



Published in final edited form as:

Hear Res. 2017 November ; 355: 23–32. doi:10.1016/j.heares.2017.09.002.

## A methodological assessment of studies that use voxel-based morphometry to study neural changes in tinnitus patients

Nicholas Scott-Wittenborn, BA, Omar A. Karadaghy, BA, Jay F. Piccirillo, MD, and Jonathan E. Peelle, PhD

Department of Otolaryngology – Head and Neck Surgery Washington University School of Medicine, Saint Louis, MO

### Abstract

**Background**—The scientific understanding of tinnitus and its etiology have transitioned from thinking of tinnitus as solely a peripheral auditory problem to an increasing awareness that cortical networks may play a critical role in tinnitus percept or bother. With this change, studies that seek to use structural brain imaging techniques to better characterize tinnitus patients have become more common. These studies include using voxel-based morphometry (VBM) to determine if there are differences in regional gray matter volume in individuals who suffer from tinnitus and those who do not. However, studies using VBM in patients with tinnitus have produced inconsistent and sometimes contradictory results.

**Objective**—This paper is a systematic review of all of the studies to date that have used VBM to study regional gray matter volume in people with tinnitus, and explores ways in which methodological differences in these studies may account for their heterogeneous results. We also aim to provide guidance on how to conduct future studies using VBM to produce more reproducible results to further our understanding of disease processes such as tinnitus.

**Methods**—Studies about tinnitus and VBM were searched for using PubMed and Embase. These returned 15 and 25 results respectively. Of these, nine met the study criteria and were included for review. An additional 5 studies were identified in the literature as pertinent to the topic at hand and were added to the review, for a total of 13 studies.

---

Corresponding Author: Jonathan E. Peelle, PhD, Department of Otolaryngology-Head and Neck Surgery, Washington University School of Medicine, 660 South Euclid Avenue, Campus Box 8115, Saint Louis, MO 63110, [jpeelle@wustl.edu](mailto:jpeelle@wustl.edu).

Author Contributions:

Nicholas Scott-Wittenborn: acquisition, analysis, and interpretation of data, drafting of the manuscript

Omar Karadaghy: analysis and interpretation of data, critical revision of the manuscript for important intellectual content

Jay F. Piccirillo: conception and design, interpretation of data, critical revision of the manuscript for important intellectual content

Jonathan E. Peelle: analysis and interpretation of data, critical revision of the manuscript for important intellectual content

Disclosures:

None of the authors have any conflicts of interest. Research reported in this publication was supported by the Washington University Institute of Clinical and Translational sciences grant UL1R000448 sub award TL1TR000449, from the National Center for Advancing Translational Sciences (NCATS). This manuscript has not been published or presented elsewhere.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Results**—There was significant heterogeneity among the studies in several areas, including inclusion and exclusion criteria, software programs, and statistical analysis. We were not able to find publicly shared data or code for any study.

**Discussion**—The differences in study design, software analysis, and statistical methodology make direct comparisons between the different studies difficult. Especially problematic are the differences in the inclusion and exclusion criteria of the study, and the statistical design of the studies, both of which could radically alter findings. Thus, heterogeneity has complicated efforts to explore the etiology of tinnitus using structural MRI.

**Conclusion**—There is a pressing need to standardize the use of VBM when evaluating tinnitus patients. While some heterogeneity is expected given the rapid advances in the field, more can be done to ensure that there is internal validity between studies.

## Keywords

Magnetic Resonance Imaging; voxel-based morphometry; tinnitus; gray matter

---

## 1.0 Introduction

Historically, tinnitus was considered a disease of the peripheral auditory system, but this view has been replaced in the last twenty years (Jastreboff 1990; Bauer 2004). There is a growing awareness that while the inception of tinnitus may be due to pathology or injury in the auditory periphery, it is fundamentally a central disease based on maladaptive neuroplasticity and misappropriated attention (Shore, Roberts, and Langguth 2016; Rauschecker, Leaver, and Mühlau 2010; Kraus and Canlon 2012; Roberts, Husain, and Eggermont 2013). To validate this theory of the pathophysiological origin of tinnitus, scientists have begun to use magnetic resonance imaging (MRI) to determine if there are any structural abnormalities in the brain of tinnitus patients. Voxel-based morphometry (VBM) (Mechelli et al. 2005) is one method used to quantify regional brain volume from structural MRI scans.

Although a number of studies have used VBM to better understand tinnitus, the results are inconsistent (Adjamian et al. 2014). One problem associated with comparing VBM studies of tinnitus is that the methods used to generate the images are complex, and a full understanding of the process requires knowledge in multiple fields of expertise, including neuroscience, neuroimaging, physics, otolaryngology, and statistics. The aim of this investigation is to perform a systematic review of the methodological rigor of studies that use VBM to evaluate patients with tinnitus. Our goal is to help interpret prior research while also providing guidance for future studies that seek to identify structural brain changes associated with tinnitus.

One larger issue that we do not address is whether there are likely to be structural brain changes associated with tinnitus that are detectable using MRI-based approaches. It may be that the underlying changes in neural processing are subtle enough that they are not detectable using macroanatomical techniques. Alternatively, the heterogeneity across different patient phenotypes may obscure effects at the group level. These are empirical

questions in the sense that if there is no true difference in brain structure between patients with tinnitus and controls, well-designed studies should point to this conclusion. However, the assumptions underlying the use of VBM to study patients with tinnitus (or any special population) are important to consider when planning or interpreting a research study.

It is also important to note that VBM is not the only approach for studying regional gray matter differences: surface-based techniques estimating cortical thickness can also be used, although the correspondence between volume- and surface-based approaches is not always straightforward (Hutton et al. 2009). We have chosen to focus on VBM in part because there are more existing studies that use VBM to study patients with tinnitus compared to other approaches.

### 1.1 A brief overview of voxel-based morphometry

VBM is an imaging technique that estimates brain tissue density or volume from structural MRI images. Typically, images from each individual are segmented into tissue classes (e.g., gray matter, white matter, cerebrospinal fluid) and spatially normalized to a common stereotactic space smoothed with a three-dimensional Gaussian kernel. The smoothed, normalized images can then be statistically assessed (for example, comparing groups of participants, or performing correlation analyses across participants).

One significant challenge that arises when analyzing VBM data is that the statistical tests are performed across many thousands of voxels: if uncontrolled for multiple comparisons, false positives are common. Several approaches have been developed for controlling false positives in neuroimaging at either the voxel or cluster level (a larger group of voxels), including using Gaussian random field theory (Friston et al. 1994), false discovery rate (Chumbley and Friston 2009; Genovese, Lazar, and Nichols 2002), and permutation testing (Nichols and Holmes 2002; Smith and Nichols 2009). If random field theory is used, for inference to be valid local adjustments to smoothness are required to control for image nonstationarity (Hayasaka et al. 2004).

