their high correlations. In addition, participants with more than 20% missing values were list-wise excluded before imputing remaining missing values with the 'missForest' function. Initially, simple logistic regression models were built for each predictor variable, and only variables significantly associated with the subgroups were included in the multivariable logistic regression models. The multivariable models were fitted using LASSO (R package glmnet; Friedman et al., 2010) (Friedman et al., 2010), which includes a penalisation for the sum of the absolute coefficients (Tibshirani, 1996). The argument lambda, which defines the penalty for the coefficients, was selected using 5-fold cross-validation, choosing the largest value for which error was within one standard error from the minimum (Breiman et al., 1984; Friedman et al., 2010). This method shrinks some predictor coefficients to zero, allowing selection of the most relevant variables. For characterisation of identified subgroups, variables that either had significantly different distributions at alpha level 0.001 or that were important for the classification models were used.

For performance evaluation, 5-fold cross-validation was used in an outer loop (lambda was selected using 5-fold cross-validation in an inner loop; nested cross-validation - see for example Varma and Simon, 2006). For each model, the following metrics were calculated (mean and standard deviation from the five cross-validation runs):

- Accuracy: the fraction of all instances that are classified correctly.
- Sensitivity (or True Positive Rate): the proportion of being correctly classified as positive of all positives.

- Specificity: the proportion of being correctly classified as negative of all negatives.
- The Area Under the Receiver Operating Characteristic (ROC) curve (AUC). The ROC curve is a plot of the True Positive Rate (sensitivity) on the y-axis and the False Positive Rate (1 – specificity) on the x-axis for different thresholds of a predictor.
- Positive Predictive Value (PPV): the proportions of being correctly classified as positive of all classified as positive.
- Negative Predictive Values (NPV): the proportion of being correctly classified as negative of all classified as negative.

Higher values for all these metrics indicate better performance. A good predictive performance would serve as a further validation that the identified subgroups have robust underlying phenotypic structure.

5.3 Results

5.3.1 Data pre-processing

5.3.1.1 PCA on audiometric thresholds

PCA on the audiometric data revealed that four components could explain almost 90% of the variance in audiometric thresholds for both the BRC and the tinnitus STOP datasets (Figure 10). The variable-loadings of these components were clinically interpretable. Specifically, principal component 1 in both cases corresponded to an average hearing loss across frequencies and ears. Principal component 2 (BRC dataset) and 3 (STOP dataset) corresponded to audiometric threshold differences between left and right ears. Principal component 3 (BRC dataset) and 2 (STOP dataset) corresponded to audiometric threshold differences between low and high frequencies. Principal component 4 in both cases corresponded to audiometric threshold differences between the mid-frequencies and the adjacent (low and high) frequencies.

5.3.1.2 FAMD on non-tinnitus-specific characteristics

For the clustering approach 2 (General Phenotypic Variables Clustering) applied to the STOP and ESIT datasets, all non-tinnitus-specific variables were used for clustering. In both cases, FAMD was applied for dimensionality reductions. For both datasets, the criterion for selecting number of components for inclusion in the clustering analysis

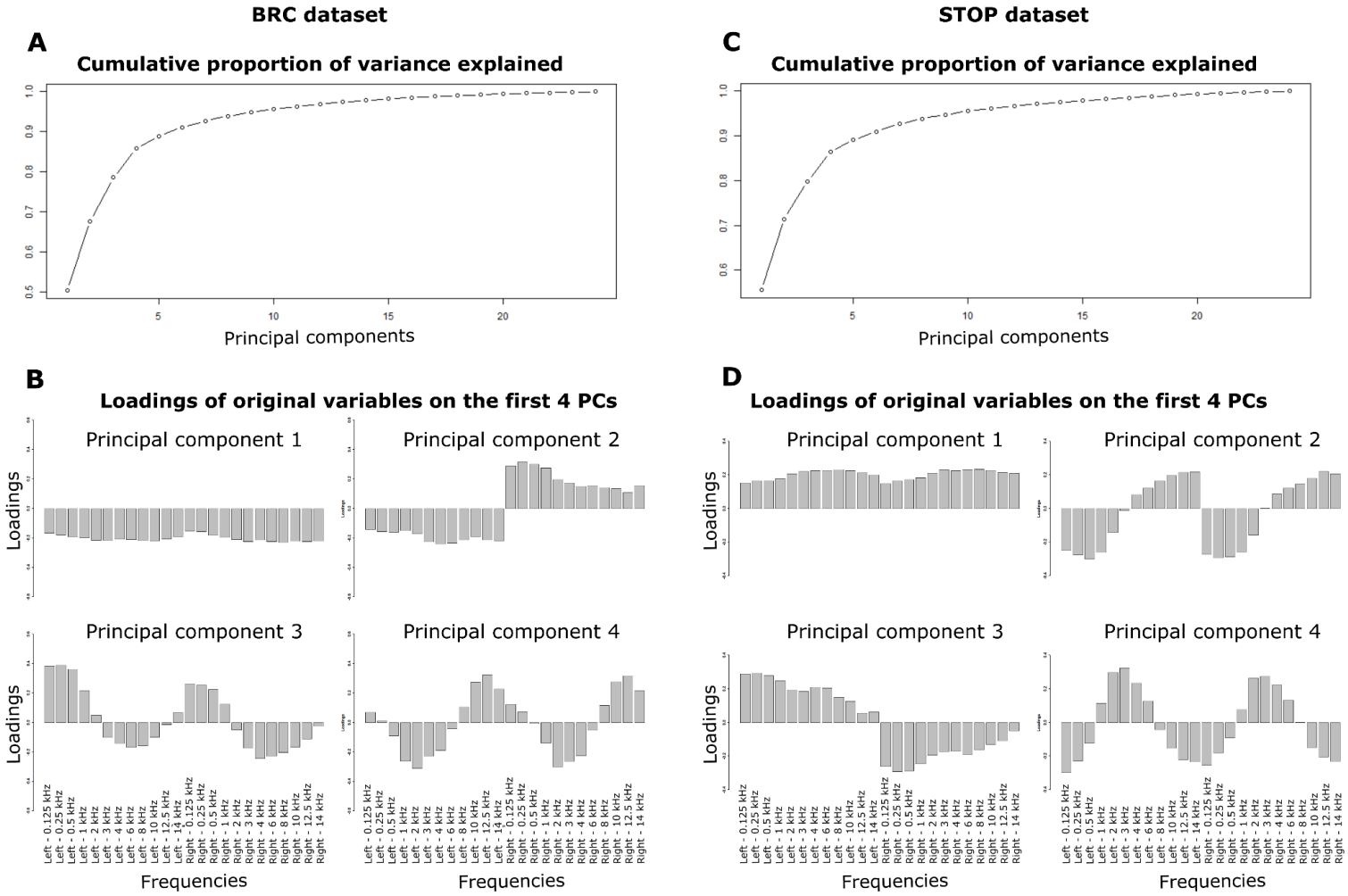

Figure 10. Results from the Principal Component Analysis (PCA) of the audiometric thresholds from the BRC (A, B) and STOP (tinnitus participants) datasets (C, D).

was that which explained at least 50% of the total variance. For the STOP dataset, 166 variables were included in the FAMD (Appendix 5.3). Principal components 1-7 cumulatively explained 52.5% of the variance (each explaining at least 2.8% of the variance) and therefore were selected for inclusion in the clustering analysis. For the ESIT dataset, there were 42 included variables (Appendix 5.4). Principal components 1-11 cumulatively explained 52.9% of the variance (each explaining at least 3% of the variance) and so were selected for inclusion in the clustering analysis.

5.3.1.3 Boruta importance for discriminating tinnitus

To select variables based on their importance for discriminating tinnitus from non-tinnitus participants (for the clustering approach 3: Tinnitus Discriminating Variables Clustering), the Boruta algorithm was applied on 167 non-tinnitus-specific variables (Appendix 5.9) from 324 participants with tinnitus and 262 participants without tinnitus during the past year. Eight out of the 332 participants with tinnitus having more than 30% missing values were excluded from this analysis. This approach identified 35 variables as important for tinnitus status classification (Appendix 5.5).

5.3.2 Clustering solutions, and validation, ranking and selection of subgroups

5.3.2.1 Overview

In this section, validation results and ranking and selection of the subgroups identified across the seven clustering analyses are presented. For each analysis, eight clustering algorithms were applied with 191 parameter combinations. The total number of discovered clusters for each analysis differed because the number of identified clusters was not prespecified for the density-based clustering algorithm. All identified clusters were ranked based on phenotypic differentiation (using data that were not used for clustering), Jaccard stability, and silhouette index. The most highly ranked clusters, with Jaccard stability being at least 0.75, were selected for further characterisation.

5.3.2.2 Clustering approach 1: Audiometric Variables Clustering

STOP dataset For this analysis, 24 and 55 variables were used for clustering and validation, respectively (Appendix 5.1). A total of 382 clustering solutions (191 applied on the raw audiometric variables and 191 applied on their first four principal components) were generating resulting in 1360 discovered subgroups. Table 10 presents the validation results of the top five subgroups. The top ranked subgroup was

Table 10. Top 5 subgroups from the Audiometric Variables Clustering approach applied to the STOP dataset (selected subgrouping in bold).

Rank	Subgrouping details	Number of significant differences	Jaccard	Silhouette	Subgroup sizes
1	**PCs_5clusters_cluster1_kmeans**	**15**	**0.91**	**0.5**	**125/207**
2	PCs_2clusters_cluster2_pam	15	0.89	0.13	184/148
3	PCs_2clusters_cluster1_pam	15	0.88	0.39	148/184
4	PCs_2clusters_cluster2_sota	15	0.85	0.41	157/175
5	PCs_3clusters_cluster1_sota	15	0.85	0.41	157/175

the first cluster identified by applying the k-means algorithm with five clusters on the first four principal components of the audiometric variables. Fifteen out of the 55 variables used for validation differed significantly between observations of this subgroup and all other observations. The Jaccard stability was 0.91 and the silhouette was 0.5. The sizes of the identified subgroup and the subgroup with all other observations were 125 and 207, respectively.

BRC dataset Thirty-two and 17 variables were used for clustering and validation, respectively (Appendix 5.2). Like for the STOP dataset, a total of 382 parameter combinations (191 applied on the raw audiometric variables and 191 applied on their first four principal components) were assessed. From these runs, 1240 subgroups were discovered. Table 11 presents the validation results of the top five subgroups. The top ranked subgroup was the first subgroup identified by applying the archetypal analysis with three archetypes on the raw audiometric variables. Three out of the 17 variables used for validation differed significantly between observations of this subgroup and all other observations. The Jaccard stability was 0.8 and the silhouette was 0.4. The sizes of the identified subgroup and the subgroup with all other observations were 83 and 105, respectively.

ESIT dataset No audiometric data were available in the ESIT dataset.

Table 11. Top 5 subgroups from the Audiometric Variables Clustering approach applied to the BRC dataset (selected subgrouping in bold).

Rank	Subgrouping details	Number of significant differences	Jaccard	Silhouette	Subgroup sizes
1	**Raw_3clusters_cluster1_archetypal**	**3**	**0.8**	**0.4**	**83/105**
2	Raw_7clusters_cluster4_hkmeans	2	0.5	0.44	28/160
3	Raw_8clusters_cluster4_kmeans	2	0.85	0.44	9/179
4	Raw_8clusters_cluster4_hkmeans	2	0.53	0.44	28/160
5	Raw_8clusters_cluster6_hkmeans	2	0.84	0.44	9/179

5.3.2.3 Clustering approach 2: General Phenotypic Variables Clustering

STOP dataset For this analysis, 166 (their first seven principal components) and 55 variables were used for clustering and validation, respectively (Appendix 5.3). The 191 parameter combinations resulted in 649 discovered subgroups in this analysis. Table 12 presents the validation results of the top five subgroups. The first two top ranked subgroups had very low Jaccard stabilities and silhouettes. Therefore, the third subgrouping was selected for further analysis. This was identified with density-based clustering by setting Eps at 1.8 and MinPts at 10. The assessed subgroup was the first cluster from a two-cluster solution. Eleven out of the 55 variables used for validation differed significantly between observations of this subgroup and all other observations. The Jaccard stability was 0.88 and the silhouette 0.3. The sizes of the identified subgroup and the subgroup with all other observations were 257 and 75, respectively.

Table 12. Top 5 subgroups from the General Phenotypic Variables Clustering approach applied to the STOP dataset (selected subgrouping in bold).

Rank	Subgrouping details	Number of significant differences	Jaccard	Silhouette	Subgroup sizes
1	3clusters_cluster3_pam	13	0.44	0.01	116/216
2	3clusters_cluster2_sota	12	0.28	0.04	63/269
3	**2clusters_cluster1_1.8eps_10minpts_dbscan**	**11**	**0.88**	**0.3**	**257/75**
4	2clusters_cluster1_1.6eps_5minpts_dbscan	11	0.88	0.29	243/89
5	2clusters_cluster1_1.7eps_7minpts_dbscan	11	0.88	0.29	255/77

BRC dataset This approach was not applied to the BRC dataset as it did not include as many non-tinnitus-specific variables as the other two datasets.

ESIT dataset For this analysis, 42 (their first 11 principal components) and 64 variables were used for clustering and validation, respectively (Appendix 5.4). The 191 parameter combinations resulted in 638 discovered subgroups. Examining subgroups with Jaccard stability 0.75 or higher, the only tinnitus-specific phenotypic differences observed were with regards to the temporal relationship of the onset of tinnitus and other comorbidities, which were not considered sufficient to define different tinnitus subphenotypes. Therefore, no subgrouping from this analysis was selected for further assessment.

5.3.2.4 Clustering approach 3: Tinnitus Discriminating Variables Clustering

STOP dataset Thirty-five and 55 variables were used for clustering and validation, respectively (Appendix 5.5). From the 191 parameter combination, 469 discovered subgroups were generated. Table 13 presents the validation results of the top five subgroups. The top ranked subgroup was the second cluster identified by applying PAM clustering with two clusters. Thirteen out of the 55 variables used for validation differed significantly between observations of this subgroup and all other observations. The Jaccard stability was 0.91 and the silhouette was 0.23. The sizes of the identified subgroup and the subgroup with all other observations were 200 and 132, respectively.

Table 13. Top 5 subgroups from the Tinnitus Discriminating Variables Clustering approach applied to the STOP dataset (selected subgrouping in bold).

Rank	Subgrouping details	Number of significant differences	Jaccard	Silhouette	Subgroup sizes
1	**2clusters_cluster2_pam**	**13**	**0.91**	**0.23**	**200/132**
2	2clusters_cluster1_pam	13	0.88	0.23	132/200
3	3clusters_cluster3_kmeans	11	0.93	0.25	127/205
4	4clusters_cluster4_hkmeans	11	0.56	0.28	76/256
5	5clusters_cluster5_hkmeans	11	0.55	0.28	76/256

BRC and ESIT datasets Information from participants without tinnitus were not available in the BRC and ESIT datasets.

5.3.2.5 Clustering approach 4: Independent Validation Clustering

STOP and BRC dataset For this analysis, eight and 25 variables that were available in both the STOP and the BRC datasets were used for clustering and validation, respectively (Appendix 5.6). Six hundred and sixty-nine discovered subgroups resulting from 191 parameter combinations were assessed. Table 14 presents the validation results of the top five subgroups. The fifth subgrouping was the first with Jaccard stability above 0.75 and was selected for further analysis. It was the second cluster of a 4-cluster solution discovered by density-based clustering with Eps set at 1.6 and MinPts at 5. Four out of the 25 variables used for validation differed significantly between observations of this subgroup and all other observations in both the discovery (STOP) and the validation (BRC) datasets (there were six total differences in the discovery dataset). The Jaccard stability was 0.93 and the silhouette was 0.3. The sizes of the identified subgroup and the subgroup with all other observations for the discovery dataset were 116 and 216, respectively. For the validation dataset, the corresponding subgroup sizes were 47 and 141.

Table 14. Top 5 subgroups from the Independent Validation Clustering approach applied to the STOP and BRC datasets (selected subgrouping in bold).

Rank	Subgrouping details	Number of variables differing significantly in both datasets (variables differing in STOP dataset)	Jaccard	Silhouette	Subgroup sizes (discovery dataset)
1	3clusters_cluster3_archetypal	5 (10)	0.48	0.22	154/178
2	7clusters_cluster7_archetypal	5 (11)	0.46	0.38	79/253
3	4clusters_cluster3_archetypal	5 (10)	0.43	0.19	140/192
4	3clusters_cluster2_archetypal	5 (8)	0.4	0.1	133/199
5	**4clusters_cluster2_1.6eps_5minpts_dbscan**	**4 (6)**	**0.93**	**0.3**	**116/216**

STOP and ESIT datasets For this final analysis, five and 75 variables that were available in both the STOP and the ESIT datasets were used for clustering and validation, respectively (Appendix 5.7). Two-thousand two-hundred and twenty-three discovered subgroups resulting from 191 parameter combinations were assessed. Table 15 presents the validation results of the top five subgroups. The top ranked subgroup was the first cluster identified by applying the k-means algorithm with 2 clusters. Seven out of the 75 variables used for validation differed significantly between observations of this subgroup and all other observations in both the discovery (STOP) and the validation (BRC) datasets (there were 17 total differences in the discovery dataset). The Jaccard stability was 0.96 and the silhouette was 0.29. The sizes of the identified subgroup and the subgroup with all other observations for the discovery dataset were 275 and 57, respectively. For the validation dataset, the corresponding subgroup sizes were 125 and 42.

Table 15. Top 5 subgroups from the Independent Validation Clustering approach applied to the STOP and ESIT datasets (selected subgrouping in bold).

Rank	Subgrouping details	Number of variables differing significantly in both datasets (variables differing in STOP dataset)	Jaccard	Silhouette	Subgroup sizes (discovery dataset)
1	**2clusters_cluster1_kmeans**	**7 (17)**	**0.96**	**0.29**	**275/57**
2	2clusters_cluster2_kmeans	7 (17)	0.87	0.11	57/275
3	8clusters_cluster8_2.8eps_7minpts_dbscan	7 (21)	0.85	0.03	54/278
4	8clusters_cluster8_2.8eps_8minpts_dbscan	7 (21)	0.77	0.03	54/278
5	10clusters_cluster10_2.8eps_6minpts_dbscan	7 (14)	0.67	0.03	38/294

5.3.2.6 Overview of selected subgroupings

Across these seven analyses, six subgroupings were selected as summarised in Table 16. Five of these come from the STOP dataset and one from the BRC dataset. The algorithms that discovered the selected subgroupings were k-means clustering, partitioning around medoids, density-based clustering, and archetypal analysis. Jaccard stabilities ranged from 0.80 to 0.96 and silhouettes from 0.23 to 0.5.

Assessing the Jaccard coefficient between the most similar subgroups of any two subgroupings from the STOP dataset revealed that some subgroups were similar (Jaccard coefficient close to 0.74) but no pair reached the 0.75 threshold (Table 17). Therefore, characteristics of all six subgroups were examined. The two most common pairs of subgroups were 1) the first subgroup of the STOP Audiometric Variables Clustering and the second subgroup of the STOP Tinnitus Discriminating Variables Clustering, and 2) the first subgroups of the STOP General Phenotypic Variables Clustering and the STOP-ESIT Independent Validation Clustering.

In section 5.3.3, each pair of subgroups is further assessed by examination of their phenotypic characteristics. To make comparisons easier, the order of subgroups from the STOP Audiometric Variable Clustering, the BRC Audiometric Variables Clustering, and the STOP-BRC Independent Validation Clustering was reversed so that subgroup 1 was always the subgroup of the largest size.

Table 16. Summary of the six selected subgroupings.

Analysis details	Subgrouping details	Number of variables differing significantly	Jaccard (of discovered subgroup)	Silhouette	Subgroup sizes
STOP Audiometric Variables Clustering Subgrouping	PCs_5clusters_cluster1_kmeans	15	0.91	0.5	125/207
BRC Audiometric Variables Clustering Subgrouping	Raw_3clusters_cluster1_archetypal	3	0.8	0.4	83/105
STOP General Phenotypic Variables Clustering Subgrouping	2clusters_cluster1_1.8eps_10minpts_dbscan	11	0.88	0.3	257/75
STOP Tinnitus Discriminating Variables Clustering Subgrouping	2clusters_cluster2_pam	13	0.91	0.23	200/132
STOP-BRC Independent Validation Clustering Subgrouping	4clusters_cluster2_1.6eps_5minpts_dbscan	4 (6)	0.93	0.3	116/216
STOP-ESIT Independent Validation Clustering Subgrouping	2clusters_cluster1_kmeans	7 (17)	0.96	0.29	275/57

Table 17. Jaccard coefficients of the most similar subgroups between any two selected subgroupings from the STOP dataset.

	STOP General Phenotypic Variables Clustering	STOP Tinnitus Discriminating Variables Clustering	STOP-BRC Independent Validation Clustering	STOP-ESIT Independent Validation Clustering
STOP Audiometric Variables Clustering	0.54	0.74	0.52	0.61
STOP General Phenotypic Variables Clustering	-	0.53	0.49	0.74
STOP Tinnitus Discriminating Variables Clustering	-	-	0.5	0.59
STOP-BRC Independent Validation Clustering	-	-	-	0.56

5.3.3 Characterisation of selected subgroups

5.3.3.1 Overview of logistic regression models

The number of variables that were important for the logistic regression models ranged from three to 24 (Table 18). Predictive performance of the LASSO models was satisfactory for five out of six models having accuracy higher than 85%. The high predictive performance of these five models further validated that the identified subgroups have distinct phenotypic structure. The model predicting subgroups from the STOP General Phenotypic Variables Clustering had the worst performance, with particularly low sensitivity (i.e., proportion of those being correctly classified as subgroup 2 from all belonging in subgroup 2) and PPV (i.e., proportion of those being correctly classified as subgroup 2 from all that were classified as subgroup 2). Down-sampling when training the algorithm (to account for the imbalance in subgroup sizes) did not result in better performance (data not shown). Therefore, results from this model are not discussed any further. In addition, it is worth noting that in the STOP-BRC Independent Validation Clustering model, sex was not significant in simple regression due to quasi-complete separation (no males in subgroup 2). However, including height and weight partially compensated for the information in this variable.

Table 18. Predictive performance evaluation measures for the six LASSO regression models (mean values and standard deviations from the five cross-validation runs). Number of variables contributing to each model (having non-zero coefficients) are also reported.

	STOP Audiometric Variables Clustering	BRC Audiometric Variables Clustering	STOP General Phenotypic Variables Clustering	STOP Tinnitus Discriminating Variables Clustering	STOP-BRC Independent Validation Clustering	STOP-ESIT Independent Validation Clustering
Number of important variables	11	3	13	12	21	24
AUC	94.13+-3.67	98.75+-0.76	71.37+-7.37	92.93+-3.38	91.12+-3.01	81.36+-4.49
Specificity	88.16+-9.34	96+-6.52	73.97+-13.56	85.64+-10.03	83.53+-7.84	92.36+-6.35
Sensitivity	89.82+-7.33	95.15+-4.97	69.33+-7.6	89.63+-6.11	89.87+-4.9	67.12+-12.74
Accuracy	88.79+-6.59	95.56+-2.48	72.91+-8.99	87.25+-5.35	85.65+-3.8	87.94+-3.42
NPV	93.73+-3.97	96.19+-3.98	89.28+-1.18	93.01+-3.58	94.35+-2.13	93.2+-2.4
PPV	82.83+-10.76	95.67+-6.89	46.58+-11.02	81.38+-9.4	74.67+-9.63	68.84+-14.45

5.3.3.2 Overview of descriptions of subgroupings

Table 19 presents selected characteristics and comparisons of subgroup pairs (subgroup 1 and 2) from the six selected subgroupings. In addition, audiometric data for all six subgroupings and ULLs for the five STOP subgroupings were plotted to gain a better understanding of the hearing profile of the discovered subgroups (Figure 11 and Figure 12). Overviews of variables that differed significantly between subgroups for each subgrouping and were important for LASSO classification models can be found in Appendices 5.10 and 5.11, respectively. Narrative descriptions of the characteristics of the two subgroups from each of the six selected subgroupings are provided in sections 5.3.3.3 – 5.3.3.5. Subgroupings with similar patterns are described in the same sections. Overall, all subgroupings included subgroups with distinct phenotypic characteristics. Therefore, they can be referred to as tinnitus subphenotypes.

Table 19. Comparisons of characteristics of subgroup pairs for the six selected subgroupings.

	STOP Audiometric Variables Clustering	BRC Audiometric Variables Clustering	STOP General Phenotypic Variables Clustering	STOP Tinnitus Discriminating Variables Clustering	STOP-BRC Independent Validation Clustering	STOP-ESIT Independent Validation Clustering
Number of differing/assessed variables	124/226	50/95	57/226	128/226	37/226	30/226
Size	332	188	332	332	332	332
Subgroup 1	207	105	257	200	216	275
Subgroup 2	125	83	75	132	116	57
Age (y)	**52 (40, 64)*^**	**61 (53, 66)*^**	52 (40, 64)	**52 (40, 64)*^**	52 (40, 64)	**52 (40, 64)^**
Age subgroup 1	**59 (50.3, 69)**	**65 (59, 71)**	52 (41, 64)	**59 (50, 69)**	52 (41, 63.5)	**52 (42, 65)**
Age subgroup 2	**39 (34, 48)**	**55 (48, 61)**	50 (39, 63.8)	**40 (35, 49)**	51 (39, 64.3)	**47 (38, 54)**
Sex (female/male)	164/168	69/113	163/168	164/168	**164/168***	**163/168*^**
Sex subgroup 1	98/109	41/59	121/136	93/107	**48/168**	**123/151**
Sex subgroup 2	66/59	28/54	42/32	71/61	**116/0**	**40/17**
Weight (kg)	75 (64.8, 85)	NA	75 (64.8, 85)	75 (64.8, 85)	**75 (64.8, 85)*^**	75 (64.8, 85)
Weight subgroup 1	75 (65, 85)	NA	75 (65, 85)	76 (65, 86)	**80 (70.8, 89)**	75 (65, 85)
Weight subgroup 2	71 (62, 83)	NA	74 (63, 82)	71.5 (62, 83)	**64 (59, 71)**	70 (63, 88)
Height (cm)	174 (167, 182)	NA	174 (167, 182)	174 (167, 182)	**174 (167, 182)*^**	174 (167, 182)
Height subgroup 1	174 (167, 182)	NA	175 (167, 182)	174 (168, 182)	**180 (172.3, 185)**	175 (168, 182)
Height subgroup 2	173 (166, 181)	NA	171.5 (165, 181.5)	173.5 (165, 182)	**167 (163, 170)**	170 (164, 176)
Education (lower/higher)	92/240	NA	92/240	92/240	**92/240^**	92/240
Education subgroup 1	60/147	NA	69/188	57/143	**69/147**	73/202
Education subgroup 2	32/93	NA	23/52	35/97	**23/93**	19/38
Alcohol (number of drinks per week)	3 (1, 6)	NA	3 (1, 6)	3 (1, 6)	**3 (1, 6)^**	3 (1, 6)
Alcohol subgroup 1	3 (1, 6)	NA	3 (1, 6)	3 (1, 6)	**3 (1, 6)**	3 (1, 6)

	STOP Audiometric Variables Clustering	BRC Audiometric Variables Clustering	STOP General Phenotypic Variables Clustering	STOP Tinnitus Discriminating Variables Clustering	STOP-BRC Independent Validation Clustering	STOP-ESIT Independent Validation Clustering
Alcohol subgroup 2	2 (1, 5)	NA	2 (0, 5)	2 (1, 5)	**2 (1, 4)**	2 (0, 6)
Mean hearing threshold both ears (dB HL)	**7.62 (1.9, 16.5)*^**	**27.68 (16.1, 36.1)*^**	**7.62 (1.9, 16.5)***	**7.62 (1.9, 16.5)*^**	**7.62 (1.9, 16.5)*^**	7.62 (1.9, 16.5)
Mean hearing subgroup 1	**12.79 (7.1, 22.1)**	**35 (30.4, 42.9)**	**6.96 (1.4, 13.6)**	**12.7 (7.2, 22.5)**	**9.34 (2.6, 19.9)**	7.93 (2.5, 16.8)
Mean hearing subgroup 2	**1 (-1.0, 5.5)**	**14.64 (10.4, 20.4)**	**14.41 (4.7, 25.5)**	**0.9 (-1.0, 4.8)**	**5.7 (1.5, 10.8)**	6.51 (0.8, 15.1)
Hearing aid use (no/yes)	**299/33***	**108/37***	**299/33***	**299/33***	**299/33*^**	299/33
Hearing aid subgroup 1	**175/32**	**35/32**	**241/16**	**168/32**	**185/31**	247/28
Hearing aid subgroup 2	**124/1**	**73/5**	**58/17**	**131/1**	**114/2**	52/5
Self-reported hearing problem (no/yes from TSCHQ; slight or no difficulty/moderate difficulty/severe difficulty/total loss from ESIT-SQ)	**117/161*^**	**41/141***	**173/108/45/2***	**117/161*^**	**117/161*^**	173/108/45/2
Hearing problem subgroup 1	**46/126**	**6/94**	**151/78/25/0**	**44/123**	**62/119**	151/87/32/2
Hearing problem subgroup 2	**71/35**	**35/47**	**22/30/20/2**	**73/38**	**55/42**	22/21/13/0
HQ score (0-42)	15 (8, 23)	12 (8.3, 19)	**15 (8, 23)***	15 (8, 23)	15 (8, 23)	**15 (8, 23)***
HQ score subgroup 1	14 (8, 22)	12 (9, 19.3)	**14 (8, 21)**	13 (8, 22.5)	15 (8.8, 23)	**14 (8, 21)**
HQ score subgroup 2	16 (9, 24)	12.5 (8, 19)	**20 (11.5, 28.5)**	16.5 (10, 24)	14 (8, 21.5)	**23 (14, 29)**
Problem with external sounds (small/moderate/big or very big)	234/71/27	NA	**234/71/27***	234/71/27	234/71/27	**234/71/27*^**
Sounds problem subgroup 1	153/39/15	NA	**197/50/10**	147/41/12	148/47/21	**208/51/16**
Sounds problem subgroup 2	81/32/12	NA	**37/21/17**	87/30/15	86/24/6	**26/20/11**
Vertigo/Dizziness (no/yes)	**215/90^**	114/68	**142/190***	**215/90^**	215/90	142/190
Vertigo/Dizziness subgroup 1	**144/47**	62/38	**123/134**	**140/45**	149/56	126/149
Vertigo/Dizziness subgroup 2	**71/43**	52/30	**19/56**	**75/45**	66/34	16/41
Presbycusis (no/yes)	303/29	NA	303/29	**303/29***	303/29	303/29

	STOP Audiometric Variables Clustering	BRC Audiometric Variables Clustering	STOP General Phenotypic Variables Clustering	STOP Tinnitus Discriminating Variables Clustering	STOP-BRC Independent Validation Clustering	STOP-ESIT Independent Validation Clustering
Presbycusis subgroup 1	181/26	NA	237/20	**174/26**	197/19	251/24
Presbycusis subgroup 2	122/3	NA	66/9	**129/3**	106/10	52/5
Acute Otitis (no/yes)	309/23	NA	309/23	309/23	309/23	**309/23^**
Otitis subgroup 1	188/19	NA	242/15	185/15	201/15	**260/15**
Otitis subgroup 2	121/4	NA	67/8	124/8	108/8	**49/8**
PSQ (0-1)	**0.3 (0.2, 0.43)***	NA	**0.3 (0.2, 0.43)***	**0.3 (0.2, 0.43)*^**	0.3 (0.2, 0.43)	**0.3 (0.2, 0.43)*^**
PSQ subgroup 1	**0.28 (0.17, 0.41)**	NA	**0.29 (0.17, 0.41)**	**0.28 (0.16, 0.4)**	0.31 (0.2, 0.46)	**0.29 (0.17, 0.39)**
PSQ subgroup 2	**0.34 (0.26, 0.49)**	NA	**0.39 (0.28, 0.53)**	**0.34 (0.26, 0.5)**	0.29 (0.18, 0.42)	**0.51 (0.34, 0.63)**
Anxiety score† (0-21 for STOP; 0-63 for BRC)	**5 (3, 8)*^t**	4 (2, 9)	**5 (3, 8)***	**5 (3, 8)*^**	5 (3, 8)	**5 (3, 8)*^t**
Anxiety subgroup 1	**5 (2, 7)**	4 (1.8, 8)	**5 (2, 7)**	**5 (2, 7)**	5 (3, 8)	**4 (2, 7)**
Anxiety subgroup 2	**7 (4, 10)**	5 (2.3, 11)	**7 (3, 10)**	**6 (3, 10)**	5 (3, 8)	**9 (6, 11)**
Depression score (0-21 for STOP; 0-63 for BRC)	2 (1, 5)	2 (0, 4)	**2 (1, 5)***	2 (1, 5)	**2 (1, 5)*^**	**2 (1, 5)*^**
Depression) subgroup 1	2 (1, 4)	1 (0, 4)	**2 (1, 4)**	2 (1, 5)	**3 (1, 5)**	**2 (1, 4)**
Depression) subgroup 2	2 (1, 5)	2 (0, 4.8)	**3 (2, 7)**	2 (1, 5)	**1.5 (1, 3)**	**6 (3, 8)**
Headaches (no/yes)	248/62	112/70	258/74	248/62	248/62	258/74
Headaches subgroup 1	164/33	57/43	210/47	159/29	172/36	222/53
Headaches subgroup 2	84/29	55/27	48/27	89/33	76/26	36/21
Neck pain (no/yes)	255/77	NA	**255/77***	255/77	255/77	**255/77*^**
Neck pain subgroup 1	166/41	NA	**213/44**	165/35	164/52	**228/47**
Neck pain subgroup 2	89/36	NA	**42/33**	90/42	91/25	**27/30**
TMJ pain (no/yes)	302/30	NA	**302/30***	**302/30^**	302/30	**302/30*^**
TMJ pain subgroup 1	192/15	NA	**245/12**	**190/10**	193/23	**261/14**
TMJ pain subgroup 2	110/15	NA	**57/18**	**112/20**	109/7	**41/16**

	STOP Audiometric Variables Clustering	BRC Audiometric Variables Clustering	STOP General Phenotypic Variables Clustering	STOP Tinnitus Discriminating Variables Clustering	STOP-BRC Independent Validation Clustering	STOP-ESIT Independent Validation Clustering
High blood pressure (no/yes)	**282/50***	NA	282/50	**282/50*^**	282/50	282/50
High blood pressure subgroup 1	**162/45**	NA	218/39	**154/46**	177/39	233/42
High blood pressure subgroup 2	**120/5**	NA	64/11	**128/4**	105/11	49/8
Low blood pressure (no/yes)	309/23	NA	309/23	309/23	309/23	**309/23*^**
Low blood pressure subgroup 1	195/12	NA	241/16	190/10	205/11	**263/12**
Low blood pressure subgroup 2	114/11	NA	68/7	119/13	104/12	**46/11**
High cholesterol (no/yes)	**311/21***	NA	311/21	311/21	**311/21^**	311/21
High cholesterol subgroup 1	**187/20**	NA	238/19	181/19	**197/19**	258/17
High cholesterol subgroup 2	**124/1**	NA	73/2	130/2	**114/2**	53/4
Problem staying asleep (no/yes)	289/43	NA	289/43	289/43	289/43	**289/43*^**
Problem staying asleep subgroup 1	178/29	NA	227/30	174/26	187/29	**251/24**
Problem staying asleep subgroup 2	111/14	NA	62/13	115/17	102/14	**38/19**
Problem falling asleep (no/yes)	297/35	NA	297/35	297/35	297/35	**297/35*^**
Problem falling asleep subgroup 1	187/20	NA	236/21	180/20	194/22	**261/14**
Problem falling asleep subgroup 2	110/15	NA	61/14	117/15	103/13	**36/21**
Dental problems (no/yes)	309/23	NA	309/23	309/23	309/23	**309/23^**
Dental problems subgroup 1	195/12	NA	242/15	185/15	199/17	**262/13**
Dental problems subgroup 2	114/11	NA	67/8	124/8	110/6	**47/10**
Dental surgery (no/yes)	233/99	NA	233/99	233/99	**233/99^**	233/99
Dental surgery subgroup 1	141/66	NA	183/74	137/63	**160/56**	197/78
Dental surgery subgroup 2	92/33	NA	50/25	96/36	**73/43**	36/21
Any procedure (no/yes)	186/146	NA	186/146	186/146	186/146	**186/146^**
Any procedure subgroup 1	114/93	NA	153/104	110/90	124/92	**161/114**
Any procedure subgroup 2	72/53	NA	33/42	76/56	62/54	**25/32**

	STOP Audiometric Variables Clustering	BRC Audiometric Variables Clustering	STOP General Phenotypic Variables Clustering	STOP Tinnitus Discriminating Variables Clustering	STOP-BRC Independent Validation Clustering	STOP-ESIT Independent Validation Clustering
Thyroid disorder (no/yes)	310/22	NA	310/22	310/22	310/22	**310/22^**
Thyroid disorder subgroup 1	190/17	NA	241/16	183/17	206/10	**261/14**
Thyroid disorder subgroup 2	120/5	NA	69/6	127/5	104/12	**49/8**
Lyme disease (no/yes)	286/46	NA	286/46	286/46	**286/46^**	286/46
Lyme disease subgroup 1	172/35	NA	226/31	169/31	**180/36**	236/39
Lyme disease subgroup 2	114/11	NA	60/15	117/15	**106/10**	50/7
WHO Quality of Life Psychological subscale (4-20)	15 (14, 17)	NA	**15 (14, 17)***	15 (14, 17)	15 (14, 17)	**15 (14, 17)*^**
Psychological subscale subgroup 1	16 (14, 17)	NA	**16 (15, 17)**	16 (14, 17)	15 (14, 17)	**16 (15, 17)**
Psychological subscale subgroup 2	15 (13.75, 17)	NA	**15 (11.5, 16)**	15 (13, 17)	16 (15, 17)	**13 (11, 15)**
WHO Quality of Life Physical subscale (4-20)	17 (15, 18)	NA	**17 (15, 18)***	17 (15, 18)	17 (15, 18)	**17 (15, 18)***
Physical subscale subgroup 1	17 (15, 18)	NA	**17 (15, 18)**	17 (15, 18)	16 (15, 18)	**17 (15, 18)**
Physical subscale subgroup 2	17 (14, 18)	NA	**15 (13, 17)**	17 (15, 18)	17 (15, 18)	**14 (13, 16)**
WHO Quality of Life Social subscale (4-20)	15 (13, 16)	NA	15 (13, 16)	15 (13, 16)	**15 (13, 16)*^**	15 (13, 16)
Social subscale subgroup 1	15 (13, 16)	NA	15 (13, 16)	15 (13, 16)	**15 (12, 16)**	15 (13, 16)
Social subscale subgroup 2	15 (13, 16)	NA	15 (12, 16)	15 (13, 16)	**16 (13, 17)**	15 (12, 16)
WHO Quality of Life Environmental subscale (4-20)	17 (15, 18)	NA	**17 (15, 18)***	17 (15, 18)	17 (15, 18)	**17 (15, 18)***
Environmental subscale subgroup 1	17 (16, 18)	NA	**17 (16, 18)**	17 (16, 18)	17 (15, 18)	**17 (16, 18)**
Environmental subscale subgroup 2	16 (15, 18)	NA	**16 (14, 17.5)**	17 (15, 18)	17 (15, 18)	**15 (13, 17)**
Tinnitus severity score‡ (0-100)	14 (6, 28)	37.2 (24.0, 50.6)	**14 (6, 28)***	14 (6, 28)	14 (6, 28)	**14 (6, 28)***
Severity subgroup 1	16 (6, 28)	37.19 (23.2, 46.7)	**12 (6, 24)**	14 (6, 26)	17 (6, 30)	**14 (6, 24)**
Severity subgroup 2	14 (6, 26)	37.1 (25.2, 52.9)	**20 (12, 33)**	16 (6, 28)	12 (6, 21.5)	**24 (10, 45)**

	STOP Audiometric Variables Clustering	BRC Audiometric Variables Clustering	STOP General Phenotypic Variables Clustering	STOP Tinnitus Discriminating Variables Clustering	STOP-BRC Independent Validation Clustering	STOP-ESIT Independent Validation Clustering
Tinnitus annoyance scale (0-100)	**10 (2, 25)***	25 (10, 50)	**10 (2, 25)***	10 (2, 25)	**10 (2, 25)^**	10 (2, 25)
Annoyance subgroup 1	**10 (5, 30)**	20 (9, 50)	**10 (2, 20)**	10 (5, 27.5)	**10 (5, 30)**	10 (2, 25)
Annoyance subgroup 2	**5 (1, 10)**	30 (12.5, 58.8)	**20 (6, 50)**	10 (1, 20)	**10 (1, 20)**	10 (5, 50)
Tinnitus worries, annoys or upsets (severely/moderately/slightly/not at all)	**32/98/136/64^**	NA	**32/98/136/64***	32/98/136/64	32/98/136/64	32/98/136/64
Impact subgroup 1	**23/67/69/46**	NA	**21/66/110/58**	22/62/75/39	26/63/87/40	21/83/111/58
Impact subgroup 2	**9/31/67/18**	NA	**11/32/26/6**	10/36/61/25	6/35/49/24	11/15/25/6
Tinnitus awareness (% of total awake time)	**20 (10, 50)***	NA	20 (10, 50)	**20 (10, 50)***	**20 (10, 50)^**	20 (10, 50)
Awareness subgroup 1	**25 (11.3, 60)**	NA	20 (10, 50)	**25 (10, 50)**	**25 (10, 60)**	20 (10, 50)
Awareness subgroup 2	**15 (7, 25)**	NA	25 (15, 65)	**15 (10, 30)**	**20 (10, 30)**	25 (10, 75)
No management for tinnitus (no/yes)	22/310	NA	**22/310***	22/310	22/310	**22/310*^**
No management subgroup 1	13/194	NA	**10/247**	11/189	16/200	**11/264**
No management subgroup 2	9/116	NA	**12/63**	11/121	6/110	**11/46**
Age at tinnitus onset (y)	**33 (19, 45.8)***	50 (36, 59)	33 (19, 45.8)	**33 (19, 45.8)***	33 (19, 45.8)	33 (19, 45.8)
Onset age subgroup 1	**43 (30, 52)**	53 (36, 62)	32 (20, 46)	**43 (29, 53)**	33 (20, 45)	33 (19.5, 48.2)
Onset age subgroup 2	**20.1 (15, 29)**	46.9 (35.3, 56)	35.5 (16.5, 45)	**21 (16, 33)**	32.5 (15.5, 47.5)	28.6 (16, 44.5)
Tinnitus duration (y)	15 (6, 21)	**7.8 (3, 20)***	15 (6, 21)	15 (6, 21)	15 (6, 21)	15 (6, 21)
Duration subgroup 1	14 (6, 21.8)	**12 (4, 20.5)**	14.38 (6, 20)	12.5 (5, 22.8)	15 (7, 22)	15 (7, 21)
Duration subgroup 2	16 (8, 21)	**5.5 (2, 10.4)**	18 (7, 22)	15.5 (8, 20)	15 (5, 20)	10 (5.5, 20)
Sound exposure at tinnitus onset (no/yes)	195/122	141/40	192/140	**195/122***	195/122	192/140
Sound exposure at onset subgroup 1	134/64	81/19	146/111	**132/59**	125/87	162/113
Sound exposure at onset subgroup 2	61/58	60/21	46/29	**63/63**	70/35	30/27
No medication at tinnitus onset (no/yes)	109/223	NA	109/223	109/223	**109/223^**	**109/223^**
No medication at onset subgroup 1	75/132	NA	87/170	70/130	**63/153**	**81/194**
No medication at onset subgroup 2	34/91	NA	22/53	39/93	**46/70**	**28/29**

	STOP Audiometric Variables Clustering	BRC Audiometric Variables Clustering	STOP General Phenotypic Variables Clustering	STOP Tinnitus Discriminating Variables Clustering	STOP-BRC Independent Validation Clustering	STOP-ESIT Independent Validation Clustering
Presence during the day (constant/intermittent)	237/95	165/15	237/95	237/95	**237/95*^**	237/95
Presence pattern subgroup 1	154/53	90/9	181/76	142/58	**169/47**	199/76
Presence pattern subgroup 2	83/42	75/6	56/19	95/37	**68/48**	38/19
Tinnitus daily presence (no/yes)	83/249	NA	83/249	83/249	**83/249***	83/249
Tinnitus daily subgroup 1	43/164	NA	70/187	44/156	**39/177**	68/207
Tinnitus daily subgroup 2	40/85	NA	13/62	39/93	**44/72**	15/42
Tinnitus spatial perception (left ear/both ears, more left/no lateralisation (both ears equally or in the head)/both ears, more right/right ear)	29/35/200/40/11	**36/38/24/34/23*^**	29/51/186/42/12	29/35/200/40/11	**29/35/200/40/11^**	29/51/186/42/12
Spatial perception subgroup 1	22/22/119/25/9	**17/22/14/10/19**	18/41/148/33/7	22/20/116/23/9	**16/31/133/25/7**	25/41/151/39/11
Spatial perception subgroup 2	7/13/81/15/2	**19/16/10/24/4**	11/10/38/9/5	7/15/84/17/2	**13/4/67/15/4**	4/10/35/3/1
Tinnitus quality (tonal/noise/other)	**177/72/83***	NA	177/72/83	177/72/83	177/72/83	177/72/83
Tinnitus quality subgroup 1	**92/55/60**	NA	142/53/62	93/49/58	117/47/52	141/64/70
Tinnitus quality subgroup 2	**85/17/23**	NA	35/19/21	84/23/25	60/25/31	36/8/13
Tinnitus pitch matching (kHz)	**9 (4.8, 12.5)*^**	**7 (5, 10)***	9 (4.8, 12.5)	**9 (4.8, 12.5)***	9 (4.8, 12.5)	9 (4.8, 12.5)
Pitch matching subgroup 1	**8 (4, 11.2)**	**7 (5, 8)**	10 (5.5, 12.5)	**8 (4, 11.2)**	10 (5.3, 12.5)	9 (4.8, 12.5)
Pitch matching subgroup 2	**12.5 (8, 14)**	**8 (6, 10)**	8 (2.8, 12.5)	**11.2 (8, 14)**	8 (4, 12.5)	10 (4.5, 12.5)
Hearing threshold at pitch matched frequency (dB HL)	**36 (18, 56)***	NA	36 (18, 56)	**36 (18, 56)*^**	**36 (18, 56)^**	**36 (18, 56)^**
Hearing at pitch frequency subgroup 1	**50 (30.5, 60)**	NA	35 (20, 56)	**50 (30, 62)**	**42.5 (21.5, 58)**	**39.5 (20, 58)**
Hearing at pitch frequency subgroup 2	**20 (2, 32)**	NA	45 (14, 58)	**18 (2, 38)**	**26 (10, 44)**	**30 (8, 46)**
Tinnitus loudness rating (0-100)	40 (20, 60)	41 (30, 50)	**40 (20, 60)***	40 (20, 60)	40 (20, 60)	40 (20, 60)
Loudness rating (0-100 subgroup 1	40 (20, 70)	39 (30, 50)	**40 (20, 50)**	40 (25, 60)	40 (22.5, 60)	40 (20, 60)
Loudness rating (0-100 subgroup 2	35 (20, 50)	42.5 (30, 52.75)	**50 (30, 70)**	37.5 (20, 50)	35 (20, 50)	40 (30, 60)

	STOP Audiometric Variables Clustering	BRC Audiometric Variables Clustering	STOP General Phenotypic Variables Clustering	STOP Tinnitus Discriminating Variables Clustering	STOP-BRC Independent Validation Clustering	STOP-ESIT Independent Validation Clustering
Tinnitus loudness matching (dB HL)	**42 (25, 60)*^**	NA	42 (25, 60)	**42 (25, 60)***	42 (25, 60)	42 (25, 60)
Loudness matching subgroup 1	**54 (37.5, 64.3)**	NA	40 (25.8, 58.5)	**54 (35.3, 65)**	46 (30, 62)	42 (26, 60)
Loudness matching subgroup 2	**25 (12, 37)**	NA	48 (24.5, 61)	**25 (12, 42)**	34 (15, 50)	35 (22, 48)
Masking narrow band noise threshold (dB HL)	**17.5 (0, 30)*^**	NA	17.5 (0, 30)	**17.5 (0, 30)*^**	**17.5 (0, 30)^**	17.5 (0, 30)
Masking threshold subgroup 1	**30 (15, 40)**	NA	15 (0, 30)	**30 (20, 40)**	**20 (5, 35)**	20 (5, 35)
Masking threshold subgroup 2	**0 (-5, 5)**	NA	20 (2.5, 40)	**0 (-5, 10)**	**10 (0, 25)**	5 (0, 30)
Varying tinnitus loudness (no/yes from day to day from TSCHQ; stable/ sometimes fluctuating/always fluctuating over a day from ESIT-SQ)	**85/219***	79/103	124/115/68	**85/219***	85/219	124/115/68
Varying loudness subgroup 1	**66/122**	43/57	95/87/52	**64/117**	64/142	103/94/55
Varying loudness subgroup 2	**19/97**	36/46	29/28/16	**21/102**	21/77	21/21/13
Tinnitus increased by loud noise (no/yes)[£]	**93/142***	103/67	228/104	**93/142*^**	93/142	228/104
Increased by noise subgroup 1	**68/71**	62/30	186/71	**70/63**	59/101	189/86
Increased by noise subgroup 2	**25/71**	41/37	42/33	**23/79**	34/41	39/18
Tinnitus increased by alcohol (no/yes)	**309/23***	NA	309/23	**309/23***	309/23	309/23
Increased by alcohol subgroup 1	**201/6**	NA	242/15	**195/5**	199/17	257/18
Increased by alcohol subgroup 2	**108/17**	NA	67/8	**114/18**	110/6	52/5
Tinnitus increased by stress (no/yes)	**206/126***	82/96	206/126	**206/126***	206/126	**206/126***
Increased by stress subgroup 1	**145/62**	53/47	170/87	**139/61**	139/77	**182/93**
Increased by stress subgroup 2	**61/64**	29/49	36/39	**67/65**	67/49	**24/33**
Tinnitus increased by jaw movement (no/yes)	**267/65*^**	NA	267/65	267/65	**267/65^**	**267/65^**
Increased by jaw movement subgroup 1	**180/27**	NA	212/45	168/32	**165/51**	**227/48**
Increased by jaw movement subgroup 2	**87/38**	NA	55/20	99/33	**102/14**	**40/17**

	STOP Audiometric Variables Clustering	BRC Audiometric Variables Clustering	STOP General Phenotypic Variables Clustering	STOP Tinnitus Discriminating Variables Clustering	STOP-BRC Independent Validation Clustering	STOP-ESIT Independent Validation Clustering
Tinnitus increased by poor quality sleep (no/yes)	231/101	NA	**231/101***	231/101	231/101	231/101
Increased by poor sleep subgroup 1	150/57	NA	**191/66**	147/53	147/69	199/76
Increased by poor sleep subgroup 2	81/44	NA	**40/35**	84/48	84/32	32/25
Thoughts of conditions related to increased tinnitus (no/yes)	**221/111***	NA	221/111	221/111	221/111	**221/111*^**
Increased conditions subgroup 1	**152/55**	NA	176/81	143/57	148/68	**200/75**
Increased conditions subgroup 2	**69/56**	NA	45/30	78/54	73/43	**21/36**
Tinnitus reduced by jaw movement (no/yes)	317/15	NA	317/15	317/15	317/15	**317/15^**
Reduced by jaw movement subgroup 1	199/8	NA	247/10	195/5	207/9	**267/8**
Reduced by jaw movement subgroup 2	118/7	NA	70/5	122/10	110/6	**50/7**
Tinnitus reduced by driving (no/yes)	310/22	NA	310/22	310/22	310/22	**310/22^**
Reduced by driving subgroup 1	194/13	NA	241/16	187/13	205/11	**262/13**
Reduced by driving subgroup 2	116/9	NA	69/6	123/9	105/11	**48/9**

Values for numeric variables: median (first, third quartile). Values for categorical variables: frequency count per category. For variables assessed both with the ESIT-SQ and the TSCHQ (e.g., presence of vertigo/dizziness), the variable from the ESIT-SQ is shown for the STOP General Phenotypic Variables Clustering and the STOP-ESIT Independent Validation Clustering, and the variable from the TSCHQ for the STOP Audiometric Variables Clustering, the STOP Tinnitus Discriminating Variables Clustering, and the STOP-BRC Independent Validation Variables Clustering. *Statistically significant difference between subgroup 1 and subgroup 2. ^Non-zero coefficient in LASSO regression. Bold: values of variables that either showed statistically significant differences between subgroup 1 and subgroup 2 (*) or that had non-zero coefficient in LASSO regression (^). †Hospital Anxiety Depression Scale for Depression and Anxiety for the STOP dataset, Beck Depression Inventory and Beck Anxiety Inventory for the BRC dataset. ‡ Tinnitus Handicap Inventory for the STOP dataset, Tinnitus Handicap Questionnaire for the BRC dataset. [£]For the STOP General Phenotypic Variables Clustering and the STOP-ESIT Independent Validation Clustering the question comes from the ESIT-SQ and asks if tinnitus is increased by loud sounds; for all other subgroupings it comes from the TSCHQ which asks if tinnitus is worsened by loud noise. HADS: Hospital Anxiety Depression Scale; HQ: Hyperacusis Questionnaire; NA: Not applicable; PSQ: Perceived Stress Questionnaire; WHO: World Health Organization.

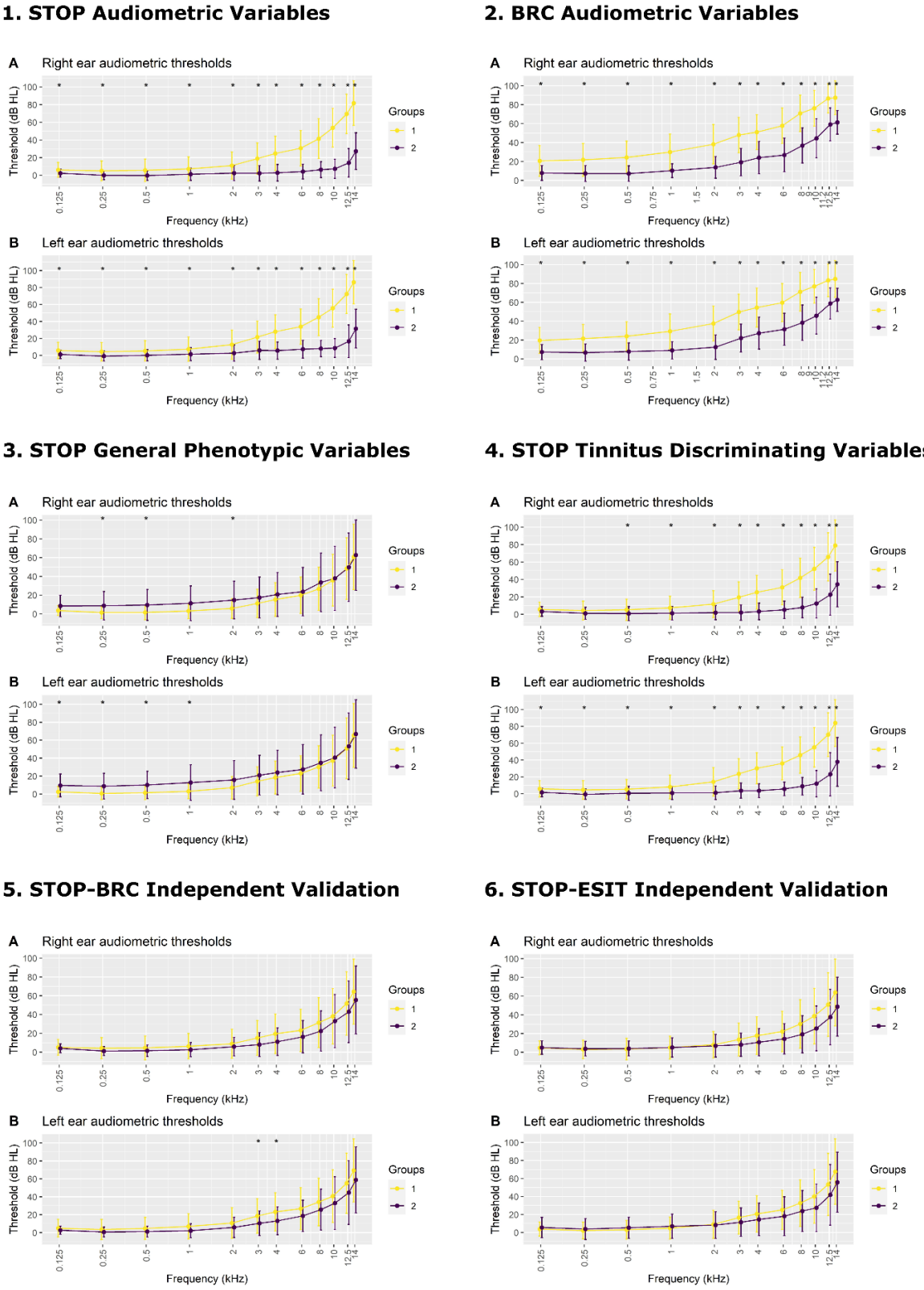

Figure 11. Mean audiometric thresholds for each subgroup of the six selected subgroupings. Upper panels (A) show thresholds from right ears and lower panels (B) thresholds from left ears. Error bars present standard deviations for each threshold. Asterisks denote statistically significant differences between subgroups ($p \leq 0.001$).

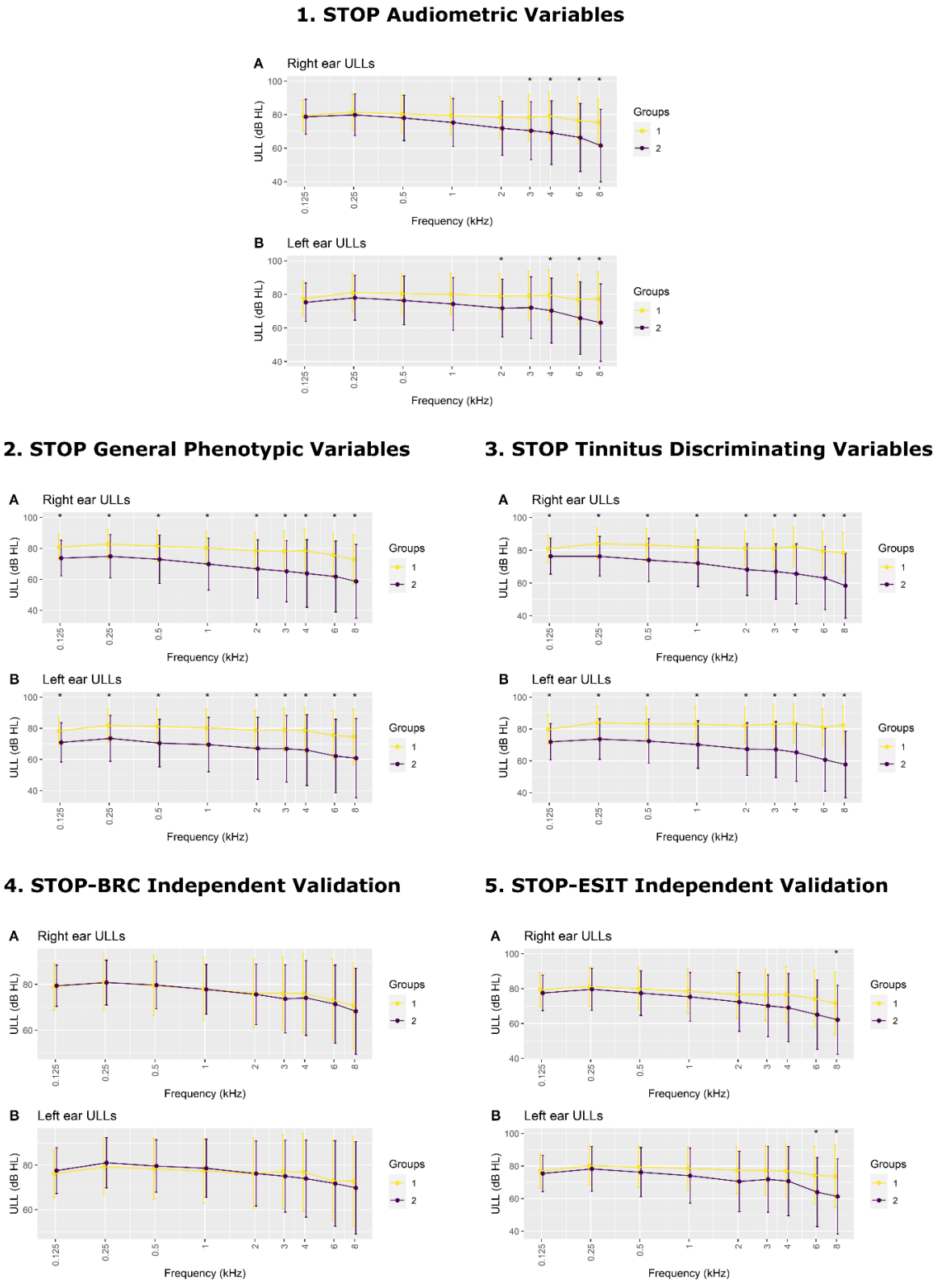

Figure 12. Uncomfortable Loudness Levels (ULL) for each subgroup of the five selected subgroupings from the STOP dataset (means and standard deviations). Upper panels (A) show levels from right ears and lower panels (B) levels from left ears. Asterisks denote statistically significant differences between subgroups (p≤0.001).

5.3.3.3 Descriptions of subgroupings from the STOP Audiometric Variables Clustering, the STOP Tinnitus Discriminating Variables Clustering, and the BRC Audiometric Variables Clustering

Similar subphenotypic patterns were found from the STOP Audiometric Variables Clustering and the STOP Tinnitus Discriminating Variables Clustering. In both cases, subgroup 1 was characterised by older age, worse overall hearing thresholds, more often use of hearing aids, self-reported hearing problems, lower stress and anxiety scores, relatively more often high blood pressure, higher percentage of time being aware of tinnitus, older age at tinnitus onset, lower tinnitus pitch-matched frequency, higher hearing threshold at pitch matched frequency, higher tinnitus matching loudness, higher tinnitus masking noise thresholds, and relatively lower frequency of tinnitus having varying loudness, and being reduced by loud noise, stress, or alcohol (Table 19). However, the information from the variables hearing aid use, tinnitus awareness, age at onset, varying tinnitus loudness, and tinnitus being increased by stress and alcohol were probably redundant as these were not important in either of the multivariable LASSO regression models. Moreover, in both cases, subgroup 1 reported relatively less often vertigo/dizziness than group 2. Although this difference was not significant at alpha level 0.001 (p=0.0193 and 0.0150 respectively), vertigo was important for classification (non-zero coefficient) in the LASSO regression models. In contrast to these similarities, subgroup 1 from the STOP Audiometric Variables Clustering had more often high cholesterol, higher tinnitus annoyance scores, relatively higher frequency of noise-like rather than tonal tinnitus, lower frequency of tinnitus being increased by jaw movements, and reported less often thought of any conditions being related to increased tinnitus than subgroup 2. Participants in this subgroup also reported relatively less often being slightly worried, annoyed, or upset by tinnitus than participants in subgroup 2 (difference not significant at alpha level 0.001 [p=0.004], but variable important for the LASSO regression model). On the other hand, subgroup 1 participants from the STOP Tinnitus Discriminating Variables Clustering reported relatively more often presbycusis and less often sound exposure at tinnitus onset than subgroup 2 participants. They also reported relatively less often TMJ pain than participants from subgroup 2. This difference was not significant at alpha level 0.001 (p=0.0028), but the variable was important for the LASSO classification model.

Audiometric profiles were similar across these two subgroupings (from the STOP Audiometric Variables Clustering and the STOP Tinnitus Discriminating Variables Clustering), with subgroup 2 having relatively unaffected hearing and subgroup 1 having overall higher hearing thresholds, increasing at higher frequencies in both cases (Figure 11, plots 1 and 4). Subgroup 2 from both approaches showed some deterioration at extended-high frequencies. ULLs were significantly higher at all frequencies for subgroup 1 from the STOP Tinnitus Discriminating Variables Clustering (Figure 12, plot 3). Subgroup 1 from the STOP Audiometric Variables Clustering also showed significantly higher levels than subgroup 2, but only for frequencies at and above 2 kHz (Figure 12, plot 1).

