The impact of age-related hearing loss on structural neuroanatomy: a meta-analysis

Kate Slade1\*†, Johannes H Reilly1†, Kamila Jablonska1, El Smith1, Lawrence D Hayes3, Christopher J Plack1,2, Helen E Nuttall1\*

1Department of Psychology, Faculty of Science and Technology, Lancaster University, Lancaster, UK

2Manchester Centre for Audiology and Deafness, School of Health Sciences, University of Manchester, Manchester, UK

3 Institute of Clinical Exercise and Health Science, School of Health and Life Sciences, University of the West of Scotland, Glasgow, UK

†These authors have contributed equally to this work and share first authorship

**\* Correspondence:**Corresponding Authors  
[k.slade2@lancaster.ac.uk](mailto:k.slade2@lancaster.ac.uk)  
[h.nuttall1@lancaster.ac.uk](mailto:h.nuttall1@lancaster.ac.uk)

Keywords: Age-related hearing loss, grey matter, structural MRI, brain volume

Abstract

This meta-analysis investigated the association between age-related hearing loss and structural neuroanatomy, specifically changes to grey matter volume. Hearing loss is associated with increased risk of cognitive decline. Hence, understanding the effects of hearing loss in older age on brain health is essential. We reviewed studies which compared older participants with hearing loss (HL) to older adults without clinical hearing loss (NH), on neuroanatomical outcomes, specifically grey matter volume as measured by magnetic resonance imaging. A total of five studies met the inclusion criteria, three of which were included in an analysis of whole-brain grey matter volume (HL-group n = 113; NH-group n = 138), and three were included in analyses of lobe wise grey matter volume (HL-group n = 139; NH-group n = 162). Effect-size seed-based d mapping software was employed for whole-brain and lobe-wise analysis of grey matter volume. The analysis indicated there was no significant difference between older adults with HL compared to those with no age-related HL in whole-brain grey matter volume. Due to lacking stereotactic coordinates, the atrophy of grey matter in specific neuroanatomical locations could only be conducted at lobe-level. These data indicate that older adults with HL show increased grey matter atrophy in the temporal lobe only (not in occipital, parietal, or frontal), compared to older adults without HL. The implications for theoretical frameworks of the HL and cognitive decline relationship are discussed in relation to the results. Overall, the findings endorse regular auditory testing for ≥60-year-olds, as HL co-occurred with atrophy in grey matter. Managing age-related HL, e.g., through hearing aids, may reduce the adverse functional effects of grey matter atrophy. This meta-analysis was pre-registered on PROSPERO (CRD42021265375).

# Introduction

The population is ageing, meaning that health issues which affect older adults become more prevalent (Office for National Statistics, 2018), impacting on the older population’s quality of life and placing increasing pressure on health care services. Two such health concerns are hearing loss and dementia. In the UK around 70% of people aged 70+ experience hearing loss (1), and around 7.1% of over 65s, rising to 14% of those over 80, are living with dementia (2). Critically, there is evidence that these health concerns may be associated. Hearing loss has been identified as the largest potentially modifiable risk factor for dementia (3). Hearing loss could account for as much as 8% of global dementia cases (4). It is likely that many risk factors are associated, and exacerbate one another leading to increased dementia prevalence in certain individuals. Considering this, understanding the neural mechanisms of hearing loss, and how they may contribute to the association between hearing loss and dementia, is a priority.

Age-related hearing loss (ARHL) is often caused by degeneration of the inner and outer hair cells within the cochlea. These cells are responsible for the transduction of sound, and their atrophy manifests in high-frequency hearing loss (5). Age-related atrophies in the peripheral auditory system can also be observed in the stria vascularis, a cochlea structure responsible for maintaining metabolic processes (6), or in degeneration of spiral ganglion cells, the initial neurons in the pathway from the ear to the brain (7). Importantly, evidence suggests that changes and atrophies in people with ARHL do not end at the peripheral auditory system, but are also evident in the auditory pathway and auditory cortex (8,9). Understanding the cortical changes, in auditory areas and beyond, would provide valuable insights into how the brain changes in people with ARHL, and provide evidence for the mechanisms that underpin the association between hearing loss and cognitive decline.

A number of potential explanations for the relation between hearing loss and dementia have been proposed. These include non-causal hypotheses such as: 1. The common cause hypothesis, which suggests that rather than hearing loss leading to dementia, there is a common neuro-degenerative pathology which underlies both conditions such as general ageing or vascular disease (10,11); or 2. The hearing bias in cognitive assessment hypothesis, which suggests that there may be an overestimation of the link between hearing loss and dementia, because untreated hearing loss could be a significant confound in clinical cognitive assessments that rely on auditory presentation (12,13). However, the relation between hearing loss and cognitive decline remains after controlling for age and vascular factors (14,15), and hearing loss has been found to be associated with poorer cognitive functioning even when the cognitive assessments do not rely on verbal communication (16). As such, a causal mechanism may be more likely. Causal hypotheses include: 1. The cognitive load (or information degradation) hypothesis, which theorises that people with ARHL are required to use more cognitive resources for speech perception leaving fewer available for general cognitive processes, which could lead to the symptoms of dementia (17); and 2. The sensory deprivation hypothesis, which postulates that reduced sensory input from the ear leads to reduced neural activation, cortical re-organisation, and atrophy across brain areas involved with speech perception (9,18). Both these causal hypotheses make suggestions with regards to functional or structural neuro-cognitive changes that might accompany ARHL, including upregulation or reorganisation of neural resources (19) or atrophy across speech perception networks (see (20) for a discussion of cortical changes). As such, a comprehensive review of the current literature on the neural consequences of ARHL is required to generate evidence to refute or support these causal hypotheses.

The first step in the systematic examination of neural consequences of ARHL is to assess the evidence for tangible neuroanatomical changes, in both auditory and wider cortices. There is evidence from longitudinal studies that individuals with ARHL display accelerated grey matter (GM) atrophy in auditory cortex compared to individuals without ARHL (21), however these group differences have not always reached statistical significance (22). In cross-sectional studies, there is also evidence for decreases in whole brain volume in those with ARHL compared to those without (9), and in specific brain areas associated with speech perception including the anterior cingulate cortex (23). The mechanism by which ARHL leads to wider brain atrophy is unclear. One potential explanation is that over-reliance on wider brain networks involved in speech perception due to impaired auditory processing contributes to neural degeneration of these areas. There is evidence that individuals with ARHL, compared to those without, display increased functional connectivity across auditory and visual sensory cortices (24), and between auditory cortex and the cingulo-opercular network after controlling for both age and cognitive function (25). The over-reliance on these additional brain networks to support speech perception in individuals with ARHL could enable neural degeneration due to glutamate excitotoxicity (23). Through this mechanism, the neurons across the up-regulated brain networks may die due to prolonged activation of glutamate receptors beyond their natural capacity (26).

Despite evidence for potential up-regulation and cortical atrophy, it is still unclear as to whether or not ARHL exacerbates the cortical changes observed in ageing. Heterogeneity in research methods, such as differences in participant age ranges, imaging techniques, or clinical definitions of hearing loss, as well as small sample sizes, can lead to ambiguity in interpretation of the results. This meta-analysis will deliver a systematic and specific analysis of the existing literature on neuroanatomical changes in ARHL, controlling for some of these confounds through appropriate inclusion criteria, and study quality assessment. We sought to investigate across cross-sectional and longitudinal evidence whether GM volume, as measured by MRI, differs in adults aged ≥60 years with ARHL compared to those without ARHL. It was hypothesised that we would observe 1. decreased whole brain GM volume, as well as 2. largest decreases in GM volume in the temporal lobe compared to other cortical lobes, in individuals with ARHL compared to those without ARHL.

# Methods

This meta-analysis was pre-registered on PROSPERO (PROSPERO 2021, CRD42021265375), available from: <https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021265375>) Additionally, all materials including: search strings; references obtained at all screening stages; screening manuals and inter-rater consistency data; extracted data; and analysis files can be found in the project’s repository on the Open Science Framework (OSF): https://osf.io/g5qcb/.