There are a multitude of choices on how to execute each of these steps. Since the introduction of VBM, advances have been made in various areas including: tissue class segmentation (Ashburner and Friston 2005), spatial normalization by using high-dimensional diffeomorphic registration approaches (Avants and Gee 2004; Ashburner 2007), and understanding of statistical analyses and covariates (Barnes et al. 2010; Malone et al. 2015; Peelle, Cusack, and Henson 2012).

## 2.0 Methodology

### 2.1 Study selection criteria

The selection criteria for articles in this review were that they be published in peer-reviewed journals in the English language and that they used VBM to assess gray matter in adults with tinnitus.

## 2.2 Literature search

The process for paper selection is shown in Figure 1. A comprehensive literature review of imaging studies concerning tinnitus was performed using PubMed and Embase with the assistance of a research librarian. The following terms constituted the Embase search: “‘tinnitus’/exp OR tinnitus AND (‘voxel based morphometry’/exp OR ‘voxel based morphometry’)” and returned 25 results. The PubMed search used the following terms: “Tinnitus and ‘voxel based morphometry’” and returned 15 results. Two of the authors (NSW and OAK) analyzed the results, removing duplicates and studies that were not primary investigations of gray matter volume assessed by VBM. This process left nine studies that addressed VBM and tinnitus. Upon further review, we identified one additional article (Leaver et al., 2011) that did not appear in the search, but was cited in the VBM/tinnitus literature; reviewers also pointed us to three additional papers (Allan et al. 2016, Krick et al. 2015, Mahoney et al. 2011). In total, 13 articles met the selection criteria of the study.

## 2.3 Data extraction and study assessment

To evaluate methodological rigor we used the principles put forth by Ridgway et al. (2008). These principles, which combine common sense research design with items focused on VBM study reproducibility, guided our own data selection. Specifically, Ridgway et al. encourage researchers to detail the procedures used for segmentation and spatial normalization.

All authors agreed on the data points to be collected prior to study initiation. Two of the authors (NSW and OAK) did independent data extractions and evaluations. We discussed discrepancies between the two extractions and came to a consensus on all discrepancies. There were no issues that necessitated including a third reviewer.

## 3.0 Results

### 3.1 Demographics and study design

The list of studies can be found in Table 1. The demographic information of these studies is in Table 2. The sample size of the studies ranged from seven to 257 with a mean of 56.1 participants and median of 24. None of the studies discussed sample size justification or power calculations. Nine of the studies explicitly stated their subjects’ tinnitus severity, while four mentioned which severity score they used, but did not include the actual score (Allen et al. 2016, Leaver et al. 2011, Leaver et al. 2012, Melcher et al. 2013).

The papers used three different, basic study designs: those without any control groups, those with only cases and controls, and those who included controls with hearing loss as a separate group for analysis. Three studies did not use any control groups and analyzed brain volume changes of tinnitus using only regression models (Schecklmann et al. 2012, Schecklmann et al. 2013, Vanneste et al. 2015). Three of the studies that used controls had controls with hearing loss and with normal hearing (Allan et al. 2016, Boyen et al. 2013, Husain et al. 2011). The other seven studies used a single control group in their VBM study (Krick et al.

2015, Landgrebe et al. 2009, Leaver et al. 2011, Leaver et al. 2012, Mahoney et al. 2011, Melcher et al. 2013, Mühlau et al. 2005).

There was also significant heterogeneity in terms of the inclusion and exclusion criteria used by the studies. Five of the studies excluded those individuals who had psychiatric disease (Husain et al. 2011, Krick et al. 2015, Landgrebe et al. 2009, Shecklmann et al. 2012, Vanneste et al. 2015). Four of the studies did not allow for any type of hearing loss (Landgrebe et al. 2009, Melcher et al. 2013, Mühlau et al. 2006, Shecklmann et al. 2013). An important consistency of the studies is that all but three (Allan et al. 2016, Boyen et al. 2013, Mahoney et al. 2011) of the studies excluded patients with anatomical causes of tinnitus, such as acoustic neuroma.

The selection criteria for controls were also variable, with four studies matching only on age and sex (Krick et al. 2015, Landgrebe et al. 2009, Mahoney et al. 2011, Mühlau et al. 2006); three studies matching controls on age, sex, and hearing loss (Allan et al. 2016, Husain et al. 2011, Melcher et al. 2013); and one study that did not describe how they chose their controls (Boyen et al. 2013). Two studies did not mention matching their controls to patients on any variables (Leaver et al. 2011, Leaver et al. 2012).

### 3.2 Technical data

Table 3 summarizes the technical details of data acquisition. All of the studies described the type and field strength of the scanner that they used. Of the 13 studies, 10 also reported the size of their voxels, three did not (Mahoney et al. 2011, Melcher et al. 2013, Mühlau et al. 2006). Similarly, 10 studies reported the size of their acquisition matrix, three did not (Allan et al. 2015, Husain et al. 2006, Melcher et al. 2013).

### 3.3 Preprocessing of image data

Of the 13 articles, 8 used Statistical Parametric Mapping (SPM; Wellcome Department of Imaging Neuroscience, University College, London) version 8, three used SPM5, and two used SPM2. One performed additional steps such as manually setting the origin prior to automated preprocessing (Melcher et al. 2013). Four of the studies described their reference space (Boyen et al. 2013, Husain et al. 2011, Landgrebe et al. 2009, Leaver et al. 2011). Eight of the studies described their tissue segmentation procedure (Allan et al. 2015, Boyen et al. 2013, Husain et al. 2011, Krick et al. 2016, Landgrebe et al. 2009, Melcher et al. 2013, Mühlau et al. 2006, Vanneste et al. 2015). Seven of the thirteen studies used an 8 mm isotropic Gaussian kernel Full Width Half Maximum smoothing, one used 12 mm (Leaver et al. 2011), three used 10 mm (Allan et al. 2015, Krick et al. 2016, Landgrebe et al. 2009), one used 6 mm (Leaver et al. 2012), and one did not state their smoothing parameter (Mahoney et al. 2011). A full review of imaging processing is in Table 4.

### 3.4 Statistical methodology

All of the papers analyzed the whole brain. The actual statistical analysis of the whole brain is in Table 5. Of note, two of the studies that used controls first used an ANOVA or ANCOVA analysis before performing t tests (Boyen et al. 2013, Husain et al. 2011). In

addition, while most of the control-based studies used a correction for multiple comparisons, two studies did not (Leaver et al. 2011, Leaver et al. 2012).

All of the papers looked at regions of interest (ROIs). Although multiple approaches were used to define ROIs, the most common was to combine Brodmann areas from Wake Forest University Pickatlas (ANSIR Laboratory Wake Forest University School of Medicine, Winston-Salem) and subcortical auditory nuclei defined by the Montreal Neurological Institute (MNI) coordinates. The specific of region of interest analyses are displayed in Table 6.