The subgrouping from the BRC Audiometric Variables Clustering showed similar subphenotypic patterns to the STOP Audiometric Variables Clustering and STOP Tinnitus Discriminating Variables Clustering subgroupings, with subgroup 1 having older age, higher overall hearing thresholds, relatively more often use of hearing aids and self-reported hearing problems, and lower tinnitus pitch-matched frequency than subgroup 2 (Table 19). This subgroup also had longer tinnitus duration, and differed in the distribution of tinnitus spatial perception, with relatively higher frequency of tinnitus in the right ear, and lower frequency of tinnitus in the left ear or in both ears but more right, compared to subgroup 2. In addition, hearing thresholds were higher than in subgroup 2 at all frequencies (Figure 11, plot 2). However, in this case hearing thresholds for subgroup 2 were also affected at both high and extended high frequencies (in contrast to subgroup 2 from the STOP subgroupings that were only somewhat affected at extended-high frequencies).

5.3.3.4 Descriptions of subgroupings from the STOP General Phenotypic Variables Clustering and the STOP-ESIT Independent Validation Clustering

The subgroupings from the STOP General Phenotypic Variables Clustering and the STOP-ESIT Independent Validation Clustering also had similar phenotypic patterns. In both subgroupings, subgroup 1 was characterised by lower hyperacusis scores and relatively less often reporting external sounds to be a moderate, big, or very big problem, lower stress, anxiety, and depression scores, relatively less often reporting neck pain or TMJ pain, higher WHO quality of life scores for the psychological, physical, and environmental subscales, lower tinnitus severity scores, and relatively more often reporting no management for tinnitus compared to subgroup 2 (Table 19).

Of these two subgroupings, the one resulting from the STOP-ESIT Independent Validation clustering was more robust as indicated by both the Jaccard stabilities and the predictive performance of the LASSO regression models. HQ score, WHO physical and environmental quality of life subscale scores, and tinnitus severity score were not important in this LASSO regression model. In addition, subgroup 1 from the STOP General Phenotypic Variables Clustering had better overall hearing thresholds, reported relatively less often use of hearing aids, having moderate or severe problems with hearing, and having vertigo, had lower tinnitus annoyance scores and loudness ratings, and reported relatively less often being moderately or severely worried, annoyed, or upset by tinnitus, and tinnitus being increased by poor sleep quality than subgroup 2. On the other hand, subgroup 1 from the STOP-ESIT Independent Validation Clustering had relatively more males, and reported relatively less often low blood pressure, problems falling or staying asleep, tinnitus being increased by stress, or any conditions related to increased tinnitus compared to subgroup 2. Participants in this subgroup were also older, reported relatively less often having acute otitis, dental problems, or thyroid disorder, or having undergone any procedure, had higher hearing thresholds at pitch-matched frequency, and reported relatively more often taking no medication at tinnitus onset and relatively less often tinnitus being increased or reduced by jaw movement, or being reduced by driving compared to subgroup 2 from the STOP-ESIT Independent Validation Clustering subgrouping. Although these differences were not significant at alpha level 0.001 ($p=0.003$, 0.039, 0.002, 0.056, 0.034, 0.046, 0.005, 0.043, 0.006, and 0.006, respectively), the variables were important for the LASSO regression model.

Looking at the audiometric profiles for the two subgroupings (from the STOP General Phenotypic Variables Clustering and the STOP-ESIT Independent Validation Clustering), these were clearly different (Figure 11, plots 3 and 6). Subgroup 2 from the STOP General Phenotypic Variables Clustering had slightly worse thresholds in low and middle frequencies (significantly different to those of subgroup 1 only for frequencies up to 2 kHz). This was not the case for the STOP-ESIT Independent Validation Clustering subgrouping, where instead subgroup 1 showed slightly worse hearing thresholds than subgroup 2 at high frequencies (differences not statistically significant). Regarding ULLs, these were significantly higher at all frequencies for subgroup 1 compared to subgroup 2 from the STOP General Phenotypic Variables Clustering (Figure 12, plot 2), despite the lower hearing thresholds at low frequencies.

Subgroup 1 from the STOP-ESIT Independent Validation Clustering, on the other hand, had higher levels only at 6 kHz for the left ear and at 8 kHz for both ears.

5.3.3.5 Descriptions of subgroupings from the STOP-BRC Independent Validation Clustering

Finally, the subgrouping from the STOP-BRC Independent Validation Clustering was rather distinct, with subgroup 1 including more male participants, and subgroup 2 including only female participants (Table 19). In addition, participants in subgroup 1 had higher height and weight, overall worse hearing thresholds, reported relatively more often use of hearing aids and self-reported hearing problems, had higher depression scores, and lower WHO quality of life scores on the social subscale, and reported more often constant (rather than intermittent) and daily (rather than non-daily) tinnitus than participants in subgroup 2. The variable about tinnitus being present daily or not daily was not important for the LASSO regression model. The same was the case for sex, but that was due to quasi-complete separation as explained above (section 5.3.3.1). This subgroup also had relatively less often higher education, consumed more alcohol, had higher hearing thresholds at pitch matched frequency, reported less often dental surgery and more often high cholesterol and Lyme disease, had relatively more often tinnitus localised in both sides but more left, had higher tinnitus annoyance, awareness, and masking noise thresholds and reported relatively more often taking no medication at tinnitus onset and tinnitus being increased by jaw movements compared to subgroup 2. These differences were not statistically significant at alpha level 0.001 (p=0.02, 0.014, 0.0013, 0.044, 0.01, 0.046, 0.033, 0.029, 0.002, 0.013, 0.066, and 0.008, respectively), but the variables were important for the LASSO regression model.

Audiometric thresholds were slightly higher for subgroup 1 compared to subgroup 2, but the difference was statistically significant only for the left ear at 3 and 4 kHz (Figure 11, plot 5). ULLs did not differ significantly between the two subgroups at any frequency (Figure 12, plot 4).

5.3.4 Three subphenotypic patterns emerging from the data modelling

5.3.4.1 Overview

As discussed in sections 5.3.3.3 and 5.3.3.4, there were similarities across the six assessed subgroupings. This allowed discovery and extraction of three main subphenotypic patterns. The first main pattern was discovered by both the STOP Audiometric Variables Clustering and the STOP Tinnitus Discriminating Variables Clustering approaches. It was extracted from Table 19 by considering variables that differed significantly among subgroups in the subgroupings (denoted by asterisks) from both these approaches. Specifically, the 17 variables that appear with an asterisk (*) in both column 1 (STOP Audiometric Variables Clustering) and 4 (STOP Tinnitus Discriminating Variables clustering) of Table 19 were used to define the phenotypic profile of this first main pattern. A similar pattern was discovered in the BRC dataset by the BRC Audiometric Variables Clustering, but only the two subgroupings from the STOP dataset (having many common variables) were used to characterise this pattern. The second main pattern was discovered by both the STOP General Phenotypic Variables Clustering and the STOP-ESIT Independent Validation Clustering approaches. It was extracted from Table 19 by considering variables that differed significantly among subgroups in both selected subgroupings from these two approaches. For this second main pattern, the 12 variables that appear with an asterisk (*) in both column 3 (STOP General Phenotypic Variables Clustering) and 6 (STOP-ESIT Independent Validation Clustering) of Table 19 were used to define its phenotypic profile. The last main pattern was the one discovered by the STOP-BRC Independent Validation Clustering approach. An overview of the characteristics of these three subphenotypic patterns is presented in Table 20.

One of the two subphenotypes from each of these patterns are described in sections 5.3.4.2 – 5.3.4.4. The other showed the exact opposite associations of characteristics. Table 21 shows the variables characterising each pattern mapped to the variable concepts introduced in Chapter 2. Across the 27 included variable concepts, those characterising more than one subphenotypic pattern were hearing ability, hearing aid use, stress-related symptoms, symptoms of anxiety, symptoms of depression, and quality of life.

Table 20. Summary of the characteristics of one subphenotype from each of the three most robust subphenotypic patterns.

Pattern 1 (discovered by the STOP Audiometric Variables Clustering and the STOP Tinnitus Discriminating Variables Clustering)	Pattern 2 (discovered by the STOP General Phenotypic Variables Clustering and the STOP-ESIT Independent Validation Clustering)	Pattern 3 (discovered by the STOP-BRC Independent Validation Clustering)
▪ Younger age ▪ Better hearing ▪ Higher stress and anxiety scores ▪ Less often high blood pressure ▪ Higher tinnitus matched pitch ▪ Lower tinnitus matched loudness ▪ Lower tinnitus masking thresholds ▪ More easily modulated tinnitus (increased by loud sounds, stress, or alcohol, varying loudness)	▪ Higher hyperacusis symptomatology ▪ Higher stress, anxiety, and depression scores ▪ Lower quality of life (psychological, physical, and environmental subscales) ▪ More often neck and TMJ problems ▪ Higher tinnitus burden (higher overall severity and more often preceding tinnitus treatments)	▪ More males ▪ Worse hearing ▪ Higher depression scores ▪ Lower quality of life (social subscale) ▪ More often constant and daily tinnitus

Table 21. Summary of variables differentiating the three main subphenotypic patterns, mapped to the variable concepts from the framework presented in Chapter 2.

	Pattern 1	Pattern 2	Pattern 3
VARIABLE CONCEPTS			
Non-tinnitus-specific			
Age	✓	-	-
Sex	-	-	✓‡
Weight*	-	-	✓
Height*	-	-	✓
Hearing ability	✓	-	✓
Hearing aid use	✓‡	-	✓
Problems with sounds	-	✓	-
Stress-related symptoms	✓	✓	-
Symptoms of anxiety	✓	✓	-
Symptoms of depression	-	✓	✓
Neck problems (neck pain)	-	✓	-
Mandible problems (TMJ pain)	-	✓	-
Circulatory system problems (high BP)	✓	-	-
Quality of life	-	✓	✓
Tinnitus-specific			
Overall severity	-	✓‡	-

	Pattern 1	Pattern 2	Pattern 3
Age at onset	✓‡	-	-
Awareness	✓‡	-	-
Preceding tinnitus treatments (any management)	-	✓	-
Presence pattern (constant or intermittent, daily or not daily)	-	-	✓
Pitch	✓	-	-
Pitch and hearing loss profile (hearing threshold at pitch matched frequency)	✓	-	-
Loudness	✓	-	-
Varying perception (varying loudness)	✓‡	-	-
External sound effect (increased by sound, masking threshold)	✓	-	-
Psychological factors effect (increased by stress)	✓‡	-	-
Substances effect (increased by alcohol)	✓‡	-	-

Variables presented were significant for both the STOP Audiometric Variables and the STOP Tinnitus Discriminating Variables Clustering subgrouping for pattern 1; for both the STOP General Phenotypic Variables and the STOP-ESIT Independent Validation Clustering subgrouping for pattern 2; for the STOP-BRC Independent Validation Clustering subgrouping for pattern 3. ‡Variables that were not important in either LASSO model (for the STOP Audiometric Variables and the STOP Tinnitus Discriminating Variables Clustering subgrouping) for pattern 1; in the LASSO model for the STOP-ESIT Independent Validation Clustering subgrouping for pattern 2; in the LASSO model for the STOP-BRC Independent Validation Clustering subgrouping for pattern 3. *Variables not included in the framework developed in Chapter 2.

5.3.4.2 Subphenotype 1: Younger age, better hearing, higher stress and anxiety, less often high blood pressure, tinnitus with higher matched pitch, lower matched loudness, lower masking thresholds, and more easily modulated

All three selected subgroupings from the STOP Audiometric Variables Clustering, the BRC Audiometric Variables Clustering, and the STOP Tinnitus Discriminating Variables Clustering included a subgroup with smaller size, younger age, better audiometric thresholds and self-reported hearing function, using less often hearing aids, and having higher tinnitus pitch matched frequency. This subgroup from both STOP datasets was also characterised by worse stress and anxiety symptomatology scores, relatively less often having high blood pressure, being aware of tinnitus for a smaller percentage of time, younger age at tinnitus onset, lower matched tinnitus loudness, lower threshold for tinnitus masking, lower hearing threshold at pitch matched frequency, tinnitus relatively more often having varying loudness, and being increased by loud sounds, alcohol, and stress.

5.3.4.3 Subphenotype 2: Higher symptomatology of hyperacusis, stress, anxiety, and depression, lower quality of life, more often neck and TMJ pain, and higher tinnitus burden

The second pattern was identified by both the STOP General Phenotypic Variables Clustering and the STOP-ESIT Independent Validation Clustering. In both cases, a subgroup with much smaller sample size, increased hyperacusis symptomatology (HQ score and ESIT-SQ question asking about problems with sounds) and mental health problems (stress, anxiety, and depression), lower WHO quality of life scores for the psychological, physical, and environmental subscales, having relatively more often neck and TMJ pain, with higher tinnitus severity scores, and reporting relatively more often having any management for tinnitus was identified.

5.3.4.4 Subphenotype 3: More males, worse hearing, greater depression symptoms, lower quality of life, and more often constant and daily tinnitus

The third pattern was discovered by the STOP-BRC Independent Validation Clustering approach. The identified subgrouping included a larger subgroup with more males, higher weight and height, with worse hearing thresholds, more often self-reported hearing problems and use of hearing aids, higher depression scores, lower WHO quality of life (social subscale), reporting relatively more often constant (rather than intermittent) tinnitus, and relatively more often daily tinnitus.

5.4 Discussion

5.4.1 Overview of results

In this study, multiple clustering analyses using different techniques were undertaken to identify robust tinnitus subphenotypes across three independent tinnitus-specific datasets. Each subgroup identified across the multiple clustering solutions was assessed independently. Subgroupings were ranked using a standardised validation protocol, leading to the selection and further assessment of six subgroupings. The algorithms that discovered these subgroupings were k-means, archetypal analysis, density-based clustering, and partitioning around medoids. Jaccard stability ranged from 0.80 to 0.96, and silhouette from 0.23 to 0.5. All subgroupings had both distinct general and tinnitus-specific characteristics. High classification accuracy of five out of six LASSO logistic regression models developed for these subgroupings further validated that they had strong phenotypic structure. Three main phenotypic patterns emerging across these

subgroupings were generalised. These are discussed further in section 5.4.2 and are related to the wider published literature.

5.4.2 Three robust subphenotypic patterns

5.4.2.1 Subphenotypic pattern 1

This subphenotypic pattern included a subphenotype reflecting those with younger age, better hearing, higher stress and anxiety symptomatology, tinnitus having higher matched pitch, lower matched loudness, lower masking thresholds, and being more easily modulated by loud sounds, stress, or alcohol. Similar patterns have been described in previous studies. In a data-driven approach, Andersson and McKenna (1998) clustered tinnitus data using four variables: depression symptomatology score, pure tone audiometry average, tinnitus loudness matching, and minimum masking level. They identified three clusters one of which had a similar phenotypic pattern with the one described here; it included younger participants, with high depression symptomatology scores and low values in hearing thresholds, loudness matching and tinnitus masking levels. Moreover, in a hypothesis-driven approach, Mores et al. (2019) grouped a tinnitus population based on hearing function (assessing thresholds up to 8 kHz) and showed that the normal hearing subphenotype had higher matched pitch, lower matched loudness, and lower minimum masking levels. One thing to investigate in future studies is if this phenotypic pattern is dependent on age (e.g., due to different environmental exposures), hearing loss, or both. Due to the high correlation of age and hearing loss, to answer this question careful selection of participants would be necessary, including sufficient numbers of younger people with hearing loss and older people with normal hearing.

The importance of differentiating tinnitus subphenotypes based on the hearing profile has been highlighted in many more studies (Jeon et al., 2016; Langguth et al., 2017b; Vanneste and De Ridder, 2016; Vielsmeier et al., 2015). Categorising audiometric profiles is a research area on its own, and many audiometric subphenotypes have been proposed (Dubno, 2019; Parthasarathy et al., 2020). However, it is not clear how many distinct hearing profiles should be differentiated when assessing people with tinnitus. An interesting audiometric subphenotype is the one characterised by asymmetric hearing loss. It was shown that the pattern of hearing loss asymmetry correlates with tinnitus laterality, indicating a relationship with tinnitus generating mechanisms

(Genitsaridi et al., 2020b). A tinnitus subphenotype with asymmetric hearing loss was discovered in a preliminary unsupervised learning analysis of this study (data not shown). However, this subphenotype did not meet the more rigorous validation thresholds set out here. This might be due to the small number of people with asymmetric hearing loss in the analysed datasets. Nevertheless, this study confirmed that the most robust audiometric tinnitus subphenotypes are based on the mean thresholds (particularly for higher frequencies) from both ears.

5.4.2.2 Subphenotypic pattern 2

The second subphenotype reflected those with higher symptomatology of hyperacusis, stress, anxiety, and depression, lower quality of life, more often neck and TMJ pain, and higher tinnitus burden. A tinnitus subphenotype defined by the presence of hyperacusis is extensively discussed in tinnitus research. Considering theoretical frameworks for tinnitus subtyping, Jastreboff and Hazell (2004) suggested managing people with hyperacusis with specialised protocols. Using an unsupervised learning approach, a similar pattern to that described here, having high hyperacusis, anxiety, depression, and tinnitus severity scores, was identified in a clustering analysis reported by Tyler et al. (2008). In addition, many hypothesis-driven studies have identified similar phenotypic patterns. Schecklmann et al. (2014) examined two subphenotypes of tinnitus; with or without hyperacusis. Hyperacusis was defined with a single questionnaire item from the TSCHQ asking whether sounds cause pain or physical discomfort. They showed an association between presence of hyperacusis and having more often TMJ complaints and neck pain, lower quality of life scores (WHO physical and psychological subscales), higher depression symptomatology and tinnitus severity scores, more often preceding tinnitus treatments, and tinnitus being more often influenced by stress. In addition, the hyperacusis subphenotype included more female participants; the difference was statistically significant but was not considered important by the authors due to its small effect size. In Schecklmann et al. (2014), presence of hyperacusis was also associated with tinnitus being more often modulated by somatic manoeuvres and noise. In a similar approach, Kojima et al. (2017) showed that people with tinnitus and hyperacusis (defined as in Schecklmann et al., 2014) had more often neck pain, higher tinnitus annoyance, and preceding treatment for tinnitus, and tinnitus was more often worsened by stress. In this study, the hyperacusis subphenotype was also characterised by having more often hearing impairment,

headache, other pain syndromes, vertigo/dizziness, and psychiatric problems, and by higher ratings for tinnitus loudness, and tinnitus being more often worsened by noise and influenced by taking a nap. Moreover, Edvall et al. (2019) investigated subphenotypes of tinnitus based on the presence or not of TMJ complaints. This study also showed an association between TMJ complaints and increased hyperacusis symptomatology, as well as female sex, neck pain, increased stress, anxiety, and depression symptomatology, reduced quality of life (physical, psychological, social, and environmental WHO subscales), increased tinnitus burden, and worsening of tinnitus by stress. Many more variables differed across the two predefined subphenotypes, such as tinnitus being affected by somatic manoeuvres. However, restricting the analysis only to people with severe tinnitus (THI score >58) the only difference between the subphenotypes with and without TMJ complaints were for psychological WHO quality of life scores, and tinnitus being affected by somatic manoeuvres and stress.

The association between higher tinnitus burden and hyperacusis was also shown in Hiller and Goebel (2007) and Cederroth et al. (2020), which makes it a consistent finding across all reviewed studies. In the first study, hyperacusis was defined based on the responses to three questions, whereas in the second it was defined as in Schecklmann et al. (2014). In addition, the association of the hyperacusis subphenotype with TMJ pain, neck pain and tinnitus more often influenced by stress was found in four studies including the current (Edvall et al., 2019; Kojima et al., 2017; Schecklmann et al., 2014). Future work should focus on further understanding the association between these variables: hyperacusis, mental health (stress, anxiety, and depression) and somatic (TMJ and neck) problems in people with tinnitus and on whether there is a better way to subphenotype tinnitus based on these.

An interesting observation from this study was that ULLs associated with self-reported hyperacusis and hearing threshold deficits in a complex way. In subgroupings where there was no differentiation of self-reported hyperacusis, ULLs seemed to be increased when hearing thresholds were increased. However, this could have been due to the fact that ULLs cannot be lower than hearing thresholds, since results by Sheldrake et al. (2015) suggested that there was no other dependence of LDLs on hearing thresholds. In accordance, subgroup 1 from STOP General phenotypic variables, with higher self-reported hyperacusis symptomatology, had higher ULLs in low frequencies than

subgroup 2, despite hearing thresholds being overall lower. Although ULLs measured in dB HL have been used previously to define hyperacusis and have been shown to be correlated with questionnaire-based measures of hyperacusis (Aazh et al., 2018; Aazh and Moore, 2017), other researchers showed that this test is not appropriate as a single test for hyperacusis (Sheldrake et al., 2015). Carefully formulated questions and questionnaires, capturing the subjective nature of the hyperacusis symptomatology (Margol-Gromada et al., 2020; Meeus et al., 2010), or a combination of ULLs and questionnaires (Aazh and Moore, 2017) might be more appropriate for assessing hyperacusis.

5.4.2.3 Subphenotypic pattern 3

The final subphenotype reflected those more likely to be male, with worse hearing, greater depression symptoms, lower quality of life (social subscale), and more often constant and daily tinnitus. The characteristics of tinnitus subphenotypes based on the temporal manifestation of tinnitus (specifically if it is constant or intermittent) have been investigated in two previous hypothesis-driven studies. Burkart et al. (2019) investigated characteristics of participants with constant or intermittent tinnitus, experiencing tinnitus during the past year 'continuously' or 'temporarily time and again', respectively. They showed that the constant tinnitus subphenotype had more often depression and social isolation, and longer tinnitus duration. In this study, constant tinnitus was also associated with higher tinnitus burden. Similarly, Koops et al. (2019) investigated characteristics of tinnitus subgroups with constant or intermittent tinnitus, defined based on the question 'Is your tinnitus continuous or intermittent?'. This study showed that the constant subphenotype included more males (most important variable for a random forest classification model), had slightly higher depression symptomatology scores (variable significant in a logistic regression model), higher tinnitus burden, and longer tinnitus duration.

It is worth noting that both these studies found a longer tinnitus duration and higher tinnitus burden for those with constant tinnitus (Burkart et al., 2019; Koops et al., 2019). In the study reported here, the subphenotype 3 with more constant tinnitus cases did show a tendency for higher tinnitus severity, but the difference did not reach statistical significance ($p=0.013$; not reported in results). As for tinnitus duration, in a preliminary analysis of the ESIT data, a small subgroup (ten participants) with intermittent tinnitus and shorter tinnitus duration and earlier age at tinnitus onset was identified (data not

shown). However, no such difference was found between the two subphenotypes identified here. This could be explained by the fact that both subphenotypes were mixed with both cases with constant and intermittent tinnitus.

5.4.3 Strengths and limitations

The use of multiple independent datasets for the unsupervised discovery of tinnitus subphenotypes is a novelty for tinnitus research. Previous studies applying unsupervised learning for tinnitus subphenotyping were reviewed in Chapter 2 (Appendix 2.3). None of the published studies have analysed data from more than one research centre. Analysing multiple tinnitus-specific datasets provided many opportunities, such as to assess replicability of subphenotypes, as done in this study. This was made possible because of the use of standardised measures for tinnitus assessment (Genitsaridi et al., 2019; Langguth et al., 2007), resulting in many common variables across datasets. However, one limitation of this study was that the analysed datasets were not collected with the aim of being analysed in this way (except for the ESIT dataset). Therefore, only a subset of the available variables was comparable across datasets. In addition, the characteristics of the sampling populations were rather distinct across the three datasets, with the STOP dataset coming from a population-based cohort and the BRC and ESIT datasets being more representative of a hearing-impaired cohort. This might have affected the Independent Dataset Validation analyses, where the most highly rated discovered subgroupings would be those less affected by the differences in sampling populations. However, this diversity in datasets was also a strength of this study allowing to investigate tinnitus heterogeneity across distinct tinnitus populations.

Another limitation of this study was that some important data that could be useful for tinnitus subphenotyping was missing from all available datasets. This included noise exposure history, and personality and coping traits, that have been shown to be relevant for tinnitus heterogeneity and can be assessed with self-report (Budd and Pugh, 1996; Norena, 2011a; Wielopolski et al., 2017). However, the ESIT-SQ does not assess these characteristics, as multiple questions are needed to accurately capture these. In addition, neuroimaging and genomic data are important for investigating tinnitus heterogeneity (Schmidt et al., 2018; Watabe et al., 2020), but were not available in this study. These would allow investigating if the identified subphenotypes were linked to differences in neural activity or in genotype, indicating that they represent distinct tinnitus endotypes. Moreover, information about responses to specific treatments would have been useful

for validating clinical significance of the identified subphenotypes but was not available.

Another novelty of this work was the application of multiple combinations of clustering and validation approaches in search of the most robust tinnitus subphenotypes. With the exceptions of van den Berge et al. (2017) and Schecklmann et al. (2012), where results when including different combinations of variables (two and three different analyses, respectively) were examined, all other studies reviewed in Chapter 2 only described the results from a single clustering analysis. In contrast, in this study multiple approaches were undertaken. These included combinations of various sets of included variables, multiple algorithms, various validation approaches, and multiple datasets. An important finding was that even though the main aim was the same, optimal solutions were discovered by different algorithms across different approaches. In the absence of any gold standard methodology the result is important because it demonstrates the need to try out different techniques for unsupervised learning (Pierre-Jean et al., 2020; Rodriguez et al., 2019).

One more novelty of this work was the choice to assess each identified subgroup separately rather than assess clustering solutions as a whole. This is important because the best clusters may arise from a method that also generates clusters with low validity. Since the aim here was to identify stable hidden subphenotypic patterns across a large set of independent data, this approach was advantageous. Had the aim been to find the best and most generalisable clustering method for analysing the datasets, clustering solution should rather have been assessed as a whole. Another benefit of the chosen approach was that results for measures that are affected by varying numbers of subgroups (e.g., the silhouette metric, Handl et al., 2005) could be comparable across different combination of algorithms and arguments.

Despite the application of multiple algorithms and approaches, the limitation of not assessing further methods should be mentioned. For example, various sets of preselected variables for approach 4 could have been assessed in a systematic way. Also, further approaches for taking into account information from participants with and without tinnitus, such as the functional random forest could have been applied (Feczko et al., 2018). Nevertheless, since novel techniques are developed constantly, assessing all possible methods would be unrealistic. Instead, choosing a subset of the available

methods covering multiple aspects, as done in this study, should suffice for a comprehensive investigation. A final limitation of this work is that only one subphenotype with optimum validation results from each approach was chosen for further characterisation. Exploring the characteristics of more identified subphenotypes could lead to the discovery of more interesting patterns. This step, however, was not automatic, requiring researcher's interpretation, thus making it rather time consuming. Having an aim that would allow automatisation, such as to identify all subphenotypes predicting a specific outcome (e.g., response to a specific treatment), would make specifying all relevant subphenotypes possible.

5.4.4 Conclusions

To conclude, this study contributed to research in tinnitus heterogeneity and subphenotyping. It proposed novel methods for investigating robust subphenotypes, inspired by and expanding on methods used in other medical fields. In addition, it provided novel insights into the characteristics of robust subphenotypic patterns, validating and building on previous findings.

5.5 Chapter contributions

All work presented in this chapter is my own contribution.

Chapter 6. Discussion

6.1 Introduction

This thesis had two main aims. The first was to contribute to the standardisation of tinnitus assessment relevant for tinnitus profiling and subphenotyping. Two objectives were set towards achieving this aim. First, to review the tinnitus subphenotyping literature and summarise variables used for this purpose. Second, to develop a tool for self-reported tinnitus assessment involving researchers from many centres. The second aim was to contribute novel insights into tinnitus heterogeneity. The objective towards this aim was to discover robust tinnitus subphenotypes using many datasets and machine learning techniques. In sections 6.2 and 6.3, I discuss where the tinnitus field is at, advances from other fields, the relevant contributions of this thesis and their importance, and future directions in relation to these topics. In section 6.4, recommendations for tinnitus assessment for future studies based on the findings of this thesis are provided. In addition, the importance of establishing multi-centre collaborations for advancing practices and knowledge of tinnitus is discussed in section 6.5.

6.2 Standardisation of tinnitus assessment

6.2.1 Literature reviews for tinnitus assessment

Considering the increasing scientific knowledge, literature reviews have become an essential component of clinical research and various review types have emerged (Grant and Booth, 2009). Regarding tinnitus research, previous reviews have highlighted the heterogeneity in tinnitus assessment methods. Examples include reviews of questions used to define tinnitus in prevalence studies (McCormack et al., 2016), instruments used or that can be used to assess treatment outcome in tinnitus clinical trials (Hall et al., 2016; Shabbir et al., 2021), objective measures of tinnitus (Jackson et al., 2019), and guidelines for tinnitus assessment (Fuller et al., 2017). Such reviews can help the design of future studies by providing summaries of practices used across the scientific community.

This thesis contributed a review of published literature on tinnitus subphenotyping (Chapter 2). An important outcome of this project was the development of a framework of variable concepts relevant for tinnitus subphenotyping, that can help researchers when designing the assessment protocols for their studies. In addition, variables that

were commonly shown to be important for subphenotyping were highlighted. These included age, sex, hearing ability, problems with sounds, depression symptomatology, mandible problems, various somatic symptoms, overall tinnitus severity, impact of tinnitus on emotion and mental health, tinnitus localisation, and somatic manoeuvres effect on tinnitus.

6.2.2 Standardised self-report tinnitus questionnaires

Self-report questionnaires are powerful tools allowing large-scale standardised data collection, especially if administered online (Evans and Mathur, 2005). For tinnitus research, numerous such questionnaires have been developed for assessing the impact and severity of tinnitus (Kuk et al., 1990; Meikle et al., 2012; Newman et al., 1996). However, there is additional information relevant to tinnitus profiling and subphenotyping that can be collected with self-report such as in relation to demographics, co-existing conditions, lifestyle, and perceptual characterisation of tinnitus. Tinnitus case history questionnaires can be used to capture this information, but few such tools have been published. The TSCHQ is a notable example that has been translated into many languages and used widely in tinnitus research (Langguth et al., 2007). However, the TSCHQ was published 14 years ago (2007) and has some limitations such as not being applicable to the non-tinnitus population. Thus, an up-to-date self-report questionnaire allowing detailed tinnitus-relevant profiling of the adult population was missing.

To fill this gap, this thesis contributed such a tool, namely the ESIT-SQ (Chapter 3). This was the result of a collaborative project among many research centres, building on previous collaboration (Cima et al., 2019; Langguth et al., 2007). The ESIT-SQ can be completed by both people with and without tinnitus. Thus, it can be used for studies needing to collect standardised information from both populations. In Chapter 4, such an application of this tool was described. Specifically, participants from a population-based cohort (the STOP project) were asked to fill in the ESIT-SQ. This enriched the database with valuable phenotypic information. In Chapter 5, I described how I used a subset of this data to initially identify variables differentiating people with tinnitus from people without tinnitus and subsequently cluster the tinnitus population using these tinnitus-discriminating variables.

The online implementation of the ESIT-SQ is a big advantage of this tool, allowing efficient data collection for internet-based studies. In Chapter 4, I described how this questionnaire was the main tool for an online survey, aiming to collect an independent dataset (the ESIT dataset) to be analysed alongside the STOP dataset. This helped identify a replicable tinnitus subphenotype across the two datasets, as described in Chapter 5.

As part of the original project, the ESIT-SQ was translated from English into six more languages: Dutch, German, Italian, Polish, Spanish, and Swedish. Since then, more translations of the questionnaire have been created by various researchers after providing them with supporting documents for the translation process. These included an Albanian, French, and Greek translation. The ESIT-SQ is thus expected to contribute a lot to standardisation in tinnitus research and results from a few studies reporting using it have already been published (Cederroth et al., 2020; Lugo et al., 2020; Trpchevska et al., 2020).

6.2.3 Future directions towards standardisation of tinnitus assessment

Achieving widely endorsed standards for tinnitus assessment would be very beneficial for tinnitus research. However, due to the multiple and diverse aspects of tinnitus, as well as the subjective nature of the condition, standardisation of assessment is rather challenging. Currently there are no clear standards on either what is important to measure (which variables or traits) or how to measure different aspects (which specific methods). Regarding what to measure, this thesis showed that there are many potential aspects that can be assessed. Considering the constraints in research (such as related to budget and time) it would be very beneficial to have a minimum set of information that would be advised as essential for assessment protocols in future studies. This thesis contributed to this by highlighting various variable concepts for their importance in tinnitus subphenotyping (see section 6.4 for a summary and further discussion). As for how to measure different aspects, this topic also deserves more attention from the scientific community. As an example, there are numerous available methods for the psychoacoustic assessment of tinnitus (Fournier et al., 2018; Henry and Meikle, 2000). Similarly, there are numerous questionnaires for assessing the impact of tinnitus (Kuk et al., 1990; Meikle et al., 2012; Newman et al., 1996), and consensus on which to use is lacking. A summary of the advantages and disadvantages of different methods with recommendation for preferred procedures would be very helpful for future research.

6.3 Understanding heterogeneity in clinical conditions

6.3.1 Clinical subphenotyping and subtyping: challenges and advances

Heterogeneity and the need for subphenotyping and subtyping are relevant for many clinical conditions (Feczko et al., 2019; Wen et al., 2020). A common challenge across different conditions is that there is no ideal method to investigate mechanisms generating these. Therefore, researchers often rely on investigating measurable traits of interest. Although there is as yet no single gold standard method to investigate phenotypic heterogeneity, various statistical and computational methods have been developed. As discussed in Chapter 1, unsupervised machine learning techniques can be applied for discovering subgroups with similar characteristics within heterogeneous clinical populations (Marquand et al., 2016; Pina et al., 2020). Such techniques have been applied for the investigation of heterogeneity in multiple conditions such as Alzheimer's disease (Gamberger et al., 2017), asthma (Kisiel et al., 2020), Meniere's disease (Frejo et al., 2017), autism (Lombardo et al., 2016), bipolar disorder (Rabelo-da-Ponte et al., 2020), coronary heart disease (Ohukainen et al., 2020), chronic fatigue syndrome (Słomko et al., 2020), chronic migraine (Woldeamanuel et al., 2020), chronic pain (Obbarius et al., 2020), chronic rhinosinusitis (Tomassen et al., 2016), COVID-19 (Wang et al., 2021), depression (Beijers et al., 2019), diabetes (Ahlqvist et al., 2018), epilepsy (Reyes et al., 2020), Parkinson's disease (Zhang et al., 2019), systemic lupus erythematosus (Lanata et al., 2019), and systemic sclerosis (Sobanski et al., 2019). An important consideration when applying these techniques is validating that any identified subgroups are relevant to the question of interest, since heterogeneity is inherently present in the human population (Feczko et al., 2019). Using variables that are known to be important for the question of interested to validate results can help overcome this issue.

A successful example of applications of unsupervised machine learning techniques is for asthma subphenotyping and endotyping. Driven by the need for personalised management for the heterogeneous manifestations of people with asthma, many studies investigated asthma phenotypic heterogeneity (Fuhlbrigge and Castro, 2020; Honkoop and Chavannes, 2020; Kaur and Chupp, 2019; Schoettler and Strek, 2020). These included multicentre studies such as the U-BIOPRED study involving centres from 11 countries and the Asthma Research Program study involving multiple centres in the United States (Lefaudeux et al., 2017; Moore et al., 2010). Data from 418 and 726

participants, respectively, were analysed in these two studies and various data analysis techniques were applied (such as for cluster stability assessment in Lefaudeux et al., 2017). Reviewing and harmonising findings across studies led to the characterisation of a few primary asthma subphenotypes (Kaur and Chupp, 2019). Age at onset, lung function, and allergic predisposition were highlighted as important variables for defining these subphenotypes (Kaur and Chupp, 2019). Across multiple identified asthma subphenotypes, some were suggested to represent distinct endotypes guiding management strategies (Hamilton and Lehman, 2020; Kaur and Chupp, 2019; Schoettler and Strek, 2020). Asthma subphenotyping based on longitudinal trajectories of features has also been investigated by applying latent transition analysis on repeated measures from 3,320 adults with asthma (Boudier et al., 2013). In addition, researchers have applied unsupervised machine learning on gene expression data to identify potential asthma endotypes (Yan et al., 2015). Despite these advances, researchers agree that novel specific biomarkers are needed in order to fully understand the endotypes driving disease expression and guide precision medicine for asthma (Fuhlbrigge and Castro, 2020; Hamilton and Lehman, 2020; Kaur and Chupp, 2019). The example from asthma research can guide research towards tinnitus subtyping and endotyping by highlighting important elements of a successful process. These elements include the utilisation of large condition-specific datasets (including multicentre and longitudinal data), the application of advanced computational techniques, the review and harmonisation of findings across independent studies, and the assessment of objective biomarkers. Future directions towards tinnitus subtyping and endotyping are discussed further in section 6.3.3.

6.3.2 Tinnitus subphenotyping

The review described in Chapter 2 showed that many studies used evidence-based methods to investigate tinnitus subphenotypes. Most of these studies took hypothesis-driven approaches. In this type of studies, researchers investigated the hypothesis that subgrouping people with tinnitus based on one or a few variables leads to distinct tinnitus subphenotypes. It was shown that many different variables were considered important and chosen for this purpose by researchers. Multiple variables were also shown to differentiate the investigated subgroups across studies. However, these studies only consider a few pre-specified variables to define subgroups. Some studies applying data-driven approaches based on unsupervised machine learning were also identified.

Although these provided useful insights into tinnitus subphenotyping, there were many methodological limitations. For example, validation of results was insufficient and none of the reviewed studies used information from more than one research or clinical centre.

In the study reported in Chapter 5, information from three tinnitus-specific datasets were analysed in order to characterise robust tinnitus subphenotypes. A comprehensive unsupervised machine learning methodology utilising various algorithms and validation techniques was applied for this purpose. Across various approaches, six ways of subgrouping people with tinnitus based on their phenotypic characteristics were described. By generalising these six subgroupings, three distinct and robust phenotypic patterns were identified and described.

The first pattern was discovered by two analyses applied to the STOP dataset: the Audiometric Variables Clustering (approach 1) and the Tinnitus Discriminating Variables Clustering (approach 3). In the first case, only audiometric thresholds were used for discovery of subgroupings (Appendix 5.1). In the second case, a set of variables that were important for discriminating tinnitus from non-tinnitus participants were used. These included mainly audiological and hyperacusis symptomatology variables (Appendix 5.5). In both cases, discovered subgroupings were externally validated and ranked based on differences in tinnitus-specific characteristics. This first pattern was characterised by the association of younger age, better hearing, higher stress and anxiety scores, higher matched tinnitus pitch, lower matched tinnitus loudness, lower tinnitus masking thresholds, more often having varying tinnitus loudness, and tinnitus more often being increased by loud sounds, stress, or alcohol. Similar patterns have been described in previous studies (Andersson and McKenna, 1998; Mores et al., 2019). Subgrouping people with tinnitus based on hearing was once again highlighted in this study, and a specific way of subphenotyping into two subgroups was discovered.

The second pattern was identified by both the STOP General Phenotypic Variables Clustering (approach 2) and the STOP-ESIT Independent Validation Clustering (approach 4) applied to the STOP dataset. In the first case, all general variables (not specific to participants with tinnitus) were used for identification of subgroupings (Appendix 5.3). Subgroupings were then externally validated and ranked based on differences in tinnitus-specific variables. The second approach used a small number of prespecified variables for the identification of subgroupings (Appendix 5.7). External

validation and ranking in this case were based on the replicability of subgroupings in the ESIT dataset. This second pattern was characterised by the association of higher hyperacusis symptomatology, higher stress, anxiety, and depression scores, lower quality of life in the psychological, physical, and environmental subscales, more often having neck and TMJ problems, higher overall tinnitus severity, and more often having preceding tinnitus treatments. Previous studies have described similar patterns to this (Edvall et al., 2019; Kojima et al., 2017; Schecklmann et al., 2014; Tyler et al., 2008). The relevance of hyperacusis, and psychological and somatic co-existing conditions for tinnitus subphenotyping is discussed a lot in tinnitus research (Cianfrone et al., 2015; Jastreboff, 2004; Schecklmann et al., 2014), however, studies do not always assess the status for these important traits.

The third pattern was identified by the STOP-BRC Independent Validation Clustering (approach 4). In this approach, clustering was applied to the STOP dataset using a small set of pre-specified variables (Appendix 5.6). Identified subgroupings were subsequently externally validated and ranked based on their replicability in the BRC dataset. This final pattern was characterised by the association of higher frequency of male sex, worse hearing, higher depression scores, lower quality of life in the social subscale, and having more often constant and daily tinnitus. Some of these associations have been described in previous studies (Burkart et al., 2019; Koops et al., 2019). Sex and temporal manifestation of tinnitus are often discussed for their relevance for tinnitus subphenotyping (Burkart et al., 2019; Koops et al., 2019; Niemann et al., 2020a; Van der Wal et al., 2020), and this study proposed a way of combining them to define two distinct tinnitus subphenotypes.

The clinical and scientific relevance of these subphenotypic patterns serves as a further validation. The first pattern differentiates people with better hearing but worse mental health from those with worse hearing but better mental health. Different management strategies would be required for these subphenotypes. In the first case, psychological interventions such as cognitive behavioural therapy (CBT) would be relevant and should focus on reducing the negative impact of tinnitus and on improving overall mental health (Hesser et al., 2011). In the second case, audiological management would be relevant, potentially focusing on the improvement of hearing with carefully fitted hearing aids (Del Bo and Ambrosetti, 2007). In addition, failing to differentiate tinnitus subphenotypes with distinct degrees of hearing loss might affect results of research

studies investigating tinnitus mechanisms as previously suggested (Vanneste and De Ridder, 2016). As for the second phenotypic pattern, the discovered subphenotype with higher burden in relation to hyperacusis symptomatology, mental health, and somatic problems was also characterised by higher tinnitus severity. This subphenotype would require specialised management approaches. The severity of each of the associated comorbidities should guide the personalisation of management protocols and the involvement of different specialised clinicians (such as audiologists, psychologists, and physiotherapists). For example, depending on the severity of hyperacusis sound therapy protocols should be carefully adjusted (Jastreboff and Hazell, 2004). This subphenotype is also very likely to have distinct underlying mechanisms driven by any of the associated comorbidities, namely hyperacusis, mental health problems, and somatic problems (Koops and van Dijk, 2021; Lau et al., 2018; Shore et al., 2007; Song et al., 2014). The third phenotypic pattern also highlighted clinically important associations of characteristics. The identified association between constant tinnitus, male sex, worse hearing, higher depression, and lower social-related quality of life suggests that specialised management protocols for people showing such combinations of characteristics should be developed. Social support and group therapies might be particularly relevant in these cases (Greenwell et al., 2016; Thompson et al., 2011).

Overall, researchers and clinicians should be aware of these subphenotypic patterns and consider them when designing their study and management protocols, respectively. In research settings, knowledge of important subphenotypes should guide planning of both participant assessment protocols (inclusion of variables that are essential for defining the subphenotypes) and data analysis (e.g. stratifying based on specific subphenotypes). As for clinical management, a tinnitus clinical expert should be responsible for evaluating phenotypic profiles of patients, and for personalising management approaches and coordinating the involvement of other clinical specialists accordingly.

6.3.3 Future directions for tinnitus subphenotyping, subtyping, and endotyping

Future studies could further evaluate the subphenotypic patterns described here or in other tinnitus subphenotyping studies by assessing their reproducibility in independent datasets. In addition, comparing response to treatments, and genetic or neuroimaging profiles between previously identified subphenotypes would provide evidence for their links to distinct underlying mechanisms and their relevance for tinnitus subtyping. This could be done either by using prospectively collected datasets from studies including

the necessary variables in their assessment protocols, or by using retrospective datasets containing the necessary information. To facilitate such efforts, it would be important to provide information on how to predict membership in any proposed subphenotypes from data-driven studies. This could be done by building and making available relevant machine learning classifiers that could be used to assign labels for the identified subphenotypes in external datasets. Such classifiers were developed in the study described in Chapter 5. A useful next step would be to create online tools for these classification models.

Another important future direction towards tinnitus subtyping would be to apply unsupervised machine learning techniques to genetic and neuroimaging data, and combinations of these with other phenotypic information. Such approaches are more likely to identify subgroups linked to distinct tinnitus endotypes. This has only been done once across the studies reviewed in this thesis (Schecklmann et al., 2012), and would be worth investigating further in larger datasets (including hundreds of participants with and without tinnitus) and using various novel methodologies such as semi-supervised machine learning. Given that such data are challenging to collect, collaborations between many research teams and standardisation in tinnitus research can help reach the large data needed.

In addition, for a future data-driven study investigating tinnitus subphenotypes, the collection of data from even larger, carefully recruited, ideally multicentre populations, would be important. This would ensure that subphenotypes with small sizes would be identifiable. It would also allow heterogeneity within predefined subgroups to be investigated. For example, clustering could be applied within subphenotypes based on the presence or absence of comorbidities such as hearing loss, hyperacusis, mental health problems, and somatic problems. Moreover, planning a study to collect repeated measures of the same information (longitudinal data) would allow assessment of the stability of identified subphenotypes or identification of subphenotypes sharing similar feature trajectories (e.g., development of comorbidities). Such approaches have been applied in other fields but not in tinnitus research (Boudier et al., 2013; Dwyer et al., 2020; Lee and Van Der Schaar, 2020; Lim et al., 2020; Stronge et al., 2019).

6.4 Putting it all together: Recommendations for tinnitus assessment

The framework proposed in Chapter 2 is indicative of the multiple traits that could be used to obtain a detailed profile of people with tinnitus. Many of these traits can be assessed with self-reported questionnaires such as the ESIT-SQ, but for others specific procedures would be necessary or more appropriate. For hearing ability an audiological assessment is essential as it provides detailed information of the hearing status. Similarly, for comorbidities that are not yet diagnosed such as TMJ problems, a clinical examination would be needed. To assess noise exposure, an under-investigated variable for tinnitus subphenotyping, various methods can be used, including detailed questionnaires, as review in Guest et al. (2018). Specific questionnaires can also be used to assess various important co-existing conditions such as hyperacusis (Margol-Gromada et al., 2020), stress (Levenstein et al., 1993), depression, and anxiety symptomatology (Beck et al., 1961; Steer and Beck, 1997; Zigmond and Snaith, 1983), as well as personality and copying-style characteristics (Budd and Pugh, 1996; Durai and Searchfield, 2016; Kirsch et al., 1989; Tellegen and Waller, 2008). In addition, for brain function and genotyping advanced neuroimaging and genomic technologies are needed (Elgoyhen et al., 2015; Szczepek et al., 2019). Treatment outcomes can also be used to subphenotype people with tinnitus, but the standardised assessment of tinnitus treatment outcomes is a challenging task (Hall et al., 2019).

Considering that it might be difficult for every study to assess all potentially relevant traits, recommendations for minimum assessment standards would be beneficial for tinnitus research. Collaborative efforts among tinnitus researchers can help achieve consensus on such standards (see section 6.5 for further discussion on the importance of multicentre collaborations). In such an attempt, the TRI consortium proposed different sets of essential (significance level A), highly recommended (level B), and of potential interest (level C) variable concepts for tinnitus assessment (Langguth et al., 2007). However, this was published many years ago. Evidence-based approaches can also help develop assessment standards, as was shown in Chapters 2 (reviewing tinnitus subphenotyping studies) and 5 (describing data-driven tinnitus subphenotyping applications). Results described in these chapters highlighted a few important variable concepts that should be considered for assessment in future tinnitus studies. These are summarised in Table 22. Examples of specific measures for assessing these variable

Table 22. Summary of variable concepts highlighted as important for tinnitus subphenotyping in Chapters 2 and 5. Examples of methods for assessing these are also provided.

Variable concept	Highlighted in Chapter 2	Highlighted in Chapter 5	Examples of methods for assessment
Non-tinnitus-specific			
Age	✓	✓	ESIT-SQ question A1
Sex	✓	✓	ESIT-SQ question A2
Hearing ability	✓	✓	Audiological assessment (e.g., pure tone audiometry), ESIT-SQ question A13
Problems with sounds	✓	✓	Audiological assessment (uncomfortable loudness levels), HQ, ESIT-SQ question A12
Stress-related symptoms	-	✓	Clinical examination, PSQ (Levenstein et al., 1993), ESIT-SQ question A16
Symptoms of anxiety	-	✓	Clinical examination, HADS-A, BDI, ESIT-SQ question A16
Symptoms of depression	✓	✓	Clinical examination, HADS-D, BAI, ESIT-SQ question A16
Neck problems (neck pain)	-	✓	Clinical examination, ESIT-SQ question A15
Mandible problems (TMJ pain)	✓	✓	Clinical examination, ESIT-SQ question A15 and A16
Various somatic symptoms	✓	-	Clinical examination, MSPQ, somatization subscale of the SCL-90-R (Delogatis, 1977), somatoform symptoms subscale of the ICD-10 Symptom Rating (Tritt et al., 2008)
Circulatory system problems (high BP)	-	✓	Clinical examination, ESIT-SQ question A15 and A16
Quality of life	-	✓	WHOQoL-BREF
Tinnitus-specific			
Objectivity (presence or not of an acoustic source generating the perception)*			Clinical examination, ESIT-SQ question B17
Overall severity	✓		TFI, THI, THQ, TQ, TRQ
Impact of tinnitus on emotion and mental health	✓	-	Emotional, psychological, and cognitive distress

Variable concept	Highlighted in Chapter 2	Highlighted in Chapter 5	Examples of methods for assessment
			score from TQ, emotional subscale from THI, TSCHQ question 17, ESIT-SQ question B4
Preceding tinnitus treatments (any management)	-	✓	ESIT-SQ question B20
Presence pattern (constant or intermittent, daily or not daily)	-	✓	ESIT-SQ question B1 and B2
Localisation	-	-	ESIT-SQ question B15
Pitch	-	✓	Audiological assessment (pitch matching)
Pitch and hearing loss profile (hearing threshold at pitch matched frequency)	-	✓	Audiological assessment
Loudness	-	✓	Audiological assessment (loudness matching)
External sound effect (increased by sound, masking threshold)	-	✓	Audiological assessment (minimum masking level), ESIT-SQ question B18 and B19
Somatic manoeuvres effect	✓		ESIT-SQ question B18 and B19

BAI: Beck Anxiety Inventory (Steer and Beck, 1997); BDI: Beck Depression Inventory (Beck et al., 1988); HQ: HADS-A: Hospital Anxiety Depression Scale for Anxiety (Zigmond and Snaith, 1983); HADS-D: Hospital Anxiety Depression Scale for Depression (Zigmond and Snaith, 1983); Hyperacusis Questionnaire (Khalfa et al., 2002); MSPQ: Modified Somatic Perception Questionnaire (Main, 1983); PSQ: Perceived Stress Questionnaire (Levenstein et al., 1993); TFI: Tinnitus Functional Index (Meikle et al., 2012); THI: Tinnitus Handicap Inventory (Newman et al., 1996); THQ: Tinnitus Handicap Questionnaire (Kuk et al., 1990); TQ: Tinnitus Questionnaire (Hallam et al., 1988); TRQ: Tinnitus Reaction Questionnaire (Wilson et al., 1991); TSCHQ: Tinnitus Sample Case History Questionnaire (Langguth et al., 2007); WHOQoL-BREF: World Health Organization's Quality of Life (Whoqol Group, 1998).

concepts are also provided. From Chapter 2, tinnitus objectivity highlighted from theoretical studies and 11 variable concepts highlighted from evidence-based studies (as summarised in Table 5) were included. From Chapter 5, variable concepts characterising the three main phenotypic patterns and that were confirmed from the multivariable analyses (LASSO regressions) were included (Table 21). As an exception, sex, rather than height and weight, was included. This is because the three variables could not be assessed simultaneously in a multivariable model (due to quasi-complete separation caused by sex). It remains to be proven in future studies whether weight and height carry any additional useful information not captured by sex. Age,

sex, hearing ability, problems with sounds, symptoms of depression, and mandible problems were highlighted in both the review and the data-driven study described in this thesis. The importance of each of these variable concepts has been widely discussed in tinnitus literature and these should be considered a minimum standard for assessment in every future tinnitus study.

6.5 The importance of international interdisciplinary collaborations

Tinnitus has proven to be a rather complex condition and its underlying mechanisms remain elusive. To accelerate findings coordinated efforts from tinnitus researchers across different centres would be very beneficial. Many recent multi-centre European Union funded projects have contributed a lot to the establishment of international interdisciplinary collaborations (ESIT, 2017; TIN-ACT, 2017; TINNET, 2014; UNITI, 2020). The ESIT project in particular made possible the collaborative development of the ESIT-SQ (Chapter 3) as well as the affiliation with the Karolinska Institute allowing analysing the STOP data (Chapter 4). Such collaborative efforts should be reinforced, including as many interested researchers as possible. An ideal aim would be to establish international teams of tinnitus researchers and clinicians from multiple centres around the world that would regularly provide tinnitus assessment protocols and renew these considering novel research findings. This would help increase overall quality of tinnitus research as it would provide a basis for study designs. Moreover, it would increase comparability across different studies potentially accelerating important discoveries.

6.6 Concluding remarks

Tinnitus subphenotyping and subtyping can significantly advance tinnitus research focusing on understanding tinnitus mechanisms and developing personalised tinnitus treatments. However, the task is rather complex, as indicated by the numerous relevant published studies. This thesis described a systematic approach to synthesise evidence and provided new methods and knowledge for tinnitus subphenotyping. Specifically, it contributed to standardisation of tinnitus assessment relevant for tinnitus subphenotyping by providing a framework of variables used in previous studies and a standardised questionnaire for self-report tinnitus-relevant assessment available in many languages. In addition, it contributed to our understanding of tinnitus heterogeneity by providing novel insights into the characteristics of robust tinnitus subphenotypic patterns. Future studies can use the methods and findings described in

this thesis to further promote tinnitus subphenotyping, towards tinnitus subtyping and endotyping.

Appendices

Appendix 2.1. Review of theoretical frameworks for tinnitus profiling or subgrouping

Reference	Details

Goodhill (1950)	Three binary categories to classify tinnitus.
	• Noise localisation
	  o Aural noises (tinnitus aurium).
	  o Head noises (tinnitus cerebri).
	• Objectivity of perceived sound
	  o Subjective (also called true, static, non-vibratory, or intrinsic) Further defined by:
	  ▪ Aetiologic diagnosis: anatomic location (external auditory canal, tympanum and middle ear, cochlea, VIII nerve, intracerebral) and pathologic physiology (e.g., anaemia, oedema, inflammation)
	  ▪ Psychosomatic diagnosis: sensitivity level and specific phobias (anxiety, conversion)
	  o Objective (also called pseudo, dynamic, vibratory, or extrinsic) Subdivided based on its origin into:
	  ▪ Vascular.
	  ▪ Muscular.
	• Ability to cope with tinnitus
	  o Compensated.
	  o Uncompensated.

Nodar (1978)	Tinnitus profile – standard information to characterise tinnitus patients.
	• Minimum of six pieces of information regarding tinnitus description
	  • Quality (ring, hiss, cricket, buzz, whine, roar, click, siren, hum).
	  • Presence (always, occasional).
	  • Pattern (continuous, pulsing).
	  • Composition (single sound, multiple sounds).
	  • Level (loud, soft).
	  • Annoyance (very annoying or not annoying).
	• Additional history information
	  • Age.
	  • Sex.
	  • Ear affected.
	  • Onset.
	  • Modulation of loudness, composition and annoyance by physical activity, emotional stress, or exposure to loud sound (9 yes or no questions).
	• Additional audiometric information
	  o Audiometric configuration (flat, rising, falling, 4K notch, saucer).
	  o Probable site of lesion (outer ear, middle ear, cochlea, neural, brain stem, central).
	  o Degree of hearing loss (normal, slight, mild, marked, severe, extreme).
	  o General tinnitus match (pure tone or a narrow band of noise based on audiometric matching), frequency match (Hz), level match (dB).

Evered and Lawrenson (1981)	Classification of tinnitus can be addressed in many ways. • Subjective tinnitus (or tinnitus) and objective tinnitus (or somatosounds, pseudo-tinnitus, vibratory tinnitus). • Classification of subjective tinnitus by: o Characteristics described by the patient

        - Number of different sounds.
        - Loudness: faint, moderately loud, very loud.
        - Pitch: low, medium, high, not identifiable, steady, warbling.
        - Temporal characteristics: long term (continuous, intermittent), short term (steady, fluctuating).
        - Localisation: left ear, right ear, bilateral, in the head.
        - Annoyance: none, mild, moderate, severe, or very severe.
        - Effect of environmental noise: tinnitus much reduced, slightly reduced, unchanged, worsened.
    - Characteristics measured with audiometry
        - Quality match: match to pure tones, narrow band of noise, broad band of noise.
        - Pitch match.
        - Loudness and relative loudness match: at 1kHz and at the pitch matched frequency, respectively.
        - Test for octave confusion.
        - Masking tests.
        - Tests of residual inhibition.
    - Probable site of lesion
        - Peripheral (external ear, middle ear, cochlea, eighth cranial nerve), central (subdivisions of the central auditory pathway), extra-auditory (vascular, nasopharyngeal, muscular, unknown).
    - Probable diagnosis: usually unknown.

Dauman and Tyler (1992)

Some considerations on the classification of tinnitus.

- Pathological or not
    - Normal (or non-pathological): lasts less than five minutes less than once a week.
    - Pathological (or abnormal): lasts for more than five minutes and is perceived more than once a week, if not continuously.
- Impact on the patient
    - Acceptable (or non-clinical tinnitus): does not bother the subject.
    - Unacceptable: disturbing to the patient
- Temporary or permanent tinnitus.
    - Short-term (or temporary): there is a return to the absence of tinnitus.
    - Permanent tinnitus: does not return, further classified into:
        - Constant.
        - Intermittent.
- Functional classification of tinnitus (according to the site of dysfunction)
    - Middle ear
        - Producing sound
            - Muscular.
            - Vascular.
        - Attenuating external sound
            - Peripheral.
            - Central.
    - Sensorineural
        - Peripheral.
        - Central.
- Aetiologic classification based on conditions causing sensorineural hearing loss

Includes: Noise-induced tinnitus, Meniere's tinnitus, ototoxic tinnitus, presbycusis tinnitus.

- Consider presence of both physiological and psychological factors increasing susceptibility or triggering tinnitus and their interactions.

Shulman (1992) — Clinical otologic and neurotologic classification.

- Otologic classification
  Based on the history and physical examination and the identification of a medical condition of the ear: classification includes diseases of the external and/or middle ear, cerumen impaction, abnormal mobility of the tympanic membrane or ossicular chain, and abnormal contractions of the muscles of the middle ear.
- Neurotologic classification
  Based on the responses of the cochleovestibular system.
  This classification allows the classification of tinnitus into one or more of the following clinical types:
  - Auditory.
  - Nonauditory.
  - Subclinical.
  - Middle ear.
  - Cochlear.
  - Neural.
  - Presbycusis type I and II.
  - Vestibular.
  - Cervical.
  - Central.
  - Contralateral.

Briner (1995) — Seven categories based on temporal patterns, perceptual parameters, behavioural responses, and severity.

- Tinnitus Type 1. Chronic subclinical (not a disorder)
  - Nearly always present in a sound-treated environment.
  - No psychological or physical distress and masked by ambient sound.
- Tinnitus Type 2. Acute physiological (not a disorder, experienced by nearly all people at some time in their lives)
  - Short term, lasting minutes to a few hours.
  - Generally tonal in nature.
  - Not accompanied by other auditory changes and not preceded by exposure to noise, ototoxic drugs, or other insults to the auditory structures.
  - No psychological or behavioural consequences other than occasional concern for the condition.
- Tinnitus Type 3. Acute clinical
  - After exposure to noise or administration of ototoxic drugs.
  - Usually accompanied by a temporary threshold shift.
  - Lasting a few hours to several days.
  - Typically not tonal in nature but rather a complex sound such as crickets or steam.
  - Can be masked by other sounds.
  - No psychological consequences other than concern for permanent damage to hearing.
  - Lasting less than six weeks.
- Tinnitus Type 4. Chronic subclinical

    - Present continuously or absent only as a part of a normal fluctuating pattern but perceived only when patient's attention is directed to it.
    - No psychological or behavioural consequences other than concern for permanent damage to hearing.
    - Present for at least six weeks.
    - Not an auditory hallucination associated with psychosis.
- Tinnitus Type 5. Chronic clinical
    - Continually perceived regardless of the amount of attention which is directed to or away from it.
    - Loudness, frequency, and annoyance level may be changed using techniques of distraction.
    - No detectable behavioural disturbance, no interference with sleep, concentration, or social activities, patient does not seek medical attention.
    - Present for at least six weeks.
    - Not an auditory hallucination associated with psychosis.
- Tinnitus Type 6. Phantom auditory pain, adaptive
    - Severe continuous tinnitus, constantly perceived, being often the chief otologic complaint.
    - Patient uses only one or two descriptors of the noxious properties of the tinnitus (e.g., shrill, deafening).
    - Adaptive behavioural changes allowing the patient to have a basically normal life and social relationships. Acceptable behaviour changes such as use of maskers or hearing aids, medically prescribed or recommended drugs, changes to a less stressful job, biofeedback therapy, and similar strategies.
    - No deterioration in the ability to conduct social relations and no obsession with tinnitus.
    - Complains about the nature of the sound or its interference with external activities (interference with sleep, concentration, listening to music or TV, etc.), but not both.
    - Able to gain pleasure from activities usually considered to be so, i.e., sex, food, and social interaction.
    - Tinnitus present for at least six months.
    - Not an auditory hallucination associated with psychosis.
- Tinnitus Type 7. Phantom auditory pain, maladaptive
    - Severe, continuous tinnitus resulting in deterioration of the patient's lifestyle or social relationships despite attempts to control the tinnitus through external (masking devices, drugs, and lifestyle changes) or internal (counselling, biofeedback, and talking it out) means.
    - Obsession with tinnitus and some type of psychological abnormality when an objective assessment tool is used.
    - Tinnitus is the chief otologic complaint.
    - Patients use three or more descriptors to describe the noxious properties of the tinnitus (e.g., piercing, thunderous, exhausting) and complain about general qualities of tinnitus (such as persistence, loudness, and annoyance) and interference with external activities, e.g., listening to music or TV, sleep, or concentration. Anhedonia is diagnostic.
    - Present for at least 6 months to avoid classifying exacerbations of Meniere's disease in this subgroup.
    - Not an auditory hallucination associated with psychosis.

Nodar (1996)

Tinno-o-gramm: standard information to characterise tinnitus patients.
Assess ABC

- A: aurium in one ear, B: binaural in both ears, C: cerebri for centred in the head
  Assess C-CLAP.
- C for cause: Stress factors that may cause or aggravate tinnitus (emotional stress, physical stress, acoustic stress, chemical stress, pathologic stress), known cause (e.g., Meniere's), presumptive cause (e.g., emotional stress).
- C for composition: ring, buzz, hiss, roar, crickets, and single or multiple sounds Plus: with questions, the clinician can determine whether the tinnitus is constant or intermittent, whether it is composed of a single sound or multiple sounds, whether it represents changes in composition, and whether the changes coincide with any events, medications, or physical activity.
- L for loudness: subjective scale or loudness matching on the audiometer.
- A for annoyance: subjective scale.
- P for pitch: subjective descriptive term (high or low) and audiometric pitch match.

Zenner (1998)

Classification of tinnitus generator mechanisms.