## Literature Search

An initial pilot search was conducted on PubMed and Prospero to: 1. confirm whether systematic reviews and meta-analyses on this topic already existed; 2. estimate the feasibility of the meta-analysis and availability of data (27); and 3. identify key papers to inform the selection of appropriate key words and criteria for the final search string. Unlike the full literature search, the pilot search was characterised by iterative searching without pre-defined search strings (28). In-depth engagement with the literature might introduce potential bias in the construction of the full literature search. Hence, engagement with the pilot search was limited to two hours.

Subsequently, the full literature search was conducted following PRISMA guidelines (29) and best practice guidelines from the “Cochrane Highly Sensitive Search Strategy” (30). Nine databases were searched (Table 1). To maximise effective retrieval of relevant papers, the research question was approached, inter alia, from medical (PubMed), nursing (CINAHL), and psychological (PsycINFO) perspectives. To ensure comprehensiveness, searches were not filtered in any way, e.g., by database internal filters, such as publication date, or by full-text articles (see Table 1 for exception in Scopus). If any full text could not be accessed the research team planned to contact the authors of the papers, and allow one week for an initial response before recontacting. A total of two weeks was granted for authors to respond and provide access to papers.

Search strings (see Supplementary Materials 1) were constructed using keywords, free-text terms, and search functions (Boolean operators, near searches, truncation, wild card symbols, quotations), to ensure specificity and sensitivity across databases. The final search strings, selected keywords, and Boolean operators were reviewed by a librarian to ensure adherence to best practice insights. To manage resource and time constraints, the search was limited to titles and abstracts.

Prior to conducting the literature search, a strategy test of sensitivity was completed in which four key papers that satisfied the inclusion criteria were identified using a database not used in the final search to avoid bias (Google Scholar). The sensitivity of the search strings was evaluated by testing how many of these four key papers indexed in the selected databases (Table 1) could be retrieved. Once the search sensitivity was acceptable, the literature search was conducted across the selected databases. All key papers were retrieved with the search indicating high likelihood that the search would successfully identify relevant articles.

## Article Screening

Articles (n = 14078) retrieved from the literature search were imported to the reference manager software CADIMA (<https://www.cadima.info/>; for a review, see (31)). An overview of the articles retained at each stage of the screening process can be found in the PRISMA flow diagram presented in Figure 1 (32). Any duplicated articles, retrieved by more than one database, were removed in a two-step process: 1. automatic de-duplication based on congruity in authors, title, and year of publication; and 2, manual de-duplication by two raters based on abstracts.

Unique articles (n = 9497) were screened by four raters for inclusion according to specific criteria (see the associated OSF repository for the full criteria used: https://osf.io/g5qcb/).in two consecutive stages: 1. title-abstract; and 2. full-text screening. Before each screening stage, the consistency between raters was checked with inter-rater reliability tests (Table 2). A proportion of articles (60 at title-abstract screening stage, and 40 at full-text screening stage) were screened by two raters in parallel until at least 80% agreement was reached for each criterion.During screening, a manual with the inclusion criteria, additional background information, and guidance for the use of CADIMA was provided (manuals are also available in the OSF repository).

After consistency checks, 10% of all titles and abstracts and 30% of full-texts were double screened by two independent raters in parallel. During this initial period in screening, inconsistencies were resolved through group discussion and if necessary, information was added to the screening manual to clarify eligibility criteria. Training and extensive guidance was provided to ensure all raters fully understood the application of eligibility criteria before the remaining articles were independently screened. Throughout this process, raters met weekly to resolve outstanding questions. Raters were instructed to include rather than exclude articles if unsure, to prevent false exclusion of papers.

The final set of articles that passed title-abstract (n = 176) and full-text screening stages (n = 14), were checked for listing in the Retraction Watch Database (<http://retractiondatabase.org/>) to ensure that only studies not retracted were included.

Articles were screened for inclusion along a set of pre-defined eligibility criteria for 1. the title-abstract and 2. the full-text screening stages. These criteria were designed in line with the PICO/PECO framework (33,34), which clarifies the review objectives and inclusion criteria across four domains: Population (P), Intervention/Exposure (I/E), Comparator (C), and Outcomes (O). To meet the inclusion criteria, articles were required to be original research, containing empirical data, and provided in English. Additionally, the articles needed to meet the following PICO/PECO criteria. (P) it was required that participants be older adults (average age of 60+ at the time of at least one study session) without clinical psychological or neurological illnesses, who either had or did not have hearing loss (HL). (I/E) HL was defined as a pure tone average (PTA) of > 25 dB HL across 0.5-2 kHz and no HL was defined as a PTA of ≤25 dB HL averaged across 0.5-2 kHz (35,36). (C) Studies needed to compare at least two groups, one group with HL and one group without HL, either longitudinally or cross-sectionally. (O) Outcome measures needed to include voxel-based morphometry data (VBM) as measured by magnetic resonance imaging (MRI) available for both groups. The outcome measures of interest were grey matter (GM) volumes for specific brain regions or for the whole brain.

## Data extraction

Data extraction was performed manually by four reviewers with identically structured Microsoft Excel (2018) forms. For each study, data were extracted by at least two independent reviewers and then checked for agreement, to decrease the possibility of manual errors (37). Any inconsistencies in the extracted data were resolved through discussion. A data extraction manual was provided (available in the OSF repository). Where data were presented visually only, means and standard deviations were read from graphs. If non-significance or significance was reported without associated exact *p*-values, the *p*-value was assumed to be *p*=.05 and *p*=.04, respectively (based on (37).

The main source of heterogeneity in analysis is likely to stem from sample characteristics, study design, and the mask applied in imaging. Therefore, data extraction included participant demographics for both HL and no HL groups (sample size, age, sex, PTA), study design (timeframe, sampling method, timescale of longitudinal measurements), details of image acquisition (MRI field strength, smoothing kernel, slice thickness, voxel size, mask, normalisation space), and outcome measures (e.g., (un)corrected *p*-values, effect sizes, mean and standard deviation of whole-brain and regional GM volumetric measurements). Any papers found not to meet the inclusion criteria at data extraction stage were excluded (for details, see the PRISMA flowchart, Figure 1). The remaining papers (n = 5) were included in this meta-analysis.

## Critical appraisal

The critical appraisal tool was based on an adapted version of the trialled AXIS appraisal framework (38) and response options were based on QualSyst (39). Individual criteria of the original AXIS tool were omitted or included based on Müller et al. (27), the STROBE statement (40), and GRADE (41,42). The critical appraisal assessed study quality and risk of bias to evaluate the reliability and validity of studies’ findings and whether findings are representative and generalisable at population-level. No automation tools were used in this process. To minimise subjectivity (27), each paper included in the analysis was appraised by two raters independently, and disagreements resolved by discussion or ultimately, a third rater. Raters were trained and received an explanatory manual. To assess homogeneity in methods and outcomes across studies, the critical appraisal accounted for whether or not research controlled for confounding factors that influence brain structure (e.g., sex, smoking status, body mass index, socio-economic status, cardio-vascular function, diabetes, co-morbidity of disorders), as well as methodological factors that could influence data interpretation (e.g., MRI field strength, sample size; the appraisal manual and method of calculation are available in the OSF repository).

## Statistical analysis

Statistical analysis was conducted using effect-size seed-based d mapping (ES-SDM) software to perform a random-effects meta-analysis (43), the software was developed to aid the meta-analysis of voxel-based data as obtained by VBM ([www.sdmproject.com](http://www.sdmproject.com)). VBM is a neuroimaging technique comparing GM concentration by mapping images onto a normalised stereotactic space and extracting GM volumes, smoothing data, and finally, comparing group GM volume differences via voxel-wise comparison (44). ES-SDM is described in detail elsewhere (43,45) and has previously been tested for reliability (46,47), including for GM volume comparison (48). ES-SDM calculates Hedges’ *g* effect sizes for mean analysis, based on group means, and standard deviations (45). Hedges’ *g* uses a pooled and weighted standard deviation based on sample size and is thus more accurate for small sample sizes (<20) than Cohen’s *d* which uses a normal standard deviation (49–51). The inclusion of non-significant findings in the analysis addresses bias towards significant overall results.