Three of the studies only used regression analysis of their tinnitus population. The full list of regressors of the different studies is in Table 7. The only commonalities between all regression models were the use of age and sex as covariates.

### 3.5 Data and code sharing

Among the studies examined, we could not find any of the datasets or analysis code publicly available. To assess whether data might be available privately, we sent the corresponding author of each of the 13 studies a link to a short, anonymous survey, and requested a response within two weeks. After two months, we received 7 responses, of which 3 indicated the data would, in principle, be available for sharing. The most common reason authors gave for not being able to share the data was the lack of IRB approval (3). Of the three authors who said data would be available, one said that the original authors would require co-authorship on any new paper as a condition of data sharing (and two others indicated collaboration and/or co-authorship would be strongly preferred). The full text of the survey and itemized responses are available from <https://osf.io/r84fb/>.

### 3.6 Results of analysis

The results of the analyzed studies are in Table 8. From this summary, it is evident that there is little agreement in the results of each of the different studies.

## 4.0 Discussion

We found considerable heterogeneity in study design and statistical methodology, making direct comparisons between studies impossible. The studies had significant differences in inclusion and exclusion criteria (especially pertaining to psychiatric disease), handling of hearing loss in matching patients and controls, and basic study designs (case control vs. cohort). The statistical methodology also varied with significant differences in the statistical treatment of data even between papers with similar study designs. We focus our discussion on what we see as the areas most likely to be useful in reaching consensus in future studies.

### 4.1 Demographics and study design

Establishing the selection (inclusion and exclusion) criteria and the use of matching criteria are among the most important study choices that an investigator makes and have a large effect on the results. These study choices determine patient characteristics, including degree

of hearing loss, presence of psychiatric disease, age, and sex, all of which are associated with differences in regional brain volume.

Hearing loss has been associated with reduced gray matter volume in auditory cortex (Peelle et al. 2011; Eckert et al. 2012). The matching of hearing loss between tinnitus patients and controls can be challenging because matching includes accounting for the degree, laterality, and frequency profile of hearing loss. However, if these effects are not properly controlled for, either through inclusion/exclusion criteria or matching, such differences could confound the results of a control group-based study.

Psychiatric disease can also affect gray matter volume. Several studies have looked at the effect that depression has on gray matter volume (van Tol et al. 2010; Soriano-Mas et al. 2011; Halford and Anderson 1991). While some studies did not show whole brain differences, a meta-analysis of 26 studies concluded that depression was associated with reductions in gray matter in several areas of the brain (Bora et al. 2012). Controlling for psychiatric disease through exclusion criteria is problematic because of the well-known connection between depression and tinnitus (Bhatt, Bhattacharyya, and Lin 2016). Through the use of validated measures of various cogent clinical phenomena related to the tinnitus experience, such as depression, anxiety, and cognitive failure, investigators can standardize the criteria for various clinical phenomena. Readers and future investigators can then interpret the impact of these clinical features on the study population. With the use of validated measures, readers and investigators can understand what a particular value means and the clinical implication of a change in the value. We do not mean to prescribe a certain measure for each cogent clinical condition, but rather we recommend that investigators select a validated measure with known psychometric and clinimetric properties that represents the best the goals of the study.

We do not mean to imply that certain subgroups of tinnitus patients should be excluded from future studies. Rather, we suggest that given the large heterogeneity in tinnitus patients and the likely impact this heterogeneity in etiology has on diagnostic test interpretation and treatment response, it is important that investigators clearly define which subgroups of tinnitus patients are included in the study, and include sufficient subjects within these unique subgroups to allow for accurate and precise analyses. Furthermore, investigators should present data in such a way that readers and future investigators can clearly see the results within a subgroup.

Age, which is associated with gray matter changes in multiple regions (Peelle, Cusack, and Henson 2012; Hutton et al. 2009; Good et al. 2001), is also a possible confounder. This is the only trait that almost all the studies used as matching criteria. Multiple studies have demonstrated that there are significant differences between gray matter volume of males and females and as such, this should be factored into matching (Vanneste, Joos, and De Ridder 2012; Seydel et al. 2013).

Another weakness of the reviewed studies is that none reported a power calculation or other analysis to determine an adequate sample size. Sample size calculations for neuroimaging studies are challenging and their omission is common. However, in recent years new tools

have been developed to facilitate this important step in study design (Mumford and Nichols 2008; Joyce and Hayasaka 2012; Durnez et al. 2016), which need not always require pilot data. In fact, access to prior study data or statistical maps to facilitate power analysis is a benefit of increased data sharing. Increased use of power analyses in the future will help to guard against underpowered studies, which may not be able to detect a true difference in the volume of gray matter and might account for some of the inconsistent results from these VBM studies.

#### 4.2 Technical data

All of the studies presented their technical data clearly and in sufficient detail to allow for easy interpretation and comparison.

#### 4.3 Processing of image data

The reporting of processing steps was sufficient in the majority of studies to allow for reproducibility. Each study noted the software they used, including version, and discussed any manual preprocessing that they carried out. Every study but one described the parameters of their smoothing.

Studies that used SPM did not always adequately describe their reference spaces, often saying they used default parameters. Although it is reasonable to infer that SPM uses ICBM152 (i.e., MNI152) space based on software defaults, the group reference space should always be explicitly noted. This is particularly important to facilitate either the interpretation of stereotactic coordinates or the use of results in future meta-analyses.

Several studies also failed to discuss tissue class segmentation adequately (Leaver et al. 2011, Schecklmann et al. 2012, Schecklmann et al. 2013). While this may not seem necessary because some authors described using the “default” parameters, the rapid pace of improvement in the both the software and the imaging techniques—including the possibility for default parameters to change—means that a more explicit description of the steps used to generate the data should be included to allow for retrospective review of the different studies’ findings.

#### 4.4 Statistical methodology

While the presentation of statistical methodology varied among the different studies, they all discussed in appropriate detail how they analyzed the data. All of the studies stated what statistical program and major version they used to evaluate the data. However, it is important to differentiate the major types of statistical analysis used in the studies. There are two major types of studies identified in this analysis, case control and cohort studies. Additionally, there were two types of case control studies: those that used multiple group comparisons vs. those that did not.

The statistical models employed were correctly used in nine of the eleven studies: two-sample t-tests to compare tinnitus patients with controls (when a control group was used), and multiple regression using some aspect of tinnitus severity as a predictor. One potential difficulty arises in the two studies (Boyen et al. 2013, Husain et al. 2011) that used an



ANOVA to identify differences between tinnitus and control groups, followed by a two-sample t-test in the areas showing a significant difference. Because these tests are not independent, care must be taken that the results of the posthoc t-tests are interpreted correctly (i.e., the corrected p values may not be accurate) (Kriegeskorte et al. 2009).