- Initially classify into:
  - Objective tinnitus.
    Examples: Glomus tumour, angiostenosis, protruding bulbus of jugular vein.
  - Subjective tinnitus.
- Further classify subjective tinnitus into:
  - Conductive tinnitus: causative mechanisms affecting transferring sound signals to the inner ear.
    Examples: Disturbance of tubal ventilation, middle ear myoclonia.
  - Sensorineural tinnitus: causative mechanisms affecting sensorineural component of hearing process.
    Further subdivided into:
    - Type I (motor tinnitus): related to the amplification of the signal by the motor of the outer hair cells.
      Examples: Hypermotility, edge-effect tinnitus, efferent tinnitus caused by regulatory disturbances of the nerves, noise trauma, ion channel disorders of the outer hair cell.
    - Type II (transduction tinnitus): related to the mechanoelectrical transduction of the signal by the inner hair cells.
      Examples: Continuous depolarization of ion channel disorders of the inner hair cells, disturbance of the stereocilia of the inner hair cells.
    - Type III (*cochleosynaptic tinnitus*): related to the synaptic transfer of the signal from the inner hair cells to the afferent nerve fibres.
      Examples: Release of transmitters, flooding with synaptic transmitters, swelling of the afferent nerve fibres, excitotoxic tinnitus.
    - Type IV (extrasensory tinnitus): related to extrasensory elements (such as the stria vascularis) involved in the sensory processing of sound.
      Examples: Disorders (e.g., of the ion channels) of the stria vascularis, circulatory disorders of the cochlea, resorption

disorders and osmolarity change of endolymph, endolymph hydrops.
    - Central tinnitus
        - Primary central tinnitus: pathogenesis in the central nervous system independent of the middle and inner ear. Examples: Brain tumours, multiple sclerosis.
        - Secondary central (centralized) tinnitus: triggered peripherally but then manifesting itself in the brain independently of the original source in the ear. Examples: Phantom tinnitus.
- Also classify based on:
    - Localisation.
    - Coping: compensated or decompensated.
    - Time course: acute, subacute, or chronic.

Jastreboff and Hazell (2004)

Categories of patients with tinnitus and/or hyperacusis for Tinnitus Retraining Therapy.

Based on the evaluation of severity of tinnitus, presence of hyperacusis, subjective significance of hearing loss and prolonged exacerbation of symptoms following sound exposure.

Five categories aiming to help design tinnitus retraining therapy:

- Category 0: Low severity of tinnitus severity with or without hearing loss.
- Category 1: High tinnitus severity without other hearing-related problems.
- Category 2: High tinnitus severity with hearing loss.
- Category 3: Hyperacusis, with or without tinnitus or hearing loss.
- Category 4: Prolonged exacerbation of tinnitus or hyperacusis following sound exposure with or without hearing loss.

Langguth et al. (2011a)

Tinnitus Research Initiative algorithm for the diagnostic and therapeutic management of tinnitus.

- Classify into:
    - Pulsatile: Venous or arterial.
    - Non-pulsatile
        - Acute with sudden hearing loss.
        - Paroxysmal.
        - Constant.

          For constant tinnitus identify presence of comorbidities (imply specific underlying aetiology):
            - Hearing loss (conductive, sensory neural).
            - Vertigo.
            - Headache.
            - Psychiatric.
            - Somatic (TMJ, neck-related).
            - Trauma.

Tunkel et al. (2014)

American Academy of Otolaryngology–Head and Neck Surgery: Clinical practice guidelines for tinnitus.

- Differentiate objective tinnitus.
- Classify subjective tinnitus based on:
- Aetiology
    - Primary (idiopathic, may or may not be associated with sensorineural hearing loss).
    - Secondary: associated with a specific underlying cause (other than sensorineural hearing loss) or an identifiable organic condition).
- Duration
    - Recent onset: less than 6 months.

- Persistent: 6 months or longer.
- Impact
  - Bothersome: distressed patient, affected quality of life, and/or functional health status, patient seeking therapy and management strategies to alleviate tinnitus.
  - Non-bothersome: does not have a significant effect on a patient's quality of life but may result in curiosity of the cause or concern about the natural history and how it might progress or change.
- Search/ask for:
  - Underlying conditions that may cause tinnitus.
  - Signs & symptoms of serious disease associated with tinnitus.
  - Presence of severe mood disturbance.
  - Unilateral tinnitus.
  - Pulsatile tinnitus.
  - Hearing difficulties.

Cianfrone et al. (2015)

Tinnitus Holistic Simplified Classification.

- Objective.
- Subjective
  - Auditory.
  - Somatosensory.
  - Psychopathology-related.
  - Combined (two or more of the above).
  - Other (hypertensive diseases not related to cochlear damage, tinnitogenic non-ototoxic effect of drugs, linked to epilepsy, endocrine, immunological disorders without auditory damage).

Levine and Oron (2015)

Perceptual characteristics and tinnitus aetiology.

Extensive classification based on tinnitus perceptual characteristics (quality, location, variability, pitch, modulations) and aetiologic factors.

(Henry, 2016)

Classify in many ways:

- Differentiate normal transient ear noise (should not be considered tinnitus).
- 2 fundamental types
  - Neurophysiologic (or sensorineural, or neural, or primary): generated within the auditory pathways and consists entirely of neural activity or just neural tinnitus.
  - Somatosounds (or secondary): generated as an acoustic signal somewhere in the head or neck (origin can be muscular, respiratory, skeletal, or vascular structures)
- Classification based on duration
  - Chronic (or permanent or persistent): present for at least 6 months.
  - Acute (or recent onset): less than 6 months duration.
- Classification based on temporal patterns of presence
  - Constant: always present even if it is not noticed or is masked by environmental sound.
  - Temporary: can be associated with noise exposure or certain medications, that usually lasts for a few days or up to a week before it resolves.
  - Intermittent: switching between being ‘‘on’’ (present) and ‘‘off’’ (absent).
- Classification based on impact of tinnitus
  - Not bothersome.

- Bothersome.
- Mild.
- Moderate.
- Severe.
- Debilitating.

Also, evaluate negative impact of tinnitus with the Tinnitus Functional Index

- Differentiate based on presence of other hearing related problems using Tinnitus and Hearing Survey (THS): Three subscales to evaluate: A. Tinnitus, B. Hearing, and C. Sound Tolerance.
- Also consider:
  - Visual Numeric Scale for loudness rating.
  - Audiological measurement of tinnitus perceptual characteristics.

6.7 Appendix 2.2. Overview of studies using hypothesis-driven approaches

Reference	Sample size	Sample characteristics	Subgroup definitions	Statistical methods	Significant level / Importance definition *
Kirsch et al. (1989)	77	Tinnitus patients assessed by an otolaryngologist	2 subgroups based on ability to cope: 1. Low copers, 2. High copers. Details: Visual analog scale asking how well participants felt they coped with the tinnitus (0 = not able to cope at all, 100 = cope extremely well) with 'High copers' being those with a rating of 60 or more and 'Low copers' those with a rating of 50 or less	chi-square test	$p \leq 0.05$
Erlandsson et al. (1991)	72	Tinnitus patients from an audiology department	2 types of classification. A. 2 subgroups based on presence of somatic comorbidities: 1. With somatic comorbidities, 2. Without somatic comorbidities B. 2 subgroups based on sex: 1. Female, 2. Male. Details: For the first categorisation, presence of self-reported symptoms of craniomandibular disorders and/or headaches was assessed.	ANOVA, Tukey's test, Mann-Whitney test	$p<0.05$
Attias et al. (1995)	100	Chronic tinnitus patients with noise-induced hearing loss (army personnel)	2 subgroups based on help seeking: 1. Help-seeking, 2. Non-help-seeking. Details: Non-help-seeking were patients that visited the clinic for routine audiological tests but were not interested in treatment.	Wilcoxon test, ANOVA and Scheffe post-hoc test	$p=0.07$ considered borderline
Andersson et al. (1999)	207	Tinnitus patients from an audiology department	2 subgroups based on tinnitus severity: 1. Grade II, 2. Grade III. Details: Based on classification system by Klockhoff and Lindblom: Grade I when tinnitus audible only in silent environments, Grade II when tinnitus audible only in ordinary acoustic environments but masked by loud environmental	Discriminant function analysis, ANCOVA	$p \leq 0.05$

			sounds, can disturb going to sleep but not sleep in general, Grade III when tinnitus audible in all acoustic environments, can disturb sleep in general and is a dominating problem affecting quality of life.		
Holgers et al. (2005)	127	Tinnitus patients from an audiology department	2 subgroups based on tinnitus severity: 1. Severe tinnitus suffering (STS), 2. Non-STS. Details: STS if during an 18-month period either absence from work more than one consecutive month, or more than three visits to a therapist or audiological physician.	Pitman and Fisher's permutation test, logistic stepwise forward regression	$p<0.05$
Hiller and Goebel (2007)	1999	People with tinnitus from a membership database of a tinnitus charity organisation	2 types of classification. A. 2 subgroups based on tinnitus loudness and severity: 1. Low tinnitus loudness and severity, 2. High tinnitus loudness and severity. B. 2 subgroups based on tinnitus severity: 1. Low severity, 2. High severity. Details: Loudness was assessed according to Klockhoff and Lindblom (1967) with the modifications from followed Scott et al. (1990): three grades of tinnitus loudness defined as: grade I (low loudness): tinnitus audible only in silent environments; grade II (excluded from comparisons): tinnitus audible in ordinary acoustic environments, but masked by loud environmental sounds; grade III (high loudness): tinnitus audible in all acoustic environments, i.e., louder than all external sounds. Tinnitus severity was assessed using the Mini-TQ (Hiller and Goebel, 2004): four categories: subgroup I (low severity): no to mild distress; subgroup II (low severity): moderate distress; subgroup III (high severity): severe distress; subgroup IV (high	t-test, chi-square test, odds ratios	$p<0.05$

			severity): very severe distress. For the second classification, only people with high loudness were included.		
Bartels et al. (2008)	265	Chronic tinnitus patients from an Otorhinolaryngology department	4 subgroups based on presence of anxiety and/or depression: 1. Type D, 2. Not Type D. Details: Assessed using a14-item Type D Scale (DS14) that consists of two subscales: negative affectivity and social inhibition. A cut-off score ≥10 on both subscales was used for definition of Type D personality	Chi-square test, Fisher's exact test, t-test, analysis of variance. Cohen's effect size, multivariable logistic regression	Bonferroni corrected $p<0.004$ (0.05/12) or $p<0.01$
Bartels et al. (2010)	265	Chronic tinnitus patients from an Otorhinolaryngology department	2 subgroups based on personality: 1. Both anxiety and depression, 2. Only anxiety, 3. Only depression, 4. Neither anxiety nor depression. Details: Definition of anxiety and depression based on HADS score with a cut-off at 8 for both subscales	t-test, chi-square test, Fisher's exact test, difference of proportions test, Cohen's d effect size	$p<0.05$ and at least small effect sizes (Cohen's $d\geq0.2$) for clinically relevant difference
Sztuka et al. (2010)	44	Tinnitus patients with normal hearing from an Otorhinolaryngology department	2 types of classification based on problems with sounds. A. 2 subgroups based on presence of hyperacusis: 1. With hyperacusis, 2. Without hyperacusis. B. 2 subgroups based on presence of hyperacusis or misophonia: 1. With hyperacusis, 2. With misophonia. Details: Hyperacusis subgroup included patients with a discomfort levels lower than 85 dB SPL for all measured frequencies. Misophonia subgroup included patients with discomfort levels lower than 85 dB only for some frequencies, for which they felt fear. The subgroup without hyperacusis	t-test, Mann–Whitney U test	$p<0.05$

			included remaining participants that did not fit in either category.		
De Ridder et al. (2011)	55	Patients from a tinnitus clinic	2 subgroups based on tinnitus severity: 1. Low distress, 2. High distress. Details: Severity subgroups defined based on the TQ score; assigned to low distress if 0–46, or to high distress if 47–84.	t-test, chi-square test	$p<0.05$
Lindblad et al. (2011)	46	Patients with tinnitus and normal hearing (or minor hearing loss) from an audiology clinic	4 subgroups based on the most likely clinical aetiology of tinnitus: 1. Acoustic trauma exposure, 2. Prolonged music exposure, 3. Hereditary tinnitus, 4. Non-auditory problems. Details: Impact of each causative factors was scored 0-5 based on its likely influence. If more than one factor had score >0, patient was assigned to the subgroup with the highest score. Also binary comparisons for each subgroup against all other.	ANOVA, MANOVA, Tukey's honestly significant difference	$p<0.05$
Vielsmeier et al. (2011)	91	Tinnitus patients from a dentistry and a tinnitus outpatient clinic	2 subgroups based on presence of temporomandibular joint complaints: 1. Yes, 2. No. Details: The first subgroup comprised patients from a dentistry outpatient clinic after confirmation of presence of both temporomandibular joint dysfunctions and tinnitus complaints. The second subgroup comprised patients from a tinnitus outpatient clinic without any subjective complaints of TMJ dysfunction. Comparisons between the second subgroup and a subpopulation from the first subgroup (those with primarily TMJ complaints, not having visited the tinnitus clinic before) were also presented. In addition comparisons between the subgroup with primarily TMJ complaints and the subgroup with TMJ complaints but primarily tinnitus complaints were presented.	t-test, chi-square test	$p<0.05$
Kreuzer et al. (2012)	1604	Tinnitus patients from worldwide tinnitus clinics before their first	5 subgroups based on presence of onset related events: 1. No onset-related event, 2. Loud blast of sound, 3. Whiplash, 4. Head trauma, 5. Other onset-related event.	ANOVA, chi-square test	Bonferroni corrected level at 0.0016 (0.05/31)

		consultation or at the baseline visit before a clinical intervention	Details: Classification based on the response to the TSCHQ question 7.		
Schecklmann et al. (2012)	286	Tinnitus patients from a specialised tinnitus clinic	3 types of classification. A. 2 subgroups based on tinnitus laterality: 1. Unilateral tinnitus, 2. Bilateral tinnitus B. 2 subgroups based on tinnitus quality: 1. Pure tone tinnitus, 2. Narrow-band tinnitus. Details: Defined during pitch-matching. C. 2 subgroups based on slope steepness of the audiogram: 1. Low steepness, 2. High steepness. Details: Based on the frequencies defining audiometric edge. Categorisation based on the subgroup median.	t-test	Probably level set at 0.05 (p-value of 0.032 referred to as near threshold, p-value of 0.056 considered as significant)
Sand et al. (2012)	240	Chronic tinnitus patients from an outpatient department	2 subgroups based on genotype: 1. Homozygous for the major allele, 2. Heterozygous or homozygous for the minor allele. Details: 5 SNPs were examined: rs2049046, rs6265, rs1110149, rs884344, rs3812047	t-test, Fisher's exact test	$p<0.05$, Bonferroni correction for multiple comparisons
Vielsmeier et al. (2012)	1204	Tinnitus patients from worldwide tinnitus clinics before their first consultation or at the baseline visit before a clinical intervention	2 subgroups based on presence of temporomandibular joint disorder: 1. Yes, 2. No. Details: Classification based on the response to the TSCHQ question 32.	t-test, chi-square test	Bonferroni corrected level at 0.0022 (0.05/23)

Heijneman et al. (2013)	80	Chronic tinnitus patients from a specialised tinnitus clinic	3 subgroups based on audiologically assessed tinnitus spectrum: 1. Monotonously increasing tinnitus likeness as a function of the test frequency, 2. Monotonously increasing tinnitus spectrum up to a certain frequency, beyond which there is a drop of at least 20 points in likeliness, 3. All remaining (typically with a spectrum decreasing with increasing frequency). Details: A 20-point downward drop between the peak of the tinnitus spectrum and the likeliness at 8 kHz was defined as a relevant decrease in likeliness. For any smaller decrease participants were allocated to subgroup 1.	Mann-Whitney U-test, t-test	$p<0.05$
Song et al. (2013)	59	Participants with narrow-band noise bilateral tinnitus recruited from a participant database	2 subgroups based on age at tinnitus onset: 1. Early onset tinnitus, 2. Late onset tinnitus Details: Mean age at tinnitus onset for late and early onset tinnitus 60.4 (± 6.9) and 29.7 (± 8.7) years, respectively.	Statistical non-parametric mapping using sLORETA-built-in voxelwise randomization tests (5000 permutations), t-tests	$p<0.05$ (corrected for multiple comparisons using sLORETA-built-in voxelwise randomization tests)
Schecklmann et al. (2014)	1713	Tinnitus patients from worldwide tinnitus clinics before their first consultation or at the baseline visit before a clinical intervention	2 subgroups based on presence of hyperacusis: 1. Yes, 2. No. Details: Classification based on the response to TSCHQ question 29.	t-test, chi-square test, Cohen's d effect size	$p<0.001$ and at least small effect sizes (Cohen's $d\geq0.2$)

Carpenter-Thompson et al. (2015)	32	Research participants with various levels of tinnitus distress	2 subgroups based on tinnitus severity: 1. Low distress, 2. High distress Details: THI score ranged from 0-18 for low distress, were > 20 for the high distress.	ANOVA, independent sample t-test	Various significance thresholds depending on the analysis (ranging from 0.05 to 0.001)
Michiels et al. (2015)	87	Chronic subjective tinnitus patients from a tinnitus clinic	2 subgroups based on the diagnosis of Cervicogenic Somatic Tinnitus (CST): 1. With CST, 2. Without CST. Details: CST if tinnitus is related to the somatosensory system of the cervical spine. The diagnosis was made when the predominant feature was the temporal coincidence of onset or increase of both neck pain and tinnitus	Fisher's exact test, t-test	$p<0.05$
Song et al. (2015)	57	Participants with bilateral pure tone tinnitus recruited from a participant database	2 types of classification. A. 2 subgroups based on age at tinnitus onset: 1. Early onset tinnitus, 2. Late onset tinnitus B. 2 subgroups based on tinnitus severity: 1. Low distress, 2. High distress. Details: Mean age at tinnitus onset for late and early onset tinnitus 52.3 (± 4.3) and 29.0 (± 10.1) years, respectively. Low distress had TQ score 0-46 and high distress 47-84.	Statistical non-parametric mapping using sLORETA-built-in voxelwise randomization tests (5000 permutations), t-test, ANOVA, discriminant analysis	$p<0.05$ (corrected for multiple comparisons using sLORETA-built-in voxelwise randomization tests)
Vielsmeier et al. (2015)	75	Chronic subjective tinnitus patients from a tertiary tinnitus clinic	2 subgroups based on high frequency hearing function: 1. Normal, 2. Hearing loss. Details: Normal was defined as ≤15 dB hearing level over all frequencies from 10 to 16 kHz. A cut-off at 20 and 25 dB was also examined.	t-test, chi-square test, Fishers exact test	$p<0.05$

		with normal conventional PTA thresholds			
Ward et al. (2015)	608	Tinnitus participants from various research studies from a hearing research unit	2 subgroups based on modulation of tinnitus with somatic manoeuvres: 1. Putative somatic tinnitus, 2. No putative somatic tinnitus. Details: Putative somatic tinnitus if they responded 'Yes' to the question 'Does any head and neck movement (e.g., moving the jaw forward or clenching the teeth), or having your arms/hands or head touched, affect your tinnitus?'	chi-square test, logistic regression	$p<0.05$
Al-Swiahb and Park (2016)	470	Tinnitus patients from a tinnitus clinic	3 subgroups based on age: 1. <40 years old, 2. 40-60 years old, 3. >60 years old	t-test, ANOVA	$p<0.05$
Jeon et al. (2016)	85	Persistent pulsatile tinnitus patients from an outpatient clinic	2 subgroups based on presence of low frequency hearing loss (LFHL) ipsilateral to tinnitus: 1. LFHL, 2. No LFHL. Details: ipsilateral LFHL if hearing thresholds greater than 10 dB HL at both 250 and 500 Hz or greater than 20 dB HL at either 250 or 500 Hz compared to the contralateral side.	Fisher's exact test	$p<0.05$
Kim et al. (2016)	121	Tinnitus patients from an outpatient clinic	3 subgroups based on audiogram configuration: 1. Flat, 2. High frequency gently sloping (HFGS), 3. High frequency steeply sloping (HFSS). Details: Originally classified into 1. Flat, 2. HFGS, 3. HFSS, 4. Low frequency ascending (LFA), 5. Mid-frequency U-curve (MFU), and 6. Mid-frequency reversed U-curve (MFRU) subgroups, but only few people had LFA, MFU, or MFRU (two in each category) and were excluded. Thresholds were assessed at 0.125, 0.250, 0.5, 1, 2, 3, 4 and 8 kHz. Flat if differences between mean threshold values at 0.250/0.5kHz, 1/2 kHz, and	One way ANOVA, chi-square test	$p<0.05$

			4/8 kHz < 15 dB. HFGS if difference between mean threshold value at 4/8 kHz and at 0.5/1 kHz 15–30 dB; HFSS if ≥30 dB.		
Moring et al. (2016)	370	People with tinnitus from a membership database of a tinnitus organisation	6 subgroups based on somatic tinnitus quality: 1. Ringing, 2. Buzzing, 3. Whooshing, 4. Combination, 5. Hissing, 6. Other. Details: Participants were asked to ‘‘Please describe your tinnitus sensation to the best of your ability. For example, responses could include ‘‘ringing’’ or ‘‘buzzing.’’. Six categories were identified using an inductive process of initial coding, using the qualitative descriptions of tinnitus quality.	MANOVA	$p<0.05$
Oostendorp et al. (2016)	122	Patients with cervicogenic tinnitus from a primary care manual therapy practice	2 subgroups based on tinnitus sensitisation: 1. With tinnitus sensitisation, 2. Without tinnitus sensitisation. Details: Tinnitus sensitisation defined by the presence of at least 5 out of 8 of the following conditions: 1. Widespread hyperalgesia and pain remote from the symptomatic region, such as shoulder pain and back pain; 2. Impairment in quality of vision; 3. Burning eyes; 4. Modulation of tinnitus by psychological stress, such as sound phobia (fear of sound); 5. Modulation of tinnitus by sensory stimulation; 6. Presence of headache; 7. Presence of dizziness; 8. Tingling in arms or legs.	t-test, chi-square test	$p\leq0.05$
Ralli et al. (2016)	310	Patients with somatosensory tinnitus (TMJ and/or neck dysfunction, and/or somatic tinnitus modulation) from a tinnitus clinic	2 subgroups based on somatic tinnitus modulation: 1. Tinnitus modulation by somatic manoeuvres, 2. No tinnitus modulation by somatic manoeuvres. Details: Assessment included 19 somatic head-and-neck and TMJ manoeuvres. Examiners asked patients to perform specific movement or to resist to a pressure applied to the head, neck, and jaw and any tinnitus modulation was reported.	t-test, z test, logistic regression (backward variable selection)	$p<0.05$

Kojima et al. (2017)	584	Tinnitus patients from an Otorhinolaryngology department	2 subgroups based on presence of hyperacusis: 1. Yes, 2. No. Details: Classification based on the response to TSCHQ question 29.	t-test, chi-square test, Cohen's d effect size	$p<0.05$ and at least small effect sizes (Cohen's $d\geq0.2$)
Langguth et al. (2017a)	958	Tinnitus patients from a tertiary tinnitus clinic	2 types of classification. A. 4 subgroups based on headache laterality: 1. Left, 2. Right, 3. Bilateral, 4. No headache. B. 6 subgroups based on headache type: 1. Non-classifiable, 2. Migraine, 3. Tension-type, 4. Tension-type and migraine, 5. Cluster, 6. No headache. Details: Classification based on a headache questionnaire (Fritsche et al., 2007) and additional questions about headache characteristics.	ANOVA, chi-square test	$p<0.001$
Schmidt et al. (2017)	57	Research participants with tinnitus and normal loudness discomfort levels	4 subgroups based on tinnitus severity and duration: 1. Mild long term tinnitus A, 2. Recent-onset tinnitus, 3. Mild long-term tinnitus B, 4. Bothersome long-term tinnitus. Details: THI score ranged from 0-18 for mild tinnitus cases and was from 18-50 for bothersome tinnitus cases. For subgroup 2, THI score ranged from 0 to 34. Tinnitus duration for the recent onset subgroup ranged from 6 months to 1 year and was more than 1 year for the long-term tinnitus subgroup. The two mild long term tinnitus subgroups had data collected on two different Siemens MRI magnets.	Post-hoc t-test after ANCOVA (main analysis included subgroups without tinnitus)	$p<0.025$ or 0.0125
Ralli et al. (2018)	226	Chronic tinnitus patients with normal hearing from an Otorhinolaryngology department	2 types of classification. A. 2 subgroups based on a clinical definition of somatic tinnitus: 1. Somatic tinnitus, 2. Not somatic tinnitus. Details: Somatic tinnitus if they met 2 criteria; a self-reported history for TMJ dysfunction and a positive modulation of tinnitus following somatic manoeuvres in the TMJ region. The first was considered positive if before the onset of tinnitus	t-test, chi-square test, multiple logistic regression	$p<0.05$

			patients reported head trauma involving TMJ region, intensive manipulation of teeth or jaw, recurrent pain episodes in the TMJ region, increase of both TMJ pain and tinnitus at the same time, or intense periods of bruxism during day or night. The second was assessed by performing five manoeuvres and reporting of changes in tinnitus loudness (increase/decrease). It was considered present if at least one somatic manoeuvre modulated tinnitus. B. 2 subgroups based on a clinical diagnosis of TMJ disorders: 1. Clinically diagnosed TMJ disorder, 2. No clinically diagnosed TMJ disorder. Details: assessed by a specialized dentist according to Diagnostic Criteria for Temporomandibular Disorders Axis I (DC/TMD).		
Schmidt et al. (2018)	65	Research participants with tinnitus and normal loudness discomfort levels	2 types of classification. A. 2 subgroups based on tinnitus duration: 1. Mild long term tinnitus A, 2. Mild recent-onset tinnitus B. 2 subgroups based on tinnitus severity: 1. Bothersome long-term tinnitus, 2. Mild long term tinnitus B. Details: Tinnitus duration for the recent onset subgroup ranged from 6 months to 1 year and was more than 1 year for the long-term tinnitus subgroups. For severity categorisation, THI score less than 18 was categorized as mild and greater than 18 as bothersome tinnitus.	ANCOVA	$p<0.05$ family-wise error corrected or $p<0.001$ uncorrected
Suzuki et al. (2018)	110	Chronic tinnitus patients from a tinnitus clinic	3 subgroups based on tinnitus quality: 1. Pure tone, 2. Noise, 3. Multiple (pure tone and noise). Details not specified.	chi-square test, Mann–Whitney test, Kruskal-Wallis test, Tukey's post-	$p\leq0.05$

				hoc test, Friedman test	
Wang et al. (2018)	207	Consecutive subjective chronic tinnitus patients from an otolaryngologic department	2 subgroups based on tinnitus severity: 1. Mild tinnitus, 2. Severe tinnitus. Details: mild tinnitus if THI ≤ 37, severe tinnitus if THI ≥ 38	t-test, chi-square test, Mann–Whitney test	p≤0.05
Yüksel et al. (2018)	215	Chronic subjective non-pulsatile tinnitus patients from an otolaryngologic department	2 subgroups based on carotid artery intima-media thickness (IMT): 1. Normal IMT, 2. Increased IMT. Details: Ultrasound evaluation of both common, internal, and external carotid arteries using 7.5-MHz linear-type B-mode probe by a radiologist blind to the medical status of the patients. Tests were performed at supine position. Normal IMT when < 1 mm, increased IMT when ≥ 1 mm.	chi-square test, t-test, Mann Whitney test	p<0.05
Burkart et al. (2019)	320	Research participants with self-report of tinnitus during the past 12 months	3 subgroups based on tinnitus manifestation over time: 1. Continuous, 2. Intermittent, 3. Single episode Details: Based or self-report response to a question about whether they experienced tinnitus 'continuously', 'temporarily time and again', or 'only once but for several days' during the past 12 months.	chi-squared test, Kruskal–Wallis H test, ANOVA, post-hoc Scheffé test	p<0.05
Edvall et al. (2019)	2482	People with tinnitus from a population-based cohort	2 subgroups based on presence of temporomandibular joint disorder: 1. Yes, 2. No. Details: Classification based on the response to the TSCHQ question 32.	chi-squared test, Wilcoxon's test	p<0.05, Benjamini and Hochberg method for p-value correction
Koops et al. (2019)	1189	Tinnitus patients consultation at the ENT	2 subgroups based on tinnitus manifestation over time: 1. Constant, 2. Intermittent.	Kolmogorov Smirnov test, binary logistic regression,	p<0.05, variables were considering important for classification if they

		department of a university clinic	Details: Classification based on the response to question 'Is your tinnitus continuous or intermittent?'	random forest classification	were in the top 5% of variable importance
Michiels et al. (2019)	1262	Participants with tinnitus completing a web-based survey	2 subgroups based on presence of somatosensory tinnitus diagnosed by a clinician: 1. Yes, 2. No	Fisher's exact test, t-test	Benjamini and Hochberg method for p-value correction
Mores et al. (2019)	31	Patients subjected to audiological evaluation at a university clinic	2 subgroups based on presence of hearing loss: 1. Yes, 2. No. Details: In the hearing loss subgroup, unilateral or bilateral mild to moderate sensorineural hearing loss was defined according to the criteria of Silman and Silverman (1997) including those with type A, As or Ad tympanogram curve and presence or absence of acoustic reflexes. The normal hearing subgroup included participants with normal results in the basic audiological evaluation (up to 20dB at 250 - 8000Hz) and type A tympanogram curve with presence of bilateral ipsilateral and contralateral acoustic reflexes.	two-proportion equality test, ANOVA	$p<0.05$
Simões et al. (2019)	388	Tinnitus patients from a tertiary tinnitus clinic	2 types of classification. A. 3 subgroups based on change in THI between two assessments: 1. Clinically improved, 2. Clinically stable, 3. Clinically worsened. B. 3 subgroups based on change in TQ between two assessments: 1. Clinically improved, 2. Clinically stable, 3. Clinically worsened. Details: For the THI, clinical improvement was defined as a decrease in more than 7 points and clinical worsening as an increase in more than 7. For the TQ, clinical improvement was defined as 5 points decrease and clinical worsening as 1 point increase.	Tests for statistically significant differences (not specified)	$p<0.05$

Boecking et al. (2020)	1238	Chronic tinnitus patients from a tinnitus centre	2 subgroups based on tinnitus severity: 1. Low burden, 2. High burden. Details: Based on the German version of the TQ with cut-off at 46 points.	ANOVA, Cohen's d effect size	$p<0.05$, Bonferroni correction
Cederroth et al. (2020)	2432	People with tinnitus from a population-based cohort	2 subgroups based on presence of hyperacusis: 1.Yes, 2. No. Details: For most patients hyperacusis was defined using the ESIT-SQ question A12 'Over the last week, have external sounds been a problem, being too loud or uncomfortable for you when they seemed normal to others around you?' with response options 'no, not a problem', 'yes, a small problem', 'yes, a moderate problem', 'yes, a big problem', 'yes, a very big problem'. For a smaller subset of included patients it was defined using the TSCHQ question 29 'Do sounds cause you pain or physical discomfort?' with response options answers were: 'yes', 'no', 'don't know'.	chi-squared test, Wilcoxon's test	$p<0.05$, Benjamini and Hochberg method for p-value correction
Han et al. (2020)	42	Patients with typewriter or middle ear myoclonus tinnitus from a tinnitus clinic	2 subgroups based on type of clinical tinnitus diagnosis: 1. Typewriter tinnitus, 2. Middle ear myoclonus tinnitus. Details: Diagnosis of typewriter tinnitus based on the response to therapy with carbamazepine. Diagnosis of middle ear myoclonus tinnitus based on resolution of symptoms after middle ear tendon resection surgery.	Fisher's exact test, Man-Whitney test	$p<0.05$
Lugo et al. (2020)	2539	People with tinnitus from a population-based cohort	2 subgroups based on presence of headache: 1. Yes, 2. No.	chi-squared test, Wilcoxon's test, Cliff's δ	$p\leq0.05$, Benjamini and Hochberg method for p-value correction
Milner et al. (2020)	67	People with tinnitus from a hearing institute	2 subgroups based on tinnitus distress: 1. Low, 2. High. Details: Based on the median score from the THI.	chi-squared test, t-test, z-score, bootstrapping	$p<0.05$, adjustment for multiple comparisons based

					on false discovery rate
Niemann et al. (2020c)	1490	Patients with tinnitus for at least three months from a tinnitus clinic	2 subgroups based on depression: 1. Subclinical depression, 2. Clinical depression. Details: Subgroups based on dichotomising the General Depression Scale Questionnaire - long form. Subclinical if score 0–15 and clinical if score 16–60.	t-test, LASSO logistic regression	$p<0.05$, variables were considering important for classification if they had non-zero coefficient in LASSO regression
Niemann et al. (2020a)	1628	Patients with tinnitus for at least three months from a tinnitus clinic	2 subgroups based on sex: 1. Female, 2. Male.	Wilcoxon rank-sum test, RIDGE regression	$p<0.05$, Benjamini and Hochberg method for p-value correction, variables were considering important for classification if they were in the top 50% of variable importance
Niemann et al. (2020b)	1414	Patients with tinnitus for at least three months from a tinnitus clinic	2 subgroups based on tinnitus severity: 1. Compensated tinnitus, 2. Decompensated tinnitus. Details: Compensated if TQ score 0-46, decompensated if TQ score 47-84.	t-test, chi squared test, Mann-Whitney, Gradient boosted trees	$p<0.05$, Bonferroni correction with $a = 0.05/10$ (0.005), variables were considering important for classification if they were included in the best model during a

					feature selection process
Peters et al. (2020)	111	Tinnitus patients from a tinnitus clinic with available MRI of the cerebellopontine angle	2 types of classification. A. 2 subgroups based on tinnitus localisation: 1. Unilateral, 2. Bilateral. Details: Unilateral if either only in the right or in the left ear. B. 2 subgroups based on tinnitus severity: 1. Mild tinnitus, 2. Moderate to very severe tinnitus. Details: mild tinnitus if THI 0 to 35, moderate to very severe tinnitus if THI 36 to 100.	chi-squared test, univariate logistic regression analysis	$p<0.05$
Van der Wal et al. (2020)	316	Tinnitus patients from a tinnitus clinic	2 subgroups based on sex: 1. Male, 2. Female.	Standard least squares, nominal logistic models	$p<0.05$
Watabe et al. (2020)	134	Patients with tinnitus for over three months from an Otolaryngology department	2 subgroups based on tinnitus severity: 1. Light or moderate, 2. Severe. Details: Light/moderate if TFI score <56, severe if score TFI score ≥58.	Fisher's exact test	$p<0.05$

Appendix 2.3. Overview of studies using data-driven approaches

Reference	Sample size	Sample characteristics	Variables for clustering	Algorithm (method for selecting number of subgroups)	Subgroup descriptions	Validation
Erlandsson et al. (1991)	42	Tinnitus patients from an audiology department	Three dimensions of the mood adjective checklist (MACL): unpleasantness/pleasantness, deactivation/activation, tension/relaxation	k-means (prespecified three subgroups)	3 subgroups: 1. Low mood 2. Moderate mood 3. High mood	Audiometric variables, aetiology of hearing loss, age, anxiety, tinnitus severity and other tinnitus characteristics
Newman et al. (1997)	51	Outpatients in the Division of Audiology at Henry Ford Hospital with gradual-onset bilateral tinnitus (all subjects with normal hearing through 1500 Hz with a sloping high-frequency loss)	Scores from self-focused attention indices (PSC and SFSC) and from somatic attention measures (MSPQ and SOM-SCL-90-R)	Not specified (four possible subgroupings of the data, each one having a different number of subgroups, were generated; each of these was examined to determine which was most intuitively appealing clinically)	2 subgroups: 1. Lower score on both self-attention and somatic attention measures (‘low self - attenders’, n=32) 2. More internally directed, higher score on the attention measures (‘high self-attenders’, n=19). On average more depressed, with greater emotional distress due to tinnitus, and with greater perceived tinnitus handicap. (Note: no differences for pitch and loudness)	Tinnitus severity measures (TRQ, THQ), depression symptomatology (BDI), tinnitus psychoacoustics (pitch and loudness matching)

Andersson and McKenna (1998)	30	Tinnitus patients at an audiology department (primary referrals for tinnitus complaints treated by a clinical psychologist)	4 variables: Beck Depression Inventory score, tinnitus loudness, Minimal Masking Level, Pure Tone Average	Hierarchical with Ward's method and squared Euclidean distances (not specified)	3 subgroups: 1. Low depression, average loudness, slightly above average Minimal Masking Level and Pure Tone Average 2. High value only in depression 3. High values in all variables	Age, duration, tinnitus grading
Rizzardo et al. (1998)	84	Patients at an ENT department with primary complaint of tinnitus, followed up at a Neurological and Psychiatric Sciences Department)	13 variables: personality traits, anxiety, depression, and illness behaviour test scores	Hierarchical complete linkage (not specified)	3 subgroups: 1. Higher scores for depression anxiety and neuroticism 2. Normal psychological tests apart from marked denial 3. One patient	Demographics, psychological or functional symptoms, clinical variables
Tyler et al. (2008)	153 (Initial sample 246, 93 excluded due to missing data)	Tinnitus patients enrolled in clinical trials	51 variables: demographics, symptomatology, case history variables, questionnaire scores, psychoacoustic tinnitus measures. Authors did not	2-step cluster analysis (tested 4-6 cluster solutions, choice of 4 because it resulted in about equal subgroup sizes)	4 subgroups: 1. Loud, persistent, and distressing tinnitus, suffering from loudness hyperacusis 2. Varying tinnitus pitch and loudness, worse in noise 3. Copers, tinnitus not influenced by touch	None

			explicitly mention all variables.		4. Copers, tinnitus worse in quiet and better in noise	
Schecklmann et al. (2012)	44	Chronic tinnitus patients enrolled in clinical trials of repetitive transcranial magnetic simulation	Three analyses using different variables: A. 3 variables: tinnitus distress, duration, and laterality B. Voxel based morphometry data C. Positron emission tomography data	Hierarchical with Ward's method and Euclidean distances (forced to 2 subgroups to have sample sizes with sufficient statistical power)	2 subgroups from each analysis: A. Unilateral versus mainly bilateral tinnitus B. Higher versus lower grey matter volume in medial superior prefrontal, cingulate, temporal, insular, orbital frontal, temporal, pre- and post-central and thalamic areas C. Higher versus lower glucose metabolism in middle and superior temporal, precuneus, and superior parietal areas	None
van den Berge et al. (2017)	A. 976, B. 761	Patients with severe tinnitus at a tertiary clinic	Two analysis using different variables: A. 8 variables: the variables with the highest loading on each of 8 components from a principal component analysis on 30 variables (demographics, tinnitus characteristics and modulating factors, tinnitus psychoacoustic measures,	2-step cluster analysis (silhouette measure)	A. 4 subgroups 1. Tinnitus not easily influenced, different hearing loss between ears 2. Gradual onset of tinnitus, easily negatively influenced by loud sounds and sleep deprivation 3. Tinnitus less loud with loud sounds, no effect from sleep deprivation or nap 4. Acute onset tinnitus, tinnitus easily negatively influenced by loud sounds or sleep deprivation B. 3 subgroups 1. Tinnitus not easily influenced, preference for noisy environments, tinnitus mostly unilateral, low tinnitus	None

			audiometric data, questionnaire scores) B. 11 variables: Demographics, tinnitus characteristics and modulating factors, audiometric data, questionnaire scores		severity scores, sounds rarely uncomfortably loud 2. Mainly males, tinnitus worse by stress, loud sounds and movement of head and neck, preference for noisy environments, sometimes sounds are uncomfortably loud, most with no or slight hearing loss, bilateral tinnitus with variable loudness 3. Tinnitus worse by loud sounds and stress, prefer silent environment, often find sounds uncomfortably loud, tinnitus often bilateral, most with no, slight, or asymmetrical hearing loss, variable loudness Note: poor silhouette measure for both analyses	
Langguth et al. (2017b)	2838	Patients with tinnitus treated at a tertiary tinnitus clinic	14 variables: hearing loss at 7 frequencies (0.125, 0.250, 0.5, 1, 2, 4, and 8 kHz) from each ear with 4 possible values (1: normal; 2: mild / moderate; 3: severe / profound; 4: no data)	Latent Class Analysis (BIC)	8 subgroups: 1. Lacking audiometry 2. Bilateral high frequency hearing loss 3. Normal hearing 4. Bilateral medium-HF hearing loss 5. Severe pantonal hearing loss 6. Left-sided pantonal medium hearing loss 7. Right-sided pantonal severe hearing loss 8. Left-sided pantonal severe hearing loss	Demographics, tinnitus severity measures (TQ, TBF-12), depression symptomatology (BDI), tinnitus characteristics (self-reported and psychoacoustic measures)

Appendix 2.4. Overview of studies using treatment response approaches

Reference	Sample size	Sample characteristics	Treatments	Outcomes and subgrouping
Côté et al. (2019)	31	Tinnitus patients (from a waiting list and a tinnitus association) with somatic component (tinnitus following cervical trauma or manipulations, simultaneous occurrence, or concurrent increase of intensity with other somatic symptoms, inadequate posture at rest, when sleeping, and at work, bruxism, or temporomandibular joint problems.	Ten physiotherapy treatments (cervical and thoracic mobilizations, muscular strengthening, stretching, postural instruction, and cervical stabilization) over six weeks on average.	Two subgroups based on improvement after treatment or not. Details: Improved if they either showed significant improvement in at least two of the three primary outcomes or indicated subjective improvement. Three primary outcomes: The Tinnitus Handicap Inventory and visual analog scales for tinnitus loudness and annoyance, measured at baseline (multiple times and averaged) and at the end. Significant improvement if post-treatment values exceeding the 95% confidence intervals of the multiple pre-treatment values. Subjective improvement was defined using the Clinical Global Improvement 7-level scale.
Kloostra et al. (2019)	61	Patients with tinnitus who received a cochlear implant from a University medical centre	Cochlear implantation	Two subgroups based on whether cochlear implantation had or had not a positive effect on tinnitus. Details: In response to the question 'Were there any changes concerning your tinnitus after the cochlear implantation?', responses 'Yes, the tinnitus I experienced before implantation disappeared after implantation.' and 'Yes, my already existing tinnitus got better after implantation' were considered positive response. Responses 'No, my already existing tinnitus remained the same after implantation' and 'Yes, my already existing tinnitus got worse after implantation' were considered no positive response.

Appendix 3.1. Sources used to develop the ESIT-SQ core and optional questions

NN	Questions/Response options	Sources
	PART A. INDIVIDUAL CHARACTERISTICS	
A1	Age (years)	
A2	At birth were you described as:	Reproduced from Balarajan et al. (2011)
	☐ Male ☐ Female ☐ Intersex ☐ Prefer not to say	Reproduced from Balarajan et al. (2011)
A3	What is your height?	Adapted from the European epidemiological survey for tinnitus question A9 (Biswas et al., 2019; Gallus et al., 2020)
A4	What is your weight?	Adapted from the European epidemiological survey for tinnitus question A10 (Biswas et al., 2019; Gallus et al., 2020)
O1	What is your handedness?	Adapted from TSCHQ question 3
	☐ Right ☐ Left ☐ Both (ambidextrous)	Adapted from TSCHQ question 3
O2	What is your country of residence?	Reproduced from STOP project questionnaire (unpublished)
O3	What was your country of birth?	Reproduced from STOP project questionnaire (unpublished)
A5	What is the highest education level you have achieved?	Adapted from the European epidemiological survey for tinnitus question A4 (Biswas et al., 2019; Gallus et al., 2020)
	☐ No school ☐ Primary (elementary school) ☐ Lower secondary (middle school) ☐ Upper secondary (high school) ☐ University or higher degree	Adapted from the European epidemiological survey for tinnitus question A4 (Biswas et al., 2019; Gallus et al., 2020)
O4	What is your marital status?	Adapted from: the European epidemiological survey for tinnitus question A6 (Biswas et al., 2019; Gallus et al., 2020); STOP project questionnaire (unpublished)

	☐ Married ☐ Living with partner ☐ Single ☐ Widow/Widower ☐ Divorced/Separated ☐ Prefer not to say	Adapted from: the European epidemiological survey for tinnitus question A6 (Biswas et al., 2019; Gallus et al., 2020); STOP project questionnaire (unpublished)
O5	How is your economic status relative to the average of the country where you leave in?	Adapted from: the European epidemiological survey for tinnitus question A7 (Biswas et al., 2019; Gallus et al., 2020)
	☐ Much higher than the average ☐ Quite higher than the average ☐ On the average ☐ Quite lower than the average ☐ Much lower than the average ☐ Prefer not to say	Adapted from: the European epidemiological survey for tinnitus question A7 (Biswas et al., 2019; Gallus et al., 2020)
O6	Which of the following describes best your current situation?	Reproduced from STOP project questionnaire (unpublished)
	☐ Employed ☐ Unemployed ☐ Running my own business/Working as a partner in a company ☐ Retired ☐ Sick leave (for more than two months) or disability pension due to illness or disability ☐ Parental leave (since two months or longer) ☐ Student ☐ Sabbatical ☐ Housewife/-Husband ☐ Other ☐ Do not know	Reproduced from STOP project questionnaire (unpublished)
O7	Have you ever worked at night (i.e., between 24:00-5:00)?	Reproduced from STOP project questionnaire (unpublished)
	☐ Yes, I do currently ☐ Yes, I have done it before ☐ No ☐ Do not know	Reproduced from STOP project questionnaire (unpublished)
A6	What is the average number of alcoholic drinks that you consume per week? (One drink equals 125 ml of wine, 330 ml of beer or 40 ml of spirits) _ _	Translated and adapted by Silvano Gallus from previous survey conducted in Italy on smoking, alcohol, and tinnitus (Asciutto et al., 2016)
A7	Which of the following options best describes your smoking status?	Translated and adapted by Silvano Gallus from previous survey conducted in Italy on smoking, alcohol, and tinnitus (Gallus et al., 2015)
	☐ Never smoker ☐ Current Smoker ☐ Ex-smoker	Translated and adapted by Silvano Gallus from previous survey conducted in Italy on smoking, alcohol, and tinnitus

		(Gallus et al., 2015; Lugo et al., 2015)
O8	How many cigarettes do you smoke per day on average? _ _	Translated and adapted by Silvano Gallus from previous survey conducted in Italy on smoking, alcohol, and tinnitus (Gallus et al., 2015; Lugo et al., 2015)
O9	How many cups of coffee do you drink per day on average? _ _	Translated and adapted by Silvano Gallus from an Italian network of case-control studies (Gallus et al., 2002)
O10	How many hours per week do you do leisure-time physical activities on average?	Translated and adapted by Silvano Gallus from an Italian network of case-control studies (Tavani et al., 2001)
	☐ Less than 2 hours per week ☐ 2 - 4 hours per week ☐ 5 - 7 hours per week ☐ More than 7 hours per week	Translated and adapted by Silvano Gallus from an Italian network of case-control studies (Tavani et al., 2001)
O11	How often do you consume meat on average?	Translated and adapted by Silvano Gallus from an ongoing Italian case-control study on tinnitus
	☐ Vegetarian/vegan (no meat) ☐ Occasional (less than 3 times per month) ☐ 1 time per week ☐ 2 - 3 times per week ☐ 4 - 5 times per week ☐ 6 or more times per week	Translated and adapted by Silvano Gallus from an ongoing Italian case-control study on tinnitus
O12	How often do you consume fish on average?	Translated and adapted by Silvano Gallus from an ongoing Italian case-control study on tinnitus
	☐ Vegetarian/vegan (no fish) ☐ Occasional (less than 3 times per month) ☐ 1 time per week ☐ 2-3 times per week ☐ 4-5 times per week ☐ 6 or more times per week	Translated and adapted by Silvano Gallus from an ongoing Italian case-control study on tinnitus
O13	How often do you consume fruits on average?	Translated and adapted by Silvano Gallus from an ongoing Italian case-control study on tinnitus
	☐ Never ☐ Occasional (less than 3 times per month) ☐ 1 - 6 times per week ☐ 1 time per day ☐ 2 - 3 times per day ☐ 4 or more times per day	Translated and adapted by Silvano Gallus from an ongoing Italian case-control study on tinnitus
O14	How often do you consume vegetables on average?	Translated and adapted by Silvano Gallus from an

		ongoing Italian case-control study on tinnitus
	☐ Never ☐ Occasional (less than 3 times per month) ☐ 1 - 6 times per week ☐ 1 time per day ☐ 2 - 3 times per day ☐ 4 or more times per day	Translated and adapted by Silvano Gallus from an ongoing Italian case-control study on tinnitus
O15	For how long do you use a mobile phone for calls on average?	Translated and adapted by Silvano Gallus from an ongoing Italian case-control study on tinnitus
	☐ No use ☐ Less than 1 hour per month ☐ Around 1 hour per month ☐ 2 - 3 hours per month ☐ Around 1 hour per week ☐ 2 - 6 hours per week ☐ 1 hour per day ☐ More than 1 hour per day	Translated and adapted by Silvano Gallus from an ongoing Italian case-control study on tinnitus
O16	For how long do you use headphones to listen to music on average?	Translated and adapted by Silvano Gallus from an ongoing Italian case-control study on tinnitus
	☐ No use ☐ Less than 1 hour per month ☐ Around 1 hour per month ☐ 2 - 3 hours per month ☐ Around 1 hour per week ☐ 2- 6 hours per week ☐ 1 hour per day ☐ More than 1 hour per day	Translated and adapted by Silvano Gallus from an ongoing Italian case-control study on tinnitus
O17	How many hours do you sleep per day on average?	Translated and adapted by Silvano Gallus from an ongoing Italian case-control study on tinnitus
	☐ Less than 6 hours per day ☐ Around 6 hours ☐ Around 7 hours ☐ Around 8 hours ☐ 9 or more hours per day	Translated and adapted by Silvano Gallus from an ongoing Italian case-control study on tinnitus
A8	How many first degree relatives (parents, children, siblings) do you know to have tinnitus or hearing loss? (please write a number next to each family member)	Adapted from TSCHQ question 4
	__Father __Mother __Brothers __Sisters __Sons __Daughters	Adapted from TSCHQ question 4
A9	Do you suffer from vertigo (sensation of spinning or tilting)?	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 31
	☐ Never ☐ Yes, less than one episode per year ☐ Yes, at least one episode per year	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 31
A10	Have you been diagnosed with any other ear condition? You can choose more than one option.	Adapted from GPCHQ questionnaire (unpublished)

	☐ Acoustic trauma (caused by loud sounds) ☐ Ear barotrauma (caused by acute change in ambient pressure) ☐ Presbycusis (aging of ears) ☐ Sudden hearing loss ☐ Other hearing loss ☐ Meniere's disease ☐ Acoustic neuroma (auditory nerve tumour) ☐ Acute otitis (ear inflammation) ☐ Serous otitis or Eustachian tube dysfunction ☐ Chronic otitis (e.g. tympanic perforation, cholesteatoma) ☐ Otosclerosis (reduced ossicles mobility) ☐ Other ear disorders. Please specify _____ ☐ No	Adapted from GPCHQ questionnaire (unpublished)
A11	Have you ever undergone any of the following procedures? You can choose more than one option.	Adapted from GPCHQ questionnaire (unpublished)
	☐ Ear surgery ☐ Dental surgery ☐ Neurosurgery ☐ Lumbar puncture ☐ Chemotherapy ☐ Head and neck radiotherapy ☐ Electroconvulsive therapy ☐ Other procedure. Please specify _____ ☐ None of these	Adapted from GPCHQ questionnaire (unpublished)
A12	Over the last week, have external sounds been a problem, being too loud or uncomfortable for you when they seemed normal to others around you? Note: external sounds refer to any sounds other than tinnitus, e.g. environmental sounds, speech, music.	Adapted from the Tinnitus and Hearing Survey (National Center for Rehabilitative Auditory Research, 2017)
	☐ No, not a problem ☐ Yes, a small problem ☐ Yes, a moderate problem ☐ Yes, a big problem ☐ Yes, a very big problem	Reproduced from the Tinnitus and Hearing Survey (National Center for Rehabilitative Auditory Research, 2017)
A13	Do you currently have any other difficulty with your hearing, such as listening to speech in a noisy situation?	Reproduced from the European epidemiological survey for tinnitus 'presence of hearing difficulty' question (Biswas et al., 2019)
	☐ Yes, cannot hear at all ☐ Yes, severe difficulty ☐ Yes, moderate difficulty ☐ Yes, slight difficulty ☐ No difficulty ☐ Do not know	Adapted from the European epidemiological survey for tinnitus 'presence of hearing difficulty' question (Biswas et al., 2019)
A14	Do you use any of the following devices? You can choose more than one option.	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 27
	☐ Hearing aid ☐ Cochlear implant ☐ Sound generator ☐ Combination device (hearing aid and sound generator in the same device) ☐ None	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 27
A15	Do you suffer from any of the following pain syndromes? You can choose more than one option.	Adapted from: GPCHQ questionnaire (unpublished);

		TSCHQ question 30; TSCHQ question 33; TSCHQ question 34
	☐ Headache ☐ Neck pain ☐ Ear pain ☐ Temporomandibular joint pain ☐ Pain in the face ☐ No ☐ Other. Please specify _____	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 30; TSCHQ question 33; TSCHQ question 34
A16	Do you have any of the following conditions that have been diagnosed by a clinician? You can choose more than one option.	Adapted from GPCHQ questionnaire (unpublished)
	Oral: ☐ Temporomandibular joint disorder ☐ Dental problems **Neurological:** ☐ Meningitis ☐ Multiple sclerosis ☐ Epilepsy ☐ Stroke ☐ Other cerebrovascular disease ☐ Dementia ☐ Other neurologic disease **Psychiatric or psychological:** ☐ Anxiety ☐ Depression ☐ Emotional trauma ☐ Excessive stress **Sleep disorders:** ☐ Difficulty falling asleep ☐ Difficulty staying asleep **Cardiovascular:** ☐ Low blood pressure ☐ High blood pressure ☐ Myocardial infraction (heart attack) **Endocrine and metabolic:** ☐ Thyroid disorder ☐ Diabetes ☐ Hyperinsulinemia ☐ Increased cholesterol **Rheumatological and immune mediated:** ☐ Rheumatoid arthritis ☐ Systemic lupus erythematosus **Otorhinolaryngological:** ☐ Chronic sinusitis ☐ Nasal septum deviation **Infectious:** ☐ Syphilis ☐ HIV ☐ Lyme disease **Other:** ☐ Anaemia ☐ Instability or other balance disorders ☐ Acid/gastroesophageal reflux ☐ Globus hystericus ☐ Other. Please specify _____ ☐ None	Adapted from GPCHQ questionnaire (unpublished)
A17	Tinnitus refers to the perception of noise in your head or ears (such as ringing or buzzing) in the absence of any corresponding source of sound external to your head. Over the past year, have you had tinnitus in your head or in one or both ears that lasts for more than five minutes at a time?	Adapted from the European epidemiological survey for tinnitus 'presence of tinnitus' question (Biswas et al., 2019)
	☐ Yes, most or all of time ☐ Yes, a lot of the time ☐ Yes, some of the time ☐No, not in the past year ☐ No, never ☐ Do not know	Adapted from the European epidemiological survey for tinnitus 'presence of tinnitus' question (Biswas et al., 2019)
	PART B. TINNITUS CHARACTERISTICS	
B1	How often do you have tinnitus on average?	Adapted from Tinnitus Questionnaire question 7 (Schechter and Henry, 2002)

	☐ Daily or almost daily ☐ Almost weekly ☐ Almost monthly ☐ Every few months ☐ Yearly	Adapted from Tinnitus Questionnaire question 7 (Schechter and Henry, 2002)
B2	What best describes your tinnitus during a day?	Adapted from: The Tinnitus Screener (National Center for Rehabilitative Auditory Research, 2017); TSCHQ question 10
	☐ Constant: you can always or usually hear it in a quiet room ☐ Intermittent: "comes and goes", cannot always hear it in a quiet room	Adapted from: The Tinnitus Screener (National Center for Rehabilitative Auditory Research, 2017); TSCHQ question 10
B3	How long ago did your tinnitus appear?	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 5
	__ months __ years ☐ Do not know	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 5
B4	Over the past year, how much does your tinnitus worry, annoy or upset you when it is at its worst?	Adapted from the European epidemiological survey for tinnitus 'tinnitus severity' question (Biswas et al., 2019)
	☐ Severely ☐ Moderately ☐ Slightly ☐ Not at all ☐ Do not know	Adapted from the European epidemiological survey for tinnitus 'tinnitus severity' question (Biswas et al., 2019)
B5	How long ago did your tinnitus start bothering you?	Adapted from GPCHQ questionnaire (unpublished)
	__ months __ years ☐ Do not know	Adapted from GPCHQ questionnaire (unpublished)
B6	Although, most patients have tinnitus of a single type, some may hear different sounds. Do you hear one or more different sounds?	Adapted from GPCHQ questionnaire (unpublished)
	☐ One sound ☐ More than one different sound	Adapted from GPCHQ questionnaire (unpublished)
B7	How was the start of your tinnitus?	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 6
	☐ Gradual ☐ Sudden ☐ Do not know	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 6
B8	If you reported any conditions/procedures in questions A9, A10, A11, A12, A13, A15 or A16, please list them here and write next to them if they	Adapted from GPCHQ questionnaire (unpublished)

	happened BEFORE, AFTER, or at about the SAME TIME as your tinnitus onset.	
B9	Was the initial onset of your tinnitus related to (you can choose more than one option):	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 7
	☐ Exposure to loud sounds ☐ Change in hearing ☐ Exposure to change in ambient pressure (e.g. flight or diving) ☐ Flu, common cold or other infection ☐ Feeling of fullness or pressure in the ears ☐ Stress ☐ Head trauma ☐ Neck trauma (e.g. whiplash) ☐ Others. Please specify _____ ☐ None	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 7
B10	Were you taking any of the medicines listed below around the time of your tinnitus onset? You can choose more than one option.	Adapted from GPCHQ questionnaire (unpublished)
	☐ Aspirin ☐ Pain killing medication. Please specify _____ ☐ Oral steroids. Please specify _____ ☐ Antibiotics. Please specify _____ ☐ Antidepressants. Please specify _____ ☐ Quinine (muscle cramps, malaria) ☐ Water tablets (diuretics). Please specify _____ ☐ Other medicines. Please specify ☐ Do not know ☐ No	Adapted from GPCHQ questionnaire (unpublished)
B11	Do you think any of the conditions mentioned before or any other conditions are related to your tinnitus onset? You can give up to 3 responses - please choose the most important.	Adapted from GPCHQ questionnaire (unpublished)
	☐ No ☐ Yes. Please specify _____	Adapted from GPCHQ questionnaire (unpublished)
B12	Is the loudness of your tinnitus stable over time or does it fluctuate over a day?	Adapted from GPCHQ questionnaire (unpublished)
	☐ Stable ☐ Sometimes fluctuating ☐ Always fluctuating ☐ Do not know	Adapted from GPCHQ questionnaire (unpublished)
B13	What does your tinnitus sound like?	Adapted from TSCHQ question 14
	☐ Tonal ☐ Noise-like ☐Music-like ☐ Crickets ☐ Other. Please specify _____	Adapted from TSCHQ question 14
B14	Please describe the pitch of your tinnitus:	Adapted from TSCHQ question 15
	☐ High pitched ☐ Medium pitched ☐ Low pitched ☐ Do not know	Adapted from TSCHQ question 15

B15	Where do you perceive your tinnitus?	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 9
	☐ Right ear ☐ Left ear ☐ Both ears, worse in left ☐ Both ears, worse in right ☐ Both ears, equally ☐ Inside the head ☐ Other. Please specify _____ ☐ Do not know	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 9
B16	Is your tinnitus rhythmic?	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 8
	☐ No ☐ Yes, following heart beat (can be checked by feeling the pulse at the same time as listening to the tinnitus) ☐ Yes, following breathing ☐ Yes, following movements of the head, neck, jaw or muscles of the face ☐ Other. Please specify _____	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 8
B17	Has a clinician ever heard your tinnitus?	Adapted from GPCHQ questionnaire (unpublished)
	☐ Yes ☐ No	Adapted from GPCHQ questionnaire (unpublished)
B18	Is your tinnitus reduced by (you can choose more than one option):	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 19; TSCHQ question 21; TSCHQ question 22; TSCHQ question 23; TSCHQ question 24; TSCHQ question 25
	☐ Very quiet environment ☐ Low intensity sounds ☐ High intensity sounds ☐ Head movements ☐ Clenching the teeth or moving the jaw ☐ Pressing your head, neck, or area around the ear ☐ Taking a nap ☐ Good sleep quality ☐ Driving ☐ Being stressed or anxious ☐ Being relaxed ☐ Drinking alcohol ☐ Drinking coffee ☐ Medications ☐ Using hearing aids ☐ Other. Please specify _____ ☐ None	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 19; TSCHQ question 21; TSCHQ question 22; TSCHQ question 24; TSCHQ question 25
B19	Is your tinnitus increased by (you can choose more than one option):	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 20; question 21; TSCHQ question 22; TSCHQ question 24; TSCHQ question 25

	☐ Very quiet environment ☐ Low intensity sounds ☐ High intensity sounds ☐ Head movements ☐ Clenching the teeth or moving the jaw ☐ Pressing your head, neck, or area around the ear ☐ Taking a nap ☐ Poor sleep quality ☐ Driving ☐ Being stressed or anxious ☐ Being relaxed ☐ Drinking alcohol ☐ Drinking coffee ☐ Medications ☐ Using hearing aids ☐ Other. Please specify _____ ☐ None	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 20; question 21; TSCHQ question 22; TSCHQ question 24; TSCHQ question 25
B20	Over the past year, have you seen your family doctor, or seen a healthcare professional at a clinic or hospital about your tinnitus?	Reproduced from the European epidemiological survey for tinnitus 'Use of healthcare resources for tinnitus' question (Biswas et al., 2019)
	☐ Yes, 5 or more visits ☐ Yes, from 2 to 4 visits ☐ Yes, just one visit ☐ Not at all ☐ Do not know	Adapted from the European epidemiological survey for tinnitus 'Use of healthcare resources for tinnitus' question (Biswas et al., 2019)
B21	Are you currently receiving any of the following types of management for your tinnitus? You can choose more than one option.	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 35
	☐ Psychiatric management ☐ Psychological management ☐ Audiological management ☐ Physiotherapy ☐ Self-management (e.g. dietary supplements, support groups, relaxation) ☐ Other. Please specify _____ ☐ No management	Adapted from: GPCHQ questionnaire (unpublished); TSCHQ question 35
B22	Do you think any of the conditions mentioned before, or any other conditions, are related to periods of increased tinnitus? You can give up to 3 responses - please choose the most important.	Adapted from GPCHQ questionnaire (unpublished)
	☐ No ☐ Yes. Please specify _____	Adapted from GPCHQ questionnaire (unpublished)

A: questions from part A; B: questions from part B; O: optional questions

TSCHQ: Tinnitus Sample Case History Questionnaire (Langguth et al., 2007); STOP: (Swedish Tinnitus Outreach Project, 2015); GPCHQ: Case history questionnaire used at the Georges Pompidou tinnitus hyperacusis clinic (Paris)

Appendix 3.2. Guide for further translations of the ESIT-SQ

Instructions for Translation Lead

The Translation Lead should be responsible for managing the resources, procedural steps and documentation and should read the good practice guide for translating and adapting hearing-related questionnaires by Hall et al. 2017 (https://doi.org/10.1080/14992027.2017.1393565).

The original ESIT-SQ developers would appreciate a copy of the completed excel template for record keeping. This is important for approval as an authorised translation of the ESIT-SQ.

Summary of translation process

1. Preparation (step 1 in Hall et al. 2017):
    - Recruit three translators, a tinnitus expert and a healthcare practitioner.
    - Inform members about the concepts of the questionnaire and the requirements of the translation.
2. Produce two independent forward translations reconciled into one final translation by the third translator (step 2 in Hall et al. 2017).
3. Conduct a Committee Review meeting (including the translators, tinnitus expert, and healthcare practitioner) to review and update translated items (step 4 in Hall et al. 2017).
4. Test out the translation with up to eight native speakers who represent the target population, at least 50% (half) with tinnitus (step 5 in Hall et al. 2017).
5. Final Review (step 6 in Hall et al. 2017):
    - Final reviewing, proofreading, formatting and finalising translation.
    - Finalise and archive translation report.

Back translation (step 3 in Hall et al. 2017) can also be included in the translation process but is not required.

Detailed instructions

1. Preparation

- Familiarise with the Translation Template (contact the source-Language developers for any questions: eleni.genitsaridi@nottingham.ac.uk):

a. The first sheet, “TranslationInfo”, is to record information about the team members (names, emails, and professional profile or Curriculum Vitae) and other general information about the translation.
b. Each other sheet has items for translation and includes the following columns:
    - Column A has the original text in English.
    - Column B has concept definitions for each item to make the meaning of the items (i.e. instructions, questions, or response options) more clear to the translators.
    - Columns C and D are to record the two independent forward translations and column E is to record the single reconciled translation.
    - Column F is to record any difficulties with the forward translations or any possible inaccuracies between the source and the target items.
    - Columns G and H are to record any comments from the healthcare practitioner or the tinnitus expert respectively.
    - Column I is to record the updated translated version after the Committee Review meeting, but before the field testing.
    - Column J is to record comments from the field testing.
    - Column K is to record the final translation.
c. Sheets “Part A” and “Part B” include the ESIT-SQ instructions and mandatory questions. Some questions have a different wording for the paper and electronic versions. Items that are not needed for the paper version are coloured in grey.
d. Sheet “Optional” includes the optional questions. It should be decided on an early stage whether these items will also be translated or not.