The analysis of GM volumes in ES-SDM was a mean analysis providing Hedges’ *g* and corresponding *z*- and *p*-values, as well as standard error, the lower and upper bounds of the effect size for each study, and a mean across studies. *Q* statistics were used to assess inter-study heterogeneity of effect sizes. The analysis followed the ES-SDM manual (available here [www.sdmproject.com](http://www.sdmproject.com)). Furthermore, we verified the analysis in RStudio (R version 4.1.0, (52), using the *metafor()* package to conduct a random-effects model meta-analysis (53), and to produce the associated forest and funnel plots. The data analysis obtained was the same in ES-SDM and R. The R code is provided in the OSF repository: https://osf.io/g5qcb.

### Calculation of missing standard deviations

Under the assumption that data were normally distributed, missing standard deviations were calculated from confidence intervals using the following formula (30):

### ES-SDM and R analysis

Sample sizes of both groups (HL vs. no-HL), means and standard deviations were entered for each study and each region of interest (ROI) into ES-SDM. Separate analyses were conducted for whole-brain and ROI GM volume using the same ES-SDM “globals” calculator as it relies on mean analysis and is, therefore, also suitable for analysis of mean ROI data. To compare ROIs, ROIs were collated into frontal, temporal, parietal, and occipital lobes. The collation was completed following the papers’ verbal labels of ROIs (e.g., superior temporal lobe was allocated to the temporal cortex) and widely accepted localisations, e.g., precentral gyrus is undisputedly considered to lie in the frontal lobe. If allocation to a lobe was unclear, a neuroanatomy textbook was consulted (54). The same data were entered into R and separate meta-analyses were conducted for whole-brain, frontal, temporal, parietal, and occipital lobe data, as was done in ES-SDM software.

# Results

Of the 9,497 articles screened, four satisfied all inclusion criteria (see Figure 1 for the PRISMA flow diagram). During title-abstract screening, a total of 413 inter-rater inconsistencies were resolved of which only 102 affected inclusion (n = 25) or exclusion (n = 77) of the article. During full-text screening there were 37 inconsistencies of which 16 affected inclusion (n = 2) or exclusion (n = 14). The number of articles that were at first included, but through discussion of inconsistencies excluded, can be explained by the instructions to screeners to be more lenient than conservative in case of uncertainty when judging whether or not the articles fulfilled screening criteria.

Across both screening stages, the criteria that caused most inconsistencies were whether or not participants were at least 60 years old and (neurologically) healthy, as well as whether or not the study made a direct comparison of neuroanatomical differences between groups.

## Heterogeneity of effect sizes and evaluation of study quality

Descriptive statistics of the meta-analysed studies are presented in Tables 3-4. Overall, critical appraisal scores did not lie below .83, indicating high methodological quality (39). In combination with the observation that all ratings fell between .83 and .89, it is unlikely that methodological inadequacies skewed results or studies formed subgroups of studies with high and low methodological quality. However, it should be noted that a source of bias might be the consistently partial fulfilment of a sampling process likely to represent the target population. All studies employed convenience sampling (recruiting from hospital settings or previous study cohorts) and acknowledged this as a limitation. The results in this meta-analysis are consequently subject to the same constraints in generalisability of results.

Across analyses, the *Q* statistic did not reach significance indicating no significant heterogeneity of effect sizes between studies (whole-brain, *Q*(2) = 3.93, *p* = .14; frontal lobe, *Q*(7) = 3.11, *p* = .87; temporal lobe, *Q*(26) = 17.59, *p* = .67; parietal lobe *Q*(2) = 2.78, *p* = .10; occipital lobe, *Q*(4) = 2.43, *p* = .66). This homogeneity, in combination with the results from the critical appraisal, suggest no significant variation in the studies’ characteristics and that the heterogeneity likely stems from sampling error alone. Thus, it is unlikely that underlying variation in methodology or participant groups between studies skewed the results (55).

## Whole-brain and lobe-wise analysis

A comprehensive overview and visualisation of results is presented in Tables 5-8, and Figures 2-6, respectively. In comparison with group no-HL, group HL showed decreased mean whole-brain GM volume. However, this atrophy was not significant, Hedges’ *g* = .12, *p* = .52. Similar to the whole-brain GM volume, group HL showed decreased GM volumes in all four lobes. This atrophy was significant in the temporal lobe (Hedges’ *g* = -0.12, *p* = .007), but was not significant in the frontal (Hedges’ *g* = -0.03, *p* = .64), parietal (Hedges’ *g* = -0.17, *p* = .52), nor occipital (Hedges’ *g* = -0.12, *p* = .14) lobes.

# Discussion

This meta-analysis sought to collate and evaluate the existing evidence for a difference in brain volume, specifically GM volume, in older adults (aged ≥60 years) with ARHL, compared to those without ARHL. We sought to include data from both cross-sectional and longitudinal study designs, in order to consolidate and analyse the available empirical evidence and provide a better understanding of cortical changes associated with ARHL. We employed ES-SDM software to conduct analysis of neuroanatomical data across the included studies (43).

Two studies, one of which took a cross-sectional approach and one of which took a longitudinal approach, which reported GM volumes for the whole brain were included in the analysis of global neuroanatomical changes associated with ARHL. This analysis served to investigate whether the entire brain displays significant GM atrophy in individuals with ARHL, compared to those without ARHL, in order to further understand how hearing loss contributes to brain ageing. The findings did not support our hypothesis that adults with ARHL would display significantly decreased whole brain GM volume, compared to those without ARHL. Despite the observation in the descriptive statistics that the ARHL group did show smaller mean GM volume than the no-ARHL group, this difference was not significant. While previous research suggests that cross-cortical and brain wide changes are associated with ARHL (9,21), this meta-analysis of collated studies suggests that changes associated with ARHL are not significantly greater than changes which occur in ageing. If this is the case, then it is possible, as suggested by the common cause hypothesis, that a neurodegenerative factor may underly the brain atrophies observed in both hearing loss and ageing.

Elevated tau protein levels could be an indicator for a potential third factor that no meta-analysed study explicitly accounted for. Tau is a protein found to abnormally aggregate in Alzheimer’s Disease and has, therefore, been considered as a viable biomarker (56). In an MRI study that considered tau protein as a confound, Belkhiria et al. (57) only found correlations between brain volume change and cognitive function in participants with ARHL (PTA >20 dB) and cochlear dysfunction, not in participants with ARHL and preserved cochlear function or in controls (all >65 years of age). Importantly, only participants with ARHL and cochlear dysfunction showed elevated tau levels. Consequently, GM atrophy in the meta-analysed studies might be caused by early-stage Alzheimer’s Disease. This may have only affected the HL-groups.

However, limiting neuroanatomical observations to whole brain analysis only may result in overlooking of essential information regarding lobe-wise cortical changes. Understanding in which cortical structures changes occur is important for establishing the role of potential underlying causal mechanisms. Three studies, two of which took a cross-sectional approach and one of which took a longitudinal approach, which reported GM volumes in specific brain areas were included in the lobe-wise analysis of GM volumes. This analysis enabled the investigation into cortical changes across brain lobes to establish whether GM atrophy extends beyond auditory cortex (situated in temporal lobe) in individuals with ARHL. The findings supported our hypothesis that the largest decreases in GM volume observed in individuals with ARHL compared to those without ARHL would occur in the temporal lobe. This is consistent with existing literature which reports increased neural atrophy in auditory cortex in individuals with ARHL, compared to those without ARHL (21). No evidence was found that declines in GM volume observed in people with ARHL extend beyond temporal lobe. Research suggests that ARHL leads to up-regulation across brain networks to support speech perception, and there is evidence from functional imaging studies to support this, showing that ARHL is associated with increased functional connectivity between auditory cortex and cognitive networks (25). Increased use of such cortical resources has been theorised to trigger neurodegeneration due to over-use of neural resources and excitotoxic cell death. Yet, our data provide no evidence for declines in GM volume beyond temporal lobe and thus do not support the hypothesis that potential compensatory activity leads to neurodegeneration. This has implications for interpretation of the causal hypotheses underlying the association between hearing loss and cognitive decline. Importantly, previous research finds that declines in cognitive functioning are also associated with greater GM volume loss in temporal regions (58). This has important implications as temporal atrophies may be an underlying mechanism in the relation between hearing loss and cognitive declines.