The level of control over false positives varied among studies for both the whole brain and ROI analysis. Two of the studies did not correct for multiple comparisons at all, which obscures the expected false positive rate and thus makes the results difficult to interpret (Leaver et al. 2011, Leaver et al. 2012).

The second statistical analysis technique that was frequently used in these studies was the use of a regression model to determine correlation between tinnitus characteristics, study demographics, and VBM findings. A problem with comparing these studies is that all three used the tinnitus questionnaire (TQ) (Goebel and Hiller 1994), while many of the other studies used the tinnitus handicap inventory (THI) (Newman, Jacobson, and Spitzer 1996), two different evaluations of tinnitus severity. There is little work on cross validity of these two different assessments of tinnitus severity (Møller et al. 2010). While they have similar psychometric properties, there is no work on how differences between the two assessments may correlate to imaging findings (Zeman et al. 2012). Thus, trying to compare results based on the scores of these two studies is difficult and may not be valid. There is also a newer measurement of tinnitus severity, the tinnitus functional index (TFI), that may be more sensitive in detecting changes to tinnitus severity (Meikle et al. 2012).

#### 4.5 Data and code sharing

There has been increasing realization in the neuroimaging community that sharing of analysis code and raw data can improve the reproducibility of research (Gorgolewski and Poldrack 2016; Pernet and Poline 2015). Reproducibility is also improved with the continuing development of tools for sharing raw data (Poldrack et al. 2013) and unthresholded statistical maps (Gorgolewski et al. 2015). One way to minimize the effect of different analysis approaches would be to reanalyze datasets using a standardized approach; unfortunately, this is impossible if the original data are unavailable or if authors do not request data from peers. In addition, data sharing would facilitate combining participants across different cohorts, which would help to increase sample size.<sup>1</sup> Future studies would do a service to the field by making their data publicly available. It is notable (though perhaps not surprising) that none of the investigators of these 13 papers used the data from a previous paper in their analysis. If data were easier to obtain, it might facilitate re-analyses, power analyses, and aggregating data across multiple studies.

#### 5.0 Conclusions

Our review of studies using VBM to investigate regional gray matter volume in patients with tinnitus revealed a large amount of heterogeneity across studies, making direct comparisons

---

<sup>1</sup>Aggregating data also raises practical challenges due to differing hardware, acquisition protocols, and patient populations at different research centers.

difficult. Future studies could improve on the situation by adopting some straightforward changes:

1. Performing power analyses prior to study initiation to determine the proper sample size.
2. Making raw data and analysis code available to promote confirmatory analyses and data aggregation.
3. Reporting effect size in addition to statistical significance (e.g., gray matter volume at the whole brain or ROI level).
4. Using the TFI as the standardized, validated, measurement of tinnitus severity to simplify comparison between studies, as well as other measures of tinnitus (THI and TQ) (time permitting) in order to ensure cross study validity.<sup>2</sup>
5. Inclusion and exclusion criteria should be standardized for hearing loss and depression to create a homogenous and representative population in all studies. Ideally, all studies would include a factorial design including participants with and without tinnitus, split by those that are or are not depressed. However, a problem with this approach is the high comorbidity of depression with tinnitus, so (a) finding patients with tinnitus who are not depressed and (b) even if they can be found, their rarity brings into question how generalizable the findings are. As a compromise, we recommend that tinnitus studies include participants with depression in order to have generalizable results. In addition, patients with tinnitus and controls should be matched on sex and hearing loss.

We believe that implementation of these recommendations will significantly improve the validity of future studies using VBM to assess tinnitus and help to resolve the inconsistency of the findings in the literature.

## 6.0 References

- Adjamian P, Hall DA, Palmer AR, Allan TW, Langers DR. Neuroanatomical abnormalities in chronic tinnitus in the human brain. *Neurosci Biobehav Rev.* 2014; 45:119–33. [PubMed: 24892904]
- Allan, Thomas W., Besle, Julien, Langers, Dave RM., Davies, Jeff, Hall, Deborah A., Palmer, Alan R., Adjamian, Peyman. Neuroanatomical Alterations in Tinnitus Assessed with Magnetic Resonance Imaging. *Frontiers in Aging Neuroscience.* 2016; 8:221. [PubMed: 27708577]
- Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage.* 2007; 38:95–113. [PubMed: 17761438]
- Ashburner J, Friston KJ. Unified segmentation. *Neuroimage.* 2005; 26:839–51. [PubMed: 15955494]
- Avants B, Gee JC. Geodesic estimation for large deformation anatomical shape averaging and interpolation. *Neuroimage.* 2004; 23(Suppl 1):S139–50. [PubMed: 15501083]
- Barnes J, Ridgway GR, Bartlett J, Henley SM, Lehmann M, Hobbs N, Clarkson MJ, MacManus DG, Ourselin S, Fox NC. Head size, age and gender adjustment in MRI studies: a necessary nuisance? *Neuroimage.* 2010; 53:1244–55. [PubMed: 20600995]
- Bauer CA. Mechanisms of tinnitus generation. *Curr Opin Otolaryngol Head Neck Surg.* 2004; 12:413–7. [PubMed: 15377954]

<sup>2</sup>Our recommendation of standardizing on the TFI reflects the advantage of having a standard measure across multiple studies. Of course, in the future, another measure may prove superior. Currently the variety of tinnitus measurements presents barriers to combining data or results across studies.