- Appoint three translators, a tinnitus expert and a healthcare practitioner with the following characteristics:
    a. The target language is their first language
    b. They have lived in the country of the target language
    c. They are highly proficient in English
    d. They are educated to University level

One person can perform both a translator's role and the tinnitus expert and/or healthcare practitioner role. Hence the Translation Team can comprise of between 3 and 5 members. The tinnitus expert should be working in the area of tinnitus (as a clinician, academic or commercial) currently and at least during the past year. The healthcare practitioner should be familiar with defining and describing common medical conditions. The two translators should independently produce a forward translation which will then be consolidated by the third translator (i.e. reviewer) into a single translation. The reviewer should be the most experienced in tinnitus and/or medical conditions.

2. Forward translations

Inform translators about the concepts of the questionnaire and the requirements of the translation and ensure that every item is translated carefully (thinking of cross-cultural adaptation and equivalence).

a. Give translators background information in case they are not familiar with tinnitus research, e.g. regarding common co-existing condition, the highly heterogeneous tinnitus population, the lack of standards for assessment and management etc.
b. Discuss with them the characteristics of the target population (such as literacy, different locations) and how the questionnaire will be administered (self-reported, either in paper or electronic version).
c. Explain them the concept of cross-cultural adaptation in translations, as described by Hall et al., 2017. You can find the relevant quotes from this article, in the end of this document.
d. Explain them how to use the translation template.
e. Two translators should produce forward translations independently of one another.
f. The two translations should be reconciled by the third translator/reviewer to create a single forward translation. Decisions should be made considering closest possible meaning and in consultation with the translators when needed. Text that is repeated (e.g. Do not know) should be translated in a common way if the target language permits it. It is important to keep consistency on terminology

used across the items within the same language. For more details read Item 2e in Hall et al., 2017.

3. Committee Review to review and update translation

- Conduct a Committee Review meeting, including translators, tinnitus expert, and healthcare practitioner, to examine whether all the translation units are accurate and whether they map to the English version. Make any necessary changes according to the committee's consensus.
- Create the translated formatted version of the questionnaire, by mapping translated items to the original English template. The translation language should be added in brackets in each page after the version number. If the optional questions are also translated, both a version with and without these questions should be created.

4. Field Testing

Field testing involves the completion of the questionnaire by a small number of people, and can include questions about the difficulty in understanding any items or response options.

- Aim to recruit eight native speakers (at least half of these with tinnitus). They should adequately represent the target population (e.g. in tinnitus severity, age, gender, education, regional dialect, socio-economic status).
- Instruct them to complete the questionnaire and make a note of the time it took them to complete it and of any word or phrase that was difficult to understand.

5. Final Review

- The translation lead should gather all results from field testing, discuss them with the translation team members and finalise translation according to their consensus.
- Check formatting and proof read for spelling and grammar mistakes.
- Finalise and archive a translation report.

Appendix 3.3. English version of the ESIT-SQ with optional questions

European School for Interdisciplinary Tinnitus Research – Screening Questionnaire (ESIT-SQ)

This questionnaire has two parts.

In part A, we ask for some general personal characteristics such as age, height, life-style questions and conditions that might concern you. Everyone can complete part A, even if you've never had tinnitus. The estimated time to complete this part is 5 minutes.

If you have experienced tinnitus during the past year, you will be asked some more tinnitus-related questions in part B. The estimated time to complete part B is between 5 and 10 minutes, depending on how you answer.

PART A. INDIVIDUAL CHARACTERISTICS

For the following questions, please give the answer that best describes you and your experiences. For some questions you can choose more than one option.

A1 Age (years)

_ _

A2 At birth were you described as:

☐ Male ☐ Female ☐ Intersex ☐ Prefer not to say

A3 What is your height?

_ _ _ cm OR _ feet _ inches

A4 What is your weight?

_ _ _ kg OR _ st _ lbs

O1 What is your handedness?

☐ Right ☐ Left ☐ Both (ambidextrous)

O2 What is your country of residence?

_____

O3 What was your country of birth?

_____

A5 What is the highest education level you have achieved?

☐ No school
☐ Primary (elementary school)
☐ Lower secondary (middle school)
☐ Upper secondary (high school)

☐ University or higher degree

O4 What is your marital status?

☐ Married
☐ Living with partner
☐ Single
☐ Widow/Widower
☐ Divorced/Separated
☐ Prefer not to say

O5 How is your economic status relative to the average of the country where you live in?

☐ Much higher than the average
☐ Quite higher than the average
☐ On the average
☐ Quite lower than the average
☐ Much lower than the average
☐ Prefer not to say

O6 Which of the following describes best your current situation?

☐ Employed
☐ Unemployed
☐ Running my own business/Working as a partner in a company
☐ Retired
☐ Sick leave (for more than two months) or disability pension due to illness or disability
☐ Parental leave (since two months or longer)
☐ Student
☐ Sabbatical
☐ Housewife/-Husband
☐ Other
☐ Do not know

O7 Have you ever worked at night (i.e. between 24:00-5:00)?

☐ Yes, I do currently
☐ Yes, I have done it before
☐ No
☐ Do not know

A6 What is the average number of alcoholic drinks that you consume per week?

1 drink equals 125 ml glass of wine, 330 ml of beer or 40 ml of spirits

_ _

A7 Which of the following options best describes your smoking status?

☐ Never smoker ☐ Current Smoker ☐ Ex-smoker

If you answered "Never smoker" or “Ex-smoker”, please go to question O9.

O8 How many cigarettes do you smoke per day on average?

_ _

O9 How many cups of coffee do you drink per day on average?

_ _

O10 How many hours per week do you do leisure-time physical activities on average?

☐ Less than 2 hours per week
☐ 2 - 4 hours per week
☐ 5 - 7 hours per week
☐ More than 7 hours per week

O11 How often do you consume meat on average?

☐ Vegetarian/vegan (no meat)
☐ Occasional (less than 3 times per month)
☐ 1 time per week
☐ 2 - 3 times per week
☐ 4 - 5 times per week
☐ 6 or more times per week

O12 How often do you consume fish on average?

☐ Vegetarian/vegan (no fish)
☐ Occasional (less than 3 times per month)
☐ 1 time per week
☐ 2 - 3 times per week
☐ 4 - 5 times per week
☐ 6 or more times per week

O13 How often do you consume fruits on average?

☐ Never
☐ Occasional (less than 3 times per month)
☐ 1 - 6 times per week
☐ 1 time per day
☐ 2 - 3 times per day
☐ 4 or more times per day

O14 How often do you consume vegetables on average?

☐ Never
☐ Occasional (less than 3 times per month)
☐ 1 - 6 times per week
☐ 1 time per day
☐ 2 - 3 times per day
☐ 4 or more times per day

O15 For how long do you use a mobile phone for calls on average?

☐ No use
☐ Less than 1 hour per month
☐ Around 1 hour per month
☐ 2 - 3 hours per month
☐ Around 1 hour per week
☐ 2 - 6 hours per week
☐ 1 hour per day
☐ More than 1 hour per day

O16 For how long do you use headphones to listen to music on average?

☐ No use
☐ Less than 1 hour per month
☐ Around 1 hour per month
☐ 2 - 3 hours per month
☐ Around 1 hour per week
☐ 2 - 6 hours per week
☐ 1 hour per day
☐ More than 1 hour per day

O17 How many hours do you sleep per day on average?

☐ Less than 6 hours per day
☐ Around 6 hours
☐ Around 7 hours
☐ Around 8 hours
☐ 9 or more hours per day

A8 How many first degree relatives (parents, children, siblings) do you know to have tinnitus or hearing loss?

Please write a number next to each family member.

__ Father __ Mother
__ Brothers __ Sisters
__ Sons __ Daughters

A9 Do you suffer from vertigo (sensation of spinning or tilting)?

☐ Never
☐ Yes, less than one episode per year
☐ Yes, at least one episode per year

A10 Have you been diagnosed with any other ear condition?

You can choose more than one option.

☐ Acoustic trauma (caused by loud sounds)
☐ Ear barotrauma (caused by acute change in ambient pressure)
☐ Presbycusis (aging of ears)
☐ Sudden hearing loss

☐ Other hearing loss
☐ Meniere's disease
☐ Acoustic neuroma (auditory nerve tumour)
☐ Acute otitis (ear inflammation)
☐ Serous otitis or Eustachian tube dysfunction
☐ Chronic otitis (e.g. tympanic perforation, cholesteatoma)
☐ Otosclerosis (reduced ossicles mobility)
☐ Other ear disorders. Please specify _____
☐ No

A11 Have you ever undergone any of the following procedures?
You can choose more than one option.

☐ Ear surgery
☐ Dental surgery
☐ Neurosurgery
☐ Lumbar puncture
☐ Chemotherapy
☐ Head and neck radiotherapy
☐ Electroconvulsive therapy
☐ Other procedure. Please specify_____
☐ None of these

A12 Over the last week, have external sounds been a problem, being too loud or uncomfortable for you when they seemed normal to others around you?
Note: external sounds refer to any sounds other than tinnitus, e.g. environmental sounds, speech, music.

☐ No, not a problem
☐ Yes, a small problem
☐ Yes, a moderate problem
☐ Yes, a big problem
☐ Yes, a very big problem

A13 Do you currently have any other difficulty with your hearing, such as listening to speech in a noisy situation?

☐ Yes, cannot hear at all
☐ Yes, severe difficulty
☐ Yes, moderate difficulty
☐ Yes, slight difficulty
☐ No difficulty
☐ Do not know

A14 Do you use any of the following devices?
You can choose more than one option.

☐ Hearing aid
☐ Cochlear implant

☐ Sound generator
☐ Combination device (hearing aid and sound generator in the same device)
☐ None

A15 Do you suffer from any of the following pain syndromes?
You can choose more than one option.
☐ Headache
☐ Neck pain
☐ Ear pain
☐ Temporomandibular joint pain
☐ Pain in the face
☐ Other. Please specify _____
☐ No

A16 Do you have any of the following conditions that have been diagnosed by a clinician?
You can choose more than one option.
Oral:
☐ Temporomandibular joint disorder
☐ Dental problems
Neurological:
☐ Meningitis
☐ Multiple sclerosis
☐ Epilepsy
☐ Stroke
☐ Other cerebrovascular disease
☐ Dementia
☐ Other neurologic disease
Psychiatric or psychological:
☐ Anxiety
☐ Depression
☐ Emotional trauma
☐ Excessive stress
Sleep disorders:
☐ Difficulty falling asleep
☐ Difficulty staying asleep
Cardiovascular:
☐ Low blood pressure
☐ High blood pressure
☐ Myocardial infraction (heart attack)
Endocrine and metabolic:
☐ Thyroid disorder
☐ Diabetes
☐ Hyperinsulinemia
☐ Increased cholesterol

Rheumatological and immune mediated:
☐ Rheumatoid arthritis
☐ Systemic lupus erythematosus
Otorhinolaryngological:
☐ Chronic sinusitis
☐ Nasal septum deviation
Infectious:
☐ Syphilis
☐ HIV
☐ Lyme disease
Other:
☐ Anaemia
☐ Instability or other balance disorders
☐ Acid/gastroesophageal reflux
☐ Globus hystericus
☐ Other. Please specify _____
☐ None

A17 Tinnitus refers to the perception of noise in your head or ears (such as ringing or buzzing) in the absence of any corresponding source of sound external to your head.
Over the past year, have you had tinnitus in your head or in one or both ears that lasts for more than five minutes at a time?

☐ Yes, most or all of time
☐ Yes, a lot of the time
☐ Yes, some of the time
☐ No, not in the past year
☐ No, never
☐ Do not know

Thank you for completing part A. If you answered "Yes" in question A17, please proceed to Part B. If you answered "No" or “Do not know” in question A17, that is the end of the questionnaire. Thank you for participating in this survey.

PART B. TINNITUS CHARACTERISTICS

Thank you for completing Part A. For the following questions, please give the answer that best describes your tinnitus and its relationship to other conditions. For some questions you can choose more than one option.

B1 How often do you have tinnitus on average?

☐ Daily or almost daily
☐ Almost weekly
☐ Almost monthly
☐ Every few months
☐ Yearly

B2 What best describes your tinnitus during a day?

☐ Constant: you can always or usually hear it in a quiet room

☐ Intermittent: "comes and goes", cannot always hear it in a quiet room

B3 How long ago did your tinnitus appear?

_ _ months

_ _ years

☐ Do not know

B4 Over the past year, how much does your tinnitus worry, annoy or upset you when it is at its worst?

☐ Severely

☐ Moderately

☐ Slightly

☐ Not at all

☐ Do not know

If you answered "Not at all" or “Do not know”, please go to question B6.

B5 How long ago did your tinnitus start bothering you?

_ _ months

_ _ years

☐ Do not know

B6 Although, most patients have tinnitus of a single type, some may hear different sounds. Do you hear one or more different sounds?

☐ One sound ☐ More than one different sound

In case you hear more than one different sound, please try to answer what best describes your most bothersome type of tinnitus in the following questions.

B7 How was the start of your tinnitus?

☐ Gradual ☐ Sudden ☐ Do not know

B8 If you reported any conditions/procedures in questions A9, A10, A11, A12, A13, A15 or A16, please list them here and write next to them if they happened BEFORE, AFTER, or at about the SAME TIME as your tinnitus onset.

_____

_____

_____

_____

B9 Was the initial onset of your tinnitus related to (you can choose more than one option):

☐ Exposure to loud sounds
☐ Change in hearing
☐ Exposure to change in ambient pressure (e.g. flight or diving)
☐ Flu, common cold or other infection
☐ Feeling of fullness or pressure in the ears
☐ Stress
☐ Head trauma
☐ Neck trauma (e.g. whiplash)
☐ Other. Please specify _____
☐ None

B10 Were you taking any of the medicines listed below around the time of your tinnitus onset?

You can choose more than one option.

☐ Aspirin
☐ Pain killing medication. Please specify _____
☐ Oral steroids. Please specify _____
☐ Antibiotics. Please specify _____
☐ Antidepressants. Please specify _____
☐ Quinine (muscle cramps, malaria)
☐ Water tables (diuretics). Please specify _____
☐ Other medicines. Please specify_____
☐ No
☐ Do not know

B11 Do you think any of the conditions mentioned before or any other conditions are related with your tinnitus onset?

You can give up to 3 responses - please choose the most important.

☐ No
☐ Yes. Please specify _____

B12 Is the loudness of your tinnitus stable over time or does it fluctuate over a day?

☐ Stable
☐ Sometimes fluctuating
☐ Always fluctuating
☐ Do not know

B13 What does your tinnitus sound like?

☐ Tonal
☐ Noise-like
☐ Music-like
☐ Crickets

☐ Other. Please specify _____

B14 Please describe the pitch of your tinnitus:

☐ High pitched
☐ Medium pitched
☐ Low pitched
☐ Do not know

B15 Where do you perceive your tinnitus?

☐ Right ear
☐ Left ear
☐ Both ears, worse in right
☐ Both ears, worse in left
☐ Both ears, equally
☐ Inside the head
☐ Other. Please specify _____
☐ Do not know

B16 Is your tinnitus rhythmic?

☐ No
☐ Yes, following heart beat (can be checked by feeling the pulse at the same time as listening to the tinnitus)
☐ Yes, following breathing
☐ Yes, following movements of the head, neck, jaw or muscles of the face
☐ Other. Please specify _____

B17 Has a clinician ever heard your tinnitus?

☐ Yes ☐ No

B18 Is your tinnitus reduced by (you can choose more than one option):

☐ Very quiet environment
☐ Low intensity sounds
☐ High intensity sounds
☐ Head movements
☐ Clenching the teeth or moving the jaw
☐ Pressing your head, neck, or area around the ear
☐ Taking a nap
☐ Good sleep quality
☐ Driving
☐ Being stressed or anxious
☐ Being relaxed
☐ Drinking alcohol

B19 Is your tinnitus increased by (you can choose more than one option):

☐ Very quiet environment
☐ Low intensity sounds
☐ High intensity sounds
☐ Head movements
☐ Clenching the teeth or moving the jaw
☐ Pressing your head, neck, or area around the ear
☐ Taking a nap
☐ Poor sleep quality
☐ Driving

☐ Drinking coffee
☐ Medications
☐ Using hearing aids
☐ Other. Please specify _____
☐ None

☐ Being stressed or anxious
☐ Being relaxed
☐ Drinking alcohol
☐ Drinking coffee
☐ Medications
☐ Using hearing aids
☐ Other. Please specify _____
☐ None

B20 Over the past year, have you seen your family doctor, or seen a healthcare professional at a clinic or hospital about your tinnitus?

☐ Yes, 5 or more visits
☐ Yes, from 2 to 4 visits
☐ Yes, just one visit
☐ Not at all
☐ Do not know

B21 Are you currently receiving any of the following types of management for your tinnitus?

You can choose more than one option.

☐ Psychiatric management
☐ Psychological management
☐ Audiological management
☐ Physiotherapy
☐ Self-management (e.g. dietary supplements, support groups, relaxation)
☐ Other. Please specify _____
☐ No management

B22 Do you think any of the conditions mentioned before, or any other conditions, are related with periods of increased tinnitus?

You can give up to 3 responses - please choose the most important.

☐ No
☐ Yes. Please specify _____

Thank you for participating in this survey.

Appendix 4.1. Details of included variables across the three datasets

Variable	STOP	BRC	ESIT	Description	Code name
Age (y)	TRUE	TRUE	TRUE	Value: Years; Question for STOP and ESIT: 'Age (years)' (ESIT-SQ A1); Question for BRC: 'Age' (TSCHQ 1).	Age
Sex (female/male) (2tschq1esitsq)	TRUE	TRUE	TRUE	Value: male/female; Question for STOP and BRC: 'Gender' (TSCHQ 2); Question for ESIT: 'At birth were you described as:'; Original response options for ESIT: 'Male', 'Female', 'Intersex', 'Prefer not to say'; Recoded to: 'Male', 'Female', since other options were never reported.	Gender_f_m_2tschq_1esitsq
Sex (female/male) (1tschq2esitsq)	TRUE	TRUE	TRUE	Value: male/male; Question for BRC: 'Gender' (TSCHQ 2); Question for STOP and ESIT: 'At birth were you described as:'; Original response options: 'Male', 'Female', 'Intersex', 'Prefer not to say'; Recoded to: 'Male', 'Female', NA, NA.	Gender_f_m_1tschq_2esitsq
Handedness (both/left/right)	TRUE	TRUE	TRUE	Value: both/left/right; Question for STOP: 'Handedness (Which hand do you use to write)?'; Question for BRC: 'Handedness' (TSCHQ 3).; Question for ESIT: 'What is your handedness?' (ESIT-SQ O1).; Response options for ESIT: 'Right', 'Left', 'Both (ambidextrous)', last one recoded to 'Both';	Handedness_2tschq_1esitsq
Family history of tinnitus (or hearing loss for ESIT) (no/yes)	TRUE	TRUE	TRUE	Value: no/yes; Question for STOP: 'Has anyone in your family had tinnitus? (You can answer with more than one alternative)'; Question for BRC: 'Family history of tinnitus complaints' (TSCHQ 4); If checked 'Parents', 'Siblings', or 'Children', variable received the value 'Yes'; Question for ESIT: How many first degree relatives (parents, children, siblings) do you know to have tinnitus or hearing loss? Please write a number next to each family member'; If at least one for 'Parents', 'Siblings', or 'Children', variable received the value 'Yes' (ESIT-SQ A8).	TinnFamilyHist_no_yes

Tinnitus duration (y)	TRUE	TRUE	TRUE	Value: years; Question for STOP and ESIT: 'How long ago did your tinnitus appear?' response in months and years (ESIT-SQ B3); For ESIT and option 'Do not know' was also available and was recoded to NA for this variable; Question for BRC: 'Initial onset: When did you first experience your tinnitus?' reported in years since onset (TSCHQ 5).	TinnitusDuration
Age at tinnitus onset (y)	TRUE	TRUE	TRUE	Value: years; Calculated by subtracting the variable 'Tinnitus duration' from 'Age'.	AgeAtOnset
Tinnitus spatial perception (left ear/both ears, more left/no lateralisation [both ears equally or in the head]/both ears, more right/right ear) (1tschq2esitsq)	TRUE	TRUE	TRUE	Value: left/bothLeft/noLat/bothRight/right; Question for STOP, ESIT and BRC: 'Where do you perceive your tinnitus?' (ESIT-SQ B15, TSCHQ 9); Original response options for ESIT and STOP: 'Right ear', 'Left ear', 'Both ears, worse in right', 'Both ears, worse in left', 'Both ears, equally', 'Inside the head', 'Other. Please specify', 'Do not know'; 'Other. Please specify' and 'Do not know' recoded to NA; 'Both ears, equally' and 'Inside the head' recoded to 'NoLat'; Original response options for BRC: 'Right ear', 'Left ear', 'Both ears, worse in right', 'Both ears, worse in left', 'Both ears, equally', 'Inside the head', 'Elsewhere'; 'Elsewhere' recoded to NA; 'Both ears, equally' and 'Inside the head' recoded to 'NoLat';	TinnLocalisation_1tschq_2esitsq
Tinnitus spatial perception (left ear/both ears, more left/no lateralisation [both ears equally or in the head]/both ears, more right/right ear) (2tschq1esitsq)	TRUE	TRUE	TRUE	See description for 'Tinnitus spatial perception (left ear/both ears, more left/no lateralisation [both ears equally or in the head]/both ears, more right/right ear) (1tschq2esitsq)'	TinnLocalisation_2tschq_1esitsq

Presence during the day (constant/intermittent)	TRUE	TRUE	TRUE	Value: constant/intermittent; Question for STOP and ESIT: 'What best describes your tinnitus during a day?' (ESIT-SQ B2); Original response options: 'Constant: you can always or usually hear it in a quiet room', 'Intermittent: 'comes and goes', cannot always hear it in a quiet room', recoded to 'Constant' and 'Intermittent', Question for BRC: 'How does your tinnitus manifest itself over time? ' (TSCHQ 10).	Const_Int_1tschq_2esitsq
Previous treatments or healthcare visits for tinnitus (no/yes)	TRUE	TRUE	TRUE	Value: no/yes; Question for ESIT and STOP: 'Over the past year, have you seen your family doctor, or seen a healthcare professional at a clinic or hospital about your tinnitus?' (ESIT-SQ B20); Original response options '5 or more visits', '2 - 4 visits', 'one visit', 'not at all', and 'do not know', recoded to 'Yes', 'Yes', 'Yes', 'No', and NA respectively; Question for BRC: 'How many different treatments have you undergone because of your tinnitus?' (TSCHQ 18); Original response options 'None', 'One', 'Several', and 'Many', recoded to 'No', 'Yes', 'Yes', and 'Yes'.	tinnitus_healthcare_no_yes
Tinnitus increased by stress (no/yes)	TRUE	TRUE	TRUE	Value: no/yes; Question for BRC: 'Does stress influence your tinnitus?' (TSCHQ 24); Original response options 'worsens my tinnitus', 'Reduces my tinnitus', and 'has no effect', recoded to 'Yes', 'No, and 'No' respectively; Question for STOP and ESIT: 'Is your tinnitus increased by (you can choose more than one option)', with checkboxes for individual conditions (ESIT-SQ B19).	increased_by_stress
Hearing aid use (no/yes)	TRUE	TRUE	TRUE	Value: no/yes; Question for ESIT and STOP: Do you use any of the following devices? You can choose more than one option.', with checkbox for Hearing aid (ESIT-SQ A14); Question for BRC: 'Do you wear hearing aids?' (TSCHQ 27); Original response options: 'Right', 'Left, 'Both', 'None'; first three recoded to 'Yes', last recoded to 'No'.	HearingAidNoYes

Headaches (no/yes) (2tschq1esitsq)	TRUE	TRUE	TRUE	Value: no/yes; Question for STOP and BRC: 'Do you suffer from headache?' (TSCHQ 7); Original response options for STOP: 'Do not know', 'No', Yes'; 'Do not know' recoded to NA; Question for ESIT: 'Do you suffer from any of the following pain syndromes? (you can choose more than one option)' (ESIT-SQ A15); Response option: 'Headache' checkbox.	Headache_pain_syndrome_2tschq_1esitsq
Vertigo (no/yes) (2tschq1esitsq)	TRUE	TRUE	TRUE	Value: no/yes; Question for STOP and BRC: 'Do you suffer from vertigo or dizziness?' (TSCHQ 31); Original response options for STOP: 'Do not know', 'No', Yes'; 'Do not know' recoded to NA; Question for ESIT: 'Do you suffer from vertigo (sensation of spinning or tilting)?' (ESIT-SQ A9); Original response option: 'Never', 'Yes, less than one episode per year', 'Yes, at least one episode per year'; last two recoded to 'Yes'.	Vertigo_noyes_2tschq_1esitsq
TMJ disorder (no/yes) (2tschq1esitsq)	TRUE	TRUE	TRUE	Value: no/yes; Question for STOP and BRC: 'Do you suffer from temporomandibular disorder? (TSCHQ 32)'; Original response options for STOP: 'Do not know', 'No', Yes'; 'Do not know' recoded to NA; Question for ESIT: Do you have any of the following conditions that have been diagnosed by a clinician? You can choose more than one option' (ESIT-SQ A16); Original response option: 'Temporomandibular joint disorder' checkbox.	tmj_disorder_tschq2_esitsq1
Change in hearing at tinnitus onset (no/yes) (2tschq1esitsq)	TRUE	TRUE	TRUE	Value: no/yes; Question for STOP: 'What do you believe caused your tinnitus?'; Question for BRC: 'Was the initial onset of your tinnitus related to:' (TSCHQ 7); Question for ESIT: 'Was the initial onset of your tinnitus related to (you can choose more than one option)' (ESIT-SQ B9); Response option: 'Change in hearing' checkbox.	onset_change_hearing_2tschq_1esitsq

Stress at tinnitus onset (no/yes) (2tschq1esitsq)	TRUE	TRUE	TRUE	Value: no/yes; Question for STOP: 'What do you believe caused your tinnitus?'; Question for BRC: 'Was the initial onset of your tinnitus related to:' (TSCHQ 7); Question for ESIT: 'Was the initial onset of your tinnitus related to (you can choose more than one option)' (ESIT-SQ B9); Response option: 'Stress' checkbox.	onset_stress_2tschq_1esitsq
Head trauma at tinnitus onset (no/yes)	TRUE	TRUE	TRUE	Value: no/yes; Question for BRC: 'Was the initial onset of your tinnitus related to:' (TSCHQ 7); Question for STOP and ESIT: 'Was the initial onset of your tinnitus related to (you can choose more than one option)' (ESIT-SQ B9); Response option: 'Head trauma' checkbox.	onset_head_trauma_1tschq_2esitsq
Sound exposure at tinnitus onset (no/yes) (2tschq1esitsq)	TRUE	TRUE	TRUE	Value: no/yes; Question for STOP: 'What do you believe caused your tinnitus?'; Question for BRC: 'Was the initial onset of your tinnitus related to:' (TSCHQ 7); Response option: 'loud blast of sound' checkbox; Question for ESIT: 'Was the initial onset of your tinnitus related to (you can choose more than one option)' (ESIT-SQ B9); Response option: 'Exposure to loud sounds' checkbox.	onset_exposure_sounds_2tschq_1esitsq
Infection at tinnitus onset (no/yes)	TRUE	TRUE	TRUE	Value: no/yes; Question for BRC: 'Was the initial onset of your tinnitus related to:' (TSCHQ 7); Response option 'Others', reporting some type of infection; Question for STOP and ESIT: 'Was the initial onset of your tinnitus related to (you can choose more than one option)' (ESIT-SQ B9); Response option: 'Flu, common cold or other infection' checkbox.	onset_infection_1tschq_2esitsq

Rhythmic tinnitus (no/yes, other/yes, with heartbeat) (2tschq1esitsq)	TRUE	TRUE	TRUE	Value: no/yesOther/yesHB; Question for STOP: 'Does your tinnitus seem to pulsate?' (TSCHQ 8); Original response options: 'Don't know', 'Yes, with heart beat', 'Yes, different from heart beat', 'No', recoded to ‘yesOther’, 'YesHB', ‘yesOther’, and 'No' respectively; Question for ESIT: 'Is your tinnitus rhythmic?' (ESIT-SQ B16; Original response options: 'No', 'Yes, following heart beat (can be checked by feeling the pulse at the same time as listening to the tinnitus)', 'Yes, following breathing', 'Yes, following movements of the head, neck, jaw or muscles of the face', 'Other. Please specify' recoded to 'No', YesHB', ‘yesOther’, ‘yesOther’, and ‘yesOther’respectively; Question for BRC: 'Does your tinnitus seem to pulsate?' (TSCHQ 8); Original response options: 'No', 'Yes, with heart beat, 'Yes, different from heart beat', and free text, recoded to 'No', 'YesHB', ‘yesOther’, and ‘yesOther’respectively.	rhythmicity_2tschq_1esitsq_4levels
Rhythmic tinnitus (no/yes, other/yes, with heartbeat) (1tschq2esitsq)	TRUE	TRUE	TRUE	Value: no/yesOther/yesHB; Question for ESIT and STOP: 'Is your tinnitus rhythmic?' (ESIT-SQ B16; Original response options: 'No', 'Yes, following heart beat (can be checked by feeling the pulse at the same time as listening to the tinnitus)', 'Yes, following breathing', 'Yes, following movements of the head, neck, jaw or muscles of the face', 'Other. Please specify' recoded to recoded to 'No', YesHB', 'YesNotHB', ‘yesOther’, and ‘yesOther’ respectively; Question for BRC: 'Does your tinnitus seem to pulsate?' (TSCHQ 8); Original response options: 'No', 'Yes, with heart beat, 'Yes, different from heart beat', and free text, recoded to 'No', 'YesHB', ‘yesOther’, and ‘yesOther’ respectively.	rhythmicity_1tschq_2esitsq_4levels

Presence of tinnitus during the past year (no/yes)	TRUE	TRUE	TRUE	Value: no/yes; Question for STOP and ESIT: 'Do you have any of the following conditions Tinnitus refers to the perception of noise in your head or ears (such as ringing or buzzing) in the absence of any corresponding source of sound external to your head. Over the past year, have you had tinnitus in your head or in one or both ears that lasts for more than five minutes at a time?; Original response options: 'Yes, most or all of time', 'Yes, a lot of the time', 'Yes, some of the time', 'No, not in the past year', 'No, never', 'Do not know'; first three recoded to 'Yes', next two to 'No', last to NA; For BRC all samples were given the value 'Yes' (ESIT-SQ A17).	TinnPastY
Change in hearing at tinnitus onset (no/yes) (1tschq2esitsq)	TRUE	TRUE	TRUE	Value: no/yes; Question for BRC: 'Was the initial onset of your tinnitus related to:' (TSCHQ 7); Question for STOP and ESIT: 'Was the initial onset of your tinnitus related to (you can choose more than one option)' (ESIT-SQ B9); Response option: 'Change in hearing' checkbox.	onset_change_hearing_1tschq_2esitsq
Sound exposure at tinnitus onset (no/yes) (1tschq2esitsq)	TRUE	TRUE	TRUE	Value: no/yes; Question for BRC: 'Was the initial onset of your tinnitus related to:' (TSCHQ 7); Response option: 'loud blast of sound' checkbox; Question for STOP and ESIT: 'Was the initial onset of your tinnitus related to (you can choose more than one option)' (ESIT-SQ B9); Response option: 'Exposure to loud sounds' checkbox.	onset_exposure_sounds_1tschq_2esitsq
Stress at tinnitus onset (no/yes) (1tschq2esitsq)	TRUE	TRUE	TRUE	Value: no/yes; Question for BRC: 'Was the initial onset of your tinnitus related to:' (TSCHQ 7); Question for STOP and ESIT: 'Was the initial onset of your tinnitus related to (you can choose more than one option)' (ESIT-SQ B9); Response option: 'Stress' checkbox.	onset_stress_1tschq_2esitsq
Headaches (no/yes) (1tschq2esitsq)	TRUE	TRUE	TRUE	Value: no/yes; Question for BRC: 'Do you suffer from headache?' (TSCHQ 7); Question for STOP and ESIT: 'Do you suffer from any of the following pain syndromes? (you can choose more than one option)' (ESIT-SQ A15); Response option: 'Headache' checkbox.	Headache_pain_syndrome_1tschq_2esitsq

Vertigo (no/yes) (1tschq2esitsq)	TRUE	TRUE	TRUE	Value: no/yes; Question for BRC: 'Do you suffer from vertigo or dizziness?' (TSCHQ 31); Question for STOP and ESIT: 'Do you suffer from vertigo (sensation of spinning or tilting)?' (ESIT-SQ A9); Original response option: 'Never', 'Yes, less than one episode per year', 'Yes, at least one episode per year'; last two recoded to 'Yes'.	Vertigo_noyes_1tschq_2esitsq
TMJ disorder (no/yes) (1tschq2esitsq)	TRUE	TRUE	TRUE	Value: no/yes; Question for BRC: 'Do you suffer from temporomandibular disorder? (TSCHQ 32)'; Question for STOP and ESIT: Do you have any of the following conditions that have been diagnosed by a clinician? You can choose more than one option' (ESIT-SQ A16); Original response option: 'Temporomandibular joint disorder' checkbox.	tmj_disorder_tschq1_esitsq2
Hearing threshold at 0.125 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	Value: dB HL; Measure for STOP: Fixed frequency Bekesy audiometry was done using the Astera 2 audiometer (Otometrics) and Sennheiser HDA 200 headphones; Measure for BRC: Pure tone audiometry was conducted manually by an examiner using a Siemens Unity 2 system and Sennheiser HDA 200 headphones.	LAudio0125
Hearing threshold at 0.25 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	LAudio025
Hearing threshold at 0.5 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	LAudio05
Hearing threshold at 0.75 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	Left750
Hearing threshold at 1 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	LAudio1
Hearing threshold at 1.5 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	Left1500

Hearing threshold at 2 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	LAudio2
Hearing threshold at 3 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	LAudio3
Hearing threshold at 4 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	LAudio4
Hearing threshold at 6 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	LAudio6
Hearing threshold at 8 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	LAudio8
Hearing threshold at 10 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	LAudio10
Hearing threshold at 12.5 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	LAudio125
Hearing threshold at 14 kHz (left ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	LAudio140
Hearing threshold at 0.125 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	RAudio0125
Hearing threshold at 0.25 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	RAudio025
Hearing threshold at 0.5 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	RAudio05
Hearing threshold at 0.75 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	Right750
Hearing threshold at 1 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	RAudio1
Hearing threshold at 1.5 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	Right1500

Hearing threshold at 2 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	RAudio2
Hearing threshold at 3 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	RAudio3
Hearing threshold at 4 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	RAudio4
Hearing threshold at 6 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	RAudio6
Hearing threshold at 8 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	RAudio8
Hearing threshold at 10 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	RAudio10
Hearing threshold at 12.5 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	RAudio125
Hearing threshold at 14 kHz (right ear) (dB HL)	TRUE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	RAudio140
Varying tinnitus loudness from day to day (no/yes)	TRUE	TRUE	FALSE	Value: no/yes; Question for BRC and STOP: 'Does the LOUDNESS of the tinnitus vary from day to day?' (TSCHQ 11).	LoudnessVar_fromdaytoday
Tinnitus annoyance scale (0-100)	TRUE	TRUE	FALSE	Value: 0-100; Question for BRC and STOP: 'What percent of your total awake time, over the last month, have you been annoyed, distressed, or irritated of your tinnitus? Please write a single number between 1 and 100' (TSCHQ 17).	TinnAnnoyance_scale100
Tinnitus worsened by loud noise (no/yes)	TRUE	TRUE	FALSE	Value: no/yes; Question for BRC and STOP: 'The presence of loud noise make your tinnitus worse?' (TSCHQ 20); Original response options: 'Yes', 'No', 'I don't know'; last recoded to NA.	NoiseWorsens_tschq
Self-reported hearing problem (no/yes)	TRUE	TRUE	FALSE	Value: no/yes; Question for BRC and STOP: 'Do you think you have a hearing problem?' (TSCHQ 26).	subjectiveHearingProb_tschq

Hyperacusis questionnaire score (0-42)	TRUE	TRUE	FALSE	Value: score 0-42; Hyperacusis questionnaire (Khalfa et al. 2002).	Hyperacusis
Tinnitus loudness rating (0-100)	TRUE	TRUE	FALSE	Value: score 0/1-100; For STOP numeric rating scale: Describe the loudness of your tinnitus using a scale of 1-100, 1 = Very faint; 100 = Very loud (TSCHQ 12); For BRC: Visual analog scale. Participants were asked to rate the perceived loudness of their tinnitus using the dial to select a position on a Borg CR100 scale with the following quasi-logarithmic anchors: 0 'extremely weak', 30 'moderate', 50 'strong', 70 'very strong, and 100 'extremely strong'.	LoundnessRating
Tinnitus pitch matching (kHz)	TRUE	TRUE	FALSE	Value: kHz; Measure for STOP: Two alternative forced choice procedure range 0.18-16 kHz; Measure for BRC: Frequency with highest tinnitus likeness rating from tinnitus tester procedure assessing 0.5-12 kHz (Roberts, Moffat et al. 2006).	TinnitusPitch
MaxDiffExt (dB)	TRUE	TRUE	FALSE	Value: dB; Calculated as the maximum mean interaural threshold difference (right ear minus left ear) of two adjacent frequencies (including thresholds at the frequency with the maximum interaural difference) as in Tsai, Sweetow et al. (2012), spanning the range of thresholds at 0.125, 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 4, 6 and 8 kHz, and including the mean difference from the available extended high frequencies (10, 12.5, and 14 kHz for the STOP dataset and 9, 10, 11.2, 12.5, and 14kHz for the BRC dataset). Thresholds at 0.75 and 1.5 kHz were not available for the STOP dataset and were calculated as the mean of the adjacent frequencies.	TSAIextasymAll Low
Mean hearing threshold both ears (dB HL)	TRUE	TRUE	FALSE	Value: dB HL; Calculated as the mean hearing loss from both ears at 0.125, 0.25, 0.5, 1, 2, 4, and 8 kHz.	MeanAllOct
Height (cm)	TRUE	FALSE	TRUE	Value: cm; Question for STOP and ESIT: What is your height?' (ESIT-SQ A3).	Height

Weight (kg)	TRUE	FALSE	TRUE	Value: kg; Question for STOP and ESIT:' What is your weight?' (ESIT-SQ A3).	Weight
Education (lower/higher)	TRUE	FALSE	TRUE	Value: lower/higher; Question for ESIT and STOP: 'What is the highest education level you have achieved?' (ESIT-SQ A5); Values other than 'University or higher' recoded to 'Lower'.	Education
Alcohol (number of drinks per week)	TRUE	FALSE	TRUE	Value: number of drinks; Question for STOP and ESIT: What is the average number of alcoholic drinks that you consume per week? 1 drink equals 125 ml glass of wine, 330 ml of beer or 40 ml of spirits' (ESIT-SQ A6).	Alcohol
Smoking (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question for ESIT and STOP: Which of the following options best describes your smoking status?' (ESIT-SQ A7).; 'Never smoker' recoded to 'No', 'Current Smoker', or 'Ex-smoker' recoded to 'Yes'	smoking_status
Any ear condition (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question: 'Have you been diagnosed with any other ear condition? You can choose more than one option.', with checkboxes for individual conditions (ESIT-SQ A10).	any_ear_condition
Acoustic trauma (no/yes)	TRUE	FALSE	TRUE	See description for 'Any ear condition (no/yes)'	acoustic_trauma
Barotrauma (no/yes)	TRUE	FALSE	TRUE	See description for 'Any ear condition (no/yes)'	barotrauma
Presbycusis (no/yes)	TRUE	FALSE	TRUE	See description for 'Any ear condition (no/yes)'	presbycusis
Sudden hearing loss (no/yes)	TRUE	FALSE	TRUE	See description for 'Any ear condition (no/yes)'	sudden_hl
Other hearing loss (no/yes)	TRUE	FALSE	TRUE	See description for 'Any ear condition (no/yes)'	other_hl
Meniere's disease (no/yes)	TRUE	FALSE	TRUE	See description for 'Any ear condition (no/yes)'	meniere
Acoustic neuroma (no/yes)	TRUE	FALSE	TRUE	See description for 'Any ear condition (no/yes)'	neuroma
Acute Otitis (no/yes)	TRUE	FALSE	TRUE	See description for 'Any ear condition (no/yes)'	acute_otitis
Serous Otitis (no/yes)	TRUE	FALSE	TRUE	See description for 'Any ear condition (no/yes)'	serous_otitis
Chronic Otitis (no/yes)	TRUE	FALSE	TRUE	See description for 'Any ear condition (no/yes)'	chronic_otitis
Otosclerosis (no/yes)	TRUE	FALSE	TRUE	See description for 'Any ear condition (no/yes)'	otosclerosis

Any procedure (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question: 'Have you ever undergone any of the following procedures? You can choose more than one option.', with checkboxes for individual conditions (ESIT-SQ A11).	any_procedure
Ear surgery (no/yes)	TRUE	FALSE	TRUE	See description for 'Any procedure (no/yes)'	ear_surg
Dental surgery (no/yes)	TRUE	FALSE	TRUE	See description for 'Any procedure (no/yes)'	dental_surg
Neurosurgery (no/yes)	TRUE	FALSE	TRUE	See description for 'Any procedure (no/yes)'	neurosurg
Lumbar puncture (no/yes)	TRUE	FALSE	TRUE	See description for 'Any procedure (no/yes)'	lumbar_punct
Chemotherapy (no/yes)	TRUE	FALSE	TRUE	See description for 'Any procedure (no/yes)'	chemo
Head and neck radiotherapy (no/yes)	TRUE	FALSE	TRUE	See description for 'Any procedure (no/yes)'	head_neck_radio
Problem with external sounds (small/moderate/big or very big)	TRUE	FALSE	TRUE	Value: small/moderate/big; Question for ESIT and STOP: 'Over the last week, have external sounds been a problem, being too loud or uncomfortable for you when they seemed normal to others around you? Note: external sounds refer to any sounds other than tinnitus, e.g. environmental sounds, speech, music.' (ESIT-SQ A12).; 'No' recoded to 'small', 'very big' recoded to 'big'	Ext_sounds_prob
Self-reported hearing difficulty (slight or no difficulty/moderate difficulty/severe difficulty/total loss)	TRUE	FALSE	TRUE	Value: light/moderate/severe/total loss; Question for ESIT and STOP: 'Do you currently have any other difficulty with your hearing, such as listening to speech in a noisy situation?' (ESIT-SQ A13); 'no difficulty' grouped with 'light difficulty', 'not know' recoded to NA.	hearing_difficulty_esitsq
Hearing device (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question: 'Do you use any of the following devices? - Any' (ESIT-SQ A14).	any_devices
Combination device (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question: 'Do you use any of the following devices?', with checkbox for Combination device (ESIT-SQ A14).	combination
Any pain syndromes (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question: 'Do you suffer from any of the following pain syndromes? You can choose more than one option.', with checkboxes for individual conditions (ESIT-SQ A15).	any_pain_syndromes
Neck pain (no/yes)	TRUE	FALSE	TRUE	See description for 'Any pain syndromes (no/yes)'	neck_pain

Ear pain (no/yes)	TRUE	FALSE	TRUE	See description for 'Any pain syndromes (no/yes)'	ear_pain
TMJ pain (no/yes)	TRUE	FALSE	TRUE	See description for 'Any pain syndromes (no/yes)'	tmj_pain
Face pain (no/yes)	TRUE	FALSE	TRUE	See description for 'Any pain syndromes (no/yes)'	face_pain
Any diagnosed condition (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question: 'Do you have any of the following conditions that have been diagnosed by a clinician? You can choose more than one option.', with checkboxes for individual conditions (ESIT-SQ A16).	diagnosed_condition_any
Dental problems (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	dental_problems
Meningitis (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	meningitis
Epilepsy (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	epilpsy
Stroke (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	stroke
Other neurologic condition (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	other_neurologic
Anxiety (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	anxiety_presence
Depression (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	depression_presence
Emotional trauma (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	emot_trauma
Stress (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	stress_presence_esitsq
Problem falling asleep (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	prob_falling_asleep
Problem staying asleep (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	prob_staying_asleep
Low BP (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	low_bp
High BP (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	high_bp
Thyroid disorder (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	thyroid_disorder
Diabetes (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	diabetes
High cholesterol (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	high_cholest

Rheumatoid arthritis (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	rheum_arthritis
Chronic sinusitis (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	chr_sinusitis
Nasal septum deviation (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	sept_deviation
HIV (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	hiv
Lyme disease (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	lyme
Anaemia (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	anaemia
Balance disorders (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	balance
Acid/gastroesophageal reflux (no/yes)	TRUE	FALSE	TRUE	See description for 'Any diagnosed condition (no/yes)'	acid_reflux
Tinnitus daily presence (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question for ESIT and STOP: 'How often do you have tinnitus on average?' (ESIT-SQ B1); All values other than 'daily' recoded to 'No'.	frequency
Tinnitus worries, annoys or upsets (severely/moderately/slightly /not at all)	TRUE	FALSE	TRUE	Value: severely/moderately/slightly/no; Question for ESIT and STOP: 'Over the past year, how much does your tinnitus worry, annoy or upset you when it is at its worst?' (ESIT-SQ B4); 'Do not know' recoded to NA.	bother_esitsq
Number of sounds (more than one/one)	TRUE	FALSE	TRUE	Value: more/one; Question for ESIT and STOP: 'Although, most patients have tinnitus of a single type, some may hear different sounds. Do you hear one or more different sounds?' (ESIT-SQ B6).	number_sounds
Change in hearing at tinnitus onset (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question: 'Was the initial onset of your tinnitus related to (you can choose more than one option)', with checkboxes for individual conditions (ESIT-SQ B9).	onset_exposure_pressure
Fullness in the ears at tinnitus onset (no/yes)	TRUE	FALSE	TRUE	See description for 'Change in hearing at tinnitus onset (no/yes)'	onset_fullness
Neck trauma at tinnitus onset (no/yes)	TRUE	FALSE	TRUE	See description for 'Change in hearing at tinnitus onset (no/yes)'	onset_neck_trauma

No event at tinnitus onset (no/yes)	TRUE	FALSE	TRUE	See description for 'Change in hearing at tinnitus onset (no/yes)'	onset_no_onset_event
Aspirin at tinnitus onset (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question: 'Were you taking any of the medicines listed below around the time of your tinnitus onset? You can choose more than one option. ', with checkboxes for individual medicines (ESIT-SQ B10).	onset_aspirin
Painkillers at tinnitus onset (no/yes)	TRUE	FALSE	TRUE	See description for 'Aspirin at tinnitus onset (no/yes)'	onset_painkillers
Antibiotics at tinnitus onset (no/yes)	TRUE	FALSE	TRUE	See description for 'Aspirin at tinnitus onset (no/yes)'	onset_antibiotics
Antidepressants at tinnitus onset (no/yes)	TRUE	FALSE	TRUE	See description for 'Aspirin at tinnitus onset (no/yes)'	onset_antidepressants
Quinine at tinnitus onset (no/yes)	TRUE	FALSE	TRUE	See description for 'Aspirin at tinnitus onset (no/yes)'	onset_quinine
Diuretics at tinnitus onset (no/yes)	TRUE	FALSE	TRUE	See description for 'Aspirin at tinnitus onset (no/yes)'	onset_diuretics
No medication at tinnitus onset (no/yes)	TRUE	FALSE	TRUE	See description for 'Aspirin at tinnitus onset (no/yes)'	onset_no_medication_at_onset
Thoughts of conditions related to tinnitus onset (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question: 'Do you think any of the conditions mentioned before or any other conditions are related with your tinnitus onset? - Yes' (ESIT-SQ B11).	thoughts_of_onset_related_condition
Varying tinnitus loudness over a day (stable/sometimes fluctuating/always fluctuating)	TRUE	FALSE	TRUE	Value: stable/sometimes/always; Question for ESIT and STOP: 'Is the loudness of your tinnitus stable over time or does it fluctuate over a day?' (ESIT-SQ B12); 'Do not know' recoded to NA.	loudness_changes_overday_ordinal

Tinnitus quality (tonal/noise/other)	TRUE	FALSE	TRUE	Value: tonal/noise/other; Question for ESIT and STOP: 'What does your tinnitus sound like?' (ESIT-SQ B13); Values 'music' and 'crickets' recoded to 'other'.	quality
Tinnitus pitch (high/medium/low)	TRUE	FALSE	TRUE	Value: high/medium/low; Question for ESIT and STOP: 'Please describe the pitch of your tinnitus:' (ESIT-SQ B14); 'Do not know' recoded to NA.	pitch
Objective tinnitus (yes/no)	TRUE	FALSE	TRUE	Value: yes/no; Question for ESIT and STOP: 'Has a clinician ever heard your tinnitus?' (ESIT-SQ B17).	objective_esitsq
Tinnitus reduced by silence (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question: 'Is your tinnitus reduced by (you can choose more than one option)', with checkboxes for individual conditions (ESIT-SQ B18).	reduced_by_silence
Tinnitus reduced by low intensity sounds (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_low_sounds
Tinnitus reduced by high intensity sounds (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_high_sounds
Tinnitus reduced by head movement (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_head_movements
Tinnitus reduced by jaw movement (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_jaw_movements
Tinnitus reduced by pressing head, neck, or area around the ear (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_head_neck_press
Tinnitus reduced by taking a nap (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_nap
Tinnitus reduced by good quality sleep (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_good_sleep

Tinnitus reduced by driving (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_driving
Tinnitus reduced by stress (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_stress
Tinnitus reduced by being relaxed (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_relaxed
Tinnitus reduced by drinking alcohol (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_alcohol
Tinnitus reduced by taking medication (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_medications
Tinnitus reduced by using hearing aids (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_hearing_aid
Tinnitus reduced by nothing (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus reduced by silence (no/yes)'	reduced_by_nothing
Tinnitus increased by silence (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question: 'Is your tinnitus increased by (you can choose more than one option)', with checkboxes for individual conditions (ESIT-SQ B19).	increased_by_silence
Tinnitus increased by low intensity sounds (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_low_sounds
Tinnitus increased by high intensity sounds (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_high_sounds
Tinnitus increased by head movement (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_head_movements
Tinnitus increased by jaw movement (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_jaw_movements

Tinnitus increased by pressing head, neck, or area around the ear (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_head_neck_press
Tinnitus increased by taking a nap (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_nap
Tinnitus increased by poor quality sleep (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_poor_sleep
Tinnitus increased by driving (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_driving
Tinnitus increased by being relaxed (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_relaxed
Tinnitus increased by drinking alcohol (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_alcohol
Tinnitus increased by drinking coffee (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_coffee
Tinnitus increased by taking medication (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_medications
Tinnitus increased by using hearing aids (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_hearing_aid
Tinnitus increased by nothing (no/yes)	TRUE	FALSE	TRUE	See description for 'Tinnitus increased by silence (no/yes)'	increased_by_nothing
Psychological management for tinnitus (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question: 'Are you currently receiving any of the following types of management for your tinnitus? You can choose more than one option.', with checkboxes for individual conditions (ESIT-SQ B21).	tinnitus_management_psychological
Audiological management for tinnitus (no/yes)	TRUE	FALSE	TRUE	See description for 'Psychological management for tinnitus (no/yes)'	tinnitus_management_audiological

Physiotherapy for tinnitus (no/yes)	TRUE	FALSE	TRUE	See description for 'Psychological management for tinnitus (no/yes)'	tinnitus_management_physiotherapy
Self-management for tinnitus (no/yes)	TRUE	FALSE	TRUE	See description for 'Psychological management for tinnitus (no/yes)'	tinnitus_management_self
No management for tinnitus (no/yes)	TRUE	FALSE	TRUE	See description for 'Psychological management for tinnitus (no/yes)'	tinnitus_management_none
Thoughts of conditions related to increased tinnitus (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question: 'Do you think any of the conditions mentioned before, or any other conditions, are related with periods of increased tinnitus? - Yes' (ESIT-SQ B22).	thoughts_of_increasing_tinnitus_condition
Night work (no/yes)	TRUE	FALSE	TRUE	Value: no/yes; Question for ESIT and STOP: 'Have you ever worked at night (i.e. between 24:00-5:00)?' (ESIT-SQ O7); 'Currently' and 'Previously' recoded to 'Yes', 'Do not know' recoded to NA.	night_work
BMI (kg/m^2)	TRUE	FALSE	TRUE	Value: kg/m^2; Calculated by dividing weight (kg) by height (m) squared.	BMI
Age at bothersome tinnitus onset (y)	TRUE	FALSE	TRUE	Value: years; Calculated by subtracting the variable 'Bothersome tinnitus duration' from 'Age'.	AgeBothOnset
Bothersome tinnitus duration (y)	TRUE	FALSE	TRUE	Value: years; Question for STOP and ESIT: 'How long ago did your tinnitus start bothering you?' response in months and years (ESIT-SQ B5); For ESIT and option 'Do not know' was also available and was recoded to NA for this variable.	BothTinnitusDuration
Gap between onset of tinnitus and bothersome tinnitus (y)	TRUE	FALSE	TRUE	Value: years; Calculated by subtracting the variable 'Bothersome tinnitus duration' from 'Tinnitus duration.	Gap_Tinn_BothTinn

Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	Value: no/before/same_time/after; Question for STOP and ESIT for all previously reported conditions/procedures: If you reported any conditions/procedures in questions A9, A10, A11, A12, A13, A15 or A16, please list them here and write next to them if they happened BEFORE, AFTER, or at about the SAME TIME as your tinnitus onset. (ESIT-SQ 8).	temp_rel_vertigo
Onset of problems with external sounds in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_ext_sounds_prob
Onset of hearing difficulties in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_hearing_difficulty
Onset of acoustic trauma in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_acoustic_trauma
Onset of barotrauma in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_barotrauma
Onset of presbycusis in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_presbycusis

Onset of sudden hearing loss in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_sudden_hl
Onset of other hearing loss in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_other_hl
Onset of Meniere's in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_meniere
Onset of acoustic neuroma in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_neuroma
Onset of acute otitis in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_acute_otitis
Onset of serous otitis in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_serous_otitis

Onset of chronic otitis in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_chronic_otitis
Onset of otosclerosis in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_otosclerosis
Onset of ear surgery in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_ear_surg
Onset of dental surgery in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_dental_surg
Onset of neurosurgery in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_neurosurg
Onset of lumbar puncture in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_lumbar_punct
Onset of chemotherapy in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_chemo

Onset of head and neck radiotherapy in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_head_neck_radio
Onset of headache in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_headache
Onset of neck pain in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_neck_pain
Onset of ear pain in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_ear_pain
Onset of TMJ pain in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_tmj_pain
Onset of face pain in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_face_pain

Onset of TMJ disorder in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_tmj_disorder
Onset of dental problems in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_dental
Onset of meningitis in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_meningitis
Onset of epilepsy in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_epilpsy
Onset of stroke in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_stroke
Onset of other neurological condition in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_other_neurologic

Onset of anxiety in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_anxiety
Onset of depression in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_depression
Onset of emotional trauma in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_emot_trauma
Onset of stress in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_stress
Onset of problems falling asleep in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_prob_falling_asleep
Onset of problems staying asleep in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_prob_staying_asleep

Onset of low BP in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_low_bp
Onset of high BP in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_high_bp
Onset of thyroid disorder in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_thyroid
Onset of diabetes in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_diabetes
Onset of increased cholesterol in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_high_cholest
Onset of rheumatoid arthritis in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_rheum_arthritis

Onset of chronic sinusitis in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_chr_sinusitis
Onset of septal deviation in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_sept_deviation
Onset of HIV in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_hiv
Onset of Lyme disease in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_lyme
Onset of anaemia in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_anaemia
Onset of balance disorders in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_balance

Onset of acid/gastroesophageal reflux in relation to tinnitus onset (never/before/at the same time/after)	TRUE	FALSE	TRUE	See description for 'Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)'	temp_rel_cid_reflux
Hearing threshold at 9 kHz (left ear) (dB HL)	FALSE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	Left9000
Hearing threshold at 11.2 kHz (left ear) (dB HL)	FALSE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	Left11200
Hearing threshold at 9 kHz (right ear) (dB HL)	FALSE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	Right9000
Hearing threshold at 11.2 kHz (right ear) (dB HL)	FALSE	TRUE	FALSE	See description for 'Hearing threshold at 0.125 kHz (left ear) (dB HL)'	Right11200
Tinnitus reduced by music or environmental sounds (no/yes)	FALSE	TRUE	FALSE	Value: no/yes; Question: 'Is your tinnitus reduced by music or by certain types of environmental sounds such as the noise of a waterfall or the noise of running water when you are standing in the shower?' (TSCHQ 19).	SoundReduces
Tinnitus affected by head or neck movement or arms/hands or head touch (no/yes)	FALSE	TRUE	FALSE	Value: no/yes; Question: 'Does any head and neck movement (e.g. moving the jaw forward or clenching the teeth), or having your arms/hands or head touched, affect your tinnitus?' (TSCHQ 21).	MovementAffects
Tinnitus during the day affected by sleep at night (no/yes)	FALSE	TRUE	FALSE	Value: no/yes; Question: 'Is there any relationship between sleep at night and your tinnitus during the day?' (TSCHQ 23).	SleepAffects
Beck Depression Inventory (0-63)	FALSE	TRUE	FALSE	Value: score 0-63; Beck Depression Inventory (Beck et al. 1988).	BDI

Beck Anxiety Inventory (0-63)	FALSE	TRUE	FALSE	Value: score 0-63; Beck Anxiety Inventory (Steer and Beck 1997).	BAI
Tinnitus Handicap Questionnaire total score (0-100)	FALSE	TRUE	FALSE	Value: 0-100 score; Tinnitus Handicap Questionnaire (Kuk et al. 1990).	THQTotal
Comfortable level of a 0.5 kHz pure tone (dB SPL)	FALSE	TRUE	FALSE	Psychoacoustic tinnitus assessment using Tinnitus Tester procedure as described in Roberts et al. (2006) and Roberts et al. (2008).	TinnTestComf05kHz
Comfortable level of a 5 kHz pure tone (dB SPL)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestComf5kHz
Tinnitus loudness matching at 0.5 kHz (dB SPL)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestLoudMatch0.5k
Tinnitus loudness matching at 1 kHz (dB SPL)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestLoudMatch1k
Tinnitus loudness matching at 2 kHz (dB SPL)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestLoudMatch2k
Tinnitus loudness matching at 3 kHz (dB SPL)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestLoudMatch3k
Tinnitus loudness matching at 4 kHz (dB SPL)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestLoudMatch4k
Tinnitus loudness matching at 5 kHz (dB SPL)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestLoudMatch5k
Tinnitus loudness matching at 6 kHz (dB SPL)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestLoudMatch6k
Tinnitus loudness matching at 7 kHz (dB SPL)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestLoudMatch7k
Tinnitus loudness matching at 8 kHz (dB SPL)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestLoudMatch8k

Tinnitus loudness matching at 10 kHz (dB SPL)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestLoudMatch10k
Tinnitus pitch matching at 0.5 kHz (likeness scale 0-100)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestPitchMatch0.5k
Tinnitus pitch matching at 1 kHz (likeness scale 0-100)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestPitchMatch1k
Tinnitus pitch matching at 2 kHz (likeness scale 0-100)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestPitchMatch2k
Tinnitus pitch matching at 3 kHz (likeness scale 0-100)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestPitchMatch3k
Tinnitus pitch matching at 4 kHz (likeness scale 0-100)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestPitchMatch4k
Tinnitus pitch matching at 5 kHz (likeness scale 0-100)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestPitchMatch5k
Tinnitus pitch matching at 6 kHz (likeness scale 0-100)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestPitchMatch6k
Tinnitus pitch matching at 7 kHz (likeness scale 0-100)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestPitchMatch7k

Tinnitus pitch matching at 8 kHz (likeness scale 0-100)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestPitchMatch8k
Tinnitus pitch matching at 10 kHz (likeness scale 0-100)	FALSE	TRUE	FALSE	See description for 'Comfortable level of a 0.5 kHz pure tone (dB SPL)'	TinnTestPitchMatch10k
Wax removal at tinnitus onset (no/yes)	FALSE	TRUE	FALSE	Value: no/yes; Question: 'Was the initial onset of your tinnitus related to:', with checkboxes for individual conditions (TSCHQ 7).	WaxRemovalOnset
Other (than loud blast of sound, whiplash, change in hearing, stress, or head trauma) event at tinnitus onset (no/yes)	FALSE	TRUE	FALSE	See description for 'Wax removal at tinnitus onset (no/yes)'	OtherOnset
Abnormal tympanogram left ear (no/yes)	TRUE	FALSE	FALSE	Immittance tympanometry for left and right ear using Otoflex 100 and OTOsuite; Type takes values A, AD, AS, B, C depending on the shape of the response as defined by an audiologist; A was considered as normal tympanogram and all other values as abnormal; TPP: Tympanometric Peak Pressure in daPA; SA: Static Admittance in mmho; TW: Tympanometric Width in daPa; ECV: Ear Canal Volume in cm3	Type_L
Tympanometric Peak Pressure (TPP) left ear (daPA)	TRUE	FALSE	FALSE	See description for 'Abnormal tympanogram left ear (no/yes)'	TPP_L
Static Admittance (SA) left ear (mmho)	TRUE	FALSE	FALSE	See description for 'Abnormal tympanogram left ear (no/yes)'	SA_L
Tympanometric Width (TW) left ear (daPa)	TRUE	FALSE	FALSE	See description for 'Abnormal tympanogram left ear (no/yes)'	TW_L
Ear Canal Volume (ECVL) left ear (cm3)	TRUE	FALSE	FALSE	See description for 'Abnormal tympanogram left ear (no/yes)'	ECV_L

Abnormal tympanogram right ear (no/yes)	TRUE	FALSE	FALSE	See description for 'Abnormal tympanogram left ear (no/yes)'	Type_R
Tympanometric Peak Pressure (TPP) right ear (daPA)	TRUE	FALSE	FALSE	See description for 'Abnormal tympanogram left ear (no/yes)'	TPP_R
Static Admittance (SA) right ear (mmho)	TRUE	FALSE	FALSE	See description for 'Abnormal tympanogram left ear (no/yes)'	SA_R
Tympanometric Width (TW) right ear (daPa)	TRUE	FALSE	FALSE	See description for 'Abnormal tympanogram left ear (no/yes)'	TW_R
Ear Canal Volume (ECVL) right ear (cm3)	TRUE	FALSE	FALSE	See description for 'Abnormal tympanogram left ear (no/yes)'	ECV_R
Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)	TRUE	FALSE	FALSE	Value: dB HL; Loudness Discomfort Level were measured OTOsuite with the following process: Starting at 1) 60dB HL, 2) 15dB below previous LDL result or 3) 5dB above pure tone threshold increasing level in 5dB steps. Store threshold when the patient reports uncomfortable loudness or no-response threshold at 100 dB HL. Pure tone presentation 1.5 seconds at audiometric frequencies used for threshold measurements.	U_125_L
Uncomfortable loudness level at 0.25 kHz (left ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_250_L
Uncomfortable loudness level at 0.5 kHz (left ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_500_L
Uncomfortable loudness level at 1 kHz (left ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_1000_L

Uncomfortable loudness level at 2 kHz (left ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_2000_L
Uncomfortable loudness level at 3 kHz (left ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_3000_L
Uncomfortable loudness level at 4 kHz (left ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_4000_L
Uncomfortable loudness level at 6 kHz (left ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_6000_L
Uncomfortable loudness level at 8 kHz (left ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_8000_L
Uncomfortable loudness level at 0.125 kHz (right ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_125_R
Uncomfortable loudness level at 0.25 kHz (right ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_250_R
Uncomfortable loudness level at 0.5 kHz (right ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_500_R
Uncomfortable loudness level at 1 kHz (right ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_1000_R