It is important to consider, with regards to both the whole-brain and lobe-wise analyses, that the included study designs varied between cross-sectional and longitudinal. First, it is possible that global GM atrophy, or atrophy across wider cortices, only occurs after prolonged sensory deprivation. In two previous longitudinal studies, a significant association between pure-tone hearing loss and reduced GM in auditory cortex was only present after at least 5 years (21,22). Hence, it is possible theoretically that atrophies extending further than auditory cortex, or temporal lobe, may only occur after prolonged up-regulation or cortical resource reallocation to assist speech perception due to ARHL. Second, in both designs, consideration of confounding factors is important, but particularly for cross-sectional research. As such, it is important to note that differences in the controlled variables across the included studies may affect the results, and create ambiguity for interpretation. By design, longitudinal research allows for increased control over individual factors which may influence data, and hence any observed neural changes are more easily interpreted as occurring due to HL, rather than ageing or another underlying neurodegenerative variable.

Further, as many studies did not report stereotactic coordinates, the data analysis options were limited to general lobe comparisons. Hence it is not possible to interpret exactly where GM atrophy occurs within the temporal lobe. Without exact cortical locations, it is difficult to draw strong conclusions regarding the underlying neural processes or systems. Additionally, all included studies employed opportunity sampling techniques. Therefore, any generalisations were limited to the targeted populations in the included studies. Importantly, these data should be interpreted with consideration of the sample size of studies included. In order to control for confounding variables and ensure heterogeneity in methods, strict inclusion criteria were used to select the studies meta-analysed, which in-turn resulted in a small sample. It has been suggested that a large sample size of individuals is required for adequate power in whole-brain meta-analysis (59). For this to be possible, there is explicit need for future large-scale longitudinal research which seeks to observe the effects of age-related hearing loss on brain morphology.

In conclusion, this meta-analysis explored the evidence for a difference in GM volume, in older adults with ARHL, compared to those without ARHL. The analysis found evidence for reduced GM volume in temporal lobes in individuals with ARHL, compared to those without ARHL. There was no evidence that GM atrophies extended to frontal, parietal, or occipital lobes, nor was there evidence for whole brain GM declines in individuals with ARHL. It is possible that significant differences in GM volume are limited to the temporal lobe, because further cortical changes only occur after a critical time period of prolonged cortical resource re-allocation. However, this finding has important implications and further longitudinal research into how neural changes across the temporal lobe in people with ARHL affects wider brain health is essential.