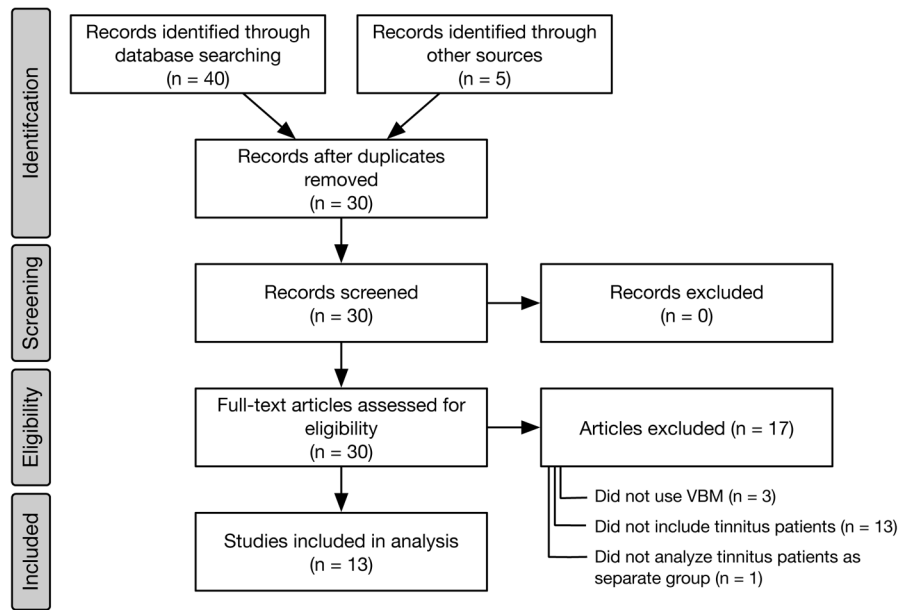
- Bhatt JM, Bhattacharyya N, Lin HW. Relationships between tinnitus and the prevalence of anxiety and depression. *Laryngoscope*. 2016
- Bora E, Fornito A, Pantelis C, Yucel M. Gray matter abnormalities in Major Depressive Disorder: a meta-analysis of voxel based morphometry studies. *J Affect Disord*. 2012; 138:9–18. [PubMed: 21511342]
- Boyen K, Langers DR, de Kleine E, van Dijk P. Gray matter in the brain: differences associated with tinnitus and hearing loss. *Hear Res*. 2013; 295:67–78. [PubMed: 22446179]
- Chumbley JR, Friston KJ. False discovery rate revisited: FDR and topological inference using Gaussian random fields. *Neuroimage*. 2009; 44:62–70. [PubMed: 18603449]
- Durnez, Joke, Degryse, Jasper, Moerkerke, Beatrijs, Seurinck, Ruth, Sochat, Vanessa, Poldrack, Russell, Nichols, Thomas. Power and sample size calculations for fMRI studies based on the prevalence of active peaks. *bioRxiv*. 2016
- Eckert, Mark A., Cute, Stephanie L., Vaden, Kenneth I., Jr, Kuchinsky, Stefanie E., Dubno, Judy R. Auditory cortex signs of age-related hearing loss. *Journal of the Association for Research in Otolaryngology*. 2012; 13:703–13. [PubMed: 22618352]
- Friston KJ, Worsley KJ, Frackowiak RS, Mazziotta JC, Evans AC. Assessing the significance of focal activations using their spatial extent. *Hum Brain Mapp*. 1994; 1:210–20. [PubMed: 24578041]
- Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*. 2002; 15:870–8. [PubMed: 11906227]
- Goebel G, Hiller W. The tinnitus questionnaire. A standard instrument for grading the degree of tinnitus. Results of a multicenter study with the tinnitus questionnaire. *Hno*. 1994; 42:166–72. [PubMed: 8175381]
- Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RS. A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage*. 2001; 14:21–36. [PubMed: 11525331]
- Gorgolewski KJ, Poldrack RA. A Practical Guide for Improving Transparency and Reproducibility in Neuroimaging Research. *PLoS Biol*. 2016; 14:e1002506. [PubMed: 27389358]
- Gorgolewski KJ, Varoquaux G, Rivera G, Schwarz Y, Ghosh SS, Maumet C, Sochat VV, Nichols TE, Poldrack RA, Poline JB, Yarkoni T, Margulies DS. NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Front Neuroinform*. 2015; 9:8. [PubMed: 25914639]
- Halford JB, Anderson SD. Anxiety and depression in tinnitus sufferers. *J Psychosom Res*. 1991; 35:383–90. [PubMed: 1920169]
- Hayasaka S, Phan KL, Liberzon I, Worsley KJ, Nichols TE. Nonstationary cluster-size inference with random field and permutation methods. *Neuroimage*. 2004; 22:676–87. [PubMed: 15193596]
- Husain FT, Medina RE, Davis CW, Szymko-Bennett Y, Simonyan K, Pajor NM, Horwitz B. Neuroanatomical changes due to hearing loss and chronic tinnitus: a combined VBM and DTI study. *Brain Res*. 2011; 1369:74–88. [PubMed: 21047501]
- Hutton C, Draganski B, Ashburner J, Weiskopf N. A comparison between voxel-based cortical thickness and voxel-based morphometry in normal aging. *Neuroimage*. 2009; 48:371–80. [PubMed: 19559801]
- Jastreboff PJ. Phantom auditory perception (tinnitus): mechanisms of generation and perception. *Neurosci Res*. 1990; 8:221–54. [PubMed: 2175858]
- Joyce KE, Hayasaka S. Development of PowerMap: a software package for statistical power calculation in neuroimaging studies. *Neuroinformatics*. 2012; 10:351–65. [PubMed: 22644868]
- Kraus KS, Canlon B. Neuronal connectivity and interactions between the auditory and limbic systems. Effects of noise and tinnitus. *Hear Res*. 2012; 288:34–46. [PubMed: 22440225]
- Krick, Christoph M., Grapp, Miriam, Daneshvar-Talebi, Jonas, Reith, Wolfgang, Plinkert, Peter K., Bolay, Hans Volker. Cortical reorganization in recent-onset tinnitus patients by the Heidelberg Model of Music Therapy. *Frontiers in Neuroscience*. 2015; 9:49. [PubMed: 25745385]
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci*. 2009; 12:535–40. [PubMed: 19396166]