Uncomfortable loudness level at 2 kHz (right ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_2000_R
Uncomfortable loudness level at 3 kHz (right ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_3000_R
Uncomfortable loudness level at 4 kHz (right ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_4000_R
Uncomfortable loudness level at 6 kHz (right ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_6000_R
Uncomfortable loudness level at 8 kHz (right ear) (dB HL)	TRUE	FALSE	FALSE	See description for 'Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)'	U_8000_R
Speech in noise hearing assessment - word score (left ear) (0-100)	TRUE	FALSE	FALSE	Speech in Noise audiometry using OTOsuite for the left and right ear. WS (word) and PS (phonemes) scores take values from 0-100. Level refers to the sound level of the presented words in dB above the hearing threshold.	WS_L
Speech in noise hearing assessment - phoneme score (left ear) (0-100)	TRUE	FALSE	FALSE	See description for 'Speech in noise hearing assessment - word score (left ear) (0-100)'	PS_L
Speech in noise hearing assessment - word score (right) (0-100)	TRUE	FALSE	FALSE	See description for 'Speech in noise hearing assessment - word score (left ear) (0-100)'	WS_R
Speech in noise hearing assessment - phoneme score (right ear) (0-100)	TRUE	FALSE	FALSE	See description for 'Speech in noise hearing assessment - word score (left ear) (0-100)'	PS_R

DPOAE at 996 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	Value: signal to noise ratio; Distortion Product Otoacoustic Emissions using OTOsuite.	X996_L
DPOAE at 1074 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1074_L
DPOAE at 1152 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1152_L
DPOAE at 1230 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1230_L
DPOAE at 1318 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1318_L
DPOAE at 1416 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1416_L
DPOAE at 1513 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1513_L
DPOAE at 1621 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1621_L
DPOAE at 1738 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1738_L
DPOAE at 1865 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1865_L
DPOAE at 2001 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X2001_L
DPOAE at 2148 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X2148_L
DPOAE at 2294 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X2294_L
DPOAE at 2460 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X2460_L

DPOAE at 2636 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X2636_L
DPOAE at 2832 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X2832_L
DPOAE at 3027 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X3027_L
DPOAE at 3251 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X3251_L
DPOAE at 3486 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X3486_L
DPOAE at 3730 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X3730_L
DPOAE at 4003 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X4003_L
DPOAE at 4287 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X4287_L
DPOAE at 4599 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X4599_L
DPOAE at 4921 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X4921_L
DPOAE at 5273 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X5273_L
DPOAE at 5654 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X5654_L
DPOAE at 6064 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X6064_L
DPOAE at 6494 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X6494_L

DPOAE at 6962 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X6962_L
DPOAE at 7460 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X7460_L
DPOAE at 7998 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X7998_L
DPOAE at 8574 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X8574_L
DPOAE at 9189 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X9189_L
DPOAE at 9853 Hz (left ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X9853_L
DPOAE at 996 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X996_R
DPOAE at 1074 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1074_R
DPOAE at 1152 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1152_R
DPOAE at 1230 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1230_R
DPOAE at 1318 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1318_R
DPOAE at 1416 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1416_R
DPOAE at 1513 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1513_R
DPOAE at 1621 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1621_R

DPOAE at 1738 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1738_R
DPOAE at 1865 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X1865_R
DPOAE at 2001 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X2001_R
DPOAE at 2148 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X2148_R
DPOAE at 2294 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X2294_R
DPOAE at 2460 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X2460_R
DPOAE at 2636 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X2636_R
DPOAE at 2832 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X2832_R
DPOAE at 3027 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X3027_R
DPOAE at 3251 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X3251_R
DPOAE at 3486 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X3486_R
DPOAE at 3730 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X3730_R
DPOAE at 4003 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X4003_R
DPOAE at 4287 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X4287_R

DPOAE at 4599 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X4599_R
DPOAE at 4921 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X4921_R
DPOAE at 5273 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X5273_R
DPOAE at 5654 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X5654_R
DPOAE at 6064 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X6064_R
DPOAE at 6494 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X6494_R
DPOAE at 6962 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X6962_R
DPOAE at 7460 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X7460_R
DPOAE at 7998 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X7998_R
DPOAE at 8574 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X8574_R
DPOAE at 9189 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X9189_R
DPOAE at 9853 Hz (right ear) (signal to noise ratio)	TRUE	FALSE	FALSE	See description for 'DPOAE at 996 Hz (left ear) (signal to noise ratio)'	X9853_R

Hearing threshold at pitch matched frequency (dB HL)	TRUE	FALSE	FALSE	Psychoacoustic tinnitus assessment for STOP data. For pitch matching (kHz) of dominant frequency component: two alternative forced choice bracketing procedure starting at 1kHz and 2kHz. Start with audiometric frequencies increasing resolution to 1/12th or 1/24th octave. Stop if the patient has difficulties distinguishing the two tones. Finish by only presenting the frequency that is the resulting pitch match and ask if this is a good match, take note of their response.; For loudness matching (dB HL): two alternative forced choice procedure, starting with 5dB steps decreasing to 1 dB steps at pitch matched frequency.; Hearing threshold at pitch matched frequency in dB HL.; Masking noise threshold using Narrow Band Noise in dB HL.; Minimal Masking Level (dB SL) starting at Masking Noise Threshold increasing level in 1dB steps every second; Time (seconds) of Residual inhibition (Narrow band noise 10dB above MML) 1 minute of masking and reported level of RI (1-4) and time for tinnitus to revert back to sound as before RI-noise. Instruct the patient and leave patient in silent for 1 minute after noise so they can focus. RI time is recorded for max 2 min (report longer RI as > 2min) due to time constraints.; Maskability categorical options: 'Complete', 'Exacerbation', 'None', 'Partial' ('Exacerbation' and 'None' were coded as maskability 'No', and the rest as 'Yes').; Residual inhibition categorical options: 'Absent' (No inhibition), 'Complete' (tinnitus gone), 'Partial' (tinnitus partially reduced), 'Reduced' (reduced inhibition, meaning tinnitus got louder).	PMF_Threshold
Tinnitus loudness matching (dB HL)	TRUE	FALSE	FALSE	See description for 'Hearing threshold at pitch matched frequency (dB HL)'	Loudness_Match_dBHL
Masking noise threshold using Narrow Band Noise (dB HL)	TRUE	FALSE	FALSE	See description for 'Hearing threshold at pitch matched frequency (dB HL)'	MN_Threshold

Maskability (no/yes)	TRUE	FALSE	FALSE	See description for 'Hearing threshold at pitch matched frequency (dB HL)'	Maskability
Residual inhibition type (absent/complete/partial/reduced)	TRUE	FALSE	FALSE	See description for 'Hearing threshold at pitch matched frequency (dB HL)'	RI_Char
Perceived Stress Questionnaire (0-1)	TRUE	FALSE	FALSE	Value: score 0-1; Perceived Stress Questionnaire (PSQ) (Levenstein et al., 1993).	PSQ.Total.score
Hospital Anxiety Depression Scale for Anxiety (0-21)	TRUE	FALSE	FALSE	Value: score 0-21; Hospital Anxiety Depression Scale for Anxiety (HADS-A) (Zigmond and Snaith, 1983).	HADS_A.Total.score
Hospital Anxiety Depression Scale for Depression (0-21)	TRUE	FALSE	FALSE	Value: score 0-21; Hospital Anxiety Depression Scale for Depression (HADS-D) (Zigmond and Snaith, 1983).	HADS_D.Total.score
World Health Organization's Quality of Life Physical subscale (4-20)	TRUE	FALSE	FALSE	Value: score 4-20; World Health Organization's Quality of Life (WHOQoL)-BREF Physical subscale (The Whoqol Group, 1998).	WHO_QoL.Physical
World Health Organization's Quality of Life Psychological subscale (4-20)	TRUE	FALSE	FALSE	Value: score 4-20; World Health Organization's Quality of Life (WHOQoL)-BREF Psychological subscale (The Whoqol Group, 1998).	WHO_QoL.Psychological
World Health Organization's Quality of Life Social subscale (4-20)	TRUE	FALSE	FALSE	Value: score 4-20; World Health Organization's Quality of Life (WHOQoL)-BREF Social subscale (The Whoqol Group, 1998).	WHO_QoL.Social
World Health Organization's Quality of Life Environmental subscale (4-20)	TRUE	FALSE	FALSE	Value: score 4-20; World Health Organization's Quality of Life (WHOQoL)-BREF Environmental subscale (The Whoqol Group, 1998).	WHO_QoL.Environment

Tinnitus Handicap Inventory (0-100)	TRUE	FALSE	FALSE	Value: score 0-100; Tinnitus Handicap Inventory (Newman et al., 1996).	THI.Total.score
Fear of Tinnitus Questionnaire (0-17)	TRUE	FALSE	FALSE	Value: score 0-17; Fear of Tinnitus Questionnaire (FTQ) (Cima et al., 2011).	FTQ.Total.Score
Tinnitus Catastrophizing Scale (0-52)	TRUE	FALSE	FALSE	Value: score 0-52; Tinnitus Catastrophizing Scale (TCS) (Cima et al., 2011).	TCS.Total.score
Tinnitus awareness (% of total awake time)	TRUE	FALSE	FALSE	Value: 0-100; Question: 'What percent of your total awake time, over the last month, have you been aware of your tinnitus ? For example, 100% would indicate that you were aware of your tinnitus all the time, and 25% would indicate that you were aware of your tinnitus ¼ of the time (Please write in a single number between 1 and 100.)' (TSCHQ 16).	NRS.Awareness
Radar plot including NRS Loudness, NRS Awareness, NRS Annoyance, FTQ , TCS, HQ, PSQ, HADS_A, HADS, WHO Physical, Psychological, Social, and Environmental subscale, and THI (area)	TRUE	FALSE	FALSE	Value: numeric - pie area; Area from radar plot including NRS Loudness, NRS Awareness, NRS Annoyance, FTQ , TCS, HQ, PSQ, HADS_A, HADS, WHO Physical subscale, WHO Psychological subscale, WHO Social subscale, WHO Environmental subscale, and THI (Schlee et al., 2017).	Pie.Size
Ever worked shifts (no/yes)	TRUE	FALSE	FALSE	Value: no/yes; Question: 'Have you ever worked shifts?'; Original response options 'Yes, I do currently' and 'Yes, I have done it before' recoded to 'Yes'.	Intro_11
Living with partner (no/yes)	FALSE	FALSE	TRUE	Value: no/yes; Question: 'What is your marital status?' (ESIT-SQ O4); Original response options 'married', 'partner', 'single', 'widow', 'divorced', and 'prefer not to say,' recoded to 'Yes', 'Yes', 'No', 'No', 'No', and NA, respectively.	esitsq_q_o4_marital_status

Employment (employed/unemployed/own job/retired/sick leave/parental leave/student/sabbatical/housework/other/do not know)	FALSE	FALSE	TRUE	Value: employed/unemployed/own_job/retired/sick_leave/parental_leave/student/sabbatical/housework/other/not_know; Question: 'Which of the following describes best your current situation?' (ESIT-SQ O6).	esitsq_q_o6_employment
Cigarettes (number per day)	FALSE	FALSE	TRUE	Value: number of cigarettes; Question: 'Which of the following describes best your current situation?' (ESIT-SQ O8).	esitsq_q_o8_cigarettes
Coffee (cups per day)	FALSE	FALSE	TRUE	Value: number of cups of coffee; Question: 'How many cups of coffee do you drink per day on average?' (ESIT-SQ O9).	esitsq_q_o9_coffee
Exercise hours per week (2/2-4/5-7/more than 7)	FALSE	FALSE	TRUE	Value: 2h_week/2_4h_week/5_7h_week/more_7h_week; Question: 'How many hours per week do you do leisure-time physical activities on average?' (ESIT-SQ O10).	esitsq_q_o10_exercise
Fruit consumption (never/less than 3 times per month/1-6 times per week/once per day/2-3 times per day/more than 4 times per day)	FALSE	FALSE	TRUE	Value: never/less_weekly/more_weekly/daily/2_3_day/more_4_day; Question: 'How often do you consume fruits on average?' (ESIT-SQ O13).	esitsq_q_o13_fruits
Vegetable consumption (never/less than 3 times per month/1-6 times per week/once per day/2-3 times per day/more than 4 times per day)	FALSE	FALSE	TRUE	Value: never/less_weekly/more_weekly/daily/2_3_day/more_4_day; Question: 'How often do you consume vegetables on average?' (ESIT-SQ O14).	esitsq_q_o14_vegetables

Mobile phone use for calls (no use/less than 1 hour per month/around 1 hour per month/2-3 hours per month/around 1 hour per week/2-6 hours per week/1 hour per day/more than 1 hour per day)	FALSE	FALSE	TRUE	Value: no/less_1h_month/1h_month/2_3h_month/1h_week/2_6h_week/1h_day/more_1h_day; Question: 'For how long do you use a mobile phone for calls on average?' (ESIT-SQ O15).	esitsq_q_o15_mobile_calls
Headphone use for music (no use/less than 1 hour per month/around 1 hour per month/2-3 hours per month/around 1 hour per week/2-6 hours per week/1 hour per day/more than 1 hour per day)	FALSE	FALSE	TRUE	Value: no/less_1h_month/1h_month/2_3h_month/1h_week/2_6h_week/1h_day/more_1h_day; Question: 'For how long do you use headphones to listen to music on average?' (ESIT-SQ O16).	esitsq_q_o16_headphones_music
Sleep per day hours (six or less/seven/eight or more)	FALSE	FALSE	TRUE	Value: six or less/seven/eight or more; Question: 'How many hours do you sleep per day on average?' (ESIT-SQ O17); Originally 5 response options (less than 6, 6, 7, 8, 9 or more) recoded into five.	esitsq_q_o17_sleep

Some variables have been assessed twice for the STOP dataset, using both the TSCHQ and the ESIT-SQ. In such cases these have been entered twice for the ESIT and BRC dataset and the descriptor ‘1tschq2esitsq’ was used to indicate that the ESIT-SQ version was used for the STOP dataset, whereas the descriptor ‘2tschq1esitsq’ is used to indicate use of the TSCHQ version.

Appendix 4.2. Common variables across the STOP and the ESIT datasets: Descriptions and comparisons

	All	ESIT	STOP	P value
All	595	200	395	-
Height (cm)	173 (165, 180), n=589	170 (164.75, 178), n=196	173 (167, 182), n=393	<0.001
Weight (kg)	75 (65, 85), n=589	77.57 (66.1, 86.82), n=195	74 (64.25, 84.75), n=394	0.047
BMI (kg/m^2)	24.49 (22.49, 27.43), n=586	26.13 (23.35, 29.06), n=194	24.1 (22.15, 26.41), n=392	<0.001
Education (lower/higher)	192/403	84/116	108/287	<0.001
Alcohol (number of drinks per week)	3 (1, 6), n=578	3 (1, 8), n=183	3 (1, 6), n=395	0.158
Smoking (no/yes)	338/256	102/97	236/159	0.054
Night work (no/yes)	372/219	119/81	253/138	0.242
Any ear condition (no/yes)	324/261	75/115	249/146	<0.001
Acoustic trauma (no/yes)	567/18	184/6	383/12	1
Barotrauma (no/yes)	582/3	189/1	393/2	1
Presbycusis (no/yes)	544/41	183/7	361/34	0.037
Sudden hearing loss (no/yes)	561/24	174/16	387/8	<0.001
Other hearing loss (no/yes)	488/97	127/63	361/34	<0.001
Meniere's disease (no/yes)	572/13	183/7	389/6	0.132
Acoustic neuroma (no/yes)	579/6	185/5	394/1	0.015
Acute Otitis (no/yes)	538/47	172/18	366/29	0.417
Serous Otitis (no/yes)	569/16	183/7	386/9	0.417
Chronic Otitis (no/yes)	575/10	184/6	391/4	0.085
Otosclerosis (no/yes)	577/8	188/2	389/6	1
Problem with external sounds (small/moderate/big or very big)	404/138/53	129/48/23	275/90/30	0.237
Self-reported hearing difficulty (slight or no difficulty/moderate difficulty/severe difficulty/total loss)	276/198/105/7	75/70/47/5	201/128/58/2	0.001
Any hearing device (no/yes)	455/139	100/99	355/40	<0.001
Combination device (hearing aid and sound generator) (no/yes)	583/11	190/9	393/2	0.001
Any procedure (no/yes)	297/278	72/108	225/170	<0.001
Ear surgery (no/yes)	530/45	155/25	375/20	<0.001
Dental surgery (no/yes)	383/192	106/74	277/118	0.01
Neurosurgery (no/yes)	561/14	172/8	389/6	0.043
Lumbar puncture (no/yes)	549/26	167/13	382/13	0.049
Chemotherapy (no/yes)	569/6	177/3	392/3	0.383
Head and neck radiotherapy (no/yes)	569/6	176/4	393/2	0.08
Any pain syndromes (no/yes)	330/247	84/98	246/149	<0.001
Neck pain (no/yes)	442/135	136/46	306/89	0.526
Ear pain (no/yes)	542/35	160/22	382/13	<0.001
TMJ pain (no/yes)	534/43	173/9	361/34	0.128
Face pain (no/yes)	546/31	163/19	383/12	<0.001

Any diagnosed condition (no/yes)	207/381	59/134	148/247	0.118
Dental problems (no/yes)	549/39	182/11	367/28	0.599
Meningitis (no/yes)	576/12	187/6	389/6	0.222
Epilepsy (no/yes)	584/4	192/1	392/3	1
Stroke (no/yes)	581/7	187/6	394/1	0.006
Other neurologic condition (no/yes)	578/10	189/4	389/6	0.736
Anxiety (no/yes)	508/80	153/40	355/40	<0.001
Depression (no/yes)	507/81	165/28	342/53	0.704
Emotional trauma (no/yes)	563/25	183/10	380/15	0.514
Stress (no/yes)	533/55	178/15	355/40	0.451
Problem falling asleep (no/yes)	523/65	174/19	349/46	0.577
Problem staying asleep (no/yes)	506/82	164/29	342/53	0.613
Low BP (no/yes)	545/43	178/15	367/28	0.739
High BP (no/yes)	488/100	154/39	334/61	0.161
Thyroid disorder (no/yes)	547/41	178/15	369/26	0.607
Diabetes (no/yes)	569/19	181/12	388/7	0.006
High cholesterol (no/yes)	540/48	170/23	370/25	0.025
Rheumatoid arthritis (no/yes)	578/10	184/9	394/1	<0.001
Chronic sinusitis (no/yes)	574/14	183/10	391/4	0.003
Nasal septum deviation (no/yes)	578/10	184/9	394/1	<0.001
HIV (no/yes)	586/2	192/1	394/1	0.549
Lyme disease (no/yes)	536/52	193/0	343/52	<0.001
Anaemia (no/yes)	564/24	186/7	378/17	0.826
Balance disorders (no/yes)	567/21	185/8	382/13	0.639
Acid/gastroesophageal reflux (no/yes)	519/69	159/34	360/35	0.003
Tinnitus daily presence (no/yes)	100/487	7/185	93/302	<0.001
Objective tinnitus (yes/no)	53/536	22/172	31/364	0.17
Age at bothersome tinnitus onset (y)	45 (30, 55.12), n=310	50.42 (39, 58), n=148	39 (22, 50), n=162	<0.001
Bothersome tinnitus duration (y)	8 (4, 18), n=311	7 (3.75, 15), n=148	9 (4, 19), n=163	0.216
Gap between onset of tinnitus and bothersome tinnitus (y)	0 (0, 47.25), n=290	0 (0, 42.75), n=142	0 (0, 48), n=148	0.814
Change in hearing at tinnitus onset (no/yes)	567/21	187/6	380/15	0.815
Fullness in the ears at tinnitus onset (no/yes)	552/36	182/11	370/25	0.856
Neck trauma at tinnitus onset (no/yes)	568/20	184/9	384/11	0.236
No event at tinnitus onset (no/yes)	443/145	146/47	297/98	0.919
Aspirin at tinnitus onset (no/yes)	562/13	173/7	389/6	0.125
Painkillers at tinnitus onset (no/yes)	545/30	167/13	378/17	0.159
Antibiotics at tinnitus onset (no/yes)	568/7	174/6	394/1	0.005
Antidepressants at tinnitus onset (no/yes)	551/24	171/9	380/15	0.506
Quinine at tinnitus onset (no/yes)	573/2	179/1	394/1	0.528

Diuretics at tinnitus onset (no/yes)	568/7	174/6	394/1	0.005
No medication at tinnitus onset (no/yes)	207/368	78/102	129/266	0.015
Thoughts of conditions related to tinnitus onset (no/yes)	321/262	90/98	231/164	0.02
Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)	257/65/58/127	91/28/22/37	166/37/36/90	0.267
Onset of problems with external sounds in relation to tinnitus onset (never/before/at the same time/after)	261/58/97/99	79/23/41/38	182/35/56/61	0.127
Onset of hearing difficulties in relation to tinnitus onset (never/before/at the same time/after)	102/83/142/163	29/42/58/48	73/41/84/115	0.001
Onset of acoustic trauma in relation to tinnitus onset (never/before/at the same time/after)	561/4/8/1	178/0/5/0	383/4/3/1	0.104
Onset of barotrauma in relation to tinnitus onset (never/before/at the same time/after)	576/2/0/1	183/1/0/0	393/1/0/1	0.676
Onset of presbycusis in relation to tinnitus onset (never/before/at the same time/after)	539/9/7/13	178/4/0/1	361/5/7/12	0.05
Onset of sudden hearing loss in relation to tinnitus onset (never/before/at the same time/after)	555/5/13/6	168/2/11/3	387/3/2/3	<0.001
Onset of other hearing loss in relation to tinnitus onset (never/before/at the same time/after)	484/26/36/20	123/17/28/13	361/9/8/7	<0.001
Onset of Meniere's in relation to tinnitus onset (never/before/at the same time/after)	566/5/5/2	177/5/2/0	389/0/3/2	0.005
Onset of acoustic neuroma in relation to tinnitus onset (never/before/at the same time/after)	573/2/4/0	179/2/3/0	394/0/1/0	0.02
Onset of acute otitis in relation to tinnitus onset (never/before/at the same time/after)	532/31/3/9	166/13/2/3	366/18/1/6	0.297
Onset of serous otitis in relation to tinnitus onset (never/before/at the same time/after)	563/9/2/5	177/3/2/2	386/6/0/3	0.234
Onset of chronic otitis in relation to tinnitus onset (never/before/at the same time/after)	569/6/0/2	178/5/0/0	391/1/0/2	0.011
Onset of otosclerosis in relation to tinnitus onset (never/before/at the same time/after)	571/1/3/3	182/0/1/1	389/1/2/2	1

Onset of ear surgery in relation to tinnitus onset (never/before/at the same time/after)	526/20/2/12	151/11/2/6	375/9/0/6	0.002
Onset of dental surgery in relation to tinnitus onset (never/before/at the same time/after)	380/78/7/64	103/37/4/20	277/41/3/44	0.001
Onset of neurosurgery in relation to tinnitus onset (never/before/at the same time/after)	557/7/1/4	168/4/0/2	389/3/1/2	0.294
Onset of lumbar puncture in relation to tinnitus onset (never/before/at the same time/after)	545/6/3/11	163/4/1/4	382/2/2/7	0.212
Onset of chemotherapy in relation to tinnitus onset (never/before/at the same time/after)	565/2/1/2	173/0/0/2	392/2/1/0	0.187
Onset of head and neck radiotherapy in relation to tinnitus onset (never/before/at the same time/after)	564/3/1/1	171/2/1/1	393/1/0/0	0.057
Onset of headache in relation to tinnitus onset (never/before/at the same time/after)	438/61/18/35	130/25/5/14	308/36/13/21	0.219
Onset of neck pain in relation to tinnitus onset (never/before/at the same time/after)	440/49/20/43	134/18/7/16	306/31/13/27	0.638
Onset of ear pain in relation to tinnitus onset (never/before/at the same time/after)	539/11/3/10	157/9/1/7	382/2/2/3	<0.001
Onset of TMJ pain in relation to tinnitus onset (never/before/at the same time/after)	530/9/7/21	169/0/1/7	361/9/6/14	0.17
Onset of face pain in relation to tinnitus onset (never/before/at the same time/after)	542/6/3/20	159/4/2/13	383/2/1/7	<0.001
Onset of TMJ disorder in relation to tinnitus onset (never/before/at the same time/after)	569/3/3/5	178/2/1/4	391/1/2/1	0.047
Onset of dental problems in relation to tinnitus onset (never/before/at the same time/after)	543/17/2/13	176/4/0/4	367/13/2/9	0.885
Onset of meningitis in relation to tinnitus onset (never/before/at the same time/after)	571/6/1/3	182/3/0/2	389/3/1/1	0.382
Onset of epilepsy in relation to tinnitus onset (never/before/at the same time/after)	578/1/1/1	186/0/1/0	392/1/0/1	0.696
Onset of stroke in relation to tinnitus onset (never/before/at the same time/after)	575/3/1/2	181/3/1/1	394/0/0/1	0.008

Onset of other neurological condition in relation to tinnitus onset (never/before/at the same time/after)	572/3/1/2	183/0/0/1	389/3/1/1	0.633
Onset of anxiety in relation to tinnitus onset (never/before/at the same time/after)	503/32/18/24	148/15/11/11	355/17/7/13	0.003
Onset of depression in relation to tinnitus onset (never/before/at the same time/after)	501/30/15/28	159/13/6/6	342/17/9/22	0.294
Onset of emotional trauma in relation to tinnitus onset (never/before/at the same time/after)	557/12/1/6	177/5/1/2	380/7/0/4	0.458
Onset of stress in relation to tinnitus onset (never/before/at the same time/after)	527/17/10/18	172/7/2/3	355/10/8/15	0.382
Onset of problems falling asleep in relation to tinnitus onset (never/before/at the same time/after)	518/26/11/22	169/8/1/7	349/18/10/15	0.488
Onset of problems staying asleep in relation to tinnitus onset (never/before/at the same time/after)	501/32/10/28	159/12/2/12	342/20/8/16	0.48
Onset of low BP in relation to tinnitus onset (never/before/at the same time/after)	540/25/3/7	173/10/1/1	367/15/2/6	0.662
Onset of high BP in relation to tinnitus onset (never/before/at the same time/after)	485/26/8/53	151/15/2/18	334/11/6/35	0.05
Onset of thyroid disorder in relation to tinnitus onset (never/before/at the same time/after)	542/18/6/12	173/10/1/3	369/8/5/9	0.164
Onset of diabetes in relation to tinnitus onset (never/before/at the same time/after)	563/5/0/11	175/2/0/7	388/3/0/4	0.058
Onset of increased cholesterol in relation to tinnitus onset (never/before/at the same time/after)	535/19/6/18	165/10/1/10	370/9/5/8	0.023
Onset of rheumatoid arthritis in relation to tinnitus onset (never/before/at the same time/after)	573/4/0/5	179/3/0/5	394/1/0/0	<0.001
Onset of chronic sinusitis in relation to tinnitus onset (never/before/at the same time/after)	569/6/1/5	178/5/0/3	391/1/1/2	0.016
Onset of septal deviation in relation to tinnitus onset (never/before/at the same time/after)	572/5/1/3	178/5/1/2	394/0/0/1	0.002

Onset of HIV in relation to tinnitus onset (never/before/at the same time/after)	580/1/0/1	186/1/0/0	394/0/0/1	0.531
Onset of Lyme disease in relation to tinnitus onset (never/before/at the same time/after)	530/21/4/20	187/0/0/0	343/21/4/20	<0.001
Onset of anaemia in relation to tinnitus onset (never/before/at the same time/after)	558/9/1/11	180/3/0/3	378/6/1/8	1
Onset of balance disorders in relation to tinnitus onset (never/before/at the same time/after)	561/6/7/7	179/3/3/1	382/3/4/6	0.518
Onset of acid/gastroesophageal reflux in relation to tinnitus onset (never/before/at the same time/after)	513/35/3/26	153/17/0/16	360/18/3/10	<0.001
Number of sounds (more than one/one)	216/372	111/82	105/290	<0.001
Varying tinnitus loudness over a day (stable/sometimes fluctuating/always fluctuating)	189/250/120	48/108/37	141/142/83	<0.001
Tinnitus quality (tonal/noise/other)	269/131/171	69/39/68	200/92/103	0.009
Tinnitus pitch (high/medium/low)	359/168/53	133/46/10	226/122/43	0.007
Tinnitus reduced by silence (no/yes)	553/33	181/10	372/23	0.85
Tinnitus reduced by low intensity sounds (no/yes)	515/71	171/20	344/51	0.421
Tinnitus reduced by high intensity sounds (no/yes)	522/64	170/21	352/43	1
Tinnitus reduced by head movement (no/yes)	571/15	189/2	382/13	0.161
Tinnitus reduced by jaw movement (no/yes)	561/25	186/5	375/20	0.197
Tinnitus reduced by pressing head, neck, or area around the ear (no/yes)	563/23	184/7	379/16	1
Tinnitus reduced by taking a nap (no/yes)	561/25	185/6	376/19	0.393
Tinnitus reduced by good quality sleep (no/yes)	452/134	163/28	289/106	0.001
Tinnitus reduced by driving (no/yes)	551/35	180/11	371/24	1
Tinnitus reduced by stress (no/yes)	584/2	190/1	394/1	0.546
Tinnitus reduced by being relaxed (no/yes)	479/107	167/24	312/83	0.016
Tinnitus reduced by drinking alcohol (no/yes)	561/25	183/8	378/17	1
Tinnitus reduced by taking medication (no/yes)	581/5	189/2	392/3	0.663

Tinnitus reduced by using hearing aids (no/yes)	539/47	157/34	382/13	<0.001
Tinnitus reduced by nothing (no/yes)	356/230	121/70	235/160	0.417
Tinnitus increased by silence (no/yes)	319/261	95/90	224/171	0.245
Tinnitus increased by low intensity sounds (no/yes)	567/13	182/3	385/10	0.764
Tinnitus increased by high intensity sounds (no/yes)	416/164	140/45	276/119	0.166
Tinnitus increased by head movement (no/yes)	560/20	179/6	381/14	1
Tinnitus increased by jaw movement (no/yes)	480/100	164/21	316/79	0.01
Tinnitus increased by pressing head, neck, or area around the ear (no/yes)	560/20	176/9	384/11	0.225
Tinnitus increased by taking a nap (no/yes)	570/10	178/7	392/3	0.014
Tinnitus increased by poor quality sleep (no/yes)	420/160	142/43	278/117	0.112
Tinnitus increased by driving (no/yes)	562/18	172/13	390/5	<0.001
Tinnitus increased by being relaxed (no/yes)	571/9	184/1	387/8	0.284
Tinnitus increased by drinking alcohol (no/yes)	536/44	170/15	366/29	0.739
Tinnitus increased by drinking coffee (no/yes)	567/13	177/8	390/5	0.032
Tinnitus increased by taking medication (no/yes)	571/9	180/5	391/4	0.153
Tinnitus increased by using hearing aids (no/yes)	571/9	178/7	393/2	0.006
Tinnitus increased by nothing (no/yes)	492/88	151/34	341/54	0.172
Thoughts of conditions related to increased tinnitus (no/yes)	380/207	121/71	259/136	0.581
Tinnitus worries, annoys or upsets (severely/moderately/slightly/not at all)	97/190/213/85	59/66/56/11	38/124/157/74	<0.001
Psychological management for tinnitus (no/yes)	579/9	189/4	390/5	0.485
Audiological management for tinnitus (no/yes)	559/29	168/25	391/4	<0.001
Physiotherapy for tinnitus (no/yes)	582/6	192/1	390/5	0.669
Self-management for tinnitus (no/yes)	544/44	163/30	381/14	<0.001
No management for tinnitus (no/yes)	80/508	54/139	26/369	<0.001

Appendix 4.3. Common variables across the STOP and the BRC datasets: Descriptions and comparisons

	All	BRC	STOP	P value
All	600	205	395	-
MaxDiffExt (dB)	-5.33 (-13.01, 8.16), n=600	-4.5 (-15, 12.5), n=205	-6.03 (-12.6, 7.42), n=395	0.673
Mean hearing threshold both ears (dB HL)	12.69 (5, 26.16), n=600	27.86 (16.07, 36.07), n=205	7.91 (2.23, 16.43), n=395	<0.001
Hearing threshold at 0.125 kHz (left ear) (dB HL)	5 (0, 10), n=600	10 (5, 20), n=205	2.1 (-1.2, 6.55), n=395	<0.001
Hearing threshold at 0.25 kHz (left ear) (dB HL)	2.6 (-2.5, 10.03), n=600	10 (5, 20), n=205	-0.3 (-3.7, 4.8), n=395	<0.001
Hearing threshold at 0.5 kHz (left ear) (dB HL)	4.55 (-2.35, 15), n=600	15 (5, 25), n=205	0 (-3.7, 7), n=395	<0.001
Hearing threshold at 0.75 kHz (left ear) (dB HL)	5 (-1.21, 15), n=600	15 (10, 30), n=205	1.1 (-2.6, 6.75), n=395	<0.001
Hearing threshold at 1 kHz (left ear) (dB HL)	5 (-0.6, 17.12), n=600	15 (5, 30), n=205	1.1 (-2.3, 8), n=395	<0.001
Hearing threshold at 1.5 kHz (left ear) (dB HL)	6.15 (0, 20.76), n=600	20 (5, 35), n=205	2.85 (-1.5, 10.93), n=395	<0.001
Hearing threshold at 2 kHz (left ear) (dB HL)	9.5 (0.5, 25), n=600	25 (10, 40), n=205	4.5 (-2.3, 16.15), n=395	<0.001
Hearing threshold at 3 kHz (left ear) (dB HL)	18.1 (5.95, 37.35), n=600	35 (20, 55), n=205	11.1 (2.55, 24.95), n=395	<0.001
Hearing threshold at 4 kHz (left ear) (dB HL)	23.65 (8.45, 45), n=600	40 (25, 60), n=205	15.3 (3.8, 32.3), n=395	<0.001
Hearing threshold at 6 kHz (left ear) (dB HL)	29.4 (10.3, 50), n=600	50 (30, 65), n=205	20.3 (6.3, 37.95), n=395	<0.001
Hearing threshold at 8 kHz (left ear) (dB HL)	38.9 (15, 61), n=600	60 (35, 75), n=205	29.3 (8.55, 51.5), n=395	<0.001
Hearing threshold at 10 kHz (left ear) (dB HL)	50.7 (20, 70), n=600	65 (50, 80), n=205	41.1 (10.6, 60.4), n=395	<0.001
Hearing threshold at 12.5 kHz (left ear) (dB HL)	65 (41.58, 80), n=600	75 (65, 85), n=205	56.1 (25.05, 72.55), n=395	<0.001
Hearing threshold at 14 kHz (left ear) (dB HL)	68.9 (51.95, 110), n=600	70 (65, 80), n=205	64.8 (41.6, 110), n=395	<0.001
Hearing threshold at 0.125 kHz (right ear) (dB HL)	5 (1.5, 10.1), n=600	10 (5, 20), n=205	3.5 (0.1, 7), n=395	<0.001
Hearing threshold at 0.25 kHz (right ear) (dB HL)	3.7 (-1.1, 10), n=600	10 (5, 25), n=205	0.5 (-2.6, 5.05), n=395	<0.001
Hearing threshold at 0.5 kHz (right ear) (dB HL)	4.05 (-1.8, 15), n=600	15 (5, 25), n=205	0.1 (-3.3, 6.8), n=395	<0.001
Hearing threshold at 0.75 kHz (right ear) (dB HL)	5 (-0.85, 15.04), n=600	15 (5, 30), n=205	0.95 (-2.55, 6.2), n=395	<0.001
Hearing threshold at 1 kHz (right ear) (dB HL)	5 (-0.5, 16.2), n=600	15 (10, 30), n=205	1.1 (-2.6, 8.6), n=395	<0.001
Hearing threshold at 1.5 kHz (right ear) (dB HL)	5.97 (0, 18.83), n=600	20 (10, 35), n=205	2.45 (-1.35, 10.18), n=395	<0.001

Hearing threshold at 2 kHz (right ear) (dB HL)	10 (0.75, 24.5), n=600	25 (10, 40), n=205	3.8 (-1.4, 12.9), n=395	<0.001
Hearing threshold at 3 kHz (right ear) (dB HL)	15 (3.5, 34.15), n=600	35 (15, 55), n=205	8 (0.2, 22.1), n=395	<0.001
Hearing threshold at 4 kHz (right ear) (dB HL)	20 (5, 41.23), n=600	35 (20, 55), n=205	12.6 (1.3, 28.7), n=395	<0.001
Hearing threshold at 6 kHz (right ear) (dB HL)	25 (7.5, 45), n=600	45 (25, 60), n=205	17.6 (3.85, 36), n=395	<0.001
Hearing threshold at 8 kHz (right ear) (dB HL)	35 (12.3, 60), n=600	60 (35, 75), n=205	24.1 (7.7, 48.2), n=395	<0.001
Hearing threshold at 10 kHz (right ear) (dB HL)	49.5 (18.48, 67.65), n=600	65 (45, 80), n=205	37.8 (9.3, 58.7), n=395	<0.001
Hearing threshold at 12.5 kHz (right ear) (dB HL)	64.1 (35, 77.42), n=600	75 (65, 85), n=205	54.1 (18.2, 70), n=395	<0.001
Hearing threshold at 14 kHz (right ear) (dB HL)	65.5 (50, 100), n=600	70 (65, 80), n=205	62.1 (34.4, 110), n=395	<0.001
Self-reported hearing problem (no/yes)	181/351	44/155	137/196	<0.001
Hyperacusis questionnaire score (0-42)	14 (9, 21), n=581	12 (9, 19), n=187	15 (9, 23), n=394	0.011
Tinnitus loudness rating (0-100)	40 (25, 60), n=574	41 (30, 50), n=199	40 (20, 60), n=375	0.376
Tinnitus pitch matching (kHz)	8 (5, 12), n=506	7 (5, 10), n=199	9 (4.61, 12.5), n=307	NA
Varying tinnitus loudness from day to day (no/yes)	186/373	88/110	98/263	<0.001
Tinnitus worsened by loud noise (no/yes)	219/247	111/74	108/173	<0.001
Tinnitus annoyance scale (0-100)	10 (5, 40), n=577	25 (10, 50), n=198	10 (5, 25), n=379	<0.001

NA: comparison not applicable because variables were assessed in substantially different ways.

Appendix 4.4. Descriptive statistics for variables only available in the STOP dataset

All	395
Ever worked shifts (no/yes)	269/124
Abnormal tympanogram left ear (no/yes)	286/80
Tympanometric Peak Pressure (TPP) left ear (daPA)	7 (1, 15), n=358
Static Admittance (SA) left ear (mmho)	0.66 (0.44, 1.03), n=357
Tympanometric Width (TW) left ear (daPa)	70 (58, 93), n=352
Ear Canal Volume (ECVL) left ear (cm3)	1.34 (1.12, 1.6), n=368
Abnormal tympanogram right ear (no/yes)	296/68
Tympanometric Peak Pressure (TPP) right ear (daPA)	7 (1, 14), n=359
Static Admittance (SA) right ear (mmho)	0.69 (0.46, 1.03), n=353
Tympanometric Width (TW) right ear (daPa)	71 (58, 94), n=350
Ear Canal Volume (ECVL) right ear (cm3)	1.39 (1.14, 1.68), n=364
Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)	75 (70, 85), n=349
Uncomfortable loudness level at 0.25 kHz (left ear) (dB HL)	80 (70, 90), n=357
Uncomfortable loudness level at 0.5 kHz (left ear) (dB HL)	80 (70, 90), n=355
Uncomfortable loudness level at 1 kHz (left ear) (dB HL)	80 (70, 90), n=359
Uncomfortable loudness level at 2 kHz (left ear) (dB HL)	80 (65, 87.5), n=363
Uncomfortable loudness level at 3 kHz (left ear) (dB HL)	80 (65, 90), n=349
Uncomfortable loudness level at 4 kHz (left ear) (dB HL)	80 (65, 90), n=348
Uncomfortable loudness level at 6 kHz (left ear) (dB HL)	75 (65, 85), n=346
Uncomfortable loudness level at 8 kHz (left ear) (dB HL)	75 (60, 85), n=325
Uncomfortable loudness level at 0.125 kHz (right ear) (dB HL)	80 (70, 85), n=341
Uncomfortable loudness level at 0.25 kHz (right ear) (dB HL)	80 (75, 90), n=354
Uncomfortable loudness level at 0.5 kHz (right ear) (dB HL)	80 (70, 90), n=358
Uncomfortable loudness level at 1 kHz (right ear) (dB HL)	80 (70, 85), n=366
Uncomfortable loudness level at 2 kHz (right ear) (dB HL)	80 (70, 85), n=364
Uncomfortable loudness level at 3 kHz (right ear) (dB HL)	75 (65, 85), n=361
Uncomfortable loudness level at 4 kHz (right ear) (dB HL)	80 (65, 87.5), n=363

Uncomfortable loudness level at 6 kHz (right ear) (dB HL)	75 (65, 85), n=353
Uncomfortable loudness level at 8 kHz (right ear) (dB HL)	75 (60, 85), n=327
Speech in noise hearing assessment - word score (left ear) (0-100)	72 (62, 78), n=387
Speech in noise hearing assessment - phoneme score (left ear) (0-100)	88.7 (82, 92), n=388
Speech in noise hearing assessment - word score (right) (0-100)	72 (64, 78), n=387
Speech in noise hearing assessment - phoneme score (right ear) (0-100)	88 (81.3, 91.3), n=388
DPOAE at 996 Hz (left ear) (signal to noise ratio)	11.9 (7.8, 17.55), n=387
DPOAE at 1074 Hz (left ear) (signal to noise ratio)	13.1 (8.2, 19.55), n=387
DPOAE at 1152 Hz (left ear) (signal to noise ratio)	13.7 (8.9, 19.95), n=387
DPOAE at 1230 Hz (left ear) (signal to noise ratio)	14.9 (9.25, 20.65), n=387
DPOAE at 1318 Hz (left ear) (signal to noise ratio)	16.4 (9.7, 21.35), n=387
DPOAE at 1416 Hz (left ear) (signal to noise ratio)	17.3 (10.6, 22.1), n=387
DPOAE at 1513 Hz (left ear) (signal to noise ratio)	18.2 (12, 23.55), n=387
DPOAE at 1621 Hz (left ear) (signal to noise ratio)	19 (11.25, 23.8), n=387
DPOAE at 1738 Hz (left ear) (signal to noise ratio)	19.1 (13.15, 24.5), n=387
DPOAE at 1865 Hz (left ear) (signal to noise ratio)	17.8 (11.65, 22.9), n=387
DPOAE at 2001 Hz (left ear) (signal to noise ratio)	18 (11.9, 23.1), n=387
DPOAE at 2148 Hz (left ear) (signal to noise ratio)	18 (12.1, 23), n=387
DPOAE at 2294 Hz (left ear) (signal to noise ratio)	18.5 (11.8, 23.8), n=387
DPOAE at 2460 Hz (left ear) (signal to noise ratio)	18.5 (12.45, 24.35), n=387
DPOAE at 2636 Hz (left ear) (signal to noise ratio)	18.5 (12.45, 23.45), n=387
DPOAE at 2832 Hz (left ear) (signal to noise ratio)	18.2 (12.8, 23.9), n=387
DPOAE at 3027 Hz (left ear) (signal to noise ratio)	18.6 (13.05, 23.75), n=387
DPOAE at 3251 Hz (left ear) (signal to noise ratio)	16.1 (11.4, 22.5), n=387
DPOAE at 3486 Hz (left ear) (signal to noise ratio)	15.9 (10.85, 20.65), n=387
DPOAE at 3730 Hz (left ear) (signal to noise ratio)	16 (10.3, 22.15), n=387
DPOAE at 4003 Hz (left ear) (signal to noise ratio)	15.9 (11, 21.7), n=387

DPOAE at 4287 Hz (left ear) (signal to noise ratio)	15.2 (11, 21.4), n=387
DPOAE at 4599 Hz (left ear) (signal to noise ratio)	15.6 (9.7, 21.6), n=387
DPOAE at 4921 Hz (left ear) (signal to noise ratio)	14.6 (7.75, 22.55), n=387
DPOAE at 5273 Hz (left ear) (signal to noise ratio)	14.4 (8, 22.2), n=387
DPOAE at 5654 Hz (left ear) (signal to noise ratio)	12.1 (6.55, 20.55), n=387
DPOAE at 6064 Hz (left ear) (signal to noise ratio)	11.8 (6.3, 18.35), n=387
DPOAE at 6494 Hz (left ear) (signal to noise ratio)	10 (6.5, 15.7), n=387
DPOAE at 6962 Hz (left ear) (signal to noise ratio)	9.4 (6.25, 13.45), n=387
DPOAE at 7460 Hz (left ear) (signal to noise ratio)	7.8 (3.4, 11.6), n=387
DPOAE at 7998 Hz (left ear) (signal to noise ratio)	8.2 (4.4, 11.75), n=387
DPOAE at 8574 Hz (left ear) (signal to noise ratio)	10.3 (6.9, 15.4), n=387
DPOAE at 9189 Hz (left ear) (signal to noise ratio)	12.3 (9.1, 16.6), n=387
DPOAE at 9853 Hz (left ear) (signal to noise ratio)	14.5 (10.9, 18.35), n=387
DPOAE at 996 Hz (right ear) (signal to noise ratio)	11.6 (7.85, 18.05), n=387
DPOAE at 1074 Hz (right ear) (signal to noise ratio)	14.2 (8.4, 19.65), n=387
DPOAE at 1152 Hz (right ear) (signal to noise ratio)	15.2 (9.15, 21.55), n=387
DPOAE at 1230 Hz (right ear) (signal to noise ratio)	16.1 (9.6, 21.95), n=387
DPOAE at 1318 Hz (right ear) (signal to noise ratio)	16.9 (10.6, 22.55), n=387
DPOAE at 1416 Hz (right ear) (signal to noise ratio)	18.7 (11.55, 24.5), n=387
DPOAE at 1513 Hz (right ear) (signal to noise ratio)	19.3 (11.2, 24.3), n=387
DPOAE at 1621 Hz (right ear) (signal to noise ratio)	19.4 (12.65, 25.05), n=387
DPOAE at 1738 Hz (right ear) (signal to noise ratio)	19.8 (13.6, 25.35), n=387
DPOAE at 1865 Hz (right ear) (signal to noise ratio)	19.4 (12.65, 24.3), n=387
DPOAE at 2001 Hz (right ear) (signal to noise ratio)	20.1 (14.1, 24.65), n=387
DPOAE at 2148 Hz (right ear) (signal to noise ratio)	19.6 (12, 24.1), n=387
DPOAE at 2294 Hz (right ear) (signal to noise ratio)	18.9 (14.15, 24.1), n=387
DPOAE at 2460 Hz (right ear) (signal to noise ratio)	19.6 (13.4, 24.55), n=387

DPOAE at 2636 Hz (right ear) (signal to noise ratio)	19.3 (12.9, 24.5), n=387
DPOAE at 2832 Hz (right ear) (signal to noise ratio)	19.8 (13.3, 24.45), n=387
DPOAE at 3027 Hz (right ear) (signal to noise ratio)	19.4 (12.95, 24.05), n=387
DPOAE at 3251 Hz (right ear) (signal to noise ratio)	17.4 (11.45, 22.8), n=387
DPOAE at 3486 Hz (right ear) (signal to noise ratio)	16.3 (11.7, 22.55), n=387
DPOAE at 3730 Hz (right ear) (signal to noise ratio)	17 (11.15, 22.7), n=387
DPOAE at 4003 Hz (right ear) (signal to noise ratio)	16.7 (11.15, 22.5), n=387
DPOAE at 4287 Hz (right ear) (signal to noise ratio)	16.8 (11.4, 22.75), n=387
DPOAE at 4599 Hz (right ear) (signal to noise ratio)	16.2 (10.75, 23), n=387
DPOAE at 4921 Hz (right ear) (signal to noise ratio)	16.4 (9.55, 23.1), n=387
DPOAE at 5273 Hz (right ear) (signal to noise ratio)	15 (8.65, 23.35), n=387
DPOAE at 5654 Hz (right ear) (signal to noise ratio)	13.7 (7.95, 22.4), n=387
DPOAE at 6064 Hz (right ear) (signal to noise ratio)	12.5 (7.4, 20.2), n=387
DPOAE at 6494 Hz (right ear) (signal to noise ratio)	10.7 (6.2, 16.85), n=387
DPOAE at 6962 Hz (right ear) (signal to noise ratio)	9.6 (7.1, 13.65), n=387
DPOAE at 7460 Hz (right ear) (signal to noise ratio)	8.3 (3.7, 11.95), n=387
DPOAE at 7998 Hz (right ear) (signal to noise ratio)	9 (6.1, 13), n=387
DPOAE at 8574 Hz (right ear) (signal to noise ratio)	10.6 (7.45, 16.65), n=387
DPOAE at 9189 Hz (right ear) (signal to noise ratio)	13 (9.35, 16.7), n=387
DPOAE at 9853 Hz (right ear) (signal to noise ratio)	15 (11.15, 18.85), n=387
Hospital Anxiety Depression Scale for Anxiety (0-21)	5 (3, 8), n=395
Hospital Anxiety Depression Scale for Depression (0-21)	2 (1, 5), n=395
World Health Organization’s Quality of Life Physical subscale (4-20)	17 (15, 18), n=391
World Health Organization’s Quality of Life Psychological subscale (4-20)	15 (14, 17), n=392
World Health Organization’s Quality of Life Social subscale (4-20)	15 (13, 16), n=395
World Health Organization’s Quality of Life Environmental subscale (4-20)	17 (15, 18), n=395
Perceived Stress Questionnaire (0-1)	0.31 (0.2, 0.44), n=395
Hearing threshold at pitch matched frequency (dB HL)	38 (18, 56), n=299

Tinnitus loudness matching (dB HL)	44 (25, 60), n=303
Masking noise threshold using Narrow Band Noise (dB HL)	20 (0, 30.5), n=296
Maskability (no/yes)	136/152
Residual inhibition type (absent/complete/partial/reduced)	69/45/124/27
Tinnitus Handicap Inventory (0-100)	16 (6, 28), n=380
Fear of Tinnitus Questionnaire (0-17)	5 (3, 6), n=377
Tinnitus Catastrophizing Scale (0-52)	11 (5, 19), n=380
Tinnitus awareness (% of total awake time)	20 (10, 50), n=379
Radar plot including NRS Loudness, NRS Awareness, NRS Annoyance, FTQ , TCS, HQ, PSQ, HADS_A, HADS, WHO Physical, Psychological, Social, and Environmental subscale, and THI (area)	2356.45 (1198.67, 4114.32), n=364

Appendix 4.5. Descriptive statistics for variables only available in the ESIT dataset

All	200
Living with partner (no/yes)	43/149
Employment (employed/unemployed/own job/retired/sick leave/parental leave/student/sabbatical/housework/other/do not know)	56/1/17/99/4/1/4/0/4/5/0
Cigarettes (number per day)	0 (0, 0), n=192
Coffee (cups per day)	2 (1, 3), n=191
Exercise hours per week (2/2-4/5-7/more than 7)	35/73/44/37
Fruit consumption (never/less than 3 times per month/1-6 times per week/once per day/2-3 times per day/more than 4 times per day)	0/13/48/47/69/13
Vegetable consumption (never/less than 3 times per month/1-6 times per week/once per day/2-3 times per day/more than 4 times per day)	1/3/40/57/71/17
Mobile phone use for calls (no use/less than 1 hour per month/around 1 hour per month/2-3 hours per month/around 1 hour per week/2-6 hours per week/1 hour per day/more than 1 hour per day)	14/75/28/20/28/15/7/4
Headphone use for music (no use/less than 1 hour per month/around 1 hour per month/2-3 hours per month/around 1 hour per week/2-6 hours per week/1 hour per day/more than 1 hour per day)	105/27/6/16/9/11/5/8
Sleep per day hours (six or less/seven/eight or more)	81/67/45

Appendix 4.6. Descriptive statistics for variables only available in the BRC dataset

All	205
Hearing threshold at 9 kHz (left ear) (dB HL)	62.16, 25.63, n=197
Hearing threshold at 11.2 kHz (left ear) (dB HL)	75 (55, 85), n=197
Hearing threshold at 9 kHz (right ear) (dB HL)	65 (40, 80), n=197
Hearing threshold at 11.2 kHz (right ear) (dB HL)	75 (60, 85), n=197
Beck Depression Inventory (0-63)	2 (0, 4), n=185
Beck Anxiety Inventory (0-63)	4 (2, 10), n=183
Wax removal at tinnitus onset (no/yes)	192/4
Other (than loud blast of sound, whiplash, change in hearing, stress, or head trauma) event at tinnitus onset (no/yes)	163/33
Comfortable level of a 0.5 kHz pure tone (dB SPL)	55.01, 15.55, n=191
Comfortable level of a 5 kHz pure tone (dB SPL)	34 (20, 50), n=191
Tinnitus loudness matching at 0.5 kHz (dB SPL)	64.25 (54, 71.88), n=190
Tinnitus loudness matching at 1 kHz (dB SPL)	67.5 (53.5, 74.88), n=190
Tinnitus loudness matching at 2 kHz (dB SPL)	62.5 (49.5, 73.38), n=190
Tinnitus loudness matching at 3 kHz (dB SPL)	52.48, 19.53, n=189
Tinnitus loudness matching at 4 kHz (dB SPL)	47.22, 19.41, n=186
Tinnitus loudness matching at 5 kHz (dB SPL)	41.97, 19, n=184
Tinnitus loudness matching at 6 kHz (dB SPL)	40.41, 19.61, n=184
Tinnitus loudness matching at 7 kHz (dB SPL)	38.15, 20.17, n=182
Tinnitus loudness matching at 8 kHz (dB SPL)	35.78, 19.56, n=175
Tinnitus loudness matching at 10 kHz (dB SPL)	27.5 (13.5, 40.75), n=151
Tinnitus pitch matching at 0.5 kHz (likeness scale 0-100)	12 (3.33, 27), n=191
Tinnitus pitch matching at 1 kHz (likeness scale 0-100)	22.33 (8, 38), n=191
Tinnitus pitch matching at 2 kHz (likeness scale 0-100)	33.67 (17.17, 48), n=191
Tinnitus pitch matching at 3 kHz (likeness scale 0-100)	40.33 (23.5, 54.67), n=191
Tinnitus pitch matching at 4 kHz (likeness scale 0-100)	50.33 (33, 63.67), n=189
Tinnitus pitch matching at 5 kHz (likeness scale 0-100)	57 (37.17, 69.83), n=191
Tinnitus pitch matching at 6 kHz (likeness scale 0-100)	56.83 (45, 72.5), n=190
Tinnitus pitch matching at 7 kHz (likeness scale 0-100)	57.66 (36.92, 75.17), n=188
Tinnitus pitch matching at 8 kHz (likeness scale 0-100)	63.67 (43.67, 78.33), n=181

Tinnitus pitch matching at 10 kHz (likeness scale 0-100)	57.67 (39.5, 81), n=159
Tinnitus reduced by music or environmental sounds (no/yes)	64/113
Tinnitus affected by head or neck movement or arms/hands or head touch (no/yes)	153/36
Tinnitus during the day affected by sleep at night (no/yes)	129/44
Tinnitus Handicap Questionnaire total score (0-100)	37.74 (24.74, 50.81), n=193

Appendix 4.7. Comparison of people with and without tinnitus from the STOP dataset

	All	No Tinnitus	Tinnitus	P value
All	657	262	395	-
Age (y)	50 (40, 62), n=656	48 (39, 60), n=262	52 (40.25, 64), n=394	0.006
Sex (female/male)*	361/293	165/95	196/198	<0.001
Handedness (both/left/right)	5/48/604	1/20/241	4/28/363	0.779
Family history of tinnitus (or hearing loss for ESIT) (no/yes)	488/169	212/50	276/119	0.001
Height (cm)	172 (167, 180), n=655	171 (166, 178), n=262	173 (167, 182), n=393	0.009
Weight (kg)	72 (64, 83.25), n=656	70 (63, 80), n=262	74 (64.25, 84.75), n=394	0.009
BMI (kg/m^2)	24.01 (21.93, 26.31), n=654	23.67 (21.72, 26.23), n=262	24.1 (22.15, 26.41), n=392	0.075
Education (lower/higher)	165/492	57/205	108/287	0.118
Alcohol (number of drinks per week)	3 (1, 5), n=656	2 (1, 5), n=261	3 (1, 6), n=395	0.418
Smoking (no/yes)	404/253	168/94	236/159	0.287
Ever worked shifts (no/yes)	436/218	167/94	269/124	0.237
Night work (no/yes)	422/230	169/92	253/138	1
MaxDiffExt (dB)	-4.52 (-11.95, 7.6), n=657	-1.47 (-10.39, 8.03), n=262	-6.03 (-12.6, 7.42), n=395	0.044
Mean hearing threshold both ears (dB HL)	6.04 (1.28, 13.55), n=657	3.72 (0.66, 8.67), n=262	7.91 (2.23, 16.43), n=395	<0.001
Hearing threshold at 0.125 kHz (left ear) (dB HL)	2.1 (-1.1, 6.3), n=657	2.3 (-0.95, 5.52), n=262	2.1 (-1.2, 6.55), n=395	0.941
Hearing threshold at 0.25 kHz (left ear) (dB HL)	-0.8 (-3.8, 4.3), n=657	-1.1 (-3.75, 3.95), n=262	-0.3 (-3.7, 4.8), n=395	0.339
Hearing threshold at 0.5 kHz (left ear) (dB HL)	-0.3 (-4.1, 6.5), n=657	-1 (-4.6, 5), n=262	0 (-3.7, 7), n=395	0.039
Hearing threshold at 0.75 kHz (left ear) (dB HL)	0.3 (-2.95, 5.65), n=657	-0.62 (-3.15, 5.09), n=262	1.1 (-2.6, 6.75), n=395	0.014
Hearing threshold at 1 kHz (left ear) (dB HL)	0.8 (-2.6, 7.3), n=657	0 (-3.1, 5.95), n=262	1.1 (-2.3, 8), n=395	0.019
Hearing threshold at 1.5 kHz (left ear) (dB HL)	1.95 (-1.9, 9.1), n=657	1.17 (-2.25, 7.3), n=262	2.85 (-1.5, 10.93), n=395	0.003
Hearing threshold at 2 kHz (left ear) (dB HL)	3.5 (-2.6, 13), n=657	2.5 (-2.8, 8.07), n=262	4.5 (-2.3, 16.15), n=395	0.006

Hearing threshold at 3 kHz (left ear) (dB HL)	8.6 (1.5, 20.1), n=657	5.5 (0.3, 14.25), n=262	11.1 (2.55, 24.95), n=395	<0.001
Hearing threshold at 4 kHz (left ear) (dB HL)	10 (1.3, 26.1), n=657	5.3 (-0.25, 16.08), n=262	15.3 (3.8, 32.3), n=395	<0.001
Hearing threshold at 6 kHz (left ear) (dB HL)	13.1 (3.6, 31), n=657	8 (1.3, 18.75), n=262	20.3 (6.3, 37.95), n=395	<0.001
Hearing threshold at 8 kHz (left ear) (dB HL)	19 (6.3, 43.8), n=657	11.55 (4.5, 27.58), n=262	29.3 (8.55, 51.5), n=395	<0.001
Hearing threshold at 10 kHz (left ear) (dB HL)	25.6 (6.3, 56), n=657	14.7 (3, 36.17), n=262	41.1 (10.6, 60.4), n=395	<0.001
Hearing threshold at 12.5 kHz (left ear) (dB HL)	47.1 (12.5, 69), n=657	30.95 (7.88, 58.12), n=262	56.1 (25.05, 72.55), n=395	<0.001
Hearing threshold at 14 kHz (left ear) (dB HL)	58.6 (32.1, 110), n=657	48.9 (24.38, 68.38), n=262	64.8 (41.6, 110), n=395	<0.001
Hearing threshold at 0.125 kHz (right ear) (dB HL)	3.5 (0.6, 7), n=657	3.6 (1.02, 7.07), n=262	3.5 (0.1, 7), n=395	0.424
Hearing threshold at 0.25 kHz (right ear) (dB HL)	0.5 (-2.3, 5.3), n=657	0.55 (-2, 5.5), n=262	0.5 (-2.6, 5.05), n=395	0.713
Hearing threshold at 0.5 kHz (right ear) (dB HL)	0.1 (-3.5, 6.5), n=657	0.2 (-3.6, 6.07), n=262	0.1 (-3.3, 6.8), n=395	0.672
Hearing threshold at 0.75 kHz (right ear) (dB HL)	0.85 (-2.5, 6.05), n=657	0.48 (-2.24, 5.83), n=262	0.95 (-2.55, 6.2), n=395	0.486
Hearing threshold at 1 kHz (right ear) (dB HL)	1.3 (-2.6, 7), n=657	1.3 (-2.3, 6.25), n=262	1.1 (-2.6, 8.6), n=395	0.472
Hearing threshold at 1.5 kHz (right ear) (dB HL)	2.05 (-1.45, 9.15), n=657	1.75 (-1.49, 6.95), n=262	2.45 (-1.35, 10.18), n=395	0.11
Hearing threshold at 2 kHz (right ear) (dB HL)	3.1 (-1.6, 11.6), n=657	2.2 (-1.8, 9.52), n=262	3.8 (-1.4, 12.9), n=395	0.037
Hearing threshold at 3 kHz (right ear) (dB HL)	6 (-0.6, 17.6), n=657	3.4 (-1.25, 11.4), n=262	8 (0.2, 22.1), n=395	<0.001
Hearing threshold at 4 kHz (right ear) (dB HL)	8.5 (0.1, 22.5), n=657	4.9 (-1.1, 13.45), n=262	12.6 (1.3, 28.7), n=395	<0.001
Hearing threshold at 6 kHz (right ear) (dB HL)	12.1 (2.8, 27.5), n=657	7.6 (1.52, 15.52), n=262	17.6 (3.85, 36), n=395	<0.001
Hearing threshold at 8 kHz (right ear) (dB HL)	16.1 (5.5, 40.6), n=657	10.8 (3.58, 25.6), n=262	24.1 (7.7, 48.2), n=395	<0.001
Hearing threshold at 10 kHz (right ear) (dB HL)	24.6 (6.6, 55), n=657	12.85 (4.1, 37.4), n=262	37.8 (9.3, 58.7), n=395	<0.001
Hearing threshold at 12.5 kHz (right ear) (dB HL)	48.5 (13.5, 67), n=657	31.95 (8.03, 60.03), n=262	54.1 (18.2, 70), n=395	<0.001

Hearing threshold at 14 kHz (right ear) (dB HL)	57.3 (28.6, 73), n=657	49.8 (22.38, 67.25), n=262	62.1 (34.4, 110), n=395	<0.001
Abnormal tympanogram left ear (no/yes)	484/122	198/42	286/80	0.214
Tympanometric Peak Pressure (TPP) left ear (daPA)	8 (1, 15), n=593	8 (1, 14.5), n=235	7 (1, 15), n=358	0.97
Static Admittance (SA) left ear (mmho)	0.7 (0.45, 1.07), n=591	0.72 (0.51, 1.14), n=234	0.66 (0.44, 1.03), n=357	0.044
Tympanometric Width (TW) left ear (daPa)	70 (55.75, 91), n=584	70 (54, 89), n=232	70 (58, 93), n=352	0.118
Ear Canal Volume (ECVL) left ear (cm3)	1.34 (1.12, 1.59), n=609	1.33 (1.12, 1.56), n=241	1.34 (1.12, 1.6), n=368	0.751
Abnormal tympanogram right ear (no/yes)	501/108	205/40	296/68	0.517
Tympanometric Peak Pressure (TPP) right ear (daPA)	8 (1, 14), n=602	8 (1, 14), n=243	7 (1, 14), n=359	0.749
Static Admittance (SA) right ear (mmho)	0.69 (0.48, 1.03), n=594	0.69 (0.53, 1), n=241	0.69 (0.46, 1.03), n=353	0.379
Tympanometric Width (TW) right ear (daPa)	70 (56.25, 90), n=586	67 (50.75, 84), n=236	71 (58, 94), n=350	0.008
Ear Canal Volume (ECVL) right ear (cm3)	1.39 (1.18, 1.66), n=610	1.39 (1.21, 1.64), n=246	1.39 (1.14, 1.68), n=364	0.987
Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)	80 (70, 85), n=535	80 (75, 90), n=186	75 (70, 85), n=349	<0.001
Uncomfortable loudness level at 0.25 kHz (left ear) (dB HL)	85 (75, 90), n=547	85 (80, 95), n=190	80 (70, 90), n=357	<0.001
Uncomfortable loudness level at 0.5 kHz (left ear) (dB HL)	80 (75, 90), n=556	85 (80, 95), n=201	80 (70, 90), n=355	<0.001
Uncomfortable loudness level at 1 kHz (left ear) (dB HL)	80 (75, 90), n=570	85 (80, 95), n=211	80 (70, 90), n=359	<0.001
Uncomfortable loudness level at 2 kHz (left ear) (dB HL)	80 (70, 90), n=579	85 (80, 90), n=216	80 (65, 87.5), n=363	<0.001
Uncomfortable loudness level at 3 kHz (left ear) (dB HL)	80 (70, 90), n=552	85 (75, 90), n=203	80 (65, 90), n=349	<0.001
Uncomfortable loudness level at 4 kHz (left ear) (dB HL)	80 (70, 90), n=552	85 (75, 90), n=204	80 (65, 90), n=348	<0.001
Uncomfortable loudness level at 6 kHz (left ear) (dB HL)	80 (65, 85), n=540	80 (70, 88.75), n=194	75 (65, 85), n=346	0.003
Uncomfortable loudness level at 8 kHz (left ear) (dB HL)	75 (65, 85), n=509	80 (70, 86.25), n=184	75 (60, 85), n=325	0.025
Uncomfortable loudness level at 0.125 kHz (right ear) (dB HL)	80 (75, 90), n=510	85 (80, 90), n=169	80 (70, 85), n=341	<0.001
Uncomfortable loudness level at 0.25 kHz (right ear) (dB HL)	85 (75, 92.5), n=547	90 (80, 95), n=193	80 (75, 90), n=354	<0.001
Uncomfortable loudness level at 0.5 kHz (right ear) (dB HL)	85 (75, 90), n=551	85 (80, 95), n=193	80 (70, 90), n=358	<0.001