**References**

1. Royal National Institute for Deaf People (RNID). Hearing Matters [Internet] (2020) [Accessed May 21, 2022]. Available from: <https://rnid.org.uk/wp-content/uploads/2020/05/Hearing-Matters-Report.pdf>
2. Wittenberg R, Hu B, Barraza-Araiza LF, Rehill A. Projections of Older People with Dementia and Costs of Dementia Care in the United Kingdom, 2019-2040. CPEC Working Paper 5 [Internet] (2019) [Accessed May 21, 2022]. Available from: <https://www.alzheimers.org.uk/sites/default/files/2019-11/cpec_report_november_2019.pdf>
3. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet (2020) 396(10248):413–46. Available from: <http://www.thelancet.com/article/S0140673620303676/fulltext>
4. Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. Lancet (2020) 390(10113):2673–734. doi: 10.1016/S0140-6736(20)30367-6
5. Liberman MC, Kujawa SG. Cochlear synaptopathy in acquired sensorineural hearing loss: Manifestations and mechanisms. Hear. Res. (2017) 349, 138–47. doi: 10.1016/j.heares.2017.01.003
6. Heeringa AN, Köppl C. The aging cochlea: Towards unraveling the functional contributions of strial dysfunction and synaptopathy. Hear. Res. (2019) 376:111–24. doi: 10.1016/j.heares.2019.02.015
7. Frisina RD, Ding B, Zhu X, Walton JP. Age-related hearing loss: Prevention of threshold declines, cell loss and apoptosis in spiral ganglion neurons. Aging (2016) 8(9):2081–99. doi: 10.18632/aging.101045
8. Grose JH, Buss E, Elmore H. Age-Related Changes in the Auditory Brainstem Response and Suprathreshold Processing of Temporal and Spectral Modulation. Trends Hear. (2019) 23:1–11. doi: 10.1177/2331216519839615
9. Rigters SC, Bos D, Metselaar M, Roshchupkin G v., Baatenburg de Jong RJ, Ikram MA, et al. Hearing Impairment Is Associated with Smaller Brain Volume in Aging. Front. Aging Neurosci. (2017) 9. doi: 10.3389/fnagi.2017.00002
10. Eckert MA, Kuchinsky SE, Vaden KI, Cute SL, Spampinato M v., Dubno JR. White Matter Hyperintensities Predict Low Frequency Hearing in Older Adults. J. Assoc. Res. Otolaryngol. (2013) 14(3):425–33. doi: 10.1007/s10162-013-0381-4
11. Lockhart SN, DeCarli C. Structural Imaging Measures of Brain Aging. Neuropsychol. Rev. (2014) 24(3):271–89. doi: 10.1007/s11065-014-9268-3
12. Füllgrabe C. On the Possible Overestimation of Cognitive Decline: The Impact of Age-Related Hearing Loss on Cognitive-Test Performance. Front. Neurosci. (2020) 14:454. doi: 10.3389/fnins.2020.00454
13. Füllgrabe C. When hearing loss masquerades as cognitive decline. J. Neurol. Neurosurg. Psychiatry. (2020) 91(12):1248–1248. doi: 10.1136/jnnp-2020-324707
14. Deal JA, Betz J, Yaffe K, Harris T, Purchase-Helzner E, Satterfield S, et al. Hearing Impairment and Incident Dementia and Cognitive Decline in Older Adults: The Health ABC Study. J. Gerontol. A Biol. Sci. Med. Sci. (2016) 72(5):703-709. doi: 10.1093/gerona/glw069
15. Gurgel RK, Ward PD, Schwartz S, Norton MC, Foster NL, Tschanz JT. Relationship of Hearing Loss and Dementia. Otol. Neurotol. (2014) 35(5):775–81. doi: 10.1097/MAO.0000000000000313
16. Lin FR. Hearing Loss and Cognition Among Older Adults in the United States. J. Gerontol. A Biol. Sci. Med. Sci. (2011) 66A(10):1131–1136. doi: 10.1093/gerona/glr115
17. Ward KM, Shen J, Souza PE, Grieco-Calub TM. Age-Related Differences in Listening Effort During Degraded Speech Recognition. Ear Hear. (2017) 38(1):74–84. doi: 10.1097/AUD.0000000000000355
18. Panouillères MTN, Möttönen R. Decline of auditory-motor speech processing in older adults with hearing loss. Neurobiol. Aging (2018) 72:89–97. doi: 10.1016/j.neurobiolaging.2018.07.013
19. Rosemann S, Thiel CM. Audio-visual speech processing in age-related hearing loss: Stronger integration and increased frontal lobe recruitment. Neuroimage (2018) 175:425–437. doi: 10.1016/j.neuroimage.2018.04.023
20. Slade K, Plack CJ, Nuttall HE. The Effects of Age-Related Hearing Loss on the Brain and Cognitive Function. Trends Neurosci. (2020) 43(10):810–821. doi: 10.1016/j.tins.2020.07.005
21. Lin FR, Ferrucci L, An Y, Goh JO, Doshi J, Metter EJ, et al. Association of hearing impairment with brain volume changes in older adults. Neuroimage (2014) 90:84–92. doi: 10.1016/j.neuroimage.2013.12.059
22. Eckert MA, Vaden KI, Dubno JR. Age-Related Hearing Loss Associations With Changes in Brain Morphology. Trends Hear. (2019) 23:1–14 doi: 10.1177/2331216519857267
23. Belkhiria C, Vergara RC, San Martín S, Leiva A, Marcenaro B, Martinez M, et al. Cingulate Cortex Atrophy Is Associated With Hearing Loss in Presbycusis With Cochlear Amplifier Dysfunction. Front. Aging Neurosci. (2019) 11:97. doi: 10.3389/fnagi.2019.00097/full
24. Puschmann S, Thiel CM. Changed crossmodal functional connectivity in older adults with hearing loss. Cortex (2017) 86:109–22. doi: 10.1016/j.cortex.2016.10.014
25. Fitzhugh MC, Hemesath A, Schaefer SY, Baxter LC, Rogalsky C. Functional Connectivity of Heschl’s Gyrus Associated With Age-Related Hearing Loss: A Resting-State fMRI Study. Front. Psychol. (2019) 10:2485. doi: 0.3389/fpsyg.2019.02485/full
26. Armada-Moreira A, Gomes JI, Pina CC, Savchak OK, Gonçalves-Ribeiro J, Rei N, et al. Going the Extra (Synaptic) Mile: Excitotoxicity as the Road Toward Neurodegenerative Diseases. Front. Cell Neurosci. (2020) 14:90. doi: 10.3389/fncel.2020.00090
27. Müller VI, Cieslik EC, Laird AR, Fox PT, Radua J, Mataix-Cols D, et al. Ten simple rules for neuroimaging meta-analysis. Neurosci. Biobehav. Rev. (2018) 84:151-161. doi: 10.1016/j.neubiorev.2017.11.012
28. Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int. J. Evid. Based Healthc. (2015) 13(3):141-146. doi: 10.1097/XEB.0000000000000050
29. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst. Rev. (2015) 4(1). doi: 10.1186/2046-4053-4-1
30. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane handbook for systematic reviews of interventions. 6.2. Cochrane (2021) Available from: [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook)
31. Kohl C, McIntosh EJ, Unger S, Haddaway NR, Kecke S, Schiemann J, et al. Online tools supporting the conduct and reporting of systematic reviews and systematic maps: a case study on CADIMA and review of existing tools. Environ. Evid. (2018) 7(8). doi: 10.1186/s13750-018-0115-5
32. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. (2021) 372:n71. doi: 10.1136/bmj.n71
33. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. Syst. Rev. (2012) 1(2) doi: 10.1186/2046-4053-1-2
34. Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. BMC Medical Informatics and Decision Making Utilization of the PICO framework to improve searching PubMed for clinical questions. BMC Med. Inform. Decis. Mak. (2007) 7(16) doi: 10.1186/1472-6947-7-16
35. British Society for Audiology (BSA). Recommended procedure. Pure-tone air-conduction and Bone-conduction threshold audiometry with and without masking [Internet] (2018) [Accessed May 21, 2022] Available from: <https://www.thebsa.org.uk/wp-content/uploads/2018/11/OD104-32-Recommended-Procedure-Pure-Tone-Audiometry-August-2018-FINAL.pdf>
36. World Health Organization (WHO). Report of the Informal Working Group on Prevention of Deafness and Hearing Impairment Programme Planning [Internet] (1991) [Accessed May 21, 2022] Available from: <https://apps.who.int/iris/bitstream/handle/10665/58839/WHO_PDH_91.1.pdf>
37. Anatürk M, Demnitz N, Ebmeier KP, Sexton CE. A systematic review and meta-analysis of structural magnetic resonance imaging studies investigating cognitive and social activity levels in older adults. Neurosci. Biobehav. Rev. (2018) 93:71-84. doi: 10.1016/j.neubiorev.2018.06.012
38. Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). BMJ Open (2016) 6(12):e011458. doi: 10.1136/bmjopen-2016-011458
39. Kmet LM, Lee RC, Cook LS. Standard Quality Assessment Criteria for Evaluating Primary Research Papers. Alberta Heritage Foundation for Medical Research. (2004) doi: 10.7939/R37M04F16
40. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. J. Clin. Epidemiol. (2008) 61(4):344-249. doi: 10.1016/j.jclinepi.2007.11.008
41. Meader N, King K, Llewellyn A, Norman G, Brown J, Rodgers M, et al. A checklist designed to aid consistency and reproducibility of GRADE assessments: development and pilot validation. Syst. Rev. (2014) 3(82). doi: 10.1186/2046-4053-3-82
42. Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill S, et al. Systems for grading the quality of evidence and the strength of recommendations I: Critical appraisal of existing approaches The GRADE Working Group. BMC Health Serv. Res. (2004) 4(1):38. doi: 10.1186/1472-6963-4-38
43. Radua J, Mataix-Cols D, Phillips ML, El-Hage W, Kronhaus DM, Cardoner N, et al. A new meta-analytic method for neuroimaging studies that combines reported peak coordinates and statistical parametric maps. Eur. Psychiatry (2012) 27(8):605-611. doi: 10.1016/j.eurpsy.2011.04.001
44. Ashburner J, Friston KJ. Voxel-Based Morphometry-The Methods. Neuroimage (2000) 11(6):805-821. doi: 10.1006/nimg.2000.0582
45. Albajes-Eizagirre A, Solanes A, Vieta E, Radua J. Voxel-based meta-analysis via permutation of subject images (PSI): Theory and implementation for SDM. Neuroimage (2019) 186:174-84. doi: 10.1016/j.neuroimage.2018.10.077
46. Radua J, Mataix-Cols D. Voxel-wise meta-analysis of grey matter changes in obsessive-compulsive disorder. Br. J. Psychiatry (2009) Nov 195(5):393–402. doi: 10.1192/bjp.bp.108.055046
47. Radua J, Rubia K, Canales-Rodríguez EJ, Pomarol-Clotet E, Fusar-Poli P, Mataix-Cols D. Anisotropic Kernels for Coordinate-Based Meta-Analyses of Neuroimaging Studies. Front. Psychiatry (2014) 5:13. doi: 10.3389/fpsyt.2014.00013
48. Wise T, Radua J, Via E, Cardoner N, Abe O, Adams TM, et al. Common and distinct patterns of grey-matter volume alteration in major depression and bipolar disorder: evidence from voxel-based meta-analysis. Mol. Psychiatry (2017) 22(10):1455-1463. doi: 10.1038/mp.2016.72
49. Hamman EA, Pappalardo P, Bence JR, Peacor SD, Osenberg CW. Bias in meta‐analyses using Hedges’ *d*. Ecosphere (2018) 9(9). doi: 10.1002/ecs2.2419
50. Wasserman S, Hedges L v., Olkin I. Statistical Methods for Meta-Analysis. J. Educ. Stat. (1988) 13(1):75-78. doi: 10.2307/1164953
51. Cooper H, Hedges L v., Valentine JC. The handbook of research synthesis and meta-analysis. 2nd ed. New York: Russell Sage Foundation (2009).
52. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. (2022) Vienna, Austria. <https://www.R-project.org/>
53. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. (2010) 36(3):1-48. doi: 10.18637/jss.v036.i03
54. Trepel M. Neuroanatomie: Struktur und Funktion. Elsevier Health Sciences. (2021)
55. Bowden J, Tierney JF, Copas AJ, Burdett S. Quantifying, displaying and accounting for heterogeneity in the meta-analysis of RCTs using standard and generalised Q statistics. BMC Med. Res. Methodol. (2011) 11(41). doi: 10.1186/1471-2288-11-41
56. Jack CR, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, et al. Hypothetical model of dynamic biomarkers of the Alzheimer’s pathological cascade. Lancet Neurol. (2010) 9(1):119–128. doi: 10.1016/S1474-4422(09)70299-6
57. Belkhiria C, Vergara RC, San Martin S, Leiva A, Martinez M, Marcenaro B, et al. Insula and Amygdala Atrophy Are Associated With Functional Impairment in Subjects With Presbycusis. Front. Aging Neurosci. (2020) 12(102). doi: 10.3389/fnagi.2020.00102
58. Armstrong NM, An Y, Shin JJ, Williams OA, Doshi J, Erus G, et al. Associations between cognitive and brain volume changes in cognitively normal older adults. Neuroimage (2020) 223:117289. doi: 10.1016/j.neuroimage.2020.117289
59. Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, et al. Reproducible brain-wide association studies require thousands of individuals. Nature (2022) 603(7902):654–60. doi: 10.1038/s41586-022-04492-9
60. Belkhiria C, Vergara RC, Martinez M, Delano PH, Delgado C. Neural links between facial emotion recognition and cognitive impairment in presbycusis. Int. J. Geriatr. Psychiatry. (2021) 36(8):1171-8. doi: 10.1002/gps.5501
61. Stephani C, Fernandez-Baca Vaca G, Maciunas R, Koubeissi M, Lüders HO. Functional neuroanatomy of the insular lobe. Brain Struct. Funct. (2011) 216(2):137-49. doi: 10.1007/s00429-010-0296-3
62. Ardila A, Bernal B, Rosselli M. The elusive role of the left temporal pole (BA38) in language: a preliminary meta-analytic connectivity study. Int. J. Neurosci. (2014) 946039:1-7 doi: 10.1155/2014/946039