- Landgrebe M, Langguth B, Rosengarth K, Braun S, Koch A, Kleinjung T, May A, de Ridder D, Hajak G. Structural brain changes in tinnitus: grey matter decrease in auditory and non-auditory brain areas. *Neuroimage*. 2009; 46:213–8. [PubMed: 19413945]
- Leaver AM, Renier L, Chevillet MA, Morgan S, Kim HJ, Rauschecker JP. Dysregulation of limbic and auditory networks in tinnitus. *Neuron*. 2011; 69:33–43. [PubMed: 21220097]
- Leaver, Amber M., Seydell-Greenwald, Anna, Turesky, Ted K., Morgan, Susan, Kim, Hung J., Rauschecker, Josef P. Cortico-limbic morphology separates tinnitus from tinnitus distress. *Frontiers in Systems Neuroscience*. 2012; 6:21. [PubMed: 22493571]
- Mahoney, Colin J., Rohrer, Jonathan D., Goll, Johanna C., Fox, Nick C., Rossor, Martin N., Warren, Jason D. Structural neuroanatomy of tinnitus and hyperacusis in semantic dementia. *Journal of Neurology, Neurosurgery & Psychiatry*. 2011
- Malone IB, Leung KK, Clegg S, Barnes J, Whitwell JL, Ashburner J, Fox NC, Ridgway GR. Accurate automatic estimation of total intracranial volume: a nuisance variable with less nuisance. *Neuroimage*. 2015; 104:366–72. [PubMed: 25255942]
- Mechelli A, Price CJ, Friston KJ, Ashburner J. Voxel-Based Morphometry of the Human Brain: Methods and Applications. *Current Medical Imaging Reviews*. 2005; 1:105–13.
- Meikle MB, Henry JA, Griest SE, Stewart BJ, Abrams HB, McArdle R, Myers PJ, Newman CW, Sandridge S, Turk DC, Folmer RL, Frederick EJ, House JW, Jacobson GP, Kinney SE, Martin WH, Nagler SM, Reich GE, Searchfield G, Sweetow R, Vernon JA. The tinnitus functional index: development of a new clinical measure for chronic, intrusive tinnitus. *Ear Hear*. 2012; 33:153–76. [PubMed: 22156949]
- Melcher JR, Knudson IM, Levine RA. Subcallosal brain structure: correlation with hearing threshold at supra-clinical frequencies (>8 kHz), but not with tinnitus. *Hear Res*. 2013; 295:79–86. [PubMed: 22504034]
- Møller, Aage R., Langguth, Berthold, DeRidder, Dirk, Kleinjung, Tobias. *Textbook of tinnitus*. Springer Science & Business Media; 2010.
- Mühlau M, Rauschecker JP, Oestreicher E, Gaser C, Rottinger M, Wohlschlagel AM, Simon F, Etgen T, Conrad B, Sander D. Structural brain changes in tinnitus. *Cereb Cortex*. 2006; 16:1283–8. [PubMed: 16280464]
- Mumford JA, Nichols TE. Power calculation for group fMRI studies accounting for arbitrary design and temporal autocorrelation. *Neuroimage*. 2008; 39:261–8. [PubMed: 17919925]
- Newman CW, Jacobson GP, Spitzer JB. Development of the Tinnitus Handicap Inventory. *Arch Otolaryngol Head Neck Surg*. 1996; 122:143–8. [PubMed: 8630207]
- Nichols TE, Holmes AP. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp*. 2002; 15:1–25. [PubMed: 11747097]
- Peelle JE, Troiani V, Grossman M, Wingfield A. Hearing loss in older adults affects neural systems supporting speech comprehension. *J Neurosci*. 2011; 31:12638–43. [PubMed: 21880924]
- Peelle, Jonathan E., Cusack, Rhodri, Henson, Richard NA. Adjusting for global effects in voxel-based morphometry: Gray matter decline in normal aging. *Neuroimage*. 2012; 60:1503–16. [PubMed: 22261375]
- Pernet C, Poline JB. Improving functional magnetic resonance imaging reproducibility. *Gigascience*. 2015; 4:15. [PubMed: 25830019]
- Poldrack RA, Barch DM, Mitchell JP, Wager TD, Wagner AD, Devlin JT, Cumba C, Koyejo O, Milham MP. Toward open sharing of task-based fMRI data: the OpenfMRI project. *Front Neuroinform*. 2013; 7:12. [PubMed: 23847528]
- Rauschecker JP, Leaver AM, Mühlau M. Tuning out the noise: limbic-auditory interactions in tinnitus. *Neuron*. 2010; 66:819–26. [PubMed: 20620868]
- Ridgway GR, Henley SM, Rohrer JD, Schill RI, Warren JD, Fox NC. Ten simple rules for reporting voxel-based morphometry studies. *Neuroimage*. 2008; 40:1429–35. [PubMed: 18314353]
- Roberts LE, Husain FT, Eggermont JJ. Role of attention in the generation and modulation of tinnitus. *Neurosci Biobehav Rev*. 2013; 37:1754–73. [PubMed: 23876286]
- Schecklmann M, Lehner A, Poepl TB, Kreuzer PM, Hajak G, Landgrebe M, Langguth B. Cluster analysis for identifying sub-types of tinnitus: a positron emission tomography and voxel-based morphometry study. *Brain Res*. 2012; 1485:3–9. [PubMed: 22613349]

- Schecklmann M, Lehner A, Poepl TB, Kreuzer PM, Rupprecht R, Rackl J, Burger J, Frank E, Hajak G, Langguth B, Landgrebe M. Auditory cortex is implicated in tinnitus distress: a voxel-based morphometry study. *Brain Struct Funct*. 2013; 218:1061–70. [PubMed: 23435735]
- Seydel C, Haupt H, Olze H, Szczepek AJ, Mazurek B. Gender and chronic tinnitus: differences in tinnitus-related distress depend on age and duration of tinnitus. *Ear Hear*. 2013; 34:661–72. [PubMed: 23439056]
- Shore, Susan E., Roberts, Larry E., Langguth, Berthold. Maladaptive plasticity in tinnitus-triggers, mechanisms and treatment. *Nature reviews Neurology*. 2016; 12:150–60. [PubMed: 26868680]
- Smith SM, Nichols TE. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*. 2009; 44:83–98. [PubMed: 18501637]
- Soriano-Mas C, Hernandez-Ribas R, Pujol J, Urretavizcaya M, Deus J, Harrison BJ, Ortiz H, Lopez-Sola M, Menchon JM, Cardoner N. Cross-sectional and longitudinal assessment of structural brain alterations in melancholic depression. *Biol Psychiatry*. 2011; 69:318–25. [PubMed: 20875637]
- van Tol MJ, van der Wee NJ, van den Heuvel OA, Nielen MM, Demenescu LR, Aleman A, Renken R, van Buchem MA, Zitman FG, Veltman DJ. Regional brain volume in depression and anxiety disorders. *Arch Gen Psychiatry*. 2010; 67:1002–11. [PubMed: 20921116]
- Vanneste S, Joos K, De Ridder D. Prefrontal cortex based sex differences in tinnitus perception: same tinnitus intensity, same tinnitus distress, different mood. *PLoS ONE*. 2012; 7:e31182. [PubMed: 22348053]
- Vanneste, Sven, Van De Heyning, Paul, De Ridder, Dirk. Tinnitus: A Large VBM-EEG Correlational Study. *PLoS ONE*. 2015; 10:e0115122. [PubMed: 25781934]
- Zeman F, Koller M, Schecklmann M, Langguth B, Landgrebe M. Tinnitus assessment by means of standardized self-report questionnaires: psychometric properties of the Tinnitus Questionnaire (TQ), the Tinnitus Handicap Inventory (THI), and their short versions in an international and multi-lingual sample. *Health Qual Life Outcomes*. 2012; 10:128. [PubMed: 23078754]

### Highlights

- We reviewed studies using voxel-based morphometry to examine regional gray matter volume in patients with tinnitus.
- Our literature search identified 13 studies that met inclusion criteria.
- The identified studies varied considerably in their methodology, including inclusion/exclusion criteria, software programs, and statistical analyses, making direct comparisons or meta-analyses impractical.
- We provide suggestions on ways future studies can improve reliability.



**Figure 1.** PRISMA diagram showing the selection process for studies included in the review. Of the 40 articles identified through database searching, 25 came from Embase and 15 from PubMed.