Uncomfortable loudness level at 1 kHz (right ear) (dB HL)	80 (75, 90), n=575	85 (80, 95), n=209	80 (70, 85), n=366	<0.001
Uncomfortable loudness level at 2 kHz (right ear) (dB HL)	80 (70, 90), n=580	85 (75, 95), n=216	80 (70, 85), n=364	<0.001
Uncomfortable loudness level at 3 kHz (right ear) (dB HL)	80 (70, 90), n=571	85 (75, 90), n=210	75 (65, 85), n=361	<0.001
Uncomfortable loudness level at 4 kHz (right ear) (dB HL)	80 (70, 90), n=573	85 (75, 90), n=210	80 (65, 87.5), n=363	<0.001
Uncomfortable loudness level at 6 kHz (right ear) (dB HL)	80 (65, 85), n=552	80 (75, 90), n=199	75 (65, 85), n=353	<0.001
Uncomfortable loudness level at 8 kHz (right ear) (dB HL)	75 (65, 85), n=517	77.5 (70, 85), n=190	75 (60, 85), n=327	0.004
Speech in noise hearing assessment - word score (left ear) (0-100)	74 (66, 80), n=641	78 (72, 82), n=254	72 (62, 78), n=387	<0.001
Speech in noise hearing assessment - phoneme score (left ear) (0-100)	89.3 (84.7, 92), n=644	90.7 (87.3, 93.3), n=256	88.7 (82, 92), n=388	<0.001
Speech in noise hearing assessment - word score (right) (0-100)	74 (66, 80), n=641	78 (72, 82), n=254	72 (64, 78), n=387	<0.001
Speech in noise hearing assessment - phoneme score (right ear) (0-100)	89.3 (84.7, 92), n=643	90.7 (87.65, 92.7), n=255	88 (81.3, 91.3), n=388	<0.001
DPOAE at 996 Hz (left ear) (signal to noise ratio)	12.5 (8, 18.3), n=649	13.6 (8.3, 18.78), n=262	11.9 (7.8, 17.55), n=387	0.116
DPOAE at 1074 Hz (left ear) (signal to noise ratio)	13.7 (8.4, 20), n=649	14.5 (9.2, 20.35), n=262	13.1 (8.2, 19.55), n=387	0.073
DPOAE at 1152 Hz (left ear) (signal to noise ratio)	14.6 (9.3, 20.8), n=649	15.55 (10.15, 21.65), n=262	13.7 (8.9, 19.95), n=387	0.011
DPOAE at 1230 Hz (left ear) (signal to noise ratio)	15.6 (9.8, 21.7), n=649	16.75 (10.62, 23), n=262	14.9 (9.25, 20.65), n=387	0.009
DPOAE at 1318 Hz (left ear) (signal to noise ratio)	17 (10.3, 21.8), n=649	17.3 (11.4, 22.28), n=262	16.4 (9.7, 21.35), n=387	0.051
DPOAE at 1416 Hz (left ear) (signal to noise ratio)	17.7 (10.9, 22.7), n=649	18.45 (11.2, 23.25), n=262	17.3 (10.6, 22.1), n=387	0.165
DPOAE at 1513 Hz (left ear) (signal to noise ratio)	18.4 (12.4, 23.8), n=649	18.8 (13.43, 24.1), n=262	18.2 (12, 23.55), n=387	0.067
DPOAE at 1621 Hz (left ear) (signal to noise ratio)	19.1 (12.3, 24.4), n=649	19.1 (13.8, 25.3), n=262	19 (11.25, 23.8), n=387	0.143
DPOAE at 1738 Hz (left ear) (signal to noise ratio)	19.3 (13.4, 24.7), n=649	19.8 (14.15, 25.35), n=262	19.1 (13.15, 24.5), n=387	0.077

DPOAE at 1865 Hz (left ear) (signal to noise ratio)	18.2 (12.1, 23.4), n=649	18.45 (13.03, 23.7), n=262	17.8 (11.65, 22.9), n=387	0.089
DPOAE at 2001 Hz (left ear) (signal to noise ratio)	18.8 (12.7, 23.5), n=649	19.6 (13.8, 23.9), n=262	18 (11.9, 23.1), n=387	0.055
DPOAE at 2148 Hz (left ear) (signal to noise ratio)	19 (12.5, 23.9), n=649	19.4 (13.7, 24.5), n=262	18 (12.1, 23), n=387	0.043
DPOAE at 2294 Hz (left ear) (signal to noise ratio)	19 (12.9, 24), n=649	19.6 (14.43, 24.1), n=262	18.5 (11.8, 23.8), n=387	0.038
DPOAE at 2460 Hz (left ear) (signal to noise ratio)	19.4 (12.9, 24.5), n=649	20.55 (14.53, 24.87), n=262	18.5 (12.45, 24.35), n=387	0.035
DPOAE at 2636 Hz (left ear) (signal to noise ratio)	19.1 (13.4, 23.8), n=649	19.75 (15.22, 25), n=262	18.5 (12.45, 23.45), n=387	0.002
DPOAE at 2832 Hz (left ear) (signal to noise ratio)	19.4 (13.9, 25), n=649	20.85 (15.3, 25.98), n=262	18.2 (12.8, 23.9), n=387	<0.001
DPOAE at 3027 Hz (left ear) (signal to noise ratio)	19.5 (13.2, 24.6), n=649	20.85 (13.95, 25.48), n=262	18.6 (13.05, 23.75), n=387	0.003
DPOAE at 3251 Hz (left ear) (signal to noise ratio)	17.3 (12.3, 23.5), n=649	18.85 (13.83, 24.78), n=262	16.1 (11.4, 22.5), n=387	<0.001
DPOAE at 3486 Hz (left ear) (signal to noise ratio)	16.9 (11.4, 22.4), n=649	18.45 (12.55, 23.67), n=262	15.9 (10.85, 20.65), n=387	<0.001
DPOAE at 3730 Hz (left ear) (signal to noise ratio)	17.7 (11.9, 23.1), n=649	19.4 (14.72, 24.58), n=262	16 (10.3, 22.15), n=387	<0.001
DPOAE at 4003 Hz (left ear) (signal to noise ratio)	17.5 (12.5, 23.1), n=649	19.7 (15.25, 25.17), n=262	15.9 (11, 21.7), n=387	<0.001
DPOAE at 4287 Hz (left ear) (signal to noise ratio)	17.3 (12.1, 23.2), n=649	19.4 (14.35, 24.5), n=262	15.2 (11, 21.4), n=387	<0.001
DPOAE at 4599 Hz (left ear) (signal to noise ratio)	17.4 (11.6, 23.2), n=649	19.9 (14.72, 25.1), n=262	15.6 (9.7, 21.6), n=387	<0.001
DPOAE at 4921 Hz (left ear) (signal to noise ratio)	17.7 (10.1, 24), n=649	21.35 (14.98, 25.2), n=262	14.6 (7.75, 22.55), n=387	<0.001
DPOAE at 5273 Hz (left ear) (signal to noise ratio)	17.4 (9.5, 24.9), n=649	21.05 (14.12, 27), n=262	14.4 (8, 22.2), n=387	<0.001
DPOAE at 5654 Hz (left ear) (signal to noise ratio)	14.8 (7.9, 22.8), n=649	18.15 (11.83, 25.4), n=262	12.1 (6.55, 20.55), n=387	<0.001

DPOAE at 6064 Hz (left ear) (signal to noise ratio)	13.7 (7.7, 20.8), n=649	15.9 (11.15, 23.7), n=262	11.8 (6.3, 18.35), n=387	<0.001
DPOAE at 6494 Hz (left ear) (signal to noise ratio)	10.8 (7.1, 17.4), n=649	12.8 (8.4, 20.67), n=262	10 (6.5, 15.7), n=387	<0.001
DPOAE at 6962 Hz (left ear) (signal to noise ratio)	9.8 (6.6, 14.2), n=649	10.75 (7.1, 16.32), n=262	9.4 (6.25, 13.45), n=387	<0.001
DPOAE at 7460 Hz (left ear) (signal to noise ratio)	8.2 (4.2, 12.1), n=649	8.75 (5.2, 12.88), n=262	7.8 (3.4, 11.6), n=387	0.011
DPOAE at 7998 Hz (left ear) (signal to noise ratio)	8.4 (4.9, 11.8), n=649	8.9 (6.03, 12.1), n=262	8.2 (4.4, 11.75), n=387	0.05
DPOAE at 8574 Hz (left ear) (signal to noise ratio)	10.2 (6.8, 15.8), n=649	10.2 (6.8, 16.17), n=262	10.3 (6.9, 15.4), n=387	0.702
DPOAE at 9189 Hz (left ear) (signal to noise ratio)	12.5 (8.7, 16.7), n=649	12.75 (8.15, 16.8), n=262	12.3 (9.1, 16.6), n=387	0.985
DPOAE at 9853 Hz (left ear) (signal to noise ratio)	14.5 (10.9, 19), n=649	14.65 (11.05, 19.78), n=262	14.5 (10.9, 18.35), n=387	0.251
DPOAE at 996 Hz (right ear) (signal to noise ratio)	12.2 (8.1, 18.1), n=649	13.1 (8.4, 18.08), n=262	11.6 (7.85, 18.05), n=387	0.13
DPOAE at 1074 Hz (right ear) (signal to noise ratio)	14.4 (9.1, 20.2), n=649	14.75 (9.4, 20.28), n=262	14.2 (8.4, 19.65), n=387	0.164
DPOAE at 1152 Hz (right ear) (signal to noise ratio)	15.3 (9.4, 21.7), n=649	15.65 (9.72, 22.08), n=262	15.2 (9.15, 21.55), n=387	0.33
DPOAE at 1230 Hz (right ear) (signal to noise ratio)	16.3 (10, 22.2), n=649	16.65 (10.33, 22.8), n=262	16.1 (9.6, 21.95), n=387	0.4
DPOAE at 1318 Hz (right ear) (signal to noise ratio)	17.3 (11, 22.7), n=649	18.3 (12.53, 22.95), n=262	16.9 (10.6, 22.55), n=387	0.046
DPOAE at 1416 Hz (right ear) (signal to noise ratio)	18.6 (11.7, 24.2), n=649	18.6 (11.83, 23.9), n=262	18.7 (11.55, 24.5), n=387	0.932
DPOAE at 1513 Hz (right ear) (signal to noise ratio)	19.3 (12.4, 24.6), n=649	19.4 (14.03, 25.17), n=262	19.3 (11.2, 24.3), n=387	0.159
DPOAE at 1621 Hz (right ear) (signal to noise ratio)	19.9 (13.9, 25.2), n=649	20.8 (14.8, 25.48), n=262	19.4 (12.65, 25.05), n=387	0.08
DPOAE at 1738 Hz (right ear) (signal to noise ratio)	20.6 (13.9, 25.6), n=649	21.25 (15.1, 26.25), n=262	19.8 (13.6, 25.35), n=387	0.13
DPOAE at 1865 Hz (right ear) (signal to noise ratio)	19.6 (12.8, 25.1), n=649	19.8 (13.03, 25.8), n=262	19.4 (12.65, 24.3), n=387	0.222

DPOAE at 2001 Hz (right ear) (signal to noise ratio)	20.4 (14.6, 25), n=649	21.05 (15.43, 25.48), n=262	20.1 (14.1, 24.65), n=387	0.073
DPOAE at 2148 Hz (right ear) (signal to noise ratio)	20.1 (13.1, 24.9), n=649	21.6 (14.33, 25.8), n=262	19.6 (12, 24.1), n=387	0.006
DPOAE at 2294 Hz (right ear) (signal to noise ratio)	19.7 (14.6, 24.8), n=649	21.3 (15.05, 25.45), n=262	18.9 (14.15, 24.1), n=387	0.009
DPOAE at 2460 Hz (right ear) (signal to noise ratio)	20.4 (14.2, 25), n=649	21.1 (15.8, 25.37), n=262	19.6 (13.4, 24.55), n=387	0.054
DPOAE at 2636 Hz (right ear) (signal to noise ratio)	19.7 (14.3, 24.9), n=649	20.35 (15.5, 25.5), n=262	19.3 (12.9, 24.5), n=387	0.032
DPOAE at 2832 Hz (right ear) (signal to noise ratio)	20.4 (14.6, 25), n=649	21.1 (16.6, 26), n=262	19.8 (13.3, 24.45), n=387	0.002
DPOAE at 3027 Hz (right ear) (signal to noise ratio)	20.1 (14, 24.6), n=649	21 (15.7, 25.78), n=262	19.4 (12.95, 24.05), n=387	0.004
DPOAE at 3251 Hz (right ear) (signal to noise ratio)	18.4 (13, 23.2), n=649	19.6 (14.95, 23.98), n=262	17.4 (11.45, 22.8), n=387	<0.001
DPOAE at 3486 Hz (right ear) (signal to noise ratio)	17.8 (12.6, 23.1), n=649	19.3 (14.83, 24.1), n=262	16.3 (11.7, 22.55), n=387	<0.001
DPOAE at 3730 Hz (right ear) (signal to noise ratio)	18.2 (12.4, 23.7), n=649	19.75 (14.17, 24.78), n=262	17 (11.15, 22.7), n=387	<0.001
DPOAE at 4003 Hz (right ear) (signal to noise ratio)	18 (12.6, 23.7), n=649	19.5 (14.1, 24.6), n=262	16.7 (11.15, 22.5), n=387	<0.001
DPOAE at 4287 Hz (right ear) (signal to noise ratio)	18.1 (12.8, 23.8), n=649	19.95 (14.8, 24.5), n=262	16.8 (11.4, 22.75), n=387	<0.001
DPOAE at 4599 Hz (right ear) (signal to noise ratio)	17.8 (11.6, 23.9), n=649	20.7 (14.03, 25.28), n=262	16.2 (10.75, 23), n=387	<0.001
DPOAE at 4921 Hz (right ear) (signal to noise ratio)	18.6 (12, 24.7), n=649	21.15 (14.8, 26.15), n=262	16.4 (9.55, 23.1), n=387	<0.001
DPOAE at 5273 Hz (right ear) (signal to noise ratio)	17.1 (10.5, 25.6), n=649	21.8 (13.5, 27.17), n=262	15 (8.65, 23.35), n=387	<0.001
DPOAE at 5654 Hz (right ear) (signal to noise ratio)	15.8 (9.3, 24.1), n=649	18.9 (11.35, 25.17), n=262	13.7 (7.95, 22.4), n=387	<0.001
DPOAE at 6064 Hz (right ear) (signal to noise ratio)	14.5 (8.6, 22.1), n=649	17.5 (10.62, 24.37), n=262	12.5 (7.4, 20.2), n=387	<0.001

DPOAE at 6494 Hz (right ear) (signal to noise ratio)	11.9 (7.1, 19), n=649	14.35 (9.25, 21.48), n=262	10.7 (6.2, 16.85), n=387	<0.001
DPOAE at 6962 Hz (right ear) (signal to noise ratio)	10.4 (7.4, 14.6), n=649	11.3 (8.3, 16.7), n=262	9.6 (7.1, 13.65), n=387	<0.001
DPOAE at 7460 Hz (right ear) (signal to noise ratio)	8.6 (4.8, 12.5), n=649	9.1 (6.43, 13.3), n=262	8.3 (3.7, 11.95), n=387	0.005
DPOAE at 7998 Hz (right ear) (signal to noise ratio)	9.2 (6.1, 13), n=649	9.25 (6.23, 12.97), n=262	9 (6.1, 13), n=387	0.982
DPOAE at 8574 Hz (right ear) (signal to noise ratio)	10.5 (7.4, 16.1), n=649	10.3 (7.12, 15.7), n=262	10.6 (7.45, 16.65), n=387	0.32
DPOAE at 9189 Hz (right ear) (signal to noise ratio)	12.8 (9.2, 16.5), n=649	12.6 (9.03, 16.2), n=262	13 (9.35, 16.7), n=387	0.398
DPOAE at 9853 Hz (right ear) (signal to noise ratio)	14.9 (11.1, 18.7), n=649	14.5 (10.7, 18.58), n=262	15 (11.15, 18.85), n=387	0.361
Hearing aid use (no/yes)	615/42	258/4	357/38	<0.001
Hyperacusis questionnaire score (0-42)	13 (7, 20), n=656	11 (6, 17), n=262	15 (9, 23), n=394	<0.001
Any ear condition (no/yes)	441/216	192/70	249/146	0.007
Acoustic trauma (no/yes)	643/14	260/2	383/12	0.055
Barotrauma (no/yes)	655/2	262/0	393/2	0.52
Presbycusis (no/yes)	616/41	255/7	361/34	0.002
Sudden hearing loss (no/yes)	646/11	259/3	387/8	0.539
Other hearing loss (no/yes)	614/43	253/9	361/34	0.009
Meniere’s disease (no/yes)	650/7	261/1	389/6	0.253
Acoustic neuroma (no/yes)	656/1	262/0	394/1	1
Acute Otitis (no/yes)	608/49	242/20	366/29	0.881
Serous Otitis (no/yes)	644/13	258/4	386/9	0.579
Chronic Otitis (no/yes)	650/7	259/3	391/4	1
Otosclerosis (no/yes)	651/6	262/0	389/6	0.086
Problem with external sounds (small/moderate/big or very big)	511/113/33	236/23/3	275/90/30	<0.001
Self-reported hearing difficulty (slight or no difficulty/moderate difficulty/severe difficulty/total loss)	405/169/67/3	204/41/9/1	201/128/58/2	<0.001
Any hearing device (no/yes)	613/44	258/4	355/40	<0.001
Combination device (hearing aid and sound generator) (no/yes)	655/2	262/0	393/2	0.52
Headaches (no/yes)*	533/124	225/37	308/87	0.011
Vertigo (no/yes)*	319/338	153/109	166/229	<0.001
TMJ disorder (no/yes)*	653/4	262/0	391/4	0.155
Any procedure (no/yes)	370/287	145/117	225/170	0.689

Ear surgery (no/yes)	628/29	253/9	375/20	0.438
Dental surgery (no/yes)	458/199	181/81	277/118	0.795
Neurosurgery (no/yes)	651/6	262/0	389/6	0.086
Lumbar puncture (no/yes)	637/20	255/7	382/13	0.818
Chemotherapy (no/yes)	649/8	257/5	392/3	0.277
Head and neck radiotherapy (no/yes)	650/7	257/5	393/2	0.122
Any pain syndromes (no/yes)	433/224	187/75	246/149	0.019
Neck pain (no/yes)	527/130	221/41	306/89	0.035
Ear pain (no/yes)	641/16	259/3	382/13	0.119
TMJ pain (no/yes)	615/42	254/8	361/34	0.005
Face pain (no/yes)	645/12	262/0	383/12	0.002
Any diagnosed condition (no/yes)	275/382	127/135	148/247	0.006
Dental problems (no/yes)	616/41	249/13	367/28	0.324
Meningitis (no/yes)	648/9	259/3	389/6	1
Epilepsy (no/yes)	653/4	261/1	392/3	1
Stroke (no/yes)	655/2	261/1	394/1	1
Other neurologic condition (no/yes)	649/8	260/2	389/6	0.487
Anxiety (no/yes)	600/57	245/17	355/40	0.12
Depression (no/yes)	582/75	240/22	342/53	0.06
Emotional trauma (no/yes)	637/20	257/5	380/15	0.246
Stress (no/yes)	599/58	244/18	355/40	0.162
Problem falling asleep (no/yes)	594/63	245/17	349/46	0.03
Problem staying asleep (no/yes)	581/76	239/23	342/53	0.081
Low BP (no/yes)	614/43	247/15	367/28	0.524
High BP (no/yes)	573/84	239/23	334/61	0.012
Thyroid disorder (no/yes)	614/43	245/17	369/26	1
Diabetes (no/yes)	646/11	258/4	388/7	1
High cholesterol (no/yes)	622/35	252/10	370/25	0.214
Rheumatoid arthritis (no/yes)	654/3	260/2	394/1	0.567
Chronic sinusitis (no/yes)	652/5	261/1	391/4	0.653
Nasal septum deviation (no/yes)	654/3	260/2	394/1	0.567
HIV (no/yes)	656/1	262/0	394/1	1
Lyme disease (no/yes)	578/79	235/27	343/52	0.327
Anaemia (no/yes)	627/30	249/13	378/17	0.706
Balance disorders (no/yes)	638/19	256/6	382/13	0.635
Acid/gastroesophageal reflux (no/yes)	607/50	247/15	360/35	0.176
Hospital Anxiety Depression Scale for Anxiety (0-21)	5 (2, 8), n=657	4 (2, 7), n=262	5 (3, 8), n=395	0.006
Hospital Anxiety Depression Scale for Depression (0-21)	2 (1, 4), n=657	2 (1, 4), n=262	2 (1, 5), n=395	0.004
World Health Organization's Quality of Life Physical subscale (4-20)	17 (15, 18), n=650	17 (15, 18), n=259	17 (15, 18), n=391	<0.001

World Health Organization's Quality of Life Psychological subscale (4-20)	16 (14, 17), n=654	16 (15, 17), n=262	15 (14, 17), n=392	0.018
World Health Organization's Quality of Life Social subscale (4-20)	15 (13, 16), n=657	15 (13, 17), n=262	15 (13, 16), n=395	0.084
World Health Organization's Quality of Life Environmental subscale (4-20)	17 (15, 18), n=657	17 (16, 18), n=262	17 (15, 18), n=395	0.015
Perceived Stress Questionnaire (0-1)	0.29 (0.17, 0.42), n=657	0.26 (0.16, 0.4), n=262	0.31 (0.2, 0.44), n=395	<0.001

*Variable assessed using the TSCHQ question for BRC dataset and the ESIT-SQ question for the STOP and ESIT datasets.

Appendix 4.8. Variables with certain percentages of missing data

	>30% missing	20-30% missing	10-20% missing	5-10% missing
STOP dataset	RI_Char, AgeAtOnset, AgeBothOnset, TinnitusDuration, BothTinnitusDuration, Gap_Tinn_BothTinn	TinnitusPitch, PMF_Threshold, Loudness_Match_dBHL, MN_Threshold, Maskability, NoiseWorsens_tschq, temp_rel_hearing_difficulty	TW_L, SA_R, TW_R, U_125_L, U_500_L, U_3000_L, U_4000_L, U_6000_L, U_8000_L, U_125_R, U_250_R, U_6000_R, U_8000_R, subjectiveHearingProb_tschq, temp_rel_vertigo, temp_rel_ext_sounds_prob	Type_L, TPP_L, SA_L, ECV_L, Type_R, TPP_R, ECV_R, U_250_L, U_1000_L, U_2000_L, U_500_R, U_1000_R, U_2000_R, U_3000_R, U_4000_R, LoundnessRating, Pie.Size, tmj_disorder_tschq2_esitsq1, loudness_changes_overday_ordinal, LoudnessVar_fromdaytoday, Headache_pain_syndrome_2tschq_1esitsq, Vertigo_noyes_2tschq_1esitsq, temp_rel_dental_surg
BRC dataset		HearingAidNoYes, TinnTestLoudMatch10k	TinnLocalisation_2tschq_1esitsq, SoundReduces, SleepAffects, TinnTestLoudMatch8k, TinnTestPitchMatch10k	TinnFamilyHist_no_yes, NoiseWorsens_tschq, MovementAffects, BDI, BAI, Hyperacusis, TinnTestLoudMatch4k, TinnTestLoudMatch5k, TinnTestLoudMatch6k, TinnTestLoudMatch7k, TinnTestPitchMatch8k
ESIT dataset	Gender_f_m_1tschq_2esitsq	BothTinnitusDuration, Gap_Tinn_BothTinn, AgeBothOnset	temp_rel_vertigo, temp_rel_ear_surg, temp_rel_dental_surg, temp_rel_neurosurg, temp_rel_lumbar_punct, temp_rel_chemo, temp_rel_head_neck_r	Alcohol, ear_surg, dental_surg, neurosurg, lumbar_punct, chemo, head_neck_radio, Headache_pain_syndrome_1tschq_2esitsq, neck_pain, ear_pain, tmj_pain, face_pain, temp_rel_acoustic_trauma, temp_rel_barotrauma, temp_rel_presbycusis, temp_rel_sudden_hl, temp_rel_other_hl, temp_rel_meniere, temp_rel_neuroma, temp_rel_acute_otitis, temp_rel_serous_otitis, temp_rel_chronic_otitis,

	adio, temp_rel_hearing_difficulty, temp_rel_headache, temp_rel_neck_pain, temp_rel_ear_pain, temp_rel_tmj_pain, temp_rel_face_pain, quality, TinnitusDuration, AgeAtOnset	temp_rel_otosclerosis, temp_rel_ext_sounds_prob, temp_rel_tmj_disorder, temp_rel_dental, temp_rel_meningitis, temp_rel_epilpsy, temp_rel_stroke, temp_rel_other_neurologic, temp_rel_anxiety, temp_rel_depression, temp_rel_emot_trauma, temp_rel_stress, temp_rel_prob_falling_asleep, temp_rel_prob_staying_asleep, temp_rel_low_bp, temp_rel_high_bp, temp_rel_thyroid, temp_rel_diabetes, temp_rel_high_cholest, temp_rel_rheum_arthritis, temp_rel_chr_sinusitis, temp_rel_sept_deviation, temp_rel_hiv, temp_rel_lyme, temp_rel_anaemia, temp_rel_balance, temp_rel_cid_reflux, onset_aspirin, onset_painkillers, onset_antibiotics, onset_antidepressants, onset_quinine, onset_diuretics, onset_no_medication_at_onset, pitch, increased_by_silence, increased_by_low_sounds, increased_by_high_sounds, increased_by_head_movements, increased_by_jaw_movements, increased_by_head_neck_press, increased_by_nap, increased_by_poor_sleep, increased_by_driving, increased_by_stress, increased_by_relaxed, increased_by_alcohol, increased_by_coffee, increased_by_medications, increased_by_hearing_aid, increased_by_nothing, TinnLocalisation_1tschq_2esitsq, any_procedure, any_pain_syndromes, thoughts_of_onset_related_condition, rhythmicity_1tschq_2esitsq_4levels

Appendix 4.9. Summary of free-text responses

ESIT-SQ question	Summarised free-text responses ('other, please specify' response option)
A10 question 'Have you been diagnosed with any other ear condition?'	Auditory processing disorder (n=1), congenital ear malformations (n=3), diplacusis (n=1), excessive ear wax (n=5), hyperacusis (n=7), balance disorders including labyrinthitis and benign positional vertigo (n=8), inner ear damage after stapedectomy (n=1), mastoiditis (n=1), tinnitus (n=10), childhood ear infection (n=1), ear drum perforation (n=4), and retracted eardrum (n=1).
A11 question 'Have you ever undergone any of the following procedures?'	Appendicectomy (n=2), abdominal hernia surgery (n=2), hip interventions (n=3), bladder and urethra interventions (n=4), scans (n=3), breast cancer interventions including lumpectomy and radiotherapy (n=3), neck fracture intervention (n=1), caesarean section (n=1), chiropractic intervention (n=1), cholecystectomy (n=2), cochlear implantation (n=2), external ear intervention including wax removal (n=4), gromets (n=2), inner ear surgery (n=1), endoscopy (n=1), heart surgery (n=3), hysterectomy (n=3), knee surgery (n=1), eye surgery (n=2), plastic surgery for burns after motor accident (n=1), parotidectomy (n=1), soft palate surgery (n=1), prostatectomy (n=1), radiosurgery (n=1), dental surgery (n=1), orthopaedic surgery including tendon repair and bone fracture repair (n=5), sinus surgery (n=2), spine surgery (n=3), tonsillectomy (n=4), and nasal surgery (n=1).
A15 question 'Do you suffer from any of the following pain syndromes?'	Fibromyalgia (n=2), other musculoskeletal pain including arthritic pain and back pain (n=20), migraine (n=5), eye pain (n=1), hemifacial spasms (n=1), specific condition causing ear pain (n=2), face pain including trigeminal neuralgia (n=2), numbness in face and neck (n=1), flank pain (n=1), jaw pain (n=1), and sinus pain (n=1).
A16 question 'Do you have any of the following conditions that have been diagnosed by a clinician?'	Adenocarcinoma (n=1), allergies (n=1), heart problems including angina and arterial fibrillation (n=4), bereavement (n=1), arthritis (n=8), osteoporosis (n=5), aspartame poisoning (n=1), asthma (n=1), oesophagus Barret (n=1), body spasms (n=1), bulging disk (n=1), chronic fatigue syndrome (n=1), kidney problems (n=1), colitis (n=2), Ehler-Danlos syndrome (n=1), enlarged prostate (n=1), falls (n=1), fibromyalgia (n=1), haemangioma (n=1), hypoglycaemia (n=2), stress history (n=1), irritable bowel syndrome (n=1), headache syndromes including migraines (n=2), anaemia (n=2), hemifacial spasms and trigeminal neuralgia (n=1), sleep apnoea (n=3), retinal tear (n=1), sarcoidosis (n=2), pleural scaring (n=1), sinus problems (n=1), transient ischemic attack in the eye (n=1), tinnitus (n=1), and transverse myelitis (n=1).

B9 question ‘Was the initial onset of your tinnitus related to (you can choose more than one option)’	Specific noise exposures such as loud creaking noise in the middle of the night, visit in noisy pub, a single episode loud sound, gas turbine engines, loud music through earphones, and whistling in ear (n=6), engine test facility (n=1), serving in royal air force (n=2), war (n=1), hot water to clean ears during childhood (n=1), wax removal (n=1), ear infections (n=1), acoustic neuroma (n=2), cochlear implantation (n=1), CROS hearing aid (n=1), vertigo (n=2), dizziness (n=1), congenital deafness (n=1), hearing loss (n=2), Meniere's (n=1), long motorcycle trip (n=1), motor accident (n=2), frozen shoulder and neck (n=1), head and neck trauma (n=2), neck stress due to intense mountain biking (n=1), swimming (n=1), quinine (n=1), hip surgery (n=1), drugs after surgery (n=2), perhaps stress (n=1), perhaps whiplash (n=1), smoking a lot over one night (n=1), toothache (n=1), recreational drugs (n=1), bereavement (n=1), and do not know or can't remember due to long duration (n=12).
B10 question ‘Were you taking any of the medicines listed below around the time of your tinnitus onset?’	Amlodipine (n=1), statins (n=6), atenolol (n=1), eye drops (n=2), bisoprolol (n=1), blood thinner (n=1), candesartan (n=1), contraceptives (n=2), diclofenac (n=1), flixonase (n=1), felodipine (n=1), hormone replacement therapy (n=2), prazoles (n=6), levothyroxine (n=5), menopace (n=1), ramipril (n=3), rapril (n=1), roaccutane (n=1), sumatriptan (n=1), symbicort (n=1), tardisc (n=1), and warfarin (n=1).
B11 question ‘Do you think any of the conditions mentioned before or any other conditions are related with your tinnitus onset? You can give up to 3 responses - please choose the most important.’ (free text from those responding ‘Yes’)	Noise exposure (n=38), trauma (n=6), musculoskeletal problems (n=3), hearing loss (n=18), cochlear implantation (n=1), acoustic neuroma (n=1), ear fullness or infection (n=7), Bell's palsy (n=1), Meniere's (n=2), labyrinthitis (n=2), vertigo (n=1), meningitis (n=1), headaches (n=1), dental problems or interventions (n=2), thyroid problems (n=1), frozen shoulder (n=1), nasal congestion (n=2), sinus problems (n=2), infection (n=1), anaemia (n=1), stroke (n=1), allergies (n=1), flight (n=1), serving in royal air force (n=3), amitriptyline (n=1), quinine (n=1), painkillers (n=1), mental condition including stress, anxiety and depression (n=15), sleep problems (n=1), repetitive strain injury (n=1), overworked (n=1), tiredness (n=1), alcohol (n=1), retirement (n=1), and night out with heavy smoking (n=1).
B13 question ‘What does your tinnitus sound like?’	Many combinations of sounds. Reported sound descriptions: Tonal (n=1), a tone with noise (n=2), whine (n=4), screeching (n=3), whistling (n=5), screaming high-pitched machines (n=1), high-toned like air or waves (n=1), wind-like (n=1), ringing (n=6), high-pitched cathode-ray tube television (n=1), high-pitched squeal (n=1), squeaking (n=2), high-pitched hiss (n=2), hissing (n=12), buzzing (n=8), fuzzy (n=1), humming (n=2), dentist drill (n=1), high-pitched electrical (n=1), swishing like washing machine (n=2), high pressure steam escaping (n=1), running water (n=3), swashing (n=1), bells (n=1), lower tonal school bell (n=1),

	aeroplane (n=6), rushing sound (n=2), roaring (n=1), whooshing (n=4), clicking (n=1), ticking (n=1), pulsatile steam train (n=1), out-of-tune radio (n=1), background radio (n=1), air-conditioning unit (n=1), (n=1), rhythmic (n=1), pulsating (n=7), beeping metallic like Morse code (n=1), noise including traffic noise and white noise (n=12), glass rim (n=1), forest sounds (n=1), music (n=4), opera singing (n=1), hammer bang (n=1), talking (n=1), nonsense talking (n=1), varying (n=3), and all of the above (n=2).
B15 question 'Where do you perceive your tinnitus?'	Free-text responses included mainly combinations of different locations. In addition, responses included back of head and neck area (n=2) and varying location (n=1).
B16 question 'Is your tinnitus rhythmic?',	Linked to eye movements (n=2), rhythmic but unspecified (n=2), constant (n=6), varying types (n=4), combinations of types (n=4), has changed over time (n=1), and do not know (n=1).
B18 question 'Is your tinnitus reduced by (you can choose more than one option)'	Hearing aids with additional notes (n=2), cochlear implant (n=1), not using CROS hearing aids (n=1), listening to music, radio or television or background music or sounds (n=14), being outdoors (n=1), distraction including focusing on other things, being busy and socialising (n=15), ignoring it (n=2), mindfulness (n=1), silence (n=2), getting on well with partner (n=1), managing stress (n=1), physical activity including swimming or skiing (n=5), good diet (n=1), sleep or nap (n=2), general anaesthesia (n=1), mental blinkers (n=1), maintaining normal blood pressure (n=1), not taking pain killers (n=1), bananas (n=1), beetroot (n=1), cannabidiol oil (n=1), gabapentin (n=1), and zopiclone (n=1).
B19 question 'Is your tinnitus increased by (you can choose more than one option)	Loud sounds or noise (n=11), using headphones (n=1), vibrations from speakers (n=1), swimming underwater (n=1), silence (n=3), anxiety for hearing related problems (n=1), being alone (n=1), being tired (n=2), focusing on it (n=3), eye movement (n=1), pressing forehead (n=1), lying in back and raising head (n=1), flying (n=1), heat and humidity (n=2), before the end of menstrual cycle (n=1), after exercise (n=1), in the morning after taking a sleeping tablet (n=1), salt intake (n=1), alcohol (n=1), and do not know (n=2).
B21 question 'Are you currently receiving any of the following types of management for your tinnitus?',	Sound exposure and auditory distractions (n=6), SoundCure (n=1), hearing aids and/or sound generator (n=4), ear plugs and avoidance of noisy environments (n=1), not wearing the CROS hearing aid, regular ear wax cleaning (n=1), one-to-one talking support from audiologist (n=1), tinnitus specialist once a year (n=1), cognitive behavioural therapy (n=3), self-care and mindfulness (n=3), yoga (n=1), chiropractor (n=1), Alexander Technique (n=1), drawing (n=1), exercise (n=1), sleep during the day (n=1), walking to improve sleep (n=1), better life-style including less drinking, better eating, and supplements (n=1), homeopathic remedy (n=1), neuromonics (n=1), hypnotherapy (n=2), and already answered (n=3).
B22 question 'Do you think any of the conditions mentioned before, or any other	Sound exposure (n=20), silence (n=4), gromets fitting (n=1), cochlear implantation (n=2), removing hearing aids (n=1), hearing loss (n=2), CROS hearing aid (n=1), vestibular problems (n=1), hyperacusis (n=1), neck

conditions, are related with periods of increased tinnitus? You can give up to 3 responses - please choose the most important.' (free text from those responding 'Yes')	trauma (n=1), jaw pain (n=1), clenching teeth (n=1), hip problems (n=2), some movements (n=1), physical tension (n=1), nap (n=2), sleep problems including lack of sleep (n=5), not getting up immediately after waking (n=1), alcohol (n=2), monosodium glutamate (n=1), aspirin overdose (n=1), focusing on tinnitus (n=1), mental health problems including anxiety and stress (n=33), arguing with partner (n=1), travelling (n=1), being alone (n=2), bad working conditions (n=1), being tired (n=3), being ill (n=2), anaemia (n=1), low blood sugar (n=1), peri-menopausal symptoms (n=1), meningitis (n=1), thyroid problems (n=1), headaches (n=1), high intensity training (n=1), barometric pressure (n=1), change in weather (n=1), hot humid weather (n=1), high blood pressure (n=1), painkillers, tylenol, and oxycodone (n=1), no (n=2), as above (n=2).

Appendix 5.1. Variables used for clustering and validation in the STOP Audiometric Variables Clustering

24 variables (raw or their first 4 principal components) used for clustering	Hearing threshold at 0.125 kHz (left ear) (dB HL), Hearing threshold at 0.25 kHz (left ear) (dB HL), Hearing threshold at 0.5 kHz (left ear) (dB HL), Hearing threshold at 1 kHz (left ear) (dB HL), Hearing threshold at 2 kHz (left ear) (dB HL), Hearing threshold at 3 kHz (left ear) (dB HL), Hearing threshold at 4 kHz (left ear) (dB HL), Hearing threshold at 6 kHz (left ear) (dB HL), Hearing threshold at 8 kHz (left ear) (dB HL), Hearing threshold at 10 kHz (left ear) (dB HL), Hearing threshold at 12.5 kHz (left ear) (dB HL), Hearing threshold at 14 kHz (left ear) (dB HL), Hearing threshold at 0.125 kHz (right ear) (dB HL), Hearing threshold at 0.25 kHz (right ear) (dB HL), Hearing threshold at 0.5 kHz (right ear) (dB HL), Hearing threshold at 1 kHz (right ear) (dB HL), Hearing threshold at 2 kHz (right ear) (dB HL), Hearing threshold at 3 kHz (right ear) (dB HL), Hearing threshold at 4 kHz (right ear) (dB HL), Hearing threshold at 6 kHz (right ear) (dB HL), Hearing threshold at 8 kHz (right ear) (dB HL), Hearing threshold at 10 kHz (right ear) (dB HL), Hearing threshold at 12.5 kHz (right ear) (dB HL), Hearing threshold at 14 kHz (right ear) (dB HL)
55 variables used for validation	Tinnitus pitch matching (kHz), Hearing threshold at pitch matched frequency (dB HL), Tinnitus loudness matching (dB HL), Masking noise threshold using Narrow Band Noise (dB HL), Maskability (no/yes), Tinnitus Handicap Inventory (0-100), Fear of Tinnitus Questionnaire (0-17), Tinnitus Catastrophizing Scale (0-52), Tinnitus loudness rating (0-100), Tinnitus awareness (% of total awake time), Tinnitus annoyance scale (0-100), Tinnitus daily presence (no/yes), Presence during the day (constant/intermittent), Tinnitus worries, annoys or upsets (severely/moderately/slightly/not at all), Number of sounds (more than one/one), Fullness in the ears at tinnitus onset (no/yes), No event at tinnitus onset (no/yes), No medication at tinnitus onset (no/yes), Thoughts of conditions related to tinnitus onset (no/yes), Tinnitus quality (tonal/noise/other), Tinnitus pitch (high/medium/low), Tinnitus reduced by silence (no/yes), Tinnitus reduced by low intensity sounds (no/yes), Tinnitus reduced by high intensity sounds (no/yes), Tinnitus reduced by jaw movement (no/yes), Tinnitus reduced by good quality sleep (no/yes), Tinnitus reduced by driving (no/yes), Tinnitus reduced by being relaxed (no/yes), Tinnitus reduced by nothing (no/yes), Tinnitus increased by silence (no/yes), Tinnitus increased by jaw movement (no/yes), Tinnitus increased by poor quality sleep (no/yes), Tinnitus increased by stress (no/yes), Tinnitus increased by drinking alcohol (no/yes), Tinnitus increased by nothing (no/yes), Previous treatments or healthcare visits for tinnitus (no/yes), No management for tinnitus (no/yes), Thoughts of conditions related to increased tinnitus (no/yes), Varying tinnitus loudness from day to day (no/yes), Rhythmic tinnitus (no/yes, other/yes, with heartbeat)*, Tinnitus worsened by loud noise (no/yes), Tinnitus spatial perception (left ear/both ears, more left/no lateralisation (both ears equally or in the head)/both ears, more right/right ear)*, Age at tinnitus onset (y), Sound exposure at tinnitus onset (no/yes)*, Stress at tinnitus onset (no/yes)*, Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after), Onset of problems with external sounds in relation to tinnitus onset (never/before/at the same time/after), Onset of hearing difficulties in relation to tinnitus onset (never/before/at the same time/after), Onset of dental surgery in relation to tinnitus onset (never/before/at the same time/after), Onset of headache in relation to tinnitus onset (never/before/at the same time/after), Onset of neck pain in relation to tinnitus onset (never/before/at the same time/after), Onset of depression in relation to tinnitus onset (never/before/at the same time/after), Onset of problems staying asleep in relation to tinnitus onset (never/before/at the same time/after), Onset of high BP in relation to tinnitus onset (never/before/at the same time/after), Onset of Lyme disease in relation to tinnitus onset (never/before/at the same time/after)

*Variable assessed using the TSCHQ question.

Appendix 5.2. Variables used for clustering and validation in the BRC Audiometric Variables Clustering

32 variables (raw or their first 4 principal components) used for clustering	Hearing threshold at 0.125 kHz (left ear) (dB HL), Hearing threshold at 0.25 kHz (left ear) (dB HL), Hearing threshold at 0.5 kHz (left ear) (dB HL), Hearing threshold at 0.75 kHz (left ear) (dB HL), Hearing threshold at 1 kHz (left ear) (dB HL), Hearing threshold at 1.5 kHz (left ear) (dB HL), Hearing threshold at 2 kHz (left ear) (dB HL), Hearing threshold at 3 kHz (left ear) (dB HL), Hearing threshold at 4 kHz (left ear) (dB HL), Hearing threshold at 6 kHz (left ear) (dB HL), Hearing threshold at 8 kHz (left ear) (dB HL), Hearing threshold at 9 kHz (left ear) (dB HL), Hearing threshold at 10 kHz (left ear) (dB HL), Hearing threshold at 11.2 kHz (left ear) (dB HL), Hearing threshold at 12.5 kHz (left ear) (dB HL), Hearing threshold at 14 kHz (left ear) (dB HL), Hearing threshold at 0.125 kHz (right ear) (dB HL), Hearing threshold at 0.25 kHz (right ear) (dB HL), Hearing threshold at 0.5 kHz (right ear) (dB HL), Hearing threshold at 0.75 kHz (right ear) (dB HL), Hearing threshold at 1 kHz (right ear) (dB HL), Hearing threshold at 1.5 kHz (right ear) (dB HL), Hearing threshold at 2 kHz (right ear) (dB HL), Hearing threshold at 3 kHz (right ear) (dB HL), Hearing threshold at 4 kHz (right ear) (dB HL), Hearing threshold at 6 kHz (right ear) (dB HL), Hearing threshold at 8 kHz (right ear) (dB HL), Hearing threshold at 9 kHz (right ear) (dB HL), Hearing threshold at 10 kHz (right ear) (dB HL), Hearing threshold at 11.2 kHz (right ear) (dB HL), Hearing threshold at 12.5 kHz (right ear) (dB HL), Hearing threshold at 14 kHz (right ear) (dB HL)
17 variables used for validation	Tinnitus duration (y), Age at tinnitus onset (y), Tinnitus spatial perception (left ear/both ears, more left/no lateralisation (both ears equally or in the head)/both ears, more right/right ear)*, Presence during the day (constant/intermittent), Varying tinnitus loudness from day to day (no/yes), Tinnitus annoyance scale (0-100), Previous treatments or healthcare visits for tinnitus (no/yes), Tinnitus worsened by loud noise (no/yes), Tinnitus increased by stress (no/yes), Tinnitus loudness rating (0-100), Change in hearing at tinnitus onset (no/yes)*, Stress at tinnitus onset (no/yes)*, Head trauma at tinnitus onset (no/yes), Sound exposure at tinnitus onset (no/yes)*, Infection at tinnitus onset (no/yes), Rhythmic tinnitus (no/yes, other/yes, with heartbeat)*, Tinnitus pitch matching (kHz)

*Variable assessed using the TSCHQ question.

Appendix 5.3. Variables used for clustering and validation in the STOP General Phenotypic Variables Clustering

166 variables (their first 7 principal components) used for clustering	Age (y), Sex (female/male)*, Abnormal tympanogram left ear (no/yes), Abnormal tympanogram right ear (no/yes), Hearing threshold at 0.125 kHz (left ear) (dB HL), Hearing threshold at 0.25 kHz (left ear) (dB HL), Hearing threshold at 0.5 kHz (left ear) (dB HL), Hearing threshold at 1 kHz (left ear) (dB HL), Hearing threshold at 2 kHz (left ear) (dB HL), Hearing threshold at 3 kHz (left ear) (dB HL), Hearing threshold at 4 kHz (left ear) (dB HL), Hearing threshold at 6 kHz (left ear) (dB HL), Hearing threshold at 8 kHz (left ear) (dB HL), Hearing threshold at 10 kHz (left ear) (dB HL), Hearing threshold at 12.5 kHz (left ear) (dB HL), Hearing threshold at 14 kHz (left ear) (dB HL), Hearing threshold at 0.125 kHz (right ear) (dB HL), Hearing threshold at 0.25 kHz (right ear) (dB HL), Hearing threshold at 0.5 kHz (right ear) (dB HL), Hearing threshold at 1 kHz (right ear) (dB HL), Hearing threshold at 2 kHz (right ear) (dB HL), Hearing threshold at 3 kHz (right ear) (dB HL), Hearing threshold at 4 kHz (right ear) (dB HL), Hearing threshold at 6 kHz (right ear) (dB HL), Hearing threshold at 8 kHz (right ear) (dB HL), Hearing threshold at 10 kHz (right ear) (dB HL), Hearing threshold at 12.5 kHz (right ear) (dB HL), Hearing threshold at 14 kHz (right ear) (dB HL), Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL), Uncomfortable loudness level at 0.25 kHz (left ear) (dB HL), Uncomfortable loudness level at 0.5 kHz (left ear) (dB HL), Uncomfortable loudness level at 1 kHz (left ear) (dB HL), Uncomfortable loudness level at 2 kHz (left ear) (dB HL), Uncomfortable loudness level at 3 kHz (left ear) (dB HL), Uncomfortable loudness level at 4 kHz (left ear) (dB HL), Uncomfortable loudness level at 6 kHz (left ear) (dB HL), Uncomfortable loudness level at 8 kHz (left ear) (dB HL), Uncomfortable loudness level at 0.125 kHz (right ear) (dB HL), Uncomfortable loudness level at 0.25 kHz (right ear) (dB HL), Uncomfortable loudness level at 0.5 kHz (right ear) (dB HL), Uncomfortable loudness level at 1 kHz (right ear) (dB HL), Uncomfortable loudness level at 2 kHz (right ear) (dB HL), Uncomfortable loudness level at 3 kHz (right ear) (dB HL), Uncomfortable loudness level at 4 kHz (right ear) (dB HL), Uncomfortable loudness level at 6 kHz (right ear) (dB HL), Uncomfortable loudness level at 8 kHz (right ear) (dB HL), Speech in noise hearing assessment - phoneme score (left ear) (0-100), Speech in noise hearing assessment - phoneme score (right ear) (0-100), DPOAE at 996 Hz (left ear) (signal to noise ratio), DPOAE at 1074 Hz (left ear) (signal to noise ratio), DPOAE at 1152 Hz (left ear) (signal to noise ratio), DPOAE at 1230 Hz (left ear) (signal to noise ratio), DPOAE at 1318 Hz (left ear) (signal to noise ratio), DPOAE at 1416 Hz (left ear) (signal to noise ratio), DPOAE at 1513 Hz (left ear) (signal to noise ratio), DPOAE at 1621 Hz (left ear) (signal to noise ratio), DPOAE at 1738 Hz (left ear) (signal to noise ratio), DPOAE at 1865 Hz (left ear) (signal to noise ratio), DPOAE at 2001 Hz (left ear) (signal to noise ratio), DPOAE at 2148 Hz (left ear) (signal to noise ratio), DPOAE at 2294 Hz (left ear) (signal to noise ratio), DPOAE at 2460 Hz (left ear) (signal to noise ratio), DPOAE at 2636 Hz (left ear) (signal to noise ratio), DPOAE at 2832 Hz (left ear) (signal to noise ratio), DPOAE at 3027 Hz (left ear) (signal to noise ratio), DPOAE at 3251 Hz (left ear) (signal to noise ratio), DPOAE at 3486 Hz (left ear) (signal to noise ratio), DPOAE at 3730 Hz (left ear) (signal to noise ratio), DPOAE at 4003 Hz (left ear) (signal to noise ratio), DPOAE at 4287 Hz (left ear) (signal to noise ratio), DPOAE at 4599 Hz (left ear) (signal to noise ratio), DPOAE at 4921 Hz (left ear) (signal to noise ratio), DPOAE at 5273 Hz (left ear) (signal to noise ratio), DPOAE at 5654 Hz (left ear) (signal to noise ratio), DPOAE at 6064 Hz (left ear) (signal to noise ratio), DPOAE at 6494 Hz (left ear) (signal to noise ratio), DPOAE at 6962 Hz (left ear) (signal to noise ratio), DPOAE at 7460 Hz (left ear) (signal to noise ratio), DPOAE at 7998 Hz (left ear) (signal to noise ratio), DPOAE at 8574 Hz (left ear) (signal to noise ratio), DPOAE at 9189 Hz (left ear) (signal to noise ratio), DPOAE at 9853 Hz (left ear) (signal to noise ratio), DPOAE at 996 Hz (right ear)

	(signal to noise ratio), DPOAE at 1074 Hz (right ear) (signal to noise ratio), DPOAE at 1152 Hz (right ear) (signal to noise ratio), DPOAE at 1230 Hz (right ear) (signal to noise ratio), DPOAE at 1318 Hz (right ear) (signal to noise ratio), DPOAE at 1416 Hz (right ear) (signal to noise ratio), DPOAE at 1513 Hz (right ear) (signal to noise ratio), DPOAE at 1621 Hz (right ear) (signal to noise ratio), DPOAE at 1738 Hz (right ear) (signal to noise ratio), DPOAE at 1865 Hz (right ear) (signal to noise ratio), DPOAE at 2001 Hz (right ear) (signal to noise ratio), DPOAE at 2148 Hz (right ear) (signal to noise ratio), DPOAE at 2294 Hz (right ear) (signal to noise ratio), DPOAE at 2460 Hz (right ear) (signal to noise ratio), DPOAE at 2636 Hz (right ear) (signal to noise ratio), DPOAE at 2832 Hz (right ear) (signal to noise ratio), DPOAE at 3027 Hz (right ear) (signal to noise ratio), DPOAE at 3251 Hz (right ear) (signal to noise ratio), DPOAE at 3486 Hz (right ear) (signal to noise ratio), DPOAE at 3730 Hz (right ear) (signal to noise ratio), DPOAE at 4003 Hz (right ear) (signal to noise ratio), DPOAE at 4287 Hz (right ear) (signal to noise ratio), DPOAE at 4599 Hz (right ear) (signal to noise ratio), DPOAE at 4921 Hz (right ear) (signal to noise ratio), DPOAE at 5273 Hz (right ear) (signal to noise ratio), DPOAE at 5654 Hz (right ear) (signal to noise ratio), DPOAE at 6064 Hz (right ear) (signal to noise ratio), DPOAE at 6494 Hz (right ear) (signal to noise ratio), DPOAE at 6962 Hz (right ear) (signal to noise ratio), DPOAE at 7460 Hz (right ear) (signal to noise ratio), DPOAE at 7998 Hz (right ear) (signal to noise ratio), DPOAE at 8574 Hz (right ear) (signal to noise ratio), DPOAE at 9189 Hz (right ear) (signal to noise ratio), DPOAE at 9853 Hz (right ear) (signal to noise ratio), Perceived Stress Questionnaire (0-1), Hospital Anxiety Depression Scale for Anxiety (0-21), Hospital Anxiety Depression Scale for Depression (0-21), Hyperacusis questionnaire score (0-42), World Health Organization's Quality of Life Physical subscale (4-20), World Health Organization's Quality of Life Psychological subscale (4-20), World Health Organization's Quality of Life Social subscale (4-20), World Health Organization's Quality of Life Environmental subscale (4-20), Height (cm), Weight (kg), Education (lower/higher), Alcohol (number of drinks per week), Smoking (no/yes), Any ear condition (no/yes), Presbycusis (no/yes), Other hearing loss (no/yes), Acute Otitis (no/yes), Any procedure (no/yes), Ear surgery (no/yes), Dental surgery (no/yes), Problem with external sounds (small/moderate/big or very big), Self-reported hearing difficulty (slight or no difficulty/moderate difficulty/severe difficulty/total loss), Any hearing device (no/yes), Hearing aid use (no/yes), Any pain syndromes (no/yes), Headaches (no/yes)*, Neck pain (no/yes), TMJ pain (no/yes), Any diagnosed condition (no/yes), TMJ disorder (no/yes)*, Dental problems (no/yes), Anxiety (no/yes), Depression (no/yes), Stress (no/yes), Problem falling asleep (no/yes), Problem staying asleep (no/yes), Low BP (no/yes), High BP (no/yes), Thyroid disorder (no/yes), High cholesterol (no/yes), Lyme disease (no/yes), Acid/gastroesophageal reflux (no/yes), Handedness (both/left/right), Ever worked shifts (no/yes), Night work (no/yes), BMI (kg/m^2), Vertigo (no/yes)*, Family history of tinnitus (or hearing loss for ESIT) (no/yes), MaxDiffExt (dB), Mean hearing threshold both ears (dB HL)
55 variables used for validation	Tinnitus pitch matching (kHz), Hearing threshold at pitch matched frequency (dB HL), Tinnitus loudness matching (dB HL), Masking noise threshold using Narrow Band Noise (dB HL), Maskability (no/yes), Tinnitus Handicap Inventory (0-100), Fear of Tinnitus Questionnaire (0-17), Tinnitus Catastrophizing Scale (0-52), Tinnitus loudness rating (0-100), Tinnitus awareness (% of total awake time), Tinnitus annoyance scale (0-100), Tinnitus daily presence (no/yes), Presence during the day (constant/intermittent), Tinnitus worries, annoys or upsets (severely/moderately/slightly/not at all), Number of sounds (more than one/one), Sound exposure at tinnitus onset (no/yes)*, Fullness in the ears at tinnitus onset (no/yes), Stress at tinnitus onset (no/yes)*, No event at tinnitus onset (no/yes), No medication at tinnitus onset (no/yes),

	Thoughts of conditions related to tinnitus onset (no/yes), Varying tinnitus loudness over a day (stable/sometimes fluctuating/always fluctuating), Tinnitus quality (tonal/noise/other), Tinnitus pitch (high/medium/low), Tinnitus reduced by silence (no/yes), Tinnitus reduced by low intensity sounds (no/yes), Tinnitus reduced by high intensity sounds (no/yes), Tinnitus reduced by jaw movement (no/yes), Tinnitus reduced by good quality sleep (no/yes), Tinnitus reduced by driving (no/yes), Tinnitus reduced by being relaxed (no/yes), Tinnitus reduced by nothing (no/yes), Tinnitus increased by silence (no/yes), Tinnitus increased by high intensity sounds (no/yes), Tinnitus increased by jaw movement (no/yes), Tinnitus increased by poor quality sleep (no/yes), Tinnitus increased by stress (no/yes), Tinnitus increased by drinking alcohol (no/yes), Tinnitus increased by nothing (no/yes), Previous treatments or healthcare visits for tinnitus (no/yes), No management for tinnitus (no/yes), Thoughts of conditions related to increased tinnitus (no/yes), Rhythmic tinnitus (no/yes, other/yes, with heartbeat)*, Tinnitus spatial perception (left ear/both ears, more left/no lateralisation (both ears equally or in the head)/both ears, more right/right ear)*, Age at tinnitus onset (y), Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after), Onset of problems with external sounds in relation to tinnitus onset (never/before/at the same time/after), Onset of hearing difficulties in relation to tinnitus onset (never/before/at the same time/after), Onset of dental surgery in relation to tinnitus onset (never/before/at the same time/after), Onset of headache in relation to tinnitus onset (never/before/at the same time/after), Onset of neck pain in relation to tinnitus onset (never/before/at the same time/after), Onset of depression in relation to tinnitus onset (never/before/at the same time/after), Onset of problems staying asleep in relation to tinnitus onset (never/before/at the same time/after), Onset of high BP in relation to tinnitus onset (never/before/at the same time/after), Onset of Lyme disease in relation to tinnitus onset (never/before/at the same time/after)

*Variable assessed using the ESIT-SQ question.

Appendix 5.4. Variables used for clustering and validation in the ESIT General Phenotypic Variables Clustering

42 variables (their first 11 principal components) used for clustering	Sex (female/male)*, Age (y), Height (cm), Weight (kg), Education (lower/higher), Alcohol (number of drinks per week), Smoking (no/yes), Sudden hearing loss (no/yes), Other hearing loss (no/yes), Acute Otitis (no/yes), Ear surgery (no/yes), Dental surgery (no/yes), Lumbar puncture (no/yes), Problem with external sounds (small/moderate/big or very big), Self-reported hearing difficulty (slight or no difficulty/moderate difficulty/severe difficulty/total loss), Hearing aid use (no/yes), Headaches (no/yes)*, Neck pain (no/yes), Ear pain (no/yes), Face pain (no/yes), Dental problems (no/yes), Anxiety (no/yes), Depression (no/yes), Emotional trauma (no/yes), Stress (no/yes), Problem falling asleep (no/yes), Problem staying asleep (no/yes), Low BP (no/yes), High BP (no/yes), Thyroid disorder (no/yes), Diabetes (no/yes), High cholesterol (no/yes), Chronic sinusitis (no/yes), Acid/gastroesophageal reflux (no/yes), Handedness (both/left/right), Night work (no/yes), BMI (kg/m^2), Any procedure (no/yes), Any hearing device (no/yes), Any pain syndromes (no/yes), Any diagnosed condition (no/yes), Vertigo (no/yes)*
64 variables used for validation	Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after), Onset of sudden hearing loss in relation to tinnitus onset (never/before/at the same time/after), Onset of other hearing loss in relation to tinnitus onset (never/before/at the same time/after), Onset of acute otitis in relation to tinnitus onset (never/before/at the same time/after), Onset of ear surgery in relation to tinnitus onset (never/before/at the same time/after), Onset of dental surgery in relation to tinnitus onset (never/before/at the same time/after), Onset of problems with external sounds in relation to tinnitus onset (never/before/at the same time/after), Onset of hearing difficulties in relation to tinnitus onset (never/before/at the same time/after), Onset of headache in relation to tinnitus onset (never/before/at the same time/after), Onset of neck pain in relation to tinnitus onset (never/before/at the same time/after), Onset of ear pain in relation to tinnitus onset (never/before/at the same time/after), Onset of face pain in relation to tinnitus onset (never/before/at the same time/after), Onset of anxiety in relation to tinnitus onset (never/before/at the same time/after), Onset of depression in relation to tinnitus onset (never/before/at the same time/after), Onset of problems staying asleep in relation to tinnitus onset (never/before/at the same time/after), Onset of low BP in relation to tinnitus onset (never/before/at the same time/after), Onset of high BP in relation to tinnitus onset (never/before/at the same time/after), Onset of thyroid disorder in relation to tinnitus onset (never/before/at the same time/after), Onset of increased cholesterol in relation to tinnitus onset (never/before/at the same time/after), Onset of acid/gastroesophageal reflux in relation to tinnitus onset (never/before/at the same time/after), Presence during the day (constant/intermittent), Tinnitus worries, annoys or upsets (severely/moderately/slightly/not at all), Number of sounds (more than one/one), Sound exposure at tinnitus onset (no/yes)*, Change in hearing at tinnitus onset (no/yes)*, Infection at tinnitus onset (no/yes), Fullness in the ears at tinnitus onset (no/yes), Stress at tinnitus onset (no/yes)*, No event at tinnitus onset (no/yes), Painkillers at tinnitus onset (no/yes), Antidepressants at tinnitus onset (no/yes), No medication at tinnitus onset (no/yes), Varying tinnitus loudness over a day (stable/sometimes fluctuating/always fluctuating), Tinnitus quality (tonal/noise/other), Tinnitus pitch (high/medium/low), Tinnitus reduced by silence (no/yes), Tinnitus reduced by low intensity sounds (no/yes), Tinnitus reduced by high intensity sounds (no/yes), Tinnitus reduced by good quality sleep (no/yes), Tinnitus reduced by driving (no/yes), Tinnitus reduced by being relaxed (no/yes), Tinnitus reduced by using hearing aids (no/yes), Tinnitus reduced by nothing (no/yes), Tinnitus increased by silence (no/yes), Tinnitus increased by high intensity sounds (no/yes), Tinnitus increased by jaw movement (no/yes), Tinnitus increased by poor quality sleep (no/yes), Tinnitus increased by driving (no/yes), Tinnitus

	increased by stress (no/yes), Tinnitus increased by drinking alcohol (no/yes), Tinnitus increased by nothing (no/yes), Previous treatments or healthcare visits for tinnitus (no/yes), Audiological management for tinnitus (no/yes), Self-management for tinnitus (no/yes), No management for tinnitus (no/yes), Tinnitus spatial perception (left ear/both ears, more left/no lateralisation (both ears equally or in the head)/both ears, more right/right ear)*, Tinnitus duration (y), Bothersome tinnitus duration (y), Gap between onset of tinnitus and bothersome tinnitus (y), Age at tinnitus onset (y), Age at bothersome tinnitus onset (y), Thoughts of conditions related to tinnitus onset (no/yes), Thoughts of conditions related to increased tinnitus (no/yes), Rhythmic tinnitus (no/yes, other/yes, with heartbeat)*

*Variable assessed using the ESIT-SQ question.