Table 1. The databases searched and the date on which the search was conducted.

|  |  |
| --- | --- |
| Database | Date of last search\* |
| Academic Search Ultimate | 05 August 2021 |
| CINAHL | 05 August 2021 |
| Embase | 05 August 2021 |
| MEDLINE Complete | 05 August 2021 |
| PsycINFO | 05 August 2021 |
| PubMed | 05 August 2021 |
| Scopus | 10 August 2021† |
| The Cochrane Library | 05 August 2021 |
| Web of Science | 05 August 2021 |

\*Articles published after termination of each search on the respective date were not included. †The final search in Scopus was delayed relative to other databases, as the Scopus search initially retrieved over 30,000 articles. The search was optimised with advice sought from a librarian. Specifically, search terms of tangential relevance, e.g., cognition, that were included in the searches of other databases were omitted in the search of Scopus. Additionally, filters were applied to limit the search to articles in the English language published after 1980, when MRI was increasingly used clinically.

Table 2. Proportional agreement of the inter-rater reliability at each screening stage.

|  |  |
| --- | --- |
| Criteria | Proportional inter-rater agreement |
| Title-Abstract Screening | Overall .95 |
| Criterion 1 | .10 |
| Criterion 2 | .95 |
| Criterion 3 | .92 |
| Criterion 4 | .99 |
| Criterion 5 | .90 |
| Full-text Screening | Overall .87 |
| Criterion 1 | .98 |
| Criterion 2 | .86 |
| Criterion 3 | .83 |
| Criterion 4 | .80 |

Table 3. Means, and standard deviations where applicable, of the data extracted for papers used in the whole-brain analysis.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Study | Design | Sample size (M/F) | | Age  (SD) | | PTA  (SD) | | NH PTA definition | MRI field strength (T) | Critical appraisal score |
| HL | NH | HL | NH | HL | NH |
| Chen et al. (2018) | CS | 22 (10/12) | 23 (11/12) | 63.59 (2.38) | 64.74 (2.65) | 34.54 (4.63) | 14.82 (1.73) | ≤ 25dB at 0.25-8kHz | 1 | .83 |
| Lin et al. (2014) | L | 51 (40/11) | 75 (36/39) | 73.80 (7.30) | 67 (6.90) | N/A | N/A | ≤ 25dB at 0.5-4kHz | 1.5 | .88 |
| Xing et al. (2021) | CS | 40 (19/21) | 40 (18/22) | 63.60 (7.07) | 61.55 (3.72) | 32.69 (3.87) | 16.17 (2.22) | ≤ 25dB at 0.25-8kHz | 1 | .88 |

*Note*: CS refers to cross-sectional and L refers to longitudinal.

Table 4. Means, and standard deviations where applicable, of the data extracted for papers used in the region of interest analyses.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Study | Design | Sample size (M/F) | | Age (SD) | | PTA score (SD) | | NH PTA definition | MRI field strength (T) | Critical appraisal score |
| HL | NH | HL | NH | HL | NH |
| Belkhiria et al. (2021) | CS | 55 (23/32) | 56 (19/37) | 75.38 (5.20) | 72.53 (5.41) | 36.27 (9.50) | 17.08 (4.80) | <25dB at 0.5-4kHz | 3 | .83 |
| Belkhiria et al. (2019) | CS | 33 (12/11) | 31 (6/25) | 73.78 (5.79) | 70.84 (4.84) | 25.68 (4.86) | 14.16 (3.15) | ≤ 20dB at 0.5-4kHz | 1 | .89 |
| Lin et al. (2014) | L | 51 (40/11) | 75 (36/39) | 73.80 (7.30) | 67 (6.90) | N/A | N/A | ≤ 25dB at 0.5-4kHz | 1.5 | .88 |

*Note*. Belkhiria et al. 2021 (60) controlled for education, cognitive abilities, visuospatial capacities, (neuro-)psychiatric symptoms. Belkhiria et al. 2019 (23) controlled for education, cognitive abilities, dementia, smoking, cardiovascular risk factors, diabetes, depression. Lin et al. 2014 (21) controlled for intracranial volume, hypertension, smoking, interactions of time with HL, age and sex. CS refers to cross-sectional and L refers to longitudinal.

Table 5. Available means and standard deviations extracted for meta-analysis of whole-brain data.

|  |  |  |  |
| --- | --- | --- | --- |
| Study | Normalisation space | GM volume (SD) | |
| HL | NH |
| Chen et al. (2018) | MNI | 564.00 (24.40) | 571.20 (20.80) |
| Lin et al. (2014) | MNI | 535.10 (40.99) | 530.30 (39.32) |
| Xing et al. (2021) | MNI | 32.3 (1.80) | 31.6 (1.40) |

Table 6. Available Means and Standard Deviations Extracted for Meta-Analysis of Region of Interest Data

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Study | MRI space | Region | GM volume (SD) | |
| HL | NH |
| Frontal lobe | | | | |
| Belkhiria et al. (2021) | TAL | LH Anterior cingulate | 2.41 (0.37) | 2.42 (0.39) |
| RH Anterior cingulate | 1.84 (0.38) | 1.79 (0.40) |
| LH Orbitofrontal | 7.32 (0.78) | 7.44 (0.72) |
| RH Orbitofrontal | 7.41 (0.80) | 7.48 (0.72) |
| Belkhiria et al. (2019) | TAL | Frontal superior | 4.91 (0.22) | 4.91 (0.18) |
| Anterior cingulate | 4.78 (0.33) | 4.81 (0.26) |
| Precentral gyrus | 4.90 (0.24) | 4.97 (0.24) |
| Lin et al. (2014) | MNI | Frontal lobe | 156.90 (14.76) | 155.10 (14.10) |
| Occipital lobe | | | | |
| Belkhiria et al. (2021) | TAL | LH Fusiform gyrus | 7.25 (1.09) | 7.57 (0.93) |
| RH Fusiform gyrus | 7.35 (1.32) | 7.36 (1.01) |
| LH Lingual gyrus | 5.77 (1.01) | 5.96 (0.87) |
| RH Lingual gyrus | 6.10 (1.00) | 6.05 (0.79) |
| Lin et al. (2014) | MNI | Occipital lobe | 74.20 (8.02) | 75.40 (7.71) |
| Temporal lobe | | | | |
| Belkhiria et al. (2021) | TAL | LH Superior temporal | 14.05 (1.71) | 13.74 (1.28) |
| RH Superior temporal | 13.05 (1.53) | 13.21 (1.13) |
| LH Transverse temporal | 0.90 (0.18) | 0.92 (0.15) |
| RH Transverse temporal | 0.72 (0.13) | 0.75 (0.13) |
| LH Middle temporal | 11.28 (1.56) | 11.29 (1.44) |
| RH Middle temporal | 11.24 (1.53) | 11.43 (1.30) |
| LH Fusiform gyrus | 7.25 (1.09) | 7.57 (0.93) |
| RH Fusiform gyrus | 7.35 (1.32) | 7.36 (1.01) |
| LH Posterior cingulate | 2.85 (0.49) | 2.83 (0.40) |
| RH Posterior cingulate | 2.74 (0.41) | 2.74 (0.48) |
| LH Insula | 5.42 (0.64) | 5.41 (0.56) |
| RH Insula | 5.55 (0.63) | 5.55 (0.51) |
| LH Hippocampus | 3.38 (0.41) | 3.53 (0.37) |
| RH Hippocampus | 3.51 (0.45) | 3.74 (0.38) |
| LH Amygdala | 1.32 (0.22) | 1.40 (0.21) |
| RH Amygdala | 1.54 (0.24) | 1.60 (0.21) |
| Belkhiria et al. (2019) | TAL | Temporal inferior | 5.52 (0.26) | 5.51 (0.19) |
| Temporal middle | 5.31 (0.21) | 5.32 (0.17) |
| Temporal superior | 5.30 (0.25) | 5.33 (0.25) |
| Posterior cingulate | 4.97 (0.25) | 4.99 (0.18) |
| Parahippocampus | 5.33 (0.48) | 5.36 (0.58) |
| Lin et al. (2014) | MNI | Temporal lobe | 114.80 (10.93) | 114.20 (10.36) |
| Parietal lobe | | | | |
| Belkhiria et al. (2019) | TAL | Postcentral gyrus | 3.99 (0.18) | 4.08 (0.21) |
| Lin et al. (2014) | MNI | Parietal lobe | 86.1 (9.47) | 85.5 (9.26) |