Table 1

## Study Overview

Author	Title	Journal	Year	Study type
Allan et al.	Neuroanatomical alterations in tinnitus assessed with magnetic resonance imaging	Frontiers in Aging Neuroscience	2016	Case control
Boyen et al.	Gray matter in the brain: Differences associated with tinnitus and hearing loss	Hearing Research	2013	Case control
Husain et al.	Neuroanatomical Changes due to Hearing Loss and Chronic Tinnitus: A Combined VBM and DTI Study	Brain Research	2011	Case control
Krick et al.	Cortical reorganization in recent-onset tinnitus patients by Heidelberg model of music theory	Frontiers in Neuroscience	2015	Case control
Landgrebe et al.	Structural brain changes in tinnitus: gray matter decrease in auditory and non-auditory brain areas	NeuroImage	2009	Case control
Leaver et al.	Dysregulation of limbic and auditory networks in tinnitus	Neuron	2011	Case control
Leaver et al.	Cortico-limbic morphology separates tinnitus from tinnitus distress	Frontiers in Systems Neuroscience	2012	Case control
Mahoney et al.	Structural neuroanatomy of tinnitus and hyperacusis in semantic dementia	Journal of Neurology, Neurosurgery and Psychiatry	2011	Case control
Melcher et al.	Subcallosal brain structure: correlation with hearing threshold at supra-clinical frequencies (>8 kHz), but not with tinnitus	Hearing Research	2013	Case control
Mühlau et al.	Structural brain changes in tinnitus	Cerebral Cortex	2005	Case control
Schecklmann et al.	Cluster analysis for identifying sub-types of tinnitus: a positron emission tomography and voxel-based morphometry study	Brain Research	2012	Cohort
Schecklmann et al.	Auditory cortex is implicated in tinnitus distress: a voxel based morphometry study	Brain Structure and Function	2013	Cohort
Vanneste et al.	Tinnitus: A Large VBM-EEG Correlational Study	PLoS ONE	2015	Cohort



**Table 2**

Demographic Information

Study	Number of subjects (male)	Age of subjects (std)	Mean Tinnitus Severity & questionnaire (std)	Hearing loss in subjects	Laterality of tinnitus	Number of controls (males)	Mean age of controls (std)
Allan et al.	73 (43)	58.3 (12.41)	THI and THQ but no data scores	Yes		55	56.9 (16.39)
Boyen et al.	31 (20)	56 (9)	29 (20) THI	Yes	22 bilateral	24 (16)	58 (6)
Husain et al.	8 (8)	56.13 (7.04)	17.25 (5.01) THI	Yes	All bilateral	11 (7)	48.09 (10.42)
Krick et al.	20 (11) 22 (13)	43.9 (10.4) 42.6 (11.50)	38.5 (15.4) 36.2 (16.8)	Only moderate hearing loss less than 40 dB	10 bilateral	20	NA
Landgrebe et al.	28 (15)	32.3 (9.4)	32.9 (13.9) Goebel and Hiller questionnaire	No	20 bilateral	28 (15)	31.2 (9.5)
Leaver et al. (2011)	11 (5)	44.4 (16)	THI (modified)	Yes	7 bilateral	21 (5)	23 (3.3)
Leaver et al. (2012)	23 (12)	47.4 (2.9)	THI but no data scores	Yes		21 (8)	49 (2.6)
Mahoney et al.	7 (4)	61.5 (5.1)	42 THI only from two patients	Yes		36 (20)	60.2 (5.9) and 66 (7.9)
Melcher et al.	24 (12)	46.9 (8.3)	TQR but no mean scores	No		24 (12)	45.8 (7.6)
Mühlau et al.	28 (13)	40	25 (16) tinnitus-fragebogen	No	All bilateral	28 (13)	39 (-)
Schecklmann et al. (2012)	44 (30)	45 (13)	35 (16) TQ	Unclear-patients with conductive hearing loss were excluded	25 bilateral	None	NA
Schecklmann et al. (2013)	257 (184)	50 (12)	39 (17) TQ	Yes	182 bilateral	None	NA
Vanneste et al.	154 (102)	50.24 (14.28)	36.02 (16.32) TQ	Yes	110 bilateral	None	NA

**Table 3**

Technical Information

Study	Scanner type	Voxel size (mm)	Matrix size (mm)	Echo time (ms)	Repetition time (ms)	Inversion time (ms)	Flip angle (°)	Number of slices
Allan et al.	3.0 T or 1.5 T Philips scanner	1×1×1		3.74	8			
Boyen et al.	3.0 T Phillips Intera MR scanner	1×1×1	256×256	3.5	9		8	170
Husain et al.	3.0 T GE Excite	.935×.935×1.3			30		12	128
Krick et al.	3.0 T Siemens Skyra	.9×.9×.9		3.74	8			
Landgrebe et al.	1.5 T Siemens Sonata	1×1×1	256×192	3.39	1900	1100	15	150
Leaver et al. (2011)	3.0 T Siemens Trio	1×1×1	256×256	2.94	2300	900	9	160
Leaver et al. (2012)	3.0 T Siemens Trio	1×1×1	256×256	3.5	2530	1100	7	176
Mahoney et al.	1.5 T scanner type unknown		256×256	5	12	650		124
Melcher et al.	3.0 T Siemens Trim Trio			3.45	2530	1100	7	
Mühlau et al.	1.5 T Siemens		256×256	3.93	1520	800	15	160
Schecklmann et al. (2012)	1.5 T Siemens Sonata	1×1×1	256×192	3.93	1900	1100	15	150
Schecklmann et al. (2013)	1.5 T Siemens Sonata	1×1×1	256×256	3.42	1880	1100	15	76
Yanneste et al.	3.0 T Siemens Trio	1×1×1	256×256	2.94	2300	900	9	160

**Table 4**

Image Processing Methodology

Study	Software	Preliminary steps	Reference space for normalization	Tissue segmentation procedure	Smoothing (FWHM)
Allan et al.	SPM8	None	ICBM152 (assumed)	Unified segmentation (assumed)	10 mm
Boyen et al.	SPM5 and VBM toolbox	None	ICBM152 (assumed)	VBM toolbox	8 mm
Husain et al.	SPM5	None	ICBM152	Unified segmentation	8 mm
Krick et al.	SPM8 and VBM8 toolbox	None	ICBM152 (assumed)	VBM8 toolbox	10 mm
Landgrebe et al.	SPM2	None	ICBM152	Unified segmentation	10 mm
Leaver et al. (2011)	SPM8	None	ICBM152 (assumed)	Unified segmentation	12 mm
Leaver et al. (2012)	SPM8	None	ICBM152	Unified segmentation	6 mm
Mahoney et al.	SPM5	None	ICBM152 (assumed)	Unified segmentation (assumed)	
Melcher et al.	SMP8	Manual rigid body transformation	ICBM152 (assumed)	Unified segmentation	8 mm
Mühlau et al.	SPM2	None	ICBM152 (assumed)	Study specific prior probability maps	8 mm
Schecklmann et al.	SPM8 and VBM toolbox	None	Not stated	Not stated	8 mm
Schecklmann et al.	SPM8 and VBM toolbox	None	Not stated	Not stated	8 mm
Vanneste et al.	SPM8	None	ICBM152 (assumed)	Unified segmentation	8 mm