Appendix 5.5. Variables used for clustering and validation in the STOP Tinnitus Discriminating Variables Clustering

35 variables used for clustering	Hearing threshold at 3 kHz (left ear) (dB HL), Hearing threshold at 4 kHz (left ear) (dB HL), Hearing threshold at 6 kHz (left ear) (dB HL), Hearing threshold at 8 kHz (left ear) (dB HL), Hearing threshold at 10 kHz (left ear) (dB HL), Hearing threshold at 12.5 kHz (left ear) (dB HL), Hearing threshold at 4 kHz (right ear) (dB HL), Hearing threshold at 6 kHz (right ear) (dB HL), Hearing threshold at 8 kHz (right ear) (dB HL), Hearing threshold at 10 kHz (right ear) (dB HL), Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL), Uncomfortable loudness level at 0.25 kHz (left ear) (dB HL), Uncomfortable loudness level at 0.5 kHz (left ear) (dB HL), Uncomfortable loudness level at 1 kHz (left ear) (dB HL), Uncomfortable loudness level at 2 kHz (left ear) (dB HL), Uncomfortable loudness level at 8 kHz (left ear) (dB HL), Uncomfortable loudness level at 0.125 kHz (right ear) (dB HL), Uncomfortable loudness level at 0.25 kHz (right ear) (dB HL), Uncomfortable loudness level at 0.5 kHz (right ear) (dB HL), Uncomfortable loudness level at 1 kHz (right ear) (dB HL), Uncomfortable loudness level at 2 kHz (right ear) (dB HL), Uncomfortable loudness level at 3 kHz (right ear) (dB HL), Uncomfortable loudness level at 4 kHz (right ear) (dB HL), Uncomfortable loudness level at 6 kHz (right ear) (dB HL), Speech in noise hearing assessment - phoneme score (left ear) (0-100), Speech in noise hearing assessment - phoneme score (right ear) (0-100), DPOAE at 3730 Hz (left ear) (signal to noise ratio), DPOAE at 4599 Hz (left ear) (signal to noise ratio), DPOAE at 4921 Hz (left ear) (signal to noise ratio), DPOAE at 5273 Hz (left ear) (signal to noise ratio), DPOAE at 5654 Hz (left ear) (signal to noise ratio), DPOAE at 6962 Hz (right ear) (signal to noise ratio), Hyperacusis questionnaire score (0-42), Weight (kg), Problem with external sounds (small/moderate/big or very big)
55 variables used for validation	Tinnitus pitch matching (kHz), Hearing threshold at pitch matched frequency (dB HL), Tinnitus loudness matching (dB HL), Masking noise threshold using Narrow Band Noise (dB HL), Maskability (no/yes), Tinnitus Handicap Inventory (0-100), Fear of Tinnitus Questionnaire (0-17), Tinnitus Catastrophizing Scale (0-52), Tinnitus loudness rating (0-100), Tinnitus awareness (% of total awake time), Tinnitus annoyance scale (0-100), Tinnitus daily presence (no/yes), Presence during the day (constant/intermittent), Tinnitus worries, annoys or upsets (severely/moderately/slightly/not at all), Number of sounds (more than one/one), Fullness in the ears at tinnitus onset (no/yes), No event at tinnitus onset (no/yes), No medication at tinnitus onset (no/yes), Thoughts of conditions related to tinnitus onset (no/yes), Tinnitus quality (tonal/noise/other), Tinnitus pitch (high/medium/low), Tinnitus reduced by silence (no/yes), Tinnitus reduced by low intensity sounds (no/yes), Tinnitus reduced by high intensity sounds (no/yes), Tinnitus reduced by jaw movement (no/yes), Tinnitus reduced by good quality sleep (no/yes), Tinnitus reduced by driving (no/yes), Tinnitus reduced by being relaxed (no/yes), Tinnitus reduced by nothing (no/yes), Tinnitus increased by silence (no/yes), Tinnitus increased by jaw movement (no/yes), Tinnitus increased by poor quality sleep (no/yes), Tinnitus increased by stress (no/yes), Tinnitus increased by drinking alcohol (no/yes), Tinnitus increased by nothing (no/yes), Previous treatments or healthcare visits for tinnitus (no/yes), No management for tinnitus (no/yes), Thoughts of conditions related to increased tinnitus (no/yes), Varying tinnitus loudness from day to day (no/yes), Rhythmic tinnitus (no/yes, other/yes, with heartbeat)*, Tinnitus worsened by loud noise (no/yes), Tinnitus spatial perception (left ear/both ears, more left/no lateralisation (both ears equally or in the head)/both ears, more right/right ear)*, Age at tinnitus onset (y), Sound exposure at tinnitus onset (no/yes)*, Stress at tinnitus onset (no/yes)*, Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after), Onset of problems with external sounds in relation to tinnitus onset (never/before/at the same time/after), Onset of hearing difficulties in relation to tinnitus onset (never/before/at the

	same time/after), Onset of dental surgery in relation to tinnitus onset (never/before/at the same time/after), Onset of headache in relation to tinnitus onset (never/before/at the same time/after), Onset of neck pain in relation to tinnitus onset (never/before/at the same time/after), Onset of depression in relation to tinnitus onset (never/before/at the same time/after), Onset of problems staying asleep in relation to tinnitus onset (never/before/at the same time/after), Onset of high BP in relation to tinnitus onset (never/before/at the same time/after), Onset of Lyme disease in relation to tinnitus onset (never/before/at the same time/after)

*Variable assessed using the TSCHQ question.

Appendix 5.6. Variables used for clustering and validation in the STOP-BRC Independent Validation Clustering

8 variables used for clustering	Age (y), Sex (female/male)**, Tinnitus annoyance scale (0-100), AudioPC_Overall[*], AudioPC_RightVLeft[£], AudioPC_LowVHigh[$], AudioPC_MiddleVLowhigh[&], Depression_score[#]
25 variables used for validation	Age (y), Sex (female/male)**, Hyperacusis questionnaire score (0-42), Tinnitus loudness rating (0-100), Tinnitus annoyance scale (0-100), Hearing aid use (no/yes), TMJ disorder (no/yes)**, Handedness (both/left/right), Varying tinnitus loudness from day to day (no/yes), Self-reported hearing problem (no/yes), Headaches (no/yes)**, Vertigo (no/yes)**, Rhythmic tinnitus (no/yes, other/yes, with heartbeat)**, Tinnitus worsened by loud noise (no/yes), Tinnitus spatial perception (left ear/both ears, more left/no lateralisation (both ears equally or in the head)/both ears, more right/right ear)**, Age at tinnitus onset (y), Sound exposure at tinnitus onset (no/yes)**, Stress at tinnitus onset (no/yes)**, MaxDiffExt (dB), Mean hearing threshold both ears (dB HL), AudioPC_Overall, AudioPC_LowVHigh, AudioPC_RightVLeft, AudioPC_MiddleVLowhigh, Depression_score

[*]principal component from PCA on audiometric variables representing overall hearing thresholds (first component for both the STOP and BRC datasets), [£]principal component representing the difference in thresholds from the right versus the left ear (second component for BRC dataset and third component for the STOP dataset),[$]principal component representing the difference in thresholds at low versus high frequencies (third component for BRC dataset and second component for the STOP dataset), [&]principal representing the difference in thresholds at middle versus low and high frequencies (fourth component for both the BRC and STOP datasets) , [#]Normalised (0-100 range) scores from the Beck Depression Inventory and the Hospital Anxiety Depression Scale for Depression scores from the BRC and the STOP datasets, respectively. **Variable assessed using the TSCHQ question.

Appendix 5.7. Variables used for clustering and validation in the STOP-ESIT Independent Validation Clustering

5 variables used for clustering	Sex (female/male)*, Tinnitus worries, annoys or upsets (severely/moderately/slightly/not at all), Tinnitus spatial perception (left ear/both ears, more left/no lateralisation (both ears equally or in the head)/both ears, more right/right ear)*, Anxiety (no/yes), Depression (no/yes)
75 variables used for validation	Age (y), Sex (female/male)*, Height (cm), Education (lower/higher), Alcohol (number of drinks per week), Smoking (no/yes), Other hearing loss (no/yes), Acute Otitis (no/yes), Any procedure (no/yes), Ear surgery (no/yes), Dental surgery (no/yes), Problem with external sounds (small/moderate/big or very big), Self-reported hearing difficulty (slight or no difficulty/moderate difficulty/severe difficulty/total loss), Any hearing device (no/yes), Hearing aid use (no/yes), Any pain syndromes (no/yes), Headaches (no/yes)*, Neck pain (no/yes), Any diagnosed condition (no/yes), Dental problems (no/yes), Anxiety (no/yes), Depression (no/yes), Stress (no/yes), Problem falling asleep (no/yes), Problem staying asleep (no/yes), Low BP (no/yes), High BP (no/yes), Thyroid disorder (no/yes), High cholesterol (no/yes), Acid/gastroesophageal reflux (no/yes), Presence during the day (constant/intermittent), Tinnitus worries, annoys or upsets (severely/moderately/slightly/not at all), Number of sounds (more than one/one), Sound exposure at tinnitus onset (no/yes)*, Fullness in the ears at tinnitus onset (no/yes), Stress at tinnitus onset (no/yes)*, No event at tinnitus onset (no/yes), No medication at tinnitus onset (no/yes), Thoughts of conditions related to tinnitus onset (no/yes), Varying tinnitus loudness over a day (stable/sometimes fluctuating/always fluctuating), Tinnitus quality (tonal/noise/other), Tinnitus pitch (high/medium/low), Tinnitus reduced by silence (no/yes), Tinnitus reduced by low intensity sounds (no/yes), Tinnitus reduced by high intensity sounds (no/yes), Tinnitus reduced by good quality sleep (no/yes), Tinnitus reduced by driving (no/yes), Tinnitus reduced by being relaxed (no/yes), Tinnitus reduced by nothing (no/yes), Tinnitus increased by silence (no/yes), Tinnitus increased by high intensity sounds (no/yes), Tinnitus increased by jaw movement (no/yes), Tinnitus increased by poor quality sleep (no/yes), Tinnitus increased by stress (no/yes), Tinnitus increased by drinking alcohol (no/yes), Tinnitus increased by nothing (no/yes), Previous treatments or healthcare visits for tinnitus (no/yes), No management for tinnitus (no/yes), Thoughts of conditions related to increased tinnitus (no/yes), Handedness (both/left/right), Night work (no/yes), Rhythmic tinnitus (no/yes, other/yes, with heartbeat)*, BMI (kg/m^2), Tinnitus spatial perception (left ear/both ears, more left/no lateralisation (both ears equally or in the head)/both ears, more right/right ear)*, Age at tinnitus onset (y), Vertigo (no/yes)*, Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after), Onset of problems with external sounds in relation to tinnitus onset (never/before/at the same time/after), Onset of hearing difficulties in relation to tinnitus onset (never/before/at the same time/after), Onset of dental surgery in relation to tinnitus onset (never/before/at the same time/after), Onset of headache in relation to tinnitus onset (never/before/at the same time/after), Onset of neck pain in relation to tinnitus onset (never/before/at the same time/after), Onset of depression in relation to tinnitus onset (never/before/at the same time/after), Onset of problems staying asleep in relation to tinnitus onset (never/before/at the same time/after), Onset of high BP in relation to tinnitus onset (never/before/at the same time/after)

*Variable assessed using the ESIT-SQ question.

Appendix 5.8. Predictor variables used for the LASSO regression models on data from the STOP and the BRC datasets. The response variable for these models was the subgroup membership from the selected subgroupings

Predictor Variables for the STOP LASSO models	Predictor Variables for the BRC LASSO model
Abnormal tympanogram left ear (no/yes)	Age (y)
Abnormal tympanogram right ear (no/yes)	Age at tinnitus onset (y)
Acid/gastroesophageal reflux (no/yes)	AudioPCLowVHigh
Acute Otitis (no/yes)	AudioPCMiddleVLowhigh
Age (y)	AudioPCOverall
Age at tinnitus onset (y)	AudioPCRightVLeft
Alcohol (number of drinks per week)	Beck Anxiety Inventory (0-63)
Anxiety (no/yes)	Beck Depression Inventory (0-63)
Any diagnosed condition (no/yes)	Change in hearing at tinnitus onset (no/yes) (2tschq1esitsq)
Any ear condition (no/yes)	Comfortable level of a 0.5 kHz pure tone (dB SPL)
Any hearing device (no/yes)	Comfortable level of a 5 kHz pure tone (dB SPL)
Any pain syndromes (no/yes)	Depressionscore
Any procedure (no/yes)	Family history of tinnitus (or HL for ESIT) (no/yes)
AudioPCLowVHigh	Gender (female/male) (2tschq1esitsq)
AudioPCMiddleVLowhigh	Handedness (both/left/right)
AudioPCOverall	Head trauma at tinnitus onset (no/yes)
AudioPCRightVLeft	Headaches (no/yes) (2tschq1esitsq)
BMI (kg/m^2)	Hearing aid use (no/yes)
Dental problems (no/yes)	Hearing threshold at 0.125 kHz (left ear) (dB HL)
Dental surgery (no/yes)	Hearing threshold at 0.125 kHz (right ear) (dB HL)
Depression (no/yes)	Hearing threshold at 0.25 kHz (left ear) (dB HL)
Depressionscore	Hearing threshold at 0.25 kHz (right ear) (dB HL)
DPOAE at 1074 Hz (left ear) (signal to noise ratio)	Hearing threshold at 0.5 kHz (left ear) (dB HL)
DPOAE at 1074 Hz (right ear) (signal to noise ratio)	Hearing threshold at 0.5 kHz (right ear) (dB HL)
DPOAE at 1152 Hz (left ear) (signal to noise ratio)	Hearing threshold at 0.75 kHz (left ear) (dB HL)
DPOAE at 1152 Hz (right ear) (signal to noise ratio)	Hearing threshold at 0.75 kHz (right ear) (dB HL)
DPOAE at 1230 Hz (left ear) (signal to noise ratio)	Hearing threshold at 1 kHz (left ear) (dB HL)

Predictor Variables for the STOP LASSO models	Predictor Variables for the BRC LASSO model
DPOAE at 1230 Hz (right ear) (signal to noise ratio)	Hearing threshold at 1 kHz (right ear) (dB HL)
DPOAE at 1318 Hz (left ear) (signal to noise ratio)	Hearing threshold at 1.5 kHz (left ear) (dB HL)
DPOAE at 1318 Hz (right ear) (signal to noise ratio)	Hearing threshold at 1.5 kHz (right ear) (dB HL)
DPOAE at 1416 Hz (left ear) (signal to noise ratio)	Hearing threshold at 10 kHz (left ear) (dB HL)
DPOAE at 1416 Hz (right ear) (signal to noise ratio)	Hearing threshold at 10 kHz (right ear) (dB HL)
DPOAE at 1513 Hz (left ear) (signal to noise ratio)	Hearing threshold at 11.2 kHz (left ear) (dB HL)
DPOAE at 1513 Hz (right ear) (signal to noise ratio)	Hearing threshold at 11.2 kHz (right ear) (dB HL)
DPOAE at 1621 Hz (left ear) (signal to noise ratio)	Hearing threshold at 12.5 kHz (left ear) (dB HL)
DPOAE at 1621 Hz (right ear) (signal to noise ratio)	Hearing threshold at 12.5 kHz (right ear) (dB HL)
DPOAE at 1738 Hz (left ear) (signal to noise ratio)	Hearing threshold at 14 kHz (left ear) (dB HL)
DPOAE at 1738 Hz (right ear) (signal to noise ratio)	Hearing threshold at 14 kHz (right ear) (dB HL)
DPOAE at 1865 Hz (left ear) (signal to noise ratio)	Hearing threshold at 2 kHz (left ear) (dB HL)
DPOAE at 1865 Hz (right ear) (signal to noise ratio)	Hearing threshold at 2 kHz (right ear) (dB HL)
DPOAE at 2001 Hz (left ear) (signal to noise ratio)	Hearing threshold at 3 kHz (left ear) (dB HL)
DPOAE at 2001 Hz (right ear) (signal to noise ratio)	Hearing threshold at 3 kHz (right ear) (dB HL)
DPOAE at 2148 Hz (left ear) (signal to noise ratio)	Hearing threshold at 4 kHz (left ear) (dB HL)
DPOAE at 2148 Hz (right ear) (signal to noise ratio)	Hearing threshold at 4 kHz (right ear) (dB HL)
DPOAE at 2294 Hz (left ear) (signal to noise ratio)	Hearing threshold at 6 kHz (left ear) (dB HL)
DPOAE at 2294 Hz (right ear) (signal to noise ratio)	Hearing threshold at 6 kHz (right ear) (dB HL)
DPOAE at 2460 Hz (left ear) (signal to noise ratio)	Hearing threshold at 8 kHz (left ear) (dB HL)
DPOAE at 2460 Hz (right ear) (signal to noise ratio)	Hearing threshold at 8 kHz (right ear) (dB HL)
DPOAE at 2636 Hz (left ear) (signal to noise ratio)	Hearing threshold at 9 kHz (left ear) (dB HL)
DPOAE at 2636 Hz (right ear) (signal to noise ratio)	Hearing threshold at 9 kHz (right ear) (dB HL)
DPOAE at 2832 Hz (left ear) (signal to noise ratio)	Hyperacusis questionnaire score (0-42)

Predictor Variables for the STOP LASSO models	Predictor Variables for the BRC LASSO model
DPOAE at 2832 Hz (right ear) (signal to noise ratio)	Infection at tinnitus onset (no/yes)
DPOAE at 3027 Hz (left ear) (signal to noise ratio)	MaxDiffExt (dB)
DPOAE at 3027 Hz (right ear) (signal to noise ratio)	Mean hearing threshold both ears (dB HL)
DPOAE at 3251 Hz (left ear) (signal to noise ratio)	Other (than loud blast of sound, whiplash, change in hearing, stress, or head trauma) event at tinnitus onset (no/yes)
DPOAE at 3251 Hz (right ear) (signal to noise ratio)	Presence during the day (constant/intermittent)
DPOAE at 3486 Hz (left ear) (signal to noise ratio)	Previous treatments or healthcare visits for tinnitus (no/yes)
DPOAE at 3486 Hz (right ear) (signal to noise ratio)	Rhythmic tinnitus (no/yes, other/yes, with heartbeat) (2tschq1esitsq)
DPOAE at 3730 Hz (left ear) (signal to noise ratio)	Self-reported hearing problem (no/yes)
DPOAE at 3730 Hz (right ear) (signal to noise ratio)	Sound exposure at tinnitus onset (no/yes) (2tschq1esitsq)
DPOAE at 4003 Hz (left ear) (signal to noise ratio)	Stress at tinnitus onset (no/yes) (2tschq1esitsq)
DPOAE at 4003 Hz (right ear) (signal to noise ratio)	Tinnitus affected by head or neck movement or arms/hands or head touch (no/yes)
DPOAE at 4287 Hz (left ear) (signal to noise ratio)	Tinnitus annoyance scale (0-100)
DPOAE at 4287 Hz (right ear) (signal to noise ratio)	Tinnitus duration (y)
DPOAE at 4599 Hz (left ear) (signal to noise ratio)	Tinnitus during the day affected by sleep at night (no/yes)
DPOAE at 4599 Hz (right ear) (signal to noise ratio)	Tinnitus Handicap Questionnaire total score (0-100)
DPOAE at 4921 Hz (left ear) (signal to noise ratio)	Tinnitus increased by stress (no/yes)
DPOAE at 4921 Hz (right ear) (signal to noise ratio)	Tinnitus loudness matching at 0.5 kHz (dB SPL)
DPOAE at 5273 Hz (left ear) (signal to noise ratio)	Tinnitus loudness matching at 1 kHz (dB SPL)
DPOAE at 5273 Hz (right ear) (signal to noise ratio)	Tinnitus loudness matching at 10 kHz (dB SPL)
DPOAE at 5654 Hz (left ear) (signal to noise ratio)	Tinnitus loudness matching at 2 kHz (dB SPL)
DPOAE at 5654 Hz (right ear) (signal to noise ratio)	Tinnitus loudness matching at 3 kHz (dB SPL)
DPOAE at 6064 Hz (left ear) (signal to noise ratio)	Tinnitus loudness matching at 4 kHz (dB SPL)

Predictor Variables for the STOP LASSO models	Predictor Variables for the BRC LASSO model
DPOAE at 6064 Hz (right ear) (signal to noise ratio)	Tinnitus loudness matching at 5 kHz (dB SPL)
DPOAE at 6494 Hz (left ear) (signal to noise ratio)	Tinnitus loudness matching at 6 kHz (dB SPL)
DPOAE at 6494 Hz (right ear) (signal to noise ratio)	Tinnitus loudness matching at 7 kHz (dB SPL)
DPOAE at 6962 Hz (left ear) (signal to noise ratio)	Tinnitus loudness matching at 8 kHz (dB SPL)
DPOAE at 6962 Hz (right ear) (signal to noise ratio)	Tinnitus loudness rating (0-100)
DPOAE at 7460 Hz (left ear) (signal to noise ratio)	Tinnitus pitch matching (kHz)
DPOAE at 7460 Hz (right ear) (signal to noise ratio)	Tinnitus pitch matching at 0.5 kHz (likeness scale 0-100)
DPOAE at 7998 Hz (left ear) (signal to noise ratio)	Tinnitus pitch matching at 1 kHz (likeness scale 0-100)
DPOAE at 7998 Hz (right ear) (signal to noise ratio)	Tinnitus pitch matching at 10 kHz (likeness scale 0-100)
DPOAE at 8574 Hz (left ear) (signal to noise ratio)	Tinnitus pitch matching at 2 kHz (likeness scale 0-100)
DPOAE at 8574 Hz (right ear) (signal to noise ratio)	Tinnitus pitch matching at 3 kHz (likeness scale 0-100)
DPOAE at 9189 Hz (left ear) (signal to noise ratio)	Tinnitus pitch matching at 4 kHz (likeness scale 0-100)
DPOAE at 9189 Hz (right ear) (signal to noise ratio)	Tinnitus pitch matching at 5 kHz (likeness scale 0-100)
DPOAE at 9853 Hz (left ear) (signal to noise ratio)	Tinnitus pitch matching at 6 kHz (likeness scale 0-100)
DPOAE at 9853 Hz (right ear) (signal to noise ratio)	Tinnitus pitch matching at 7 kHz (likeness scale 0-100)
DPOAE at 996 Hz (left ear) (signal to noise ratio)	Tinnitus pitch matching at 8 kHz (likeness scale 0-100)
DPOAE at 996 Hz (right ear) (signal to noise ratio)	Tinnitus reduced by music or environmental sounds (no/yes)
Ear surgery (no/yes)	Tinnitus spatial perception (left ear/both ears, more left/no lateralisation (both ears equally or in the head)/both ears, more right/right ear) (2tschq1esitsq)
Education (lower/higher)	Tinnitus worsened by loud noise (no/yes)
Ever worked shifts (no/yes)	TMJ disorder (no/yes) (2tschq1esitsq)
Family history of tinnitus (or HL for ESIT) (no/yes)	Varying tinnitus loudness from day to day (no/yes)
Fear of Tinnitus Questionnaire (0-17)	Vertigo (no/yes) (2tschq1esitsq)
Fullness in the ears at tinnitus onset (no/yes)	
Gender (female/male) (2tschq1esitsq)	
Handedness (both/left/right)	
Headaches (no/yes) (2tschq1esitsq)	

Predictor Variables for the STOP LASSO models	Predictor Variables for the BRC LASSO model
Hearing aid use (no/yes)	
Hearing threshold at 0.125 kHz (left ear) (dB HL)	
Hearing threshold at 0.125 kHz (right ear) (dB HL)	
Hearing threshold at 0.25 kHz (left ear) (dB HL)	
Hearing threshold at 0.25 kHz (right ear) (dB HL)	
Hearing threshold at 0.5 kHz (left ear) (dB HL)	
Hearing threshold at 0.5 kHz (right ear) (dB HL)	
Hearing threshold at 1 kHz (left ear) (dB HL)	
Hearing threshold at 1 kHz (right ear) (dB HL)	
Hearing threshold at 10 kHz (left ear) (dB HL)	
Hearing threshold at 10 kHz (right ear) (dB HL)	
Hearing threshold at 12.5 kHz (left ear) (dB HL)	
Hearing threshold at 12.5 kHz (right ear) (dB HL)	
Hearing threshold at 14 kHz (left ear) (dB HL)	
Hearing threshold at 14 kHz (right ear) (dB HL)	
Hearing threshold at 2 kHz (left ear) (dB HL)	
Hearing threshold at 2 kHz (right ear) (dB HL)	
Hearing threshold at 3 kHz (left ear) (dB HL)	
Hearing threshold at 3 kHz (right ear) (dB HL)	
Hearing threshold at 4 kHz (left ear) (dB HL)	
Hearing threshold at 4 kHz (right ear) (dB HL)	
Hearing threshold at 6 kHz (left ear) (dB HL)	
Hearing threshold at 6 kHz (right ear) (dB HL)	
Hearing threshold at 8 kHz (left ear) (dB HL)	

Predictor Variables for the STOP LASSO models	Predictor Variables for the BRC LASSO model
Hearing threshold at 8 kHz (right ear) (dB HL)	
Hearing threshold at pitch matched frequency (dB HL)	
Height (cm)	
High BP (no/yes)	
High cholesterol (no/yes)	
Hospital Anxiety Depression Scale for Anxiety (0-21)	
Hospital Anxiety Depression Scale for Depression (0-21)	
Hyperacusis questionnaire score (0-42)	
Low BP (no/yes)	
Lyme disease (no/yes)	
Maskability (no/yes)	
Masking noise threshold using Narrow Band Noise (dB HL)	
MaxDiffExt (dB)	
Mean hearing threshold both ears (dB HL)	
Neck pain (no/yes)	
Night work (no/yes)	
No event at tinnitus onset (no/yes)	
No management for tinnitus (no/yes)	
No medication at tinnitus onset (no/yes)	
Number of sounds (more than one/one)	
Onset of dental surgery in relation to tinnitus onset (never/before/at the same time/after)	
Onset of depression in relation to tinnitus onset (never/before/at the same time/after)	
Onset of headache in relation to tinnitus onset (never/before/at the same time/after)	
Onset of hearing difficulties in relation to tinnitus onset (never/before/at the same time/after)	
Onset of high BP in relation to tinnitus onset (never/before/at the same time/after)	
Onset of Lyme disease in relation to tinnitus onset (never/before/at the same time/after)	
Onset of neck pain in relation to tinnitus onset (never/before/at the same time/after)	

Predictor Variables for the STOP LASSO models	Predictor Variables for the BRC LASSO model
Onset of problems staying asleep in relation to tinnitus onset (never/before/at the same time/after)	
Onset of problems with external sounds in relation to tinnitus onset (never/before/at the same time/after)	
Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)	
Other HL (no/yes)	
Perceived Stress Questionnaire (0-1)	
Presbycusis (no/yes)	
Presence during the day (constant/intermittent)	
Previous treatments or healthcare visits for tinnitus (no/yes)	
Problem falling asleep (no/yes)	
Problem staying asleep (no/yes)	
Problem with external sounds (small/moderate/big or very big)	
Rhythmic tinnitus (no/yes, other/yes, with heartbeat) (2tschq1esitsq)	
Self-reported hearing problem (no/yes)	
Smoking (no/yes)	
Sound exposure at tinnitus onset (no/yes) (2tschq1esitsq)	
Speech in noise hearing assessment - phoneme score (left ear) (0-100)	
Speech in noise hearing assessment - phoneme score (right ear) (0-100)	
Stress (no/yes)	
Stress at tinnitus onset (no/yes) (2tschq1esitsq)	
Thoughts of conditions related to increased tinnitus (no/yes)	
Thoughts of conditions related to tinnitus onset (no/yes)	
Thyroid disorder (no/yes)	
Tinnitus annoyance scale (0-100)	
Tinnitus awareness (% of total awake time)	
Tinnitus Catastrophizing Scale (0-52)	
Tinnitus daily presence (no/yes)	
Tinnitus duration (y)	
Tinnitus Handicap Inventory (0-100)	
Tinnitus increased by drinking alcohol (no/yes)	

Predictor Variables for the STOP LASSO models	Predictor Variables for the BRC LASSO model
Tinnitus increased by jaw movement (no/yes)	
Tinnitus increased by nothing (no/yes)	
Tinnitus increased by poor quality sleep (no/yes)	
Tinnitus increased by silence (no/yes)	
Tinnitus increased by stress (no/yes)	
Tinnitus loudness matching (dB HL)	
Tinnitus loudness rating (0-100)	
Tinnitus pitch (high/medium/low)	
Tinnitus pitch matching (kHz)	
Tinnitus quality (tonal/noise/other)	
Tinnitus reduced by being relaxed (no/yes)	
Tinnitus reduced by driving (no/yes)	
Tinnitus reduced by good quality sleep (no/yes)	
Tinnitus reduced by high intensity sounds (no/yes)	
Tinnitus reduced by jaw movement (no/yes)	
Tinnitus reduced by low intensity sounds (no/yes)	
Tinnitus reduced by nothing (no/yes)	
Tinnitus reduced by silence (no/yes)	
Tinnitus spatial perception (left ear/both ears, more left/no lateralisation (both ears equally or in the head)/both ears, more right/right ear) (2tschq1esitsq)	
Tinnitus worries, annoys or upsets (severely/moderately/slightly/not at all)	
Tinnitus worsened by loud noise (no/yes)	
TMJ disorder (no/yes) (2tschq1esitsq)	
TMJ pain (no/yes)	
Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)	
Uncomfortable loudness level at 0.125 kHz (right ear) (dB HL)	
Uncomfortable loudness level at 0.25 kHz (left ear) (dB HL)	
Uncomfortable loudness level at 0.25 kHz (right ear) (dB HL)	
Uncomfortable loudness level at 0.5 kHz (left ear) (dB HL)	
Uncomfortable loudness level at 0.5 kHz (right ear) (dB HL)	
Uncomfortable loudness level at 1 kHz (left ear) (dB HL)	

Predictor Variables for the STOP LASSO models	Predictor Variables for the BRC LASSO model
Uncomfortable loudness level at 1 kHz (right ear) (dB HL)	
Uncomfortable loudness level at 2 kHz (left ear) (dB HL)	
Uncomfortable loudness level at 2 kHz (right ear) (dB HL)	
Uncomfortable loudness level at 3 kHz (left ear) (dB HL)	
Uncomfortable loudness level at 3 kHz (right ear) (dB HL)	
Uncomfortable loudness level at 4 kHz (left ear) (dB HL)	
Uncomfortable loudness level at 4 kHz (right ear) (dB HL)	
Uncomfortable loudness level at 6 kHz (left ear) (dB HL)	
Uncomfortable loudness level at 6 kHz (right ear) (dB HL)	
Uncomfortable loudness level at 8 kHz (left ear) (dB HL)	
Uncomfortable loudness level at 8 kHz (right ear) (dB HL)	
Varying tinnitus loudness from day to day (no/yes)	
Vertigo (no/yes) (2tschq1esitsq)	
Weight (kg)	
World Health Organization's Quality of Life Environmental subscale (4-20)	
World Health Organization's Quality of Life Physical subscale (4-20)	
World Health Organization's Quality of Life Psychological subscale (4-20)	
World Health Organization's Quality of Life Social subscale (4-20)	

Appendix 5.9. Importance of variables for tinnitus classification based on the Boruta algorithm

	Mean Imp	Median Imp	Min Imp	Max Imp	Norm Hits	decision
U_2000_R	9.36	9.39	5.96	11.85	1	Confirmed
U_250_R	8.7	8.75	6.11	10.82	1	Confirmed
LAudio6	8.01	8.06	5.2	10.91	1	Confirmed
RAudio6	7.32	7.34	5.11	10.31	1	Confirmed
U_500_R	7.16	7.19	4.78	9.74	1	Confirmed
U_1000_R	7.03	7.07	4.53	9.3	1	Confirmed
U_125_R	6.39	6.42	3.84	8.62	0.99	Confirmed
U_3000_R	6.33	6.37	3.61	8.72	0.99	Confirmed
LAudio4	6.12	6.16	2.95	8.09	0.99	Confirmed
LAudio125	6.11	6.12	3.52	8.39	0.99	Confirmed
LAudio10	6.09	6.08	3.72	8.46	0.99	Confirmed
LAudio8	6.01	6.05	3.48	8.63	0.99	Confirmed
U_500_L	5.89	5.93	3.35	8.01	0.98	Confirmed
X4921_L	5.75	5.83	2.94	8.16	0.97	Confirmed
RAudio8	5.39	5.4	2.44	8.33	0.96	Confirmed
RAudio10	5.31	5.32	3.19	7.49	0.97	Confirmed
LAudio3	5.26	5.36	1.22	7.7	0.94	Confirmed
U_2000_L	5.24	5.27	2.1	7.67	0.96	Confirmed
U_1000_L	4.71	4.76	2.07	7.25	0.94	Confirmed
PS_R	4.59	4.66	1.76	7.68	0.89	Confirmed
PS_L	4.33	4.34	1.45	7.08	0.86	Confirmed
X5654_L	4.2	4.21	1.9	6.64	0.84	Confirmed
Weight	4.18	4.27	0.01	7.8	0.8	Confirmed
Ext_sounds_prob	4.13	4.2	1.58	6.11	0.83	Confirmed
X5273_L	4.08	4.07	1.13	6.51	0.83	Confirmed
U_4000_R	3.95	3.97	1.06	6.69	0.79	Confirmed
RAudio4	3.94	3.97	0.09	6.4	0.81	Confirmed
Hyperacusis	3.84	3.79	1.24	6.53	0.77	Confirmed
U_6000_R	3.82	3.86	1.19	6.16	0.76	Confirmed
U_8000_L	3.57	3.61	0.79	7.19	0.7	Confirmed
U_250_L	3.52	3.53	0.87	5.74	0.71	Confirmed
X4599_L	3.41	3.5	-0.41	6.42	0.65	Confirmed
X6962_R	3.37	3.38	-0.04	6.87	0.62	Confirmed
X3730_L	3.28	3.29	-0.03	5.92	0.62	Confirmed
U_125_L	3.06	3.12	0.01	5.58	0.59	Confirmed
RAudio140	3.2	3.2	0.44	5.96	0.56	Rejected
U_8000_R	3	3.03	0.41	5.61	0.51	Rejected
X6064_L	2.9	2.87	-0.56	5.7	0.48	Rejected
WHO_QoL.Physical	2.89	2.94	0.01	5.28	0.49	Rejected
U_3000_L	2.84	2.88	-0.38	5.18	0.48	Rejected
BMI	2.83	2.79	-0.26	6.28	0.46	Rejected
LAudio140	2.64	2.69	0.02	5.04	0.26	Rejected

X6494_R	2.49	2.55	-0.11	5.39	0.23	Rejected
MeanAllOct	2.49	2.47	-0.3	4.58	0.07	Rejected
X5273_R	2.48	2.53	0.01	4.38	0.18	Rejected
U_6000_L	2.43	2.37	0.44	4.78	0.06	Rejected
RAudio125	2.39	2.5	0.07	4.56	0.05	Rejected
subjectiveHearingProb_tschq	2.31	2.39	0.1	4.43	0.06	Rejected
X4921_R	2.17	2.12	0.21	4.68	0.03	Rejected
X5654_R	2.16	2.15	0.01	5.06	0.03	Rejected
RAudio3	2.1	2.12	-0.19	4.64	0.02	Rejected
X4287_L	1.98	2.06	0.6	3.59	0.02	Rejected
U_4000_L	1.97	1.85	-0.07	4.07	0.05	Rejected
X4003_L	1.88	1.75	0.77	3.59	0.01	Rejected
X7998_R	1.72	1.88	0.21	3.48	0	Rejected
X4003_R	1.71	1.62	0.25	3.66	0	Rejected
Height	1.61	1.65	-0.81	4.44	0.01	Rejected
Gender_f_m_2tschq_1esitsq	1.59	1.71	-0.97	3.16	0	Rejected
X3730_R	1.54	1.6	-0.33	3.45	0.01	Rejected
X7998_L	1.52	1.51	-0.6	3.71	0	Rejected
TSAIextasymAllLow	1.47	1.73	-0.21	2.72	0	Rejected
X6494_L	1.43	1.38	-0.32	3.09	0.01	Rejected
X6064_R	1.33	1.68	-0.27	2.48	0	Rejected
HADS_A.Total.score	1.33	1.5	-0.95	3.09	0	Rejected
X2636_L	1.18	0.98	-0.24	2.81	0	Rejected
X4599_R	1.14	1.25	-0.53	2.57	0	Rejected
X4287_R	1.14	1.31	-0.74	2.71	0	Rejected
X2832_R	1.06	1.13	0.12	1.96	0	Rejected
X3027_R	1.05	1.37	-2.38	3.66	0	Rejected
RAudio0125	1.03	1.23	-0.86	2.43	0	Rejected
any_devices	0.98	1.08	-1.15	2.56	0	Rejected
PSQ.Total.score	0.95	1.16	-1.22	3.15	0.01	Rejected
Type_L	0.95	0.67	-0.03	2.15	0	Rejected
X2832_L	0.95	1.03	-0.13	2.03	0	Rejected
X3027_L	0.92	0.92	-0.92	2.39	0	Rejected
X3251_L	0.83	0.73	-0.63	3	0	Rejected
X996_R	0.83	0.57	-0.93	3.09	0	Rejected
HearingAidNoYes	0.82	0.91	-0.93	2.12	0	Rejected
X2001_L	0.81	0.99	-1.69	2.8	0.01	Rejected
X7460_R	0.81	0.48	-0.71	2.52	0	Rejected
RAudio025	0.8	0.71	-0.52	2.03	0	Rejected
stress_presence_esitsq	0.75	0.93	-0.82	1.74	0	Rejected
depression_presence	0.75	0.99	-0.79	1.91	0	Rejected
X2636_R	0.68	0.76	-1.22	2.39	0	Rejected
X9853_R	0.65	0.28	-1.22	2.83	0	Rejected
other_hl	0.63	0.51	-1.53	2.13	0	Rejected
X6962_L	0.59	0.49	-1.24	3.03	0	Rejected

X3251_R	0.58	0.35	-0.95	2.79	0	Rejected
X9853_L	0.53	0.52	-1.56	2.31	0	Rejected
Alcohol	0.5	0.58	-1.47	1.63	0	Rejected
LAudio2	0.47	0.42	-0.59	2.38	0	Rejected
X3486_L	0.46	0.37	-1.44	2.13	0	Rejected
presbycusis	0.44	0.57	-1	1.62	0	Rejected
X1621_L	0.44	0.22	-0.82	2.04	0	Rejected
X1865_L	0.44	0.75	-1.55	2.08	0	Rejected
LAudio0125	0.41	0.4	-0.73	1.44	0	Rejected
WHO_QoL.Environment	0.41	0.65	-1.19	1.84	0	Rejected
RAudio05	0.4	0.38	-0.93	1.18	0	Rejected
LAudio1	0.37	0.4	-1.26	1.66	0	Rejected
Right750	0.37	0.33	-0.93	1.48	0	Rejected
X1152_L	0.34	0.34	-0.73	1.19	0	Rejected
RAudio1	0.34	-0.15	-0.88	2.22	0	Rejected
HADS_D.Total.score	0.31	0.53	-1.71	1.45	0	Rejected
low_bp	0.29	0.4	-1.02	1.93	0	Rejected
Left750	0.28	0.38	-0.91	2.01	0	Rejected
tmj_pain	0.26	0.27	-1.34	2.01	0	Rejected
prob_staying_asleep	0.25	0.18	-1.47	1.68	0	Rejected
X2294_L	0.23	0.33	-1.42	2.33	0	Rejected
anxiety_presence	0.23	0.08	-1.14	2.45	0	Rejected
X2148_L	0.22	-0.27	-1.24	2.78	0	Rejected
X8574_R	0.21	0.29	-1.81	1.58	0	Rejected
acute_otitis	0.2	0.38	-1.52	1.45	0	Rejected
any_pain_syndromes	0.2	0.26	-1.24	2.06	0	Rejected
X1074_R	0.19	0.03	-1.26	1.64	0	Rejected
X7460_L	0.18	-0.01	-1.5	2.27	0	Rejected
LAudio025	0.16	0.34	-0.94	0.96	0	Rejected
neck_pain	0.14	-0.26	-1.08	1.8	0	Rejected
X1738_L	0.13	0.27	-1.54	1.43	0	Rejected
X2294_R	0.13	0.12	-2.39	1.96	0	Rejected
prob_falling_asleep	0.12	0.38	-2.01	1.34	0	Rejected
diagnosed_condition_any	0.12	0.21	-1.85	1.48	0	Rejected
Left1500	0.11	-0.17	-0.78	1.36	0	Rejected
smoking_status	0.11	0.19	-2.68	2.13	0	Rejected
X1230_L	0.11	-0.08	-1.63	1.95	0	Rejected
Education	0.1	0	-1.08	1.64	0	Rejected
high_bp	0.09	0.41	-2.24	1.29	0	Rejected
Type_R	0.08	0.44	-1.41	2.33	0	Rejected
WHO_QoL.Psychological	0.07	-0.03	-0.94	1.65	0	Rejected
X1865_R	0.04	0.02	-2.08	1.56	0	Rejected
X2460_L	0.04	0.12	-1.42	1.75	0	Rejected
X2148_R	0.01	0.17	-2.09	1.25	0	Rejected

X1318_L	0.01	0.11	-1.74	1.66	0	Rejected
Age	0	0.17	-2.75	1.37	0	Rejected
Intro_11	-0.02	-0.13	-1.66	1.54	0	Rejected
X1416_R	-0.02	-0.04	-1.98	1.96	0	Rejected
TinnFamilyHist_no_yes	-0.02	-0.17	-1.81	1.73	0	Rejected
X1318_R	-0.03	-0.42	-1.64	1.75	0	Rejected
X1152_R	-0.03	-0.17	-1.59	1.73	0	Rejected
X8574_L	-0.04	-0.29	-1.3	1.59	0	Rejected
LAudio05	-0.04	0.12	-2.23	1.83	0	Rejected
X1513_L	-0.06	-0.05	-1.39	2.24	0	Rejected
X9189_R	-0.08	0.09	-1.15	0.95	0	Rejected
lyme	-0.09	-0.04	-1.54	1.92	0	Rejected
WHO_QoL.Social	-0.11	-0.22	-1.53	1.28	0	Rejected
X2001_R	-0.12	0.02	-1.53	1.72	0	Rejected
any_ear_condition	-0.13	-0.43	-1.6	1.49	0	Rejected
X3486_R	-0.14	0.2	-2.21	0.91	0	Rejected
any_procedure	-0.16	-0.07	-2.02	2.13	0	Rejected
X1738_R	-0.16	-0.29	-1.41	1.59	0	Rejected
night_work	-0.17	-0.39	-2.52	1.91	0	Rejected
X1416_L	-0.18	-0.1	-2.08	1.08	0	Rejected
ear_surg	-0.18	0.09	-2.12	1.89	0	Rejected
acid_reflux	-0.2	-0.02	-1.74	1.53	0	Rejected
X2460_R	-0.21	-0.12	-2.06	1.57	0	Rejected
dental_surg	-0.22	0	-1.96	1.35	0	Rejected
X1513_R	-0.24	-0.25	-1.58	0.84	0	Rejected
dental_problems	-0.24	-0.01	-2.13	1.65	0	Rejected
X1074_L	-0.24	-0.31	-1.76	1.66	0	Rejected
X1230_R	-0.26	-0.01	-2.34	1.31	0	Rejected
X9189_L	-0.35	-0.63	-1.55	2.18	0	Rejected
high_cholest	-0.35	-0.44	-1.42	1.42	0	Rejected
X1621_R	-0.36	-0.46	-1.89	1.6	0	Rejected
thyroid_disorder	-0.63	-0.72	-2.31	1.51	0	Rejected
X996_L	-0.63	-0.54	-2.88	0.82	0	Rejected
RAudio2	-0.65	-0.79	-1.34	1.22	0	Rejected
Right1500	-0.68	-0.45	-3.65	0.99	0	Rejected
Handedness_2tschq_1esitsq	-0.72	-0.54	-2.35	0.91	0	Rejected

Appendix 5.10. Variables differing significantly among subgroups for the six selected subgroupings

	STOP Audiometric Variables Clustering	BRC Audiometric Variables Clustering	STOP General Phenotypic Variables Clustering	STOP Tinnitus Discriminating Variables Clustering	STOP-BRC Independent Validation Clustering	STOP-ESIT Independent Validation Clustering
Age (y)	***	***	NS	***	NS	NS
Mean hearing threshold both ears (dB HL)	***	***	***	***	***	NS
Masking noise threshold using Narrow Band Noise (dB HL)	***	NA	NS	***	NS	NS
Tinnitus loudness matching (dB HL)	***	NA	NS	***	NS	NS
Tinnitus increased by jaw movement (no/yes)	***	NA	NS	NS	NS	NS
Tinnitus pitch matching (kHz)	***	***	NS	***	NS	NS
Self-reported hearing problem (no/yes)	***	***	NA	***	***	NA
Hospital Anxiety Depression Scale for Anxiety (0-21)	***	NA	***	***	NS	***
Thoughts of conditions related to increased tinnitus (no/yes)	***	NA	NS	NS	NS	***
Age at tinnitus onset (y)	***	NS	NS	***	NS	NS
Perceived Stress Questionnaire (0-1)	***	NA	***	***	NS	***
High cholesterol (no/yes)	***	NA	NS	NS	NS	NS
Tinnitus awareness (% of total awake time)	***	NA	NS	***	NS	NS
Tinnitus increased by stress (no/yes)	***	NS	NS	***	NS	***
Varying tinnitus loudness from day to day (no/yes)	***	NS	NA	***	NS	NA

Hearing threshold at pitch matched frequency (dB HL)	***	NA	NS	***	NS	NS
Any hearing device (no/yes)	***	NA	***	***	***	NS
Hearing aid use (no/yes)	***	***	***	***	***	NS
High BP (no/yes)	***	NA	NS	***	NS	NS
Tinnitus annoyance scale (0-100)	***	NS	***	NS	NS	NS
Tinnitus worsened by loud noise (no/yes)	***	NS	NA	***	NS	NA
Tinnitus increased by drinking alcohol (no/yes)	***	NA	NS	***	NS	NS
Tinnitus quality (tonal/noise/other)	***	NA	NS	NS	NS	NS
Any ear condition (no/yes)	***	NA	NS	NS	NS	NS
Hearing threshold at 12.5 kHz (right ear) (dB HL)	***	***	NS	***	NS	NS
Hearing threshold at 12.5 kHz (left ear) (dB HL)	***	***	NS	***	NS	NS
Hearing threshold at 10 kHz (right ear) (dB HL)	***	***	NS	***	NS	NS
Hearing threshold at 14 kHz (right ear) (dB HL)	***	***	NS	***	NS	NS
Hearing threshold at 10 kHz (left ear) (dB HL)	***	***	NS	***	NS	NS
AudioPCOverall	***	***	NS	***	***	NS
Hearing threshold at 14 kHz (left ear) (dB HL)	***	***	NS	***	NS	NS
Hearing threshold at 8 kHz (left ear) (dB HL)	***	***	NS	***	NS	NS
Hearing threshold at 8 kHz (right ear) (dB HL)	***	***	NS	***	NS	NS
Hearing threshold at 6 kHz (right ear) (dB HL)	***	***	NS	***	NS	NS

Hearing threshold at 6 kHz (left ear) (dB HL)	***	***	NS	***	NS	NS
Hearing threshold at 4 kHz (right ear) (dB HL)	***	***	NS	***	NS	NS
Hearing threshold at 4 kHz (left ear) (dB HL)	***	***	NS	***	***	NS
DPOAE at 5654 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 5654 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 5273 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
Hearing threshold at 3 kHz (right ear) (dB HL)	***	***	NS	***	NS	NS
Speech in noise hearing assessment - phoneme score (left ear) (0-100)	***	NA	NS	***	NS	NS
DPOAE at 4599 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 6064 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 6494 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
Speech in noise hearing assessment - phoneme score (right ear) (0-100)	***	NA	NS	***	NS	NS
DPOAE at 4921 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 6064 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	***	NS
AudioPCLowVHigh	***	NS	***	***	NS	***

DPOAE at 4921 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 4287 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 5273 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 4287 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 6494 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 3730 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 6962 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 4003 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	***	NS
Hearing threshold at 3 kHz (left ear) (dB HL)	***	***	NS	***	***	NS
DPOAE at 4599 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 4003 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 3730 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 6962 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 7460 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 7460 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS

DPOAE at 2636 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 3486 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 3486 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	***	NS
Hearing threshold at 2 kHz (left ear) (dB HL)	***	***	NS	***	NS	NS
AudioPCMiddleVLowhigh	***	NS	NS	NS	NS	NS
Hearing threshold at 2 kHz (right ear) (dB HL)	***	***	***	***	NS	NS
DPOAE at 3027 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 2832 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 2001 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 2832 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 1152 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 1230 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 996 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 1865 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 3251 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	***	NS
Uncomfortable loudness level at 8 kHz (right ear) (dB HL)	***	NA	***	***	NS	***

Hearing threshold at 0.125 kHz (left ear) (dB HL)	***	***	***	***	NS	NS
DPOAE at 2460 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 996 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 1621 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 1416 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 2001 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 2294 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 2294 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 1074 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 3251 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	***	NS
Uncomfortable loudness level at 8 kHz (left ear) (dB HL)	***	NA	***	***	NS	***
DPOAE at 2636 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 1074 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 1230 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 2148 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS

Hearing threshold at 0.25 kHz (left ear) (dB HL)	***	***	***	***	NS	NS
DPOAE at 3027 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	***	NS
DPOAE at 1621 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
Uncomfortable loudness level at 4 kHz (right ear) (dB HL)	***	NA	***	***	NS	NS
Hearing threshold at 0.125 kHz (right ear) (dB HL)	***	***	NS	NS	NS	NS
DPOAE at 1416 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 1513 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 1318 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
Hearing threshold at 1 kHz (right ear) (dB HL)	***	***	NS	***	NS	NS
Hearing threshold at 1 kHz (left ear) (dB HL)	***	***	***	***	NS	NS
DPOAE at 2460 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 1513 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
Uncomfortable loudness level at 6 kHz (left ear) (dB HL)	***	NA	***	***	NS	***
DPOAE at 1738 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	***	NS
Hearing threshold at 0.25 kHz (right ear) (dB HL)	***	***	***	NS	NS	NS

Uncomfortable loudness level at 6 kHz (right ear) (dB HL)	***	NA	***	***	NS	NS
DPOAE at 2148 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
Hearing threshold at 0.5 kHz (right ear) (dB HL)	***	***	***	***	NS	NS
DPOAE at 1865 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 1152 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
Uncomfortable loudness level at 4 kHz (left ear) (dB HL)	***	NA	***	***	NS	NS
DPOAE at 1318 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
Hearing threshold at 0.5 kHz (left ear) (dB HL)	***	***	***	***	NS	NS
Uncomfortable loudness level at 3 kHz (right ear) (dB HL)	***	NA	***	***	NS	NS
DPOAE at 1738 Hz (right ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 7998 Hz (left ear) (signal to noise ratio)	***	NA	NS	***	NS	NS
DPOAE at 7998 Hz (right ear) (signal to noise ratio)	***	NA	***	NS	NS	NS
Uncomfortable loudness level at 2 kHz (left ear) (dB HL)	***	NA	***	***	NS	NS
Onset of high BP in relation to tinnitus onset (never/before/at the same time/after)	***	NA	NS	***	NS	NS

Tinnitus spatial perception (left ear/both ears, more left/no lateralisation (both ears equally or in the head)/both ears, more right/right ear)**	NS	***	NA	NS	NS	NA
Tinnitus duration (y)	NA	***	NA	NA	NA	NA
Comfortable level of a 5 kHz pure tone (dB SPL)	NA	***	NA	NA	NA	NA
Hearing threshold at 11.2 kHz (left ear) (dB HL)	NA	***	NA	NA	NA	NA
Hearing threshold at 9 kHz (left ear) (dB HL)	NA	***	NA	NA	NA	NA
Hearing threshold at 11.2 kHz (right ear) (dB HL)	NA	***	NA	NA	NA	NA
Hearing threshold at 9 kHz (right ear) (dB HL)	NA	***	NA	NA	NA	NA
Hearing threshold at 0.75 kHz (left ear) (dB HL)	NA	***	NA	NA	NA	NA
Hearing threshold at 1.5 kHz (left ear) (dB HL)	NA	***	NA	NA	NA	NA
Hearing threshold at 1.5 kHz (right ear) (dB HL)	NA	***	NA	NA	NA	NA
Hearing threshold at 0.75 kHz (right ear) (dB HL)	NA	***	NA	NA	NA	NA
Depression (no/yes)	NS	NA	***	NS	NS	***
Neck pain (no/yes)	NS	NA	***	NS	NS	***
Problem with external sounds (small/moderate/big or very big)	NS	NA	***	NS	NS	***
Tinnitus increased by poor quality sleep (no/yes)	NS	NA	***	NS	NS	NS
TMJ pain (no/yes)	NS	NA	***	NS	NS	***

Tinnitus loudness rating (0-100)	NS	NS	***	NS	NS	NS
World Health Organization's Quality of Life Environmental subscale (4-20)	NS	NA	***	NS	NS	***
Anxiety (no/yes)	NS	NA	***	NS	NS	***
No management for tinnitus (no/yes)	NS	NA	***	NS	NS	***
Hospital Anxiety Depression Scale for Depression (0-21)	NS	NA	***	NS	***	***
Vertigo (no/yes)*	NA	NA	***	NA	NA	NS
Tinnitus worries, annoys or upsets (severely/moderately/slightly/not at all)	NS	NA	***	NS	NS	NS
Any pain syndromes (no/yes)	NS	NA	***	NS	NS	NS
Stress (no/yes)	NS	NA	***	NS	NS	***
World Health Organization's Quality of Life Physical subscale (4-20)	NS	NA	***	NS	NS	***
Hyperacusis questionnaire score (0-42)	NS	NS	***	NS	NS	***
World Health Organization's Quality of Life Psychological subscale (4-20)	NS	NA	***	NS	NS	***
Tinnitus Handicap Inventory (0-100)	NS	NA	***	NS	NS	***
Self-reported hearing difficulty (slight or no difficulty/moderate difficulty/severe difficulty/total loss)	NA	NA	***	NA	NA	NS
Uncomfortable loudness level at 0.5 kHz (left ear) (dB HL)	NS	NA	***	***	NS	NS
Uncomfortable loudness level at 0.125 kHz (right ear) (dB HL)	NS	NA	***	***	NS	NS

Uncomfortable loudness level at 0.125 kHz (left ear) (dB HL)	NS	NA	***	***	NS	NS
Uncomfortable loudness level at 2 kHz (right ear) (dB HL)	NS	NA	***	***	NS	NS
Uncomfortable loudness level at 1 kHz (right ear) (dB HL)	NS	NA	***	***	NS	NS
Uncomfortable loudness level at 1 kHz (left ear) (dB HL)	NS	NA	***	***	NS	NS
Uncomfortable loudness level at 0.5 kHz (right ear) (dB HL)	NS	NA	***	***	NS	NS
Uncomfortable loudness level at 0.25 kHz (left ear) (dB HL)	NS	NA	***	***	NS	NS
Uncomfortable loudness level at 0.25 kHz (right ear) (dB HL)	NS	NA	***	***	NS	NS
Uncomfortable loudness level at 3 kHz (left ear) (dB HL)	NS	NA	***	***	NS	NS
Onset of vertigo in relation to tinnitus onset (never/before/at the same time/after)	NS	NA	***	NS	NS	NS
Onset of problems with external sounds in relation to tinnitus onset (never/before/at the same time/after)	NS	NA	***	NS	NS	***
Onset of neck pain in relation to tinnitus onset (never/before/at the same time/after)	NS	NA	***	NS	NS	***
Onset of depression in relation to tinnitus onset (never/before/at the same time/after)	NS	NA	***	NS	NS	***

Onset of hearing difficulties in relation to tinnitus onset (never/before/at the same time/after)	NS	NA	***	NS	NS	NS
TMJ disorder (no/yes)**	NS	NS	NA	***	NS	NA
Presbycusis (no/yes)	NS	NA	NS	***	NS	NS
Sound exposure at tinnitus onset (no/yes)**	NS	NS	NA	***	NS	NA
Onset of headache in relation to tinnitus onset (never/before/at the same time/after)	NS	NA	NS	***	NS	NS
Height (cm)	NS	NA	NS	NS	***	NS
World Health Organization's Quality of Life Social subscale (4-20)	NS	NA	NS	NS	***	NS
Weight (kg)	NS	NA	NS	NS	***	NS
Presence during the day (constant/intermittent)	NS	NS	NS	NS	***	NS
Tinnitus daily presence (no/yes)	NS	NA	NS	NS	***	NS
Sex (female/male)**	NS	NS	NA	NS	***	NA
BMI (kg/m^2)	NS	NA	NS	NS	***	NS
DPOAE at 9189 Hz (left ear) (signal to noise ratio)	NS	NA	NS	NS	***	NS
Problem staying asleep (no/yes)	NS	NA	NS	NS	NS	***
Problem falling asleep (no/yes)	NS	NA	NS	NS	NS	***
Low BP (no/yes)	NS	NA	NS	NS	NS	***
Sex (female/male)*	NA	NA	NS	NA	NA	***
Any diagnosed condition (no/yes)	NS	NA	NS	NS	NS	***
Onset of problems staying asleep in relation to tinnitus onset (never/before/at the same time/after)	NS	NA	NS	NS	NS	***

*Variable assessed using the TSCHQ question for BRC dataset and the ESIT-SQ question for the STOP dataset.
**Variable assessed using the TSCHQ question for the BRC and the STOP datasets.
***: Variable showed statistically significantly different distribution among the two subgroups at alpha level 0.001. NS: not statistically significant difference. NA: not applicable (variable not assessed).

Appendix 5.11. Variables with non-zero coefficients from multivariable LASSO regression models for the six selected subgroupings

	STOP Audiometric Variables Clustering	BRC Audiometric Variables Clustering	STOP General Phenotypic Variables Clustering	STOP Tinnitus Discriminating Variables Clustering	STOP-BRC Independent Validation Clustering	STOP-ESIT Independent Validation Clustering
Mean hearing threshold both ears (dB HL)	***	***	***	***	***	NA
Age (y)	***	***	NA	***	NA	***
Depression (no/yes)	***	NA	***	Z	NA	NA
Vertigo (no/yes)**	***	NA	NA	***	NA	NA
Tinnitus increased by jaw movement (no/yes)	***	NA	NA	NA	***	***
Tinnitus loudness matching (dB HL)	***	NA	NA	Z	Z	NA
Masking noise threshold using Narrow Band Noise (dB HL)	***	NA	Z	***	***	Z
Hospital Anxiety Depression Scale for Anxiety (0-21)	***	NA	Z	***	NA	***
Tinnitus worries, annoys or upsets (severely/moderately/slightly/not at all)	***	NA	Z	NA	NA	NA
Self-reported hearing problem (no/yes)	***	Z	NA	***	***	NA
Tinnitus pitch matching (kHz)	***	Z	Z	Z	NA	NA
Tinnitus spatial perception (left ear/both ears, more left/no lateralisation (both ears equally or in the head)/both ears, more right/right ear)**	NA	***	NA	NA	***	NA
TMJ pain (no/yes)	NA	NA	***	***	NA	***
Problem with external sounds (small/moderate/big or very big)	NA	NA	***	NA	NA	***
No management for tinnitus (no/yes)	NA	NA	***	NA	NA	***
Tinnitus increased by high intensity sounds (no/yes)	NA	NA	***	NA	NA	NA
Tinnitus pitch (high/medium/low)	NA	NA	***	NA	NA	NA
Neck pain (no/yes)	NA	NA	***	Z	NA	***
Tinnitus increased by poor quality sleep (no/yes)	NA	NA	***	Z	NA	Z

No medication at tinnitus onset (no/yes)	NA	NA	NA	NA	***	***
Dental surgery (no/yes)	NA	NA	NA	NA	***	NA
Weight (kg)	NA	NA	NA	NA	***	NA
Education (lower/higher)	NA	NA	NA	NA	***	NA
Presence during the day (constant/intermittent)	NA	NA	NA	NA	***	NA
Height (cm)	NA	NA	NA	NA	***	Z
Problem staying asleep (no/yes)	NA	NA	NA	NA	NA	***
Thyroid disorder (no/yes)	NA	NA	NA	NA	NA	***
Low BP (no/yes)	NA	NA	NA	NA	NA	***
Sex (female/male)*	NA	NA	NA	NA	NA	***
Tinnitus reduced by driving (no/yes)	NA	NA	NA	NA	NA	***
Dental problems (no/yes)	NA	NA	NA	NA	NA	***
Acute Otitis (no/yes)	NA	NA	NA	NA	NA	***
Tinnitus reduced by jaw movement (no/yes)	NA	NA	NA	Z	NA	***
World Health Organization's Quality of Life Social subscale (4-20)	NA	NA	Z	NA	***	Z
Stress (no/yes)	NA	NA	Z	NA	NA	***
Problem falling asleep (no/yes)	NA	NA	Z	NA	NA	***
Any procedure (no/yes)	NA	NA	Z	NA	NA	***
Perceived Stress Questionnaire (0-1)	Z	NA	***	***	NA	***
World Health Organization's Quality of Life Environmental subscale (4-20)	Z	NA	***	NA	NA	Z
Tinnitus loudness rating (0-100)	Z	NA	***	NA	Z	NA
Anxiety (no/yes)	Z	NA	***	Z	NA	NA
Hearing threshold at pitch matched frequency (dB HL)	Z	NA	NA	***	***	***
High BP (no/yes)	Z	NA	NA	***	NA	NA
Lyme disease (no/yes)	Z	NA	NA	NA	***	NA
Alcohol (number of drinks per week)	Z	NA	NA	Z	***	NA
High cholesterol (no/yes)	Z	NA	NA	Z	***	NA

Thoughts of conditions related to increased tinnitus (no/yes)	Z	NA	NA	Z	NA	***
Hospital Anxiety Depression Scale for Depression (0-21)	Z	NA	Z	NA	***	***
Tinnitus annoyance scale (0-100)	Z	NA	Z	NA	***	Z
Any hearing device (no/yes)	Z	NA	Z	Z	***	NA
Tinnitus awareness (% of total awake time)	Z	NA	Z	Z	***	NA
World Health Organization's Quality of Life Psychological subscale (4-20)	Z	NA	Z	Z	NA	***
Tinnitus worsened by loud noise (no/yes)	Z	Z	NA	***	NA	NA
TMJ disorder (no/yes)**	Z	Z	NA	***	NA	NA
Hearing aid use (no/yes)	Z	Z	Z	Z	***	NA

*Variable assessed using the TSCHQ question for BRC dataset and the ESIT-SQ question for the STOP dataset.

**Variable assessed using the TSCHQ question for the BRC and the STOP datasets.

***: Variable was important for the LASSO regression model (had nonzero coefficient). Z: Variable was not important for the LASSO regression model (had zero coefficient). NA: not applicable (variable not included in the model).

Bibliography

Aazh, H., Mcferran, D. & Moore, B. C., 2018. Uncomfortable loudness levels among children and adolescents seeking help for tinnitus and/or hyperacusis. International journal of audiology, 57**,** 618-623.

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

Ahlqvist, E., Storm, P., Käräjämäki, A., Martinell, M., Dorkhan, M., Carlsson, A., Vikman, P., Prasad, R. B., Aly, D. M. & Almgren, P., 2018. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. The lancet Diabetes & endocrinology, 6**,** 361-369.

Al-Swiahb, J. & Park, S. N., 2016. Characterization of tinnitus in different age groups: A retrospective review. Noise & health, 18**,** 214.

Alelyani, S., Tang, J. & Liu, H., 2013. Feature selection for clustering: a review. Data clustering: algorithms and applications, 29**,** 144.

Anderson, G. P., 2008. Endotyping asthma: new insights into key pathogenic mechanisms in a complex, heterogeneous disease. The Lancet, 372**,** 1107-1119.

Andersson, G., Lyttkens, L. & Larsen, H., 1999. Distinguishing levels of tinnitus distress. Clinical Otolaryngology & Allied Sciences, 24**,** 404-410.

Andersson, G. & Mckenna, L., 1998. Tinnitus masking and depression. Audiology, 37**,** 174-82.

Asciutto, R., Lugo, A., Pacifici, R., Colombo, P., Rota, M., La Vecchia, C. & Gallus, S., 2016. The particular story of Italians' relation with alcohol: trends in individuals' consumption by age and beverage type. Alcohol and Alcoholism, 51**,** 347-353.

Attias, J., Shemesh, Z., Bleich, A., Solomon, Z., Bar-Or, G., Alster, J. & Sohmer, H., 1995. Psychological profile of help-seeking and non-help-seeking tinnitus patients. Scandinavian Audiology, 24**,** 13-18.

Auguie, B., Antonov, A. & Auguie, M. B., 2017. Package 'gridExtra'. Miscellaneous Functions for "Grid" Graphics.

Baguley, D., Mcferran, D. & Hall, D., 2013. Tinnitus. The Lancet, 382**,** 1600-1607.