*Note*. The original naming convention of regions of studies was maintained. Belkhiria et al. (60) separated ROI data into left and right hemisphere data (LH and RH, respectively). This separation was maintained in analysis. The insula was included in the temporal lobe because although it is covered by both the frontal and temporal lobes (61), it connects strongly to cortical and subcortical structures in the temporal lobe, e.g., the superior temporal sulcus and limbic structures (62). The fusiform gyrus data, reported by Belkhiria et al. (60), were included in both the temporal and occipital lobes, because the gyrus spans across the basal surface of both lobes. Due to a lack of reported coordinates, the regions were allocated to lobes using their verbal labels, commonly accepted allocations, and a neuroanatomy textbook (54).

Table 7. Mean analysis results for whole-brain data.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Study | Hedges' *g* | Standard error | *z*-value | *p*-value | Confidence interval | |
| Low | High |
| Chen et al. (2018) | -0.31 | 0.30 | -1.04 | .30 | -0.90 | 0.28 |
| Lin et al. (2014) | 0.12 | 0.18 | 0.66 | .51 | -0.24 | 0.48 |
| Xing et al. (2021) | 0.43 | 0.23 | 1.90 | .06 | -0.01 | 0.87 |
| Overall effect | 0.12 | 0.19 | 0.64 | .52 | -0.24 | 0.48 |

Table 8. Mean analysis results for region of interest data

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Study | Region | Hedges' *g* | Standard error | *z*-value | *p*-value | CI | |
| Low | High |
| Frontal lobe | | | | | | | |
| Belkhiria et al. (2021) | LH Anterior cingulate | -0.03 | 0.19 | -0.14 | .89 | -0.40 | 0.35 |
| RH Anterior cingulate | 0.13 | 0.19 | 0.67 | .50 | -0.25 | 0.50 |
| LH Orbitofrontal | -0.16 | 0.19 | -0.84 | .40 | -0.53 | 0.21 |
| RH Orbitofrontal | 0.09 | 0.19 | -0.48 | .63 | -0.46 | 0.28 |
| Belkhiria et al. (2019) | Frontal superior | 0.00 | 0.25 | 0.00 | >.99 | -0.49 | 0.49 |
| Anterior cingulate | -0.10 | 0.25 | -0.40 | .69 | -0.59 | 0.39 |
| Precentral gyrus | -0.29 | 0.25 | -1.15 | .25 | -0.78 | 0.21 |
| Lin et al. (2014) | Frontal lobe | 0.13 | 0.18 | 0.69 | .49 | -0.23 | 0.48 |
| Overall effect |  | -0.03 | 0.07 | -0.47 | .64 | -0.18 | 0.11 |
| Temporal lobe | | | | | | | |
| Belkhiria et al. (2021) | LH Superior temporal | 0.20 | 0.19 | 1.07 | .28 | -0.17 | 0.58 |
| RH Superior temporal | -0.12 | 0.19 | -0.62 | .53 | -0.49 | 0.25 |
| LH Transverse temporal | -0.12 | 0.19 | -0.63 | .53 | -0.49 | 0.25 |
| RH Transverse temporal | -0.23 | 0.19 | -1.20 | .23 | -0.60 | 0.14 |
| LH Middle temporal | -0.01 | 0.19 | -0.04 | .97 | -0.38 | 0.37 |
| RH Middle temporal | -0.13 | 0.19 | -0.70 | .48 | -0.51 | 0.24 |
| LH Fusiform gyrus | -0.13 | 0.19 | -1.64 | .10 | -0.69 | 0.06 |
| RH Fusiform gyrus | -0.01 | 0.19 | -0.05 | .96 | -0.38 | 0.36 |
| LH Posterior cingulate | 0.04 | 0.19 | 0.23 | .81 | -0.33 | 0.42 |
| RH Posterior cingulate | 0.00 | 0.19 | 0.00 | >.99 | -0.37 | 0.37 |
| LH Insula | 0.02 | 0.19 | 0.09 | .93 | -0.36 | 0.39 |
| RH Insula | 0.00 | 0.19 | 0.00 | >.99 | -0.37 | 0.37 |
| LH Hippocampus | -0.38 | 0.19 | -1.20 | .05 | -0.76 | -0.01 |
| RH Hippocampus | -0.55 | 0.19 | 2.84 | .01 | -0.93 | -0.17 |
| LH Amygdala | -0.37 | 0.19 | -1.93 | .05 | -0.75 | 0.01 |
| RH Amygdala | -0.26 | 0.19 | -1.39 | .17 | -0.64 | 0.11 |
| Belkhiria et al. (2019) | Temporal inferior | 0.04 | 0.25 | 0.17 | .86 | -0.45 | 0.53 |
|  | Temporal middle | -0.05 | 0.25 | -0.21 | .84 | -0.54 | 0.44 |
|  | Temporal superior | -0.12 | 0.25 | -0.47 | .64 | -0.61 | 0.37 |
|  | Posterior cingulate | 0.09 | 0.25 | -0.36 | .72 | -0.58 | 0.40 |
|  | Parahippocampus | -0.06 | 0.25 | -0.22 | .82 | -0.55 | 0.43 |
| Lin et al. (2014) | Temporal lobe | 0.06 | 0.18 | 0.31 | .76 | -0.30 | 0.41 |
| Overall effect |  | -0.12 | 0.04 | -2.70 | .01 | -0.20 | -0.03 |
| Parietal lobe | | | | | | | |
| Belkhiria et al. (2019) | Postcentral gyrus | -0.46 | 0.25 | -1.80 | .07 | -0,95 | 0.04 |
| Lin et al. (2014) | Parietal lobe | 0.06 | 0.18 | 0.35 | .73 | -0.29 | 0.42 |
| Overall effect |  | -0.17 | 0.26 | -0.64 | .52 | -0.67 | 0.34 |
| Occipital lobe | | | | | | | |
| Belkhiria et al. (2021) | LH Fusiform gyrus | -0.31 | 0.19 | -1.64 | .10 | -0.69 | 0.06 |
| RH Fusiform gyrus | -0.01 | 0.19 | -0.05 | .96 | -0.38 | 0.36 |
| LH Lingual gyrus | -0.20 | 0.19 | -1.05 | .29 | -0.57 | 0.17 |
| RH Lingual gyrus | 0.06 | 0.19 | 0.29 | .77 | -0.32 | 0.43 |
| Lin et al. (2014) | Occipital lobe | -0.15 | 0.18 | -0.84 | .40 | -0.51 | 0.20 |
| Overall effect |  | -0.12 | 0.08 | -1.47 | .14 | -0.29 | 0.04 |
| *Note*. LH and RH refer to left and right hemisphere, respectively. The insula was included in the temporal lobe, and the fusiform gyrus was included in both the occipital and temporal lobes (see note in table 6 for explanations). | | | | | | | |

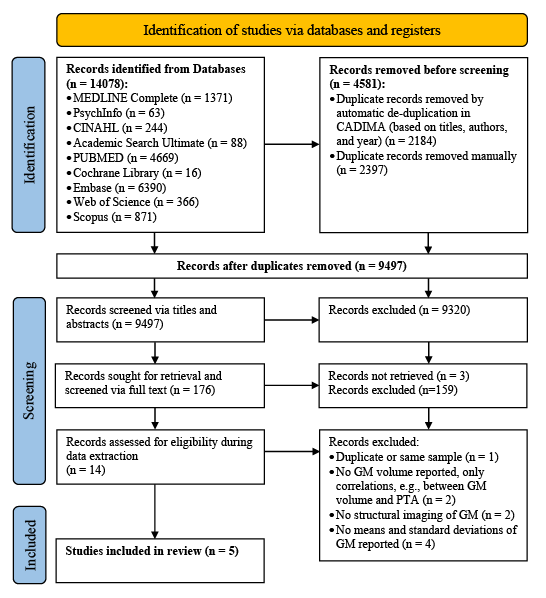


Figure 1. PRISMA flow diagram detailing the number of articles selected at each review stage. Note: GM refers to grey matter.