**Table 5**

**Whole Brain Analysis for Studies with Controls**

Study	Statistical Analysis of Whole Brain	Confounding Covariates entered	Significance Threshold	Non-Stationarity Adjustment	Correction for Variance	Implicit Masking
Allen et al.	Analysis of covariance (ANCOVA) regression model	Whole brain volume and white matter volume	P<0.05 FWE corrected at the cluster level (assumed)	None	NA	NA
Boyen et al.	First a one-way ANCOVA regression model for equality between groups than a two sample t-test between groups	Age	P<0.05 FWE corrected at the cluster level	None	Non-sphericity correction	Excluded all voxels with a GM value <0.2
Husain et al.	First an ANOVA for equality between groups than a two sample t-test between groups	Whole brain volume	P<0.001 uncorrected at voxel level and P<0.05 FWE and FDR corrected at cluster level	None	Non-sphericity correction	Excluded all voxels with a GM value <0.2
Krick et al.	Two-way ANCOVA	None	P<.05 FWE corrected at the cluster level	None	NA	NA
Landgrebe et al.	Two sample t-test	Whole brain volume	P<0.05 FDR corrected at voxel level	NA	Non-sphericity correction	Excluded all voxels with a GM value <0.2
Leaver et al. (2011)	Two sample t-test	Age, total gray or white matter volume	P<0.0001 uncorrected at the voxel level	NA	NA	Excluded all voxels with a GM value <0.2
Leaver et al. (2012)	t-tests between groups	None	P<.002 uncorrected at the voxel level	NA	Non-sphericity correction	Excluded all voxels with a GM value <0.2
Mahoney et al.	T test between groups(assumed)	Whole brain volume, age, gender	P<.05 FWE corrected at the voxel level	NA	NA	Included all voxels with intensity >0.01 in >70% of subjects
Melcher et al.	Two sample t-test	None	P<0.05 FWE corrected at the voxel level and P<0.001 uncorrected at the cluster level	None	Non-sphericity correction	Excluded all voxels with a GM value <0.2
Mühlau et al.	Two sample t-tests	Whole brain volume	P<0.05 FDR correction at voxel level and P<0.05 FDR corrected at the cluster level	None	Non-sphericity correction	Excluded all voxels with a GM value <0.2

Table 6

## Region of Interest Analysis

Study	How ROI was defined	Statistical test used on ROI	Significance threshold
Allan et al.	Brodmann areas were defined by WFU_pickatlas, auditory nuclei were defined via MNI coordinates	Two sample t-test	
Boyen et al.	Brodmann areas were defined by WFU_pickatlas, auditory nuclei were defined via MNI coordinates, left and right cerebellum defined according to the WFU_pickatlas	ANCOVA regression model and then two sample t-test	P<0.05 corrected for FDR for cluster
Husain et al.	Brodmann areas were defined by WFU_pickatlas, auditory nuclei were defined via MNI coordinates	ANOVA regression model and then two sample t-test	P<0.005 uncorrected at cluster level for one test. Also did second test with P<0.001 uncorrected and P<0.05 for voxel and cluster level
Landgrebe et al.	Brodmann areas were defined by WFU_pickatlas, auditory nuclei were defined via MNI coordinates	Voxelwise t-tests	P<0.05 FDR corrected for voxel and cluster level
Leaver et al. (2011)	Auditory cortex was defined by a silence vs. sound contrast. MGN was defined by WFU_pickatlas	Two sample t-test	p<0.0001 uncorrected for voxel level *For superior temporal cortex p<0.01
Leaver et al. (2012)	ROIs were defined by MNI coordinates and by prior templates with manually identified structures	Two sample t-test	
Melcher et al.	Auditory nuclei were defined by MNI coordinates	Two sample t-test	P<0.001 uncorrected for cluster level and P<0.05 unclear correction for cluster level
Mühlau et al.	Brodmann areas were defined by WFU_pickatlas, auditory nuclei were defined via MNI coordinates	Voxelwise t-tests	P<0.05 FDR correction at voxel and cluster level
Vanneste et al.	Brodmann areas were defined by WFU_pickatlas	Split-half analysis and then 3 simple regression analyses	P<0.001 uncorrected and p<0.05 FDR correction at voxel and cluster level

**Table 7**

Regression Covariates

Study	Age	Gender	Tinnitus type	Tinnitus laterality	Tinnitus related distress	Tinnitus volume	Tinnitus duration	Tinnitus frequency	Tinnitus sensation	Hearing level
Schecklmann et al. (2012)	X	X								X
Schecklmann et al. (2013)	X	X	X	X		X				X
Vanneste et al.	X	X	X	X	X	X		X	X	

Table 8

## Results

Study	Whole Brain Analysis	ROI analysis
Allan et al.	Decrease in gray matter for all patients in left superior frontal gyrus and right precuneus No significant finding for the severe tinnitus group Increase in WM in left Heschl's Gyrus and cochlear nucleus for the tinnitus patients with normal hearing group	Increase in gray matter in the superior olivary complex when looking at all patients For the matched cohorts, there were no significant findings No significant findings for tinnitus patients with normal hearing vs. controls
Boyen et al.	No significant finding	Increases in gray matter in the left primary auditory cortex Decreases in gray matter in both inferior temporal areas
Husain et al.	No significant finding	No significant finding
Krick et al.	Tinnitus patients had significant GM increase with musical therapy in Heschl's Gyrus, and the Rolandic operculum when compared to active controls	
Landgrebe et al.	No significant finding	
Leaver et al. (2011)	Tinnitus patients had significantly less in gray matter of subcallosal region, vmPFC. Increase in vmPFC white matter in tinnitus patients	Decrease in gray matter in the right inferior colliculus, and the left hippocampus
Leaver et al. (2012)	Tinnitus patients less GM volume in vmPFC, dmPFC, left supramarginal gyrus	Tinnitus patients had significantly less in gray matter of subcallosal region, vmPFC. Increase in vmPFC white matter in tinnitus patients
Mahoney et al.	Tinnitus patients had increased gray matter in the right posterior superior temporal gyrus and sulcus and reduced gray matter in bilateral orbitofrontal cortices	Decrease in gray matter bilaterally in ventromedial prefrontal cortex
Melcher et al.	No significant findings	No significant finding
Mühlau et al.	Decrease in gray matter in the subcallosal area	Increase in gray matter in the right geniculate body
Schecklmann et al. (2012)	Cluster analysis of VBM data showed that groups differed in gray matter volume in medial frontal, cingulate, temporal, insular, pre- and postcentral, and thalamic areas	
Schecklmann et al. (2013)	Decrease of gray matter in bilateral middle and superior temporal cortex (including Heschel's gyrus and insula) associated with tinnitus severity	
Yanneste et al.	Decrease in gray matter correlated with tinnitus distress loudness and duration	Reduced gray matter density for the unilateral tinnitus patients in comparison to the bilateral tinnitus patients in the right primary auditory cortex which is in association with tinnitus lateralization