Balarajan, M., Gray, M. & Mitchell, M. 2011. Monitoring equality: Developing a gender identity question, Equality and Human Rights Commission London.

Bartels, H., Middel, B. L., Van Der Laan, B. F., Staal, M. J. & Albers, F. W., 2008. The additive effect of co-occurring anxiety and depression on health status, quality of life and coping strategies in help-seeking tinnitus sufferers. Ear and hearing, 29**,** 947-956.

Bartels, H., Pedersen, S., Van Der Laan, B., Staal, M., Albers, F. & Middel, B., 2010. The impact of type D personality on health-related quality of life in tinnitus patients is mainly mediated by anxiety and depression. Otology & Neurotology, 31**,** 11-18.

Beck, A. T., Steer, R. A. & Carbin, M. G., 1988. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clinical psychology review, 8**,** 77-100.

Beck, A. T., Ward, C., Mendelson, M., Mock, J. & Erbaugh, J., 1961. Beck depression inventory (BDI). Arch Gen Psychiatry, 4**,** 561-571.

Beijers, L., Wardenaar, K. J., Van Loo, H. M. & Schoevers, R. A., 2019. Data-driven biological subtypes of depression: systematic review of biological approaches to depression subtyping. Molecular psychiatry, 24**,** 888-900.

Biswas, R., Lugo, A., Gallus, S., Akeroyd, M. A. & Hall, D. A., 2019. Standardized questions in English for estimating tinnitus prevalence and severity, hearing difficulty and usage of healthcare resources, and their translation into 11 European languages. Hearing research, 377**,** 330-338.

Boecking, B., Von Sass, J., Sieveking, A., Schaefer, C., Brueggemann, P., Rose, M. & Mazurek, B., 2020. Tinnitus-related distress and pain perceptions in patients with chronic tinnitus–Do psychological factors constitute a link? PloS one, 15**,** e0234807.

Boudier, A., Curjuric, I., Basagaña, X., Hazgui, H., Anto, J. M., Bousquet, J., Bridevaux, P. O., Dupuis-Lozeron, E., Garcia-Aymerich, J. & Heinrich, J., 2013. Ten-year follow-up of cluster-based asthma phenotypes in adults. A pooled analysis of three cohorts. American journal of respiratory and critical care medicine, 188**,** 550-560.

Breiman, L., Friedman, J., Olshen, R. & Stone, C. J. 1984. Classification and regression trees, Florida, CRC Press.

Briner, W., 1995. A behavioral nosology for tinnitus. Psychological Reports, 77**,** 27-34.

Brock, G., Pihur, V., Datta, S. & Datta, S., 2011. clValid, an R package for cluster validation. Journal of Statistical Software (Brock et al., March 2008).

Budd, R. J. & Pugh, R., 1996. Tinnitus coping style and its relationship to tinnitus severity and emotional distress. Journal of Psychosomatic Research, 41**,** 327-335.

Burkart, M., Brueggemann, P., Szczepek, A., Frank, D. & Mazurek, B., 2019. Intermittent tinnitus—an empirical description. HNO, 67**,** 51-58.

Caballé, N. C., Castillo-Sequera, J. L., Gómez-Pulido, J. A., Gómez-Pulido, J. M. & Polo-Luque, M. L., 2020. Machine learning applied to diagnosis of human diseases: A systematic review. Applied Sciences, 10**,** 5135.

Carpenter-Thompson, J. R., Schmidt, S., Mcauley, E. & Husain, F. T., 2015. Increased frontal response may underlie decreased tinnitus severity. PloS one, 10.

Cederroth, C. R., Lugo, A., Edvall, N. K., Lazar, A., Lopez-Escamez, J.-A., Bulla, J., Uhlen, I., Hoare, D. J., Baguley, D. M. & Canlon, B., 2020. Association between Hyperacusis and Tinnitus. Journal of Clinical Medicine, 9**,** 2412.

Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A. & Charrad, M. M., 2014. Package 'nbclust'. Journal of statistical software, 61**,** 1-36.

Cianfrone, G., Mazzei, F., Salviati, M., Turchetta, R., Orlando, M. P., Testugini, V., Carchiolo, L., Cianfrone, F. & Altissimi, G., 2015. Tinnitus Holistic Simplified Classification (THoSC) A New Assessment for Subjective Tinnitus, With Diagnostic and Therapeutic Implications. Annals of Otology, Rhinology & Laryngology, 124**,** 550-560.

Cima, R., Mazurek, B., Haider, H., Kikidis, D., Lapira, A., Noreña, A. & Hoare, D., 2019. A multidisciplinary European guideline for tinnitus: diagnostics, assessment, and treatment. Hno, 67**,** 10-42.

Cima, R. F., Crombez, G. & Vlaeyen, J. W., 2011. Catastrophizing and fear of tinnitus predict quality of life in patients with chronic tinnitus. Ear and hearing, 32**,** 634-641.

Côté, C., Baril, I., Morency, C.-È., Montminy, S., Couture, M., Leblond, J., Roos, M. & Roy, J.-S., 2019. Long-Term Effects of a Multimodal Physiotherapy Program on the Severity of Somatosensory Tinnitus and Identification of Clinical

Indicators Predicting Favorable Outcomes of the Program. Journal of the American Academy of Audiology, 30**,** 720-730.

D'arcy, S., Hamilton, C., Hughes, S., Hall, D. A., Vanneste, S., Langguth, B. & Conlon, B., 2017. Bi-modal stimulation in the treatment of tinnitus: a study protocol for an exploratory trial to optimise stimulation parameters and patient subtyping. BMJ open, 7.

Dahl, A. & Zaitlen, N., 2020. Genetic Influences on Disease Subtypes. Annual Review of Genomics and Human Genetics, 21**,** 413-435.

Dauman, R. & Tyler, R. Some considerations on the classification of tinnitus. Proceedings of the fourth international tinnitus seminar, Bordeaux, France, 1992. Kugler & Ghedini Publications The Netherlands, 225-229.

Dauman, R., Tyler, R. & Aran, J. Some considerations on the classification of tinnitus. Proceedings of the fourth international tinnitus seminar, Bordeaux, France, 1991. Kugler., Amsterdam, 225-229.

Davies, J., Gander, P. E., Andrews, M. & Hall, D. A., 2014. Auditory network connectivity in tinnitus patients: a resting-state fMRI study. International Journal of Audiology, 53**,** 192-198.

De Almeida, M. D. V., Graça, P., Afonso, C., D'amicis, A., Lappalainen, R. & Damkjaer, S., 1999. Physical activity levels and body weight in a nationally representative sample in the European Union. Public health nutrition, 2**,** 105-114.

De Ridder, D., Schlee, W., Vanneste, S., Londero, A., Weisz, N., Kleinjung, T., Shekhawat, G. S., Elgoyhen, A. B., Song, J.-J. & Andersson, G., 2021. Tinnitus and tinnitus disorder: Theoretical and operational definitions (an international multidisciplinary proposal). Progress in Brain Research.

De Ridder, D., Vanneste, S. & Congedo, M., 2011. The distressed brain: a group blind source separation analysis on tinnitus. PloS one, 6**,** e24273.

De Ridder, D., Vanneste, S., Weisz, N., Londero, A., Schlee, W., Elgoyhen, A. B. & 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.

Del Bo, L. & Ambrosetti, U., 2007. Hearing aids for the treatment of tinnitus. Progress in brain research, 166**,** 341-345.

Delogatis, L., 1977. SCL-90-R administration, scoring and procedures manual. Clinical Psychometric Research.

Dubno, J. R., 2019. New insights on age-related hearing loss. Journal of the South Carolina Academy of Science, 17**,** 3.

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

Dwyer, P., Saron, C. D. & Rivera, S. M., 2020. Identification of Longitudinal Sensory Subtypes in Typical Development and Autism Spectrum Development Using Growth Mixture Modelling. Research in Autism Spectrum Disorders, 78**,** 101645.

Edvall, N. K., Gunan, E., Genitsaridi, E., Lazar, A., Mehraei, G., Billing, M., Tullberg, M., Bulla, J., Whitton, J. & Canlon, B., 2019. Impact of temporomandibular joint complaints on tinnitus-related distress. Frontiers in neuroscience, 13**,** 879.

Eggermont, J. J. & Roberts, L. E., 2004. The neuroscience of tinnitus. Trends in neurosciences, 27**,** 676-682.

Elgoyhen, A. B., Langguth, B., De Ridder, D. & Vanneste, S., 2015. Tinnitus: perspectives from human neuroimaging. Nature Reviews Neuroscience, 16**,** 632-642.

Elhamifar, E. & Vidal, R., 2013. Sparse subspace clustering: Algorithm, theory, and applications. IEEE transactions on pattern analysis and machine intelligence, 35**,** 2765-2781.

Erlandsson, S. I., Rubinstein, B., Axelsson, A. & Carlsson, S. G., 1991. Psychological dimensions in patients with disabling tinnitus and craniomandibular disorders. Br J Audiol, 25**,** 15-24.

Esit. 2017. European School for Interdisciplinary Tinnitus Research (ESIT) - Homepage [Online]. Available: https://esit.tinnitusresearch.net/ [Accessed 30 June 2020].

Esit. 2018. European School for Interdisciplinary Tinnitus Research (ESIT) Database - Homepage [Online]. Available: https://www.esit-database.com/ [Accessed 22 June 2020].

Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd, 1996. 226-231.

Eugster, M. & Leisch, F., 2009. From spider-man to hero-archetypal analysis in R.

Evans, J. R. & Mathur, A., 2005. The value of online surveys. Internet research.

Evered, D. & Lawrenson, G., 1981. Tinnitus: Ciba Foundation Symposium 85. Pitman, London.

Feczko, E., Balba, N., Miranda-Dominguez, O., Cordova, M., Karalunas, S., Irwin, L., Demeter, D., Hill, A., Langhorst, B. & Painter, J. G., 2018. Subtyping cognitive profiles in autism spectrum disorder using a functional random forest algorithm. Neuroimage, 172**,** 674-688.

Feczko, E., Miranda-Dominguez, O., Marr, M., Graham, A. M., Nigg, J. T. & Fair, D. A., 2019. The heterogeneity problem: approaches to identify psychiatric subtypes. Trends in cognitive sciences, 23**,** 584-601.

Ferreira, D., Nordberg, A. & Westman, E., 2020. Biological subtypes of Alzheimer disease: A systematic review and meta-analysis. Neurology, 94**,** 436-448.

Folmer, R. L., Martin, W. H. & Shi, Y., 2004. Tinnitus: questions to reveal the cause, answers to provide relief. Journal of Family Practice, 53**,** 532-541.

Forte, J. C., Perner, A. & Van Der Horst, I. C., 2019. The use of clustering algorithms in critical care research to unravel patient heterogeneity. Springer.

Fournier, P., Cuvillier, A. F., Gallego, S., Paolino, F., Paolino, M., Quemar, A., Londero, A. & Norena, A., 2018. A New Method for Assessing Masking and Residual Inhibition of Tinnitus. Trends Hear, 22**,** 2331216518769996.

Fraley, C. & Raftery, A. E., 2007. Bayesian regularization for normal mixture estimation and model-based clustering. Journal of classification, 24**,** 155-181.

Frejo, L., Martin-Sanz, E., Teggi, R., Trinidad, G., Soto-Varela, A., Santos-Perez, S., Manrique, R., Perez, N., Aran, I. & Almeida-Branco, M., 2017. Extended phenotype and clinical subgroups in unilateral Meniere disease: a cross-sectional study with cluster analysis. Clinical Otolaryngology, 42**,** 1172-1180.

Friedman, J., Hastie, T. & Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. Journal of statistical software, 33**,** 1.

Fritsche, G., Hueppe, M., Kukava, M., Dzagnidze, A., Schuerks, M., Yoon, M. S., Diener, H. C. & Katsarava, Z., 2007. Validation of a german language questionnaire for screening for migraine, tension-type headache, and trigeminal autonomic cephalgias. Headache, 47**,** 546-51.

Fuhlbrigge, A. L. & Castro, M., 2020. Precision medicine in asthma—using phenotypes to understand endotypes that lead us to new therapeutic options. The Journal of Allergy and Clinical Immunology: In Practice, 8**,** 496-497.

Fuller, T. E., Haider, H. F., Kikidis, D., Lapira, A., Mazurek, B., Norena, A., Rabau, S., Lardinois, R., Cederroth, C. R., Edvall, N. K., Brueggemann, P. G., Rosing, S. N., Kapandais, A., Lungaard, D., Hoare, D. J. & Cima, R. F. F., 2017. Different Teams, Same Conclusions? A Systematic Review of Existing Clinical Guidelines for the Assessment and Treatment of Tinnitus in Adults. Frontiers in Psychology, 8**,** 206.

Galili, T., Mitelpunkt, A., Shachar, N., Marcus-Kalish, M. & Benjamini, Y. Categorize, cluster, and classify: a 3-c strategy for scientific discovery in the medical informatics platform of the human brain project. International Conference on Discovery Science, 2014. Springer, 73-86.

Gallus, S., Bertuzzi, M., Tavani, A., Bosetti, C., Negri, E., La Vecchia, C., Lagiou, P. & Trichopoulos, D., 2002. Does coffee protect against hepatocellular carcinoma? British journal of cancer, 87**,** 956-959.

Gallus, S., Lugo, A., Garavello, W., Bosetti, C., Santoro, E., Colombo, P., Perin, P., La Vecchia, C. & Langguth, B., 2015. Prevalence and determinants of tinnitus in the Italian adult population. Neuroepidemiology, 45**,** 12-19.

Gallus, S., Lugo, A., Liu, X., Behrakis, P., Boffi, R., Bosetti, C., Carreras, G., Chatenoud, L., Clancy, L. & Continente, X., 2020. Who smokes in Europe? data from 12 European countries in the TACKSHS survey (2017-2018). Journal of Epidemiology**,** JE20190344.

Gamberger, D., Lavrač, N., Srivatsa, S., Tanzi, R. E. & Doraiswamy, P. M., 2017. Identification of clusters of rapid and slow decliners among subjects at risk for Alzheimer's disease. Scientific reports, 7**,** 1-12.

Garnier, S., 2018. viridis: Default Color Maps from "matplotlib". R package version 0.5. 1. 2018.

Genitsaridi, E., Hoare, D. J., Kypraios, T. & Hall, D. A., 2020a. A Review and a Framework of Variables for Defining and Characterizing Tinnitus Subphenotypes. Brain Sciences, 10**,** 938.

Genitsaridi, E., Kypraios, T., Edvall, N. K., Trpchevska, N., Canlon, B., Hoare, D. J., Cederroth, C. & Hall, D. A., 2020b. The spatial percept of tinnitus is associated with hearing asymmetry: subgroup comparisons. medRxiv.

Genitsaridi, E., Partyka, M., Gallus, S., Lopez-Escamez, J. A., Schecklmann, M., Mielczarek, M., Trpchevska, N., Santacruz, J. L., Schoisswohl, S., Riha, C., Lourenco, M., Biswas, R., Liyanage, N., Cederroth, C. R., Perez-Carpena, P., Devos, J., Fuller, T., Edvall, N. K., Hellberg, M. P., D'antonio, A., Gerevini, S., Sereda, M., Rein, A., Kypraios, T., Hoare, D. J., Londero, A., Pryss, R., Schlee, W. & Hall, D. A., 2019. Standardised profiling for tinnitus research: The European School for Interdisciplinary Tinnitus Research Screening Questionnaire (ESIT-SQ). Hear Res, 377**,** 353-359.

Gibrin, P. C. D., Melo, J. J. & Marchiori, L. Prevalence of tinnitus complaints and probable association with hearing loss, diabetes mellitus and hypertension in elderly. CoDAS, 2013. 176-80.

Gohel, D., 2020. flextable: Functions for Tabular Reporting.

Goodhill, V., 1950. The management of tinnitus. The Laryngoscope, 60**,** 442-450.

Grant, M. J. & Booth, A., 2009. A typology of reviews: an analysis of 14 review types and associated methodologies. Health information & libraries journal, 26**,** 91-108.

Grasgruber, P., Cacek, J., Kalina, T. & Sebera, M., 2014. The role of nutrition and genetics as key determinants of the positive height trend. Economics & Human Biology, 15**,** 81-100.

Green, M. J., Girshkin, L., Kremerskothen, K., Watkeys, O. & Quidé, Y., 2019. A systematic review of studies reporting data-driven cognitive subtypes across the psychosis spectrum. Neuropsychology Review**,** 1-15.

Greenwell, K., Sereda, M., Coulson, N., El Refaie, A. & Hoare, D. J., 2016. A systematic review of techniques and effects of self-help interventions for tinnitus: Application of taxonomies from health psychology. International journal of audiology, 55**,** S79-S89.

Guest, H., Dewey, R. S., Plack, C. J., Couth, S., Prendergast, G., Bakay, W. & Hall, D. A., 2018. The noise exposure structured interview (NESI): An instrument for the comprehensive estimation of lifetime noise exposure. Trends in hearing, 22**,** 2331216518803213.

Habtewold, T. D., Rodijk, L. H., Liemburg, E. J., Sidorenkov, G., Boezen, H. M., Bruggeman, R. & Alizadeh, B. Z., 2020. A systematic review and narrative synthesis of data-driven studies in schizophrenia symptoms and cognitive deficits. Translational Psychiatry, 10**,** 1-24.

Halkidi, M., Batistakis, Y. & Vazirgiannis, M., 2001. On clustering validation techniques. Journal of intelligent information systems, 17**,** 107-145.

Hall, D. A., Fackrell, K., Li, A. B., Thavayogan, R., Smith, S., Kennedy, V., Tinoco, C., Rodrigues, E. D., Campelo, P. & Martins, T. D., 2018. A narrative synthesis of research evidence for tinnitus-related complaints as reported by patients and their significant others. Health and quality of life outcomes, 16**,** 61.

Hall, D. A., Haider, H., Szczepek, A. J., Lau, P., Rabau, S., Jones-Diette, J., Londero, A., Edvall, N. K., Cederroth, C. R. & Mielczarek10, M., 2016. Systematic review of outcome domains and instruments used in clinical trials of tinnitus treatments in adults.

Hall, D. A., Hibbert, A., Smith, H., Haider, H. F., Londero, A., Mazurek, B., Fackrell, K. & Initiative, C. O. M. I. T., 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.

Hallam, R., Jakes, S. & Hinchcliffe, R., 1988. Cognitive variables in tinnitus annoyance. British Journal of Clinical Psychology, 27**,** 213-222.

Hallam, R., Rachman, S. & Hinchcliffe, R., 1984. Psychological aspects of tinnitus. Contributions to medical psychology, 3**,** 31-53.

Hamilton, D. & Lehman, H., 2020. Asthma phenotypes as a guide for current and future biologic therapies. Clinical reviews in allergy & immunology, 59**,** 160-174.

Han, J., Park, J., Park, S., Vidal, J., Ashaikh, H., Kim, D. & Park, S., 2020. Typewriter tinnitus: An investigative comparison with middle ear myoclonic tinnitus and its long-term therapeutic response to carbamazepine. Auris, Nasus, Larynx.

Handl, J., Knowles, J. & Kell, D. B., 2005. Computational cluster validation in post-genomic data analysis. Bioinformatics, 21**,** 3201-3212.

Hebert, S., Canlon, B., Hasson, D., Hanson, L. L. M., Westerlund, H. & Theorell, T., 2012. Tinnitus severity is reduced with reduction of depressive mood–a prospective population study in Sweden. PloS one, 7**,** e37733.

Heijneman, K. M., De Kleine, E., Wiersinga-Post, E. & Van Dijk, P., 2013. Can the tinnitus spectrum identify tinnitus subgroups? Noise and Health, 15**,** 101.

Hennig, C. 2004. A general robustness and stability theory for cluster analysis, Schwerpunkt Mathematische Statistik und Stochastische Prozesse, Fachbereich ….

Hennig, C., 2007. Cluster-wise assessment of cluster stability. Computational Statistics & Data Analysis, 52**,** 258-271.

Hennig, C., 2020. fpc: Flexible Procedures for Clustering. R package version 2.2-8 (2020).

Henry, J. A., 2016. "Measurement" of tinnitus. Otology & Neurotology, 37**,** e276-e285.

Henry, J. A., Dennis, K. C. & Schechter, M. A., 2005. General review of tinnitus. Journal of speech, language, and hearing research.

Henry, J. A. & Meikle, M. B., 2000. Psychoacoustic measures of tinnitus. Journal of the American Academy of Audiology, 11**,** 138-155.

Herrero, J., Valencia, A. & Dopazo, J., 2001. A hierarchical unsupervised growing neural network for clustering gene expression patterns. Bioinformatics, 17**,** 126-136.

Hesse, G., 2016. Evidence and evidence gaps in tinnitus therapy. GMS current topics in otorhinolaryngology, head and neck surgery, 15.

Hesser, H., Weise, C., Westin, V. Z. & Andersson, G., 2011. A systematic review and meta-analysis of randomized controlled trials of cognitive–behavioral therapy for tinnitus distress. Clinical psychology review, 31**,** 545-553.

Hiller, W. & Goebel, G., 2004. Rapid assessment of tinnitus-related psychological distress using the Mini-TQ. Int J Audiol, 43**,** 600-604.

Hiller, W. & Goebel, G., 2007. When tinnitus loudness and annoyance are discrepant: audiological characteristics and psychological profile. Audiol Neurootol, 12**,** 391-400.

Hoare, D. J., Kowalkowski, V. L. & Hall, D. A., 2012. Effects of frequency discrimination training on tinnitus: results from two randomised controlled trials. Journal of the Association for Research in Otolaryngology, 13**,** 543-559.

Hoare, D. J., Van Labeke, N., Mccormack, A., Sereda, M., Smith, S., Al Taher, H., Kowalkowski, V. L., Sharples, M. & Hall, D. A., 2014. Gameplay as a source of intrinsic motivation in a randomized controlled trial of auditory training for tinnitus. PloS one, 9.

Hoekstra, C. E., Wesdorp, F. & Van Zanten, G., 2014. Socio-Demographic, Health, and Tinnitus Related Variables Affecting Tinnitus Severity. Ear and Hearing, 35**,** 544-554.

Holgers, K.-M., Zöger, S. & Svedlund, K., 2005. Predictive factors for development of severe tinnitus suffering-further characterisation: Factores predictivos para el desarrollo de tinitus severo que sufren una caracterización adicional. International journal of audiology, 44**,** 584-592.

Honkoop, P. J. & Chavannes, N. H., 2020. Asthma phenotypes in primary care. Nature Publishing Group.

Horne, E., Tibble, H., Sheikh, A. & Tsanas, A., 2020. Challenges of Clustering Multimodal Clinical Data: Review of Applications in Asthma Subtyping. JMIR Medical Informatics, 8**,** e16452.

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

Jastreboff, P., 2004. Tinnitus retraining therapy (TRT): clinical implementation of the model. Tinnitus retraining therapy: implementing the neurophysiological model**,** 63-144.

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

Jastreboff, P. J. & Hazell, J. W. 2004. Tinnitus retraining therapy: Implementing the neurophysiological model, Cambridge University Press.

Jeon, H. W., Kim, S. Y., Choi, B. S., Bae, Y. J., Koo, J.-W. & Song, J.-J., 2016. Pseudo-low frequency hearing loss and its improvement after treatment may be objective signs of significant vascular pathology in patients with pulsatile tinnitus. Otology & Neurotology, 37**,** 1344-1349.

Kassambara, A. 2017. Practical guide to principal component methods in R: PCA, M (CA), FAMD, MFA, HCPC, factoextra, STHDA.

Kassambara, A., 2018. ggpubr:"ggplot2" based publication ready plots. R package version 0.1, 7.

Kassambara, A. & Mundt, F., 2017. Package 'factoextra'. Extract and visualize the results of multivariate data analyses, 76.

Kaufman, L. & Rousseeuw, P. J. 1990. Finding groups in data: an introduction to cluster analysis, John Wiley & Sons.

Kaur, R. & Chupp, G., 2019. Phenotypes and endotypes of adult asthma: moving toward precision medicine. Journal of Allergy and Clinical Immunology, 144**,** 1-12.

Khalfa, S., Dubal, S., Veuillet, E., Perez-Diaz, F., Jouvent, R. & Collet, L., 2002. Psychometric normalization of a hyperacusis questionnaire. Orl, 64**,** 436-442.

Khan, A., 2018. collapsibleTree: Interactive Collapsible Tree Diagrams using 'D3.js'.

Kim, H.-J., Lee, H.-J., An, S.-Y., Sim, S., Park, B., Kim, S. W., Lee, J. S., Hong, S. K. & Choi, H. G., 2015. Analysis of the prevalence and associated risk factors of tinnitus in adults. PLoS One, 10**,** e0127578.

Kim, S. I., Kim, M. G., Kim, S. S., Byun, J. Y., Park, M. S. & Yeo, S. G., 2016. Evaluation of tinnitus patients by audiometric configuration. American journal of otolaryngology, 37**,** 1-5.

Kim, Y., Park, Y., Han, K., Vu, D., Cho, K. & Lee, S., 2018. Prevalence of tinnitus according to temporomanibular joint disorders and dental pain: the Korean National Population-based Study. Journal of oral rehabilitation.

Kirsch, C. A., Blanchard, E. B. & Parnes, S. M., 1989. Psychological characteristics of individuals high and low in their ability to cope with tinnitus. Psychosom Med, 51**,** 209-17.

Kisiel, M. A., Zhou, X., Sundh, J., Ställberg, B., Lisspers, K., Malinovschi, A., Sandelowsky, H., Montgomery, S., Nager, A. & Janson, C., 2020. Data-driven questionnaire-based cluster analysis of asthma in Swedish adults. NPJ primary care respiratory medicine, 30**,** 1-9.

Klockhoff, I. & Lindblom, U., 1967. Menière's disease and hydrochlorothiazide (Dichlotride®)—a critical analysis of symptoms and therapeutic effects. Acta oto-laryngologica, 63**,** 347-365.

Kloostra, F., Arnold, R., Hofman, R., Burgerhof, J. & Van Dijk, P., 2019. Models to predict positive and negative effects of cochlear implantation on tinnitus. Laryngoscope Investigative Otolaryngology, 4**,** 138-142.

Kojima, T., Kanzaki, S., Oishi, N. & Ogawa, K., 2017. Clinical characteristics of patients with tinnitus evaluated with the Tinnitus Sample Case History Questionnaire in Japan: A case series. PloS one, 12**,** e0180609.

Kontis, V., Bennett, J. E., Mathers, C. D., Li, G., Foreman, K. & Ezzati, M., 2017. Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble. The Lancet, 389**,** 1323-1335.

Koops, E. & Van Dijk, P., 2021. Hyperacusis in tinnitus patients relates to enlarged subcortical and cortical responses to sound except at the tinnitus frequency. Hearing Research, 401**,** 108158.

Koops, E. A., Husain, F. T. & Van Dijk, P., 2019. Profiling intermittent tinnitus: a retrospective review. International journal of audiology, 58**,** 1-7.

Kreuzer, P. M., Landgrebe, M., Schecklmann, M., Staudinger, S. & Langguth, B., 2012. Trauma-associated tinnitus: audiological, demographic and clinical characteristics. PLoS One, 7**,** e45599.

Krittanawong, C., Zhang, H., Wang, Z., Aydar, M. & Kitai, T., 2017. Artificial intelligence in precision cardiovascular medicine. Journal of the American College of Cardiology, 69**,** 2657-2664.

Kuhn, M., 2015. Caret: classification and regression training. Astrophysics Source Code Library.

Kuk, F. K., Tyler, R. S., Russell, D. & Jordan, H., 1990. The psychometric properties of a tinnitus handicap questionnaire. Ear and hearing, 11**,** 434-445.

Kursa, M. B. & Rudnicki, W. R., 2010. Feature selection with the Boruta package. J Stat Softw, 36**,** 1-13.

Lanata, C. M., Paranjpe, I., Nititham, J., Taylor, K. E., Gianfrancesco, M., Paranjpe, M., Andrews, S., Chung, S. A., Rhead, B. & Barcellos, L. F., 2019. A phenotypic and genomics approach in a multi-ethnic cohort to subtype systemic lupus erythematosus. Nature communications, 10**,** 1-13.

Landgrebe, M., Zeman, F., Koller, M., Eberl, Y., Mohr, M., Reiter, J., Staudinger, S., Hajak, G. & 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**,** 42.

Langguth, B., Biesinger, E., Del Bo, L., De Ridder, D., Goodey, R., Herraiz, C., Kleinjung, T., Lainez, M. J., Landgrebe, M. & Paolino, M. 2011a. Algorithm for the diagnostic and therapeutic management of tinnitus. Textbook of tinnitus. Springer.

Langguth, B., Goodey, R., Azevedo, A., Bjorne, A., Cacace, A., Crocetti, A., Del Bo, L., De Ridder, D., Diges, I. & Elbert, T., 2007. Consensus for tinnitus patient assessment and treatment outcome measurement: Tinnitus Research Initiative meeting, Regensburg, July 2006. Progress in brain research, 166**,** 525-536.

Langguth, B., Hund, V., Landgrebe, M. & Schecklmann, M., 2017a. Tinnitus Patients with Comorbid Headaches: The Influence of Headache Type and Laterality on Tinnitus Characteristics. Frontiers in Neurology, 8**,** 440.

Langguth, B., Landgrebe, M., Kleinjung, T., Sand, G. P. & Hajak, G., 2011b. Tinnitus and depression. The world journal of biological psychiatry, 12**,** 489-500.

Langguth, B., Landgrebe, M., Schlee, W., Schecklmann, M., Vielsmeier, V., Steffens, T., Staudinger, S., Frick, H. & Frick, U., 2017b. Different Patterns of hearing loss among Tinnitus Patients: a latent class analysis of a large sample. Frontiers in neurology, 8.

Lau, P., Wollbrink, A., Wunderlich, R., Engell, A., Löhe, A., Junghöfer, M. & Pantev, C., 2018. Targeting heterogeneous findings in neuronal oscillations in tinnitus: Analyzing meg novices and mental health comorbidities. Frontiers in psychology, 9**,** 235.

Lê, S., Josse, J. & Husson, F., 2008. FactoMineR: an R package for multivariate analysis. Journal of statistical software, 25**,** 1-18.

Lee, C. & Van Der Schaar, M. Temporal Phenotyping using Deep Predictive Clustering of Disease Progression. International Conference on Machine Learning, 2020. PMLR, 5767-5777.

Lefaudeux, D., De Meulder, B., Loza, M. J., Peffer, N., Rowe, A., Baribaud, F., Bansal, A. T., Lutter, R., Sousa, A. R. & Corfield, J., 2017. U-BIOPRED clinical adult

asthma clusters linked to a subset of sputum omics. Journal of Allergy and Clinical Immunology, 139**,** 1797-1807.

Levenstein, S., Prantera, C., Varvo, V., Scribano, M. L., Berto, E., Luzi, C. & Andreoli, A., 1993. Development of the Perceived Stress Questionnaire: a new tool for psychosomatic research. Journal of psychosomatic research, 37**,** 19-32.

Levine, R. A. & Oron, Y. 2015. Tinnitus. Handbook of clinical neurology. Elsevier.

Lim, K., Smucny, J., Barch, D. M., Lam, M., Keefe, R. S. & Lee, J., 2020. Cognitive Subtyping in Schizophrenia: A Latent Profile Analysis. Schizophrenia Bulletin.

Lindblad, A.-C., Hagerman, B. & Rosenhall, U., 2011. Noise-induced tinnitus: a comparison between four clinical groups without apparent hearing loss. Noise and Health, 13**,** 423.

Lockwood, A. H., Salvi, R. J. & Burkard, R. F., 2002. Tinnitus. New England Journal of Medicine, 347**,** 904-910.

Lombardo, M. V., Lai, M.-C., Auyeung, B., Holt, R. J., Allison, C., Smith, P., Chakrabarti, B., Ruigrok, A. N., Suckling, J. & Bullmore, E. T., 2016. Unsupervised data-driven stratification of mentalizing heterogeneity in autism. Scientific Reports, 6**,** 1-15.

Lötvall, J., Akdis, C. A., Bacharier, L. B., Bjermer, L., Casale, T. B., Custovic, A., Lemanske Jr, R. F., Wardlaw, A. J., Wenzel, S. E. & Greenberger, P. A., 2011. Asthma endotypes: a new approach to classification of disease entities within the asthma syndrome. Journal of Allergy and Clinical Immunology, 127**,** 355-360.

Lugo, A., Asciutto, R., Pacifici, R., Colombo, P., La Vecchia, C. & Gallus, S., 2015. Smoking in Italy 2013-2014, with a focus on the young. Tumori Journal, 101**,** 529-534.

Lugo, A., Edvall, N. K., Lazar, A., Mehraei, G., Lopez-Escamez, J.-A., Bulla, J., Uhlen, I., Canlon, B., Gallus, S. & Cederroth, C. R., 2020. Relationship between headaches and tinnitus in a Swedish study. Scientific Reports, 10**,** 1-12.

Luo, Y., Eran, A., Palmer, N., Avillach, P., Levy-Moonshine, A., Szolovits, P. & Kohane, I. S., 2020. A multidimensional precision medicine approach identifies an autism subtype characterized by dyslipidemia. Nature Medicine, 26**,** 1375-1379.

Maas, I. L., Bruggemann, P., Requena, T., Bulla, J., Edvall, N. K., Hjelmborg, J. V. B., Szczepek, A. J., Canlon, B., Mazurek, B., Lopez-Escamez, J. A. & Cederroth, C. R., 2017. Genetic susceptibility to bilateral tinnitus in a Swedish twin cohort. Genet Med, 19**,** 1007-1012.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K., 2012. Cluster: cluster analysis basics and extensions. R package version, 1**,** 56.

Mahboubi, H., Oliaei, S., Kiumehr, S., Dwabe, S. & Djalilian, H. R., 2013. The prevalence and characteristics of tinnitus in the youth population of the United States. The Laryngoscope, 123**,** 2001-2008.

Main, C. J., 1983. The modified somatic perception questionnaire (MSPQ). Journal of Psychosomatic Research, 27**,** 503-514.

Margol-Gromada, M., Sereda, M. & Baguley, D. M., 2020. Readability assessment of self-report hyperacusis questionnaires. International journal of audiology, 59**,** 506-512.

Mariette, J. & Villa-Vialaneix, N., 2018. Unsupervised multiple kernel learning for heterogeneous data integration. Bioinformatics, 34**,** 1009-1015.

Marquand, A. F., Wolfers, T., Mennes, M., Buitelaar, J. & Beckmann, C. F., 2016. Beyond lumping and splitting: a review of computational approaches for stratifying psychiatric disorders. Biological psychiatry: cognitive neuroscience and neuroimaging, 1**,** 433-447.

Martines, F., Sireci, F., Cannizzaro, E., Costanzo, R., Martines, E., Mucia, M., Plescia, F. & Salvago, P., 2015. Clinical observations and risk factors for tinnitus in a Sicilian cohort. Eur Arch Otorhinolaryngol, 272**,** 2719-29.

Martinez, C., Wallenhorst, C., Mcferran, D. & Hall, D. A., 2015. Incidence rates of clinically significant tinnitus: 10-year trend from a cohort study in England. Ear and hearing, 36**,** e69.

Mathers, C. D., Stevens, G. A., Boerma, T., White, R. A. & Tobias, M. I., 2015. Causes of international increases in older age life expectancy. The Lancet, 385**,** 540-548.

Mccormack, A., Edmondson-Jones, M., Somerset, S. & Hall, D., 2016. A systematic review of the reporting of tinnitus prevalence and severity. Hearing research, 337**,** 70-79.

Meeus, O. M., Spaepen, M., Ridder, D. D. & Heyning, P. H. V. D., 2010. Correlation between hyperacusis measurements in daily ENT practice. International Journal of Audiology, 49**,** 7-13.

Meikle, M. B., Henry, J. A., Griest, S. E., Stewart, B. J., Abrams, H. B., Mcardle, R., Myers, P. J., Newman, C. W., Sandridge, S. & Turk, D. C., 2012. The tinnitus functional index: development of a new clinical measure for chronic, intrusive tinnitus. Ear and hearing, 33**,** 153-176.

Michiels, S., De Hertogh, W., Truijen, S. & Van De Heyning, P., 2015. Cervical spine dysfunctions in patients with chronic subjective tinnitus. Otology & Neurotology, 36**,** 741-745.

Michiels, S., Ganz Sanchez, T., Oron, Y., Gilles, A., Haider, H. F., Erlandsson, S., Bechter, K., Vielsmeier, V., Biesinger, E., Nam, E. C., Oiticica, J., De Medeiros, I. R. T., Bezerra Rocha, C., Langguth, B., Van De Heyning, P., De Hertogh, W. & Hall, D. A., 2018. Diagnostic Criteria for Somatosensory Tinnitus: A Delphi Process and Face-to-Face Meeting to Establish Consensus. Trends Hear, 22**,** 2331216518796403.

Michiels, S., Harrison, S., Vesala, M. & Schlee, W., 2019. The Presence of Physical Symptoms in Patients With Tinnitus: International Web-Based Survey. Interactive journal of medical research, 8**,** e14519.

Milner, R., Lewandowska, M., Ganc, M., Nikadon, J., Niedziałek, I., Jędrzejczak, W. W. & Skarżyński, H., 2020. Electrophysiological correlates of focused attention on low-and high-distressed tinnitus. PLoS ONE, 15.

Moffat, G., Adjout, K., Gallego, S., Thai-Van, H., Collet, L. & Norena, A. J., 2009. Effects of hearing aid fitting on the perceptual characteristics of tinnitus. Hear Res, 254**,** 82-91.

Moore, W. C., Meyers, D. A., Wenzel, S. E., Teague, W. G., Li, H., Li, X., D'agostino Jr, R., Castro, M., Curran-Everett, D. & Fitzpatrick, A. M., 2010. Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. American journal of respiratory and critical care medicine, 181**,** 315-323.

Mores, J. T., Bozza, A., Magni, C., Casali, R. L. & Do Amaral, M. I. R., 2019. Clinical profile and implications of tinnitus in individuals with and without hearing loss. CEP, 13083**,** 887.

Moring, J., Bowen, A., Thomas, J. & Bira, L., 2016. The emotional and functional impact of the type of tinnitus sensation. Journal of Clinical Psychology in Medical Settings, 23**,** 310-8.

Moseley, A., Elkins, M., Herbert, R., Maher, C. & Sherrington, C., 2009. Cochrane reviews used more rigorous methods than non-Cochrane reviews: survey of systematic reviews in physiotherapy. Journal of clinical epidemiology, 62**,** 1021.

Murtagh, F. & Legendre, P., 2011. Ward's hierarchical clustering method: Clustering criterion and agglomerative algorithm. arXiv preprint arXiv:1111.6285.

National Center for Rehabilitative Auditory Research. 2017. Tinnitus Questionnaires [Online]. Available: https://www.ncrar.research.va.gov/Education/Documents/TinnitusDocuments/TinnitusQuestionnaires.asp [Accessed 17 June 2020].

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

Newman, C. W., Wharton, J. A. & Jacobson, G. P., 1997. Self-focused and somatic attention in patients with tinnitus. JOURNAL-AMERICAN ACADEMY OF AUDIOLOGY, 8**,** 143-149.

Nicolas-Puel, C., Faulconbridge, R. L., Guitton, M., Puel, J.-L., Mondain, M. & Uziel, A., 2002. Characteristics of tinnitus and etiology of associated hearing loss: a study of 123 patients. The international tinnitus journal, 8**,** 37-44.

Niemann, U., Boecking, B., Brueggemann, P., Mazurek, B. & Spiliopoulou, M., 2020a. Gender-Specific Differences in Patients With Chronic Tinnitus—Baseline Characteristics and Treatment Effects. Frontiers in Neuroscience, 14**,** 487.

Niemann, U., Boecking, B., Brueggemann, P., Mebus, W., Mazurek, B. & Spiliopoulou, M., 2020b. Tinnitus-related distress after multimodal treatment can be characterized using a key subset of baseline variables. PloS one, 15**,** e0228037.

Niemann, U., Brueggemann, P., Boecking, B., Mazurek, B. & Spiliopoulou, M., 2020c. Development and internal validation of a depression severity prediction model for tinnitus patients based on questionnaire responses and socio-demographics. Scientific reports, 10**,** 1-9.

Nodar, R. H., 1978. Tinnitus aurium: An approach to classification. Otolaryngology, 86**,** ORL-40-ORL-46.

Nodar, R. H., 1996. Tinnitus reclassified: new oil in an old lamp. Otolaryngology--Head and Neck Surgery, 114**,** 582-585.

Norena, A., 2011a. An integrative model of tinnitus based on a central gain controlling neural sensitivity. Neurosci Biobehav Rev, 35**,** 1089-109.

Norena, A., Cransac, H. & Chery-Croze, S., 1999. Towards an objectification by classification of tinnitus. Clinical Neurophysiology, 110**,** 666-675.

Norena, A. J., 2011b. An integrative model of tinnitus based on a central gain controlling neural sensitivity. Neurosci Biobehav Rev, 35**,** 1089-109.

Obbarius, A., Fischer, F., Liegl, G., Obbarius, N., Van Bebber, J., Hofmann, T. & Rose, M., 2020. A Step Towards a Better Understanding of Pain Phenotypes: Latent Class Analysis in Chronic Pain Patients Receiving Multimodal Inpatient Treatment. Journal of Pain Research, 13**,** 1023.

Ohukainen, P., Kuusisto, S., Kettunen, J., Perola, M., Järvelin, M.-R., Mäkinen, V.-P. & Ala-Korpela, M., 2020. Data-driven multivariate population subgrouping via lipoprotein phenotypes versus apolipoprotein B in the risk assessment of coronary heart disease. Atherosclerosis, 294**,** 10-15.

Oostendorp, R. A., Bakker, I., Elvers, H., Mikolajewska, E., Michiels, S., De Hertogh, W. & Samwel, H., 2016. Cervicogenic somatosensory tinnitus: An indication for manual therapy plus education? Part 2: A pilot study. Manual Therapy, 23**,** 106-113.

Parsons, L., Haque, E. & Liu, H., 2004. Subspace clustering for high dimensional data: a review. Acm sigkdd explorations newsletter, 6**,** 90-105.

Parthasarathy, A., Pinto, S. R., Lewis, R. M., Goedicke, W. & Polley, D. B., 2020. Data-driven segmentation of audiometric phenotypes across a large clinical cohort. Scientific reports, 10**,** 1-12.

Peters, T. T., Van Den Berge, M. J., Free, R. H., Van Der Vliet, A. M., Knoppel, H., Van Dijk, P. & Hofman, R., 2020. The Relation Between Tinnitus and a Neurovascular Conflict of the Cochleovestibular Nerve on Magnetic Resonance Imaging. Otology & Neurotology, 41**,** e124-e131.

Pierre-Jean, M., Deleuze, J.-F., Le Floch, E. & Mauger, F., 2020. Clustering and variable selection evaluation of 13 unsupervised methods for multi-omics data integration. Briefings in bioinformatics, 21**,** 2011-2030.

Pina, A., Macedo, M. P. & Henriques, R. 2020. Clustering Clinical Data in R. Mass Spectrometry Data Analysis in Proteomics. Springer.

Planey, K. 2015. Unsupervised Learning Across Multiple Datasets. Stanford University.

Prayuenyong, P., Kasbekar, A. V. & Baguley, D. M., 2020. The efficacy of statins as otoprotective agents: A systematic review. Clinical Otolaryngology, 45**,** 21-31.

Probst, T., Pryss, R. C., Langguth, B., Spiliopoulou, M., Landgrebe, M., Vesala, M., Harrison, S., Schobel, J., Reichert, M. & Stach, M., 2017. Outpatient Tinnitus Clinic, Self-Help Web Platform, or Mobile Application to Recruit Tinnitus Study Samples? Frontiers in aging neuroscience, 9**,** 113.

R Core Team, 2020. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rabelo-Da-Ponte, F. D., Lima, F. M., Martinez-Aran, A., Kapczinski, F., Vieta, E., Rosa, A. R., Kunz, M. & Czepielewski, L. S., 2020. Data-driven cognitive phenotypes in subjects with bipolar disorder and their clinical markers of severity. Psychological Medicine**,** 1-8.

Ralli, M., Altissimi, G., Turchetta, R., Mazzei, F., Salviati, M., Cianfrone, F., Orlando, M. P., Testugini, V. & Cianfrone, G., 2016. Somatosensory tinnitus: Correlation between cranio-cervico-mandibular disorder history and somatic modulation. Audiology and Neurotology, 21**,** 372-382.

Ralli, M., Greco, A., Boccassini, A., Altissimi, G., Di Paolo, C., Falasca, V., De Virgilio, A., Polimeni, A., Cianfrone, G. & De Vincentiis, M., 2018. Subtyping patients with somatic tinnitus: Modulation of tinnitus and history for somatic dysfunction help identify tinnitus patients with temporomandibular joint disorders. Plos one, 13**,** e0202050.

Ren, K., 2016. Rlist: A toolbox for non-tabular data manipulation. R package version 0.4.

Rendón, E., Abundez, I., Arizmendi, A. & Quiroz, E. M., 2011. Internal versus external cluster validation indexes. International Journal of computers and communications, 5**,** 27-34.

Reyes, A., Kaestner, E., Ferguson, L., Jones, J. E., Seidenberg, M., Barr, W. B., Busch, R. M., Hermann, B. P. & Mcdonald, C. R., 2020. Cognitive phenotypes in temporal lobe epilepsy utilizing data-and clinically driven approaches: Moving toward a new taxonomy. Epilepsia, 61**,** 1211-1220.

Rizzardo, R., Savastano, M., Maron, M. B., Mangialaio, M. & Salvadori, L., 1998. Psychological distress in patients with tinnitus. J Otolaryngol, 27**,** 21-5.

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

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C. & Müller, M., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC bioinformatics, 12**,** 77.

Rodriguez, M. Z., Comin, C. H., Casanova, D., Bruno, O. M., Amancio, D. R., Costa, L. D. F. & Rodrigues, F. A., 2019. Clustering algorithms: A comparative approach. PloS one, 14**,** e0210236.

Rousseeuw, P. J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 20**,** 53-65.

Rueda, A. & Krishnan, S., 2018. Clustering Parkinson's and age-related voice impairment signal features for unsupervised learning. Advances in Data Science and Adaptive Analysis, 10**,** 1840007.

Ryan, J. M., Corry, J. R., Attewell, R. & Smithson, M. J., 2002. A comparison of an electronic version of the SF-36 General Health Questionnaire to the standard paper version. Quality of life research, 11**,** 19-26.

Sand, P. G., Langguth, B., Schecklmann, M. & Kleinjung, T., 2012. GDNF and BDNF gene interplay in chronic tinnitus. International Journal of Molecular Epidemiology and Genetics: IJMEG, 3**,** 245.

Sanders, P. J., Doborjeh, Z. G., Doborjeh, M. G., Kasabov, N. K. & Searchfield, G. D., 2021. Prediction of Acoustic Residual Inhibition of Tinnitus Using a Brain-Inspired Spiking Neural Network Model. Brain Sciences, 11**,** 52.

Saria, S. & Goldenberg, A., 2015. Subtyping: What it is and its role in precision medicine. IEEE Intelligent Systems, 30**,** 70-75.

Schechter, M. A. & Henry, J. A., 2002. Assessment and Treatment of Tinnitus Patients Using a "Masking Approach". Journal of the American Academy of Audiology, 13**,** 545-558.

Schecklmann, M., Landgrebe, M., Langguth, B. & Group, T. D. S., 2014. Phenotypic characteristics of hyperacusis in tinnitus. PLoS One, 9**,** e86944.

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

Schlee, W., Hall, D. A., Canlon, B., Cima, R. F. F., De Kleine, E., Hauck, F., Huber, A., Gallus, S., Kleinjung, T., Kypraios, T., Langguth, B., Lopez-Escamez, J. A., Lugo, A., Meyer, M., Mielczarek, M., Norena, A., Pfiffner, F., Pryss, R. C., Reichert, M., Requena, T., Schecklmann, M., Van Dijk, P., Van De Heyning, P., Weisz, N. & Cederroth, C. R., 2018. Innovations in Doctoral Training and Research on Tinnitus: The European School on Interdisciplinary Tinnitus Research (ESIT) Perspective. Frontiers in Aging Neuroscience, 9**,** 447.

Schmidt, S. A., Carpenter-Thompson, J. & Husain, F. T., 2017. Connectivity of precuneus to the default mode and dorsal attention networks: A possible invariant marker of long-term tinnitus. NeuroImage: Clinical, 16**,** 196-204.

Schmidt, S. A., Zimmerman, B., Medina, R. O. B., Carpenter-Thompson, J. R. & Husain, F. T., 2018. Changes in gray and white matter in subgroups within the tinnitus population. Brain research, 1679**,** 64-74.

Schoettler, N. & Strek, M. E., 2020. Recent advances in severe asthma: from phenotypes to personalized medicine. Chest, 157**,** 516-528.

Scott, B., Lindberg, P., Melin, L. & Lyttkens, L., 1990. Predictors of tinnitus discomfort, adaptation and subjective loudness. British journal of audiology, 24**,** 51-62.

Scrucca, L., Fop, M., Murphy, T. B. & Raftery, A. E., 2016. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. The R journal, 8**,** 289.

Shabbir, M., Akeroyd, M. A. & Hall, D. A., 2021. A comprehensive literature search to identify existing measures assessing "concentration" as a core outcome domain for sound-based interventions for chronic subjective tinnitus in adults.

Shargorodsky, J., Curhan, G. C. & Farwell, W. R., 2010. Prevalence and characteristics of tinnitus among US adults. The American journal of medicine, 123**,** 711-718.

Sheldrake, J., Diehl, P. U. & Schaette, R., 2015. Audiometric characteristics of hyperacusis patients. Frontiers in neurology, 6**,** 105.

Shore, S., Zhou, J. & Koehler, S., 2007. Neural mechanisms underlying somatic tinnitus. Progress in brain research, 166**,** 107-548.

Shulman, A., 1991. Classification of tinnitus. Tinnitus, Diagnosis and Treatment**,** 248-252.

Shulman, A., 1992. Subjective idiopathic tinnitus: a unified plan of management. American journal of otolaryngology, 13**,** 63-74.

Shulman, A. & Goldstein, B., 2010. Principles of tinnitology: tinnitus diagnosis and treatment a tinnitus-targeted therapy. International Tinnitus Journal, 16**,** 73-85.

Silman, S. & Silverman, C. A., 1997. Basic audiologic testing. Silman S, Silverman CA. Auditory diagnosis: principles and applications. San Diego: Singular Publishing Group**,** 44-52.

Simões, J., Schlee, W., Schecklmann, M., Langguth, B., Farahmand, D. & Neff, P., 2019. Big five personality traits are Associated with tinnitus improvement over time. Scientific Reports, 9**,** 1-9.

Simoes, J. P., Neff, P., Schoisswohl, S., Bulla, J., Schecklmann, M., Harrison, S. S., Vesala, M., Langguth, B. & Schlee, W., 2019. Towards personalized tinnitus treatment: An exploratory study based on internet crowdsensing. Frontiers in public health, 7**,** 157.

Sismanis, A. 2005. Tinnitus: advances in diagnosis and management, American Academy of Otolaryngology--Head And Neck Surgery Foundation.

Słomko, J., Estévez-López, F., Kujawski, S., Zawadka-Kunikowska, M., Tafil-Klawe, M., Klawe, J. J., Morten, K. J., Szrajda, J., Murovska, M. & Newton, J. L., 2020. Autonomic Phenotypes in Chronic Fatigue Syndrome (CFS) Are Associated with Illness Severity: A Cluster Analysis. Journal of clinical medicine, 9**,** 2531.

Sobanski, V., Giovannelli, J., Allanore, Y., Riemekasten, G., Airò, P., Vettori, S., Cozzi, F., Distler, O., Matucci-Cerinic, M. & Denton, C., 2019. Phenotypes determined by cluster analysis and their survival in the prospective European Scleroderma Trials and Research cohort of patients with systemic sclerosis. Arthritis & Rheumatology, 71**,** 1553-1570.

Song, J.-J., De Ridder, D., Schlee, W., Van De Heyning, P. & Vanneste, S., 2013. "Distressed aging": the differences in brain activity between early-and late-onset tinnitus. Neurobiology of aging, 34**,** 1853-1863.

Song, J.-J., De Ridder, D., Weisz, N., Schlee, W., Van De Heyning, P. & Vanneste, S., 2014. Hyperacusis-associated pathological resting-state brain oscillations in the tinnitus brain: a hyperresponsiveness network with paradoxically inactive auditory cortex. Brain Structure and Function, 219**,** 1113-1128.

Song, J.-J., Vanneste, S., Schlee, W., Van De Heyning, P. & De Ridder, D., 2015. Onset-related differences in neural substrates of tinnitus-related distress: the anterior cingulate cortex in late-onset tinnitus, and the frontal cortex in early-onset tinnitus. Brain Structure and function, 220**,** 571-584.

Steer, R. A. & Beck, A. T., 1997. Beck Anxiety Inventory.

Stekhoven, D. J., 2015. missForest: Nonparametric missing value imputation using random forest. ascl**,** ascl: 1505.011.

Stouffer, J. L. & Tyler, R. S., 1990. Characterization of tinnitus by tinnitus patients. J Speech Hear Disord, 55**,** 439-53.

Stronge, S., Cichocka, A. & Sibley, C. G., 2019. The heterogeneity of self-regard: A latent transition analysis of self-esteem and psychological entitlement. Journal of Research in Personality, 82**,** 103855.

Suzuki, F. a. D. B., Suzuki, F. A., Onishi, E. T. & Penido, N. O., 2018. Psychoacoustic classification of persistent tinnitus. Brazilian journal of otorhinolaryngology, 84**,** 583-590.

Swedish Tinnitus Outreach Project. 2015. STOP Homepage [Online]. Available: https://stop.ki.se [Accessed 17 June 2020].

Szczepek, A. J., Frejo, L., Vona, B., Trpchevska, N., Cederroth, C. R., Caria, H. & Lopez-Escamez, J. A., 2019. Recommendations on collecting and storing samples for genetic studies in hearing and tinnitus research. Ear and hearing, 40**,** 219.

Sztuka, A., Pospiech, L., Gawron, W. & Dudek, K., 2010. DPOAE in estimation of the function of the cochlea in tinnitus patients with normal hearing. Auris Nasus Larynx, 37**,** 55-60.

Tavani, A., Gallus, S., La Vecchia, C., Dal Maso, L., Negri, E., Pelucchi, C., Montella, M., Conti, E., Carbone, A. & Franceschi, S., 2001. Physical activity and risk of ovarian cancer: an Italian case-control study. International journal of cancer, 91**,** 407-411.

Tellegen, A. & Waller, N. G., 2008. Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. The SAGE handbook of personality theory and assessment, 2**,** 261-292.

Theodoridis, S. & Koutroumbas, K., 2009. Pattern recognition. Fourth ed.: Academic Press.

Thompson, P., Pryce, H. & Refaie, E., 2011. Group or individual tinnitus therapy: What matters to participants? Audiological Medicine, 9**,** 110-116.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58**,** 267-288.

Tin-Act. 2017. TINnitus - Assessment, Causes and Treatments (TIN-ACT) - Homepage [Online]. Available: https://www.tinact.eu/ [Accessed 03 March 2021].

Tinnet. 2014. TINnitus research NETwork (TINNET) - Homepage [Online]. Available: https://tinnet.tinnitusresearch.net/ [Accessed 3 March 2021].

Tomassen, P., Vandeplas, G., Van Zele, T., Cardell, L.-O., Arebro, J., Olze, H., Förster-Ruhrmann, U., Kowalski, M. L., Olszewska-Ziąber, A. & Holtappels, G., 2016. Inflammatory endotypes of chronic rhinosinusitis based on cluster analysis of biomarkers. Journal of Allergy and Clinical Immunology, 137**,** 1449-1456. e4.

Tritt, K., Von Heymann, F., Zaudig, M., Zacharias, I., Söllner, W. & Loew, T., 2008. Development of the" ICD-10-Symptom-Rating"(ISR) questionnaire. Zeitschrift fur Psychosomatische Medizin und Psychotherapie, 54**,** 409-418.

Trpchevska, N., Bulla, J., Prada Hellberg, M., Edvall, N. K., Lazar, A., Mehraei, G., Uhlen, I., Schlee, W., Canlon, B. & Gallus, S., 2020. Sex-Dependent Aggregation of Tinnitus in Swedish Families. Journal of clinical medicine, 9**,** 3812.

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., 2014. Clinical practice guideline: tinnitus. Otolaryngology–Head and Neck Surgery, 151**,** S1-S40.

Tyler, R., Coelho, C., Tao, P., Ji, H., Noble, W., Gehringer, A. & Gogel, S., 2008. Identifying tinnitus subgroups with cluster analysis. Am J Audiol, 17**,** S176-84.

Uniti. 2020. Unification of Treatments and Interventions for Tinnitus Patients (UNITI) - Homepage [Online]. Available: https://uniti.tinnitusresearch.net/ [Accessed].

Van Den Berge, M. J., Free, R. H., Arnold, R., De Kleine, E., Hofman, R., Van Dijk, J. M. C. & Van Dijk, P., 2017. Cluster analysis to identify possible subgroups in tinnitus patients. Frontiers in neurology, 8**,** 115.

Van Der Wal, A., Luyten, T., Cardon, E., Jacquemin, L., Vanderveken, O. M., Topsakal, V., Van De Heyning, P., De Hertogh, W., Van Looveren, N. & Van Rompaey, V., 2020. Sex Differences in the Response to Different Tinnitus Treatment. Frontiers in Neuroscience, 14**,** 422.

Vandekerckhove, O., De Buck, E. & Van Wijngaerden, E., 2019. Lyme disease in Western Europe: an emerging problem? A systematic review. Acta Clinica Belgica**,** 1-9.

Vanneste, S. & De Ridder, D., 2016. Deafferentation-based pathophysiological differences in phantom sound: tinnitus with and without hearing loss. Neuroimage, 129**,** 80-94.

Varma, S. & Simon, R., 2006. Bias in error estimation when using cross-validation for model selection. BMC bioinformatics, 7**,** 91.

Vega-Pons, S. & Ruiz-Shulcloper, J., 2011. A survey of clustering ensemble algorithms. International Journal of Pattern Recognition and Artificial Intelligence, 25**,** 337-372.

Vernon, J. A. & Press, L. S., 1994. Characteristics of tinnitus induced by head injury. Archives of Otolaryngology–Head & Neck Surgery, 120**,** 547-551.

Vielsmeier, V., Kleinjung, T., Strutz, J., Bürgers, R., Kreuzer, P. M. & Langguth, B., 2011. Tinnitus with temporomandibular joint disorders: a specific entity of tinnitus patients? Otolaryngology--Head and Neck Surgery, 145**,** 748-752.

Vielsmeier, V., Lehner, A., Strutz, J., Steffens, T., Kreuzer, P. M., Schecklmann, M., Landgrebe, M., Langguth, B. & Kleinjung, T., 2015. The relevance of the high frequency audiometry in tinnitus patients with normal hearing in conventional pure-tone audiometry. BioMed Research International, 2015**,** 1-5.

Vielsmeier, V., Strutz, J., Kleinjung, T., Schecklmann, M., Kreuzer, P. M., Landgrebe, M. & Langguth, B., 2012. Temporomandibular joint disorder complaints in tinnitus: further hints for a putative tinnitus subtype. PLoS One, 7**,** e38887.

Wang, X., Jehi, L., Ji, X. & Mazzone, P. J., 2021. Phenotypes and Subphenotypes of Patients with COVID-19: a Latent Class Modeling Analysis. Chest.

Wang, Y., Zhang, J. N., Hu, W., Li, J. J., Zhou, J. X., Zhang, J. P., Shi, G. F., He, P., Li, Z. W. & Li, M., 2018. The characteristics of cognitive impairment in subjective chronic tinnitus. Brain and behavior, 8**,** e00918.

Ward, J., Vella, C., Hoare, D. J. & Hall, D. A., 2015. Subtyping Somatic Tinnitus: A Cross-Sectional UK Cohort Study of Demographic, Clinical and Audiological Characteristics. PLoS One, 10**,** e0126254.

Watabe, T., Kanzaki, S., Sato, N., Matsunaga, T., Muramatsu, M. & Ogawa, K., 2020. Single nucleotide polymorphisms in tinnitus patients exhibiting severe distress. Scientific Reports, 10**,** 1-10.

Wen, J., Varol, E., Chand, G., Sotiras, A. & Davatzikos, C. MAGIC: Multi-scale Heterogeneity Analysis and Clustering for Brain Diseases. International Conference on Medical Image Computing and Computer-Assisted Intervention, 2020. Springer, 678-687.

White, T., Van Der Ende, J. & Nichols, T. E., 2019. Beyond Bonferroni revisited: concerns over inflated false positive research findings in the fields of conservation genetics, biology, and medicine. Conservation Genetics, 20**,** 927-937.

Whoqol Group, 1998. Development of the World Health Organization WHOQOL-BREF quality of life assessment. Psychological medicine, 28**,** 551-558.

Wickham, H. 2016. ggplot2: elegant graphics for data analysis, springer.

Widiger, T. A. & Samuel, D. B., 2005. Diagnostic categories or dimensions? A question for the Diagnostic and statistical manual of mental disorders. Journal of abnormal psychology, 114**,** 494.

Wielopolski, J., Kleinjung, T., Koch, M., Peter, N., Meyer, M., Rufer, M. & Weidt, S., 2017. Alexithymia Is Associated with Tinnitus Severity. Frontiers in Psychiatry, 8**,** 223.

Wilson, P. H., Henry, J., Bowen, M. & Haralambous, G., 1991. Tinnitus reaction questionnaire: psychometric properties of a measure of distress associated with tinnitus. Journal of Speech, Language, and Hearing Research, 34**,** 197-201.

Woldeamanuel, Y. W., Sanjanwala, B. M., Peretz, A. M. & Cowan, R. P., 2020. exploring natural clusters of chronic Migraine phenotypes: A cross-Sectional clinical Study. Scientific reports, 10**,** 1-7.

Wolfers, T., Floris, D. L., Dinga, R., Van Rooij, D., Isakoglou, C., Kia, S. M., Zabihi, M., Llera, A., Chowdanayaka, R. & Kumar, V. J., 2019. From pattern classification to stratification: towards conceptualizing the heterogeneity of Autism Spectrum Disorder. Neuroscience & Biobehavioral Reviews, 104**,** 240-254.

Wright, K. B., 2005. Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. Journal of computer-mediated communication, 10**,** JCMC1034.

Xu, D. & Tian, Y., 2015. A comprehensive survey of clustering algorithms. Annals of Data Science, 2**,** 165-193.

Yan, X., Chu, J.-H., Gomez, J., Koenigs, M., Holm, C., He, X., Perez, M. F., Zhao, H., Mane, S. & Martinez, F. D., 2015. Noninvasive analysis of the sputum transcriptome discriminates clinical phenotypes of asthma. American journal of respiratory and critical care medicine, 191**,** 1116-1125.

Yu, J.-N., Nam, G. E., Han, K., Kim, J.-S., Kim, Y.-H., Cho, K. H., Kang, G. & Roh, Y. K., 2019. Association between menstrual cycle irregularity and tinnitus: a nationwide population-based study. Scientific reports, 9**,** 1-6.

Yu, S., Tranchevent, L., Liu, X., Glanzel, W., Suykens, J. A., De Moor, B. & Moreau, Y., 2011. Optimized data fusion for kernel k-means clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34**,** 1031-1039.

Yüksel, F., Karataş, D., Türkdoğan, F. T. & Yüksel, Ö., 2018. Increased Atherosclerosis Correlates with Subjective Tinnitus Severity. Indian Journal of Otolaryngology & Head & Neck Surgery, 70**,** 119-124.

Zenner, H., 1998. A systematic classification of tinnitus generator mechanisms. The international tinnitus journal, 4**,** 109.

Zhang, X., Chou, J., Liang, J., Xiao, C., Zhao, Y., Sarva, H., Henchcliffe, C. & Wang, F., 2019. Data-driven subtyping of Parkinson's disease using longitudinal clinical records: a cohort study. Scientific reports, 9**,** 1-12.

Zigmond, A. S. & Snaith, R. P., 1983. The hospital anxiety and depression scale. Acta psychiatrica scandinavica, 67**,** 361-370.

Zimmerman, B. J., Abraham, I., Schmidt, S. A., Baryshnikov, Y. & Husain, F. T., 2018. Dissociating tinnitus patients from healthy controls using resting-state cyclicity analysis and clustering. Network Neuroscience, 3**,** 67-89.