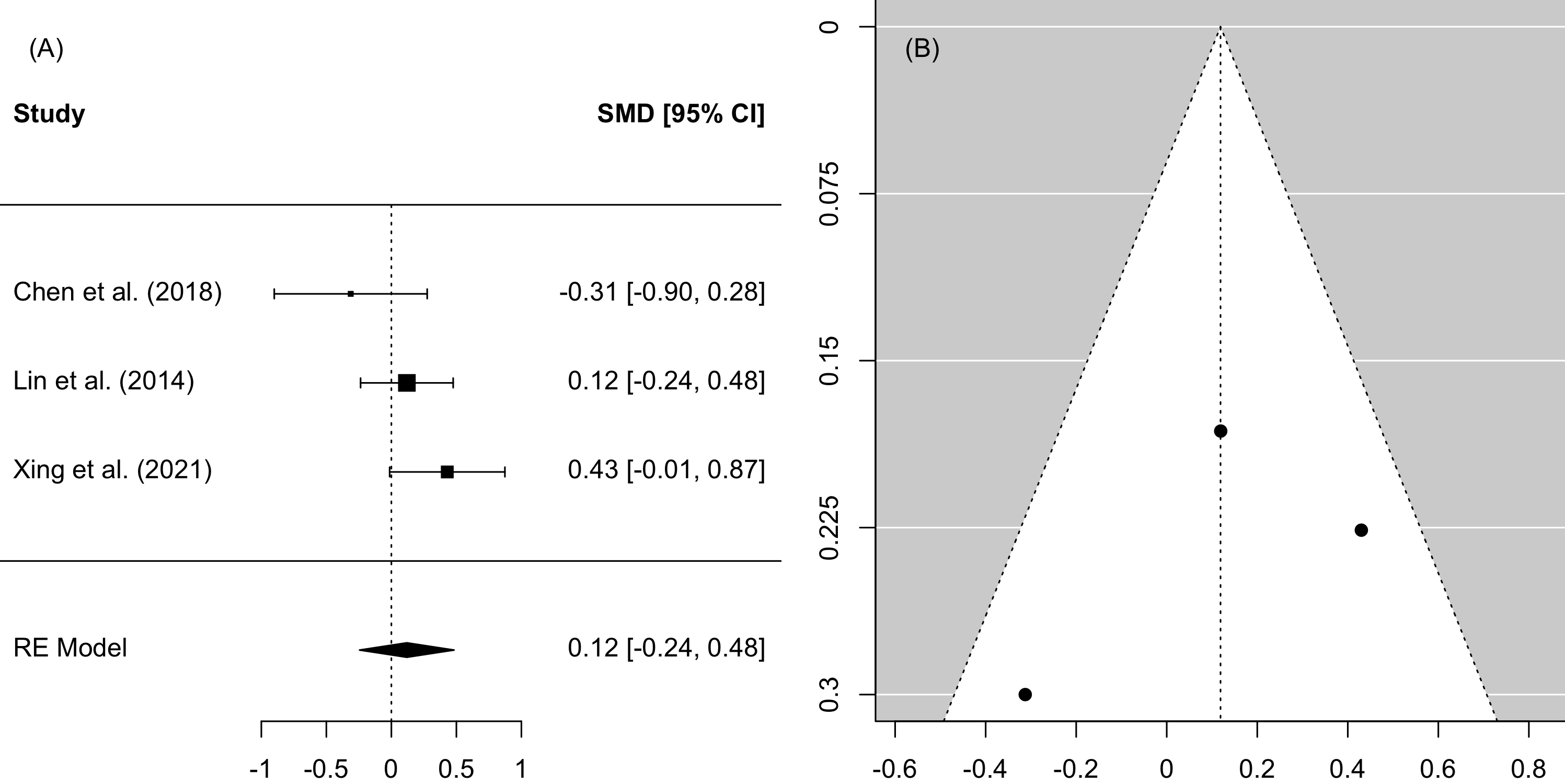


Figure 2. Forest plot (A) and funnel plot (B) of the whole-brain Analysis. Note: In the forest plot (A) negative values on the x-axis indicate grey matter atrophy.

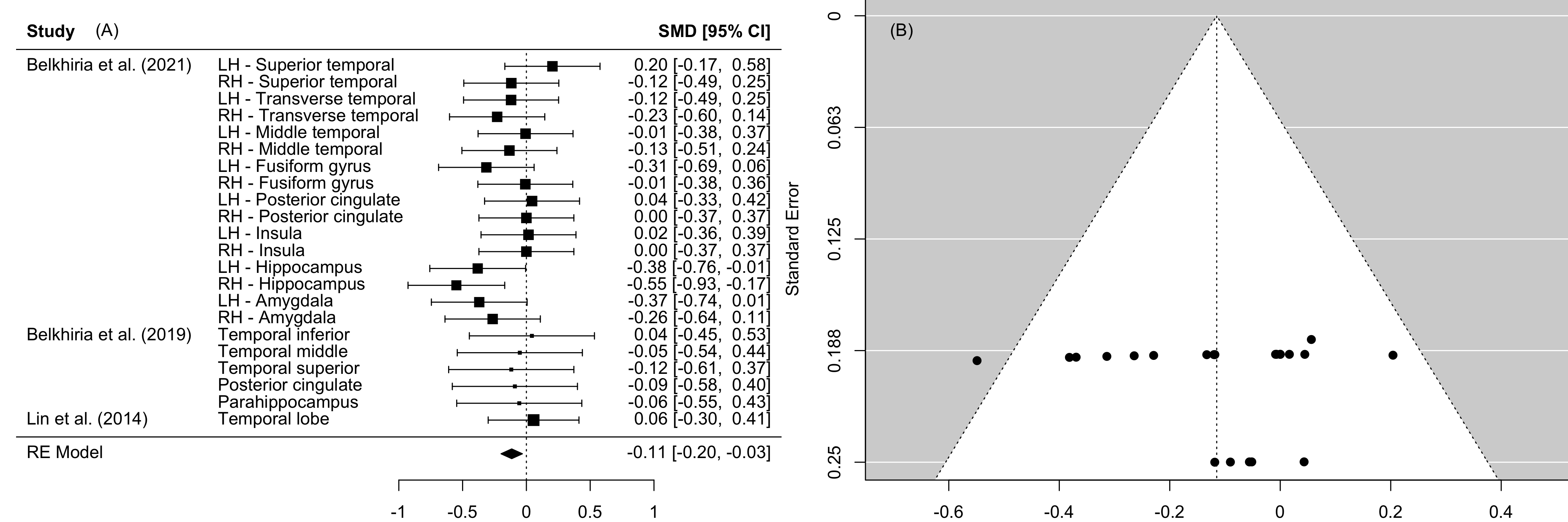


Figure 3. Forest plot (A) and funnel plot (B) of the temporal lobe analysis. Note: In the forest plot (A) negative values on the x-axis indicate grey matter atrophy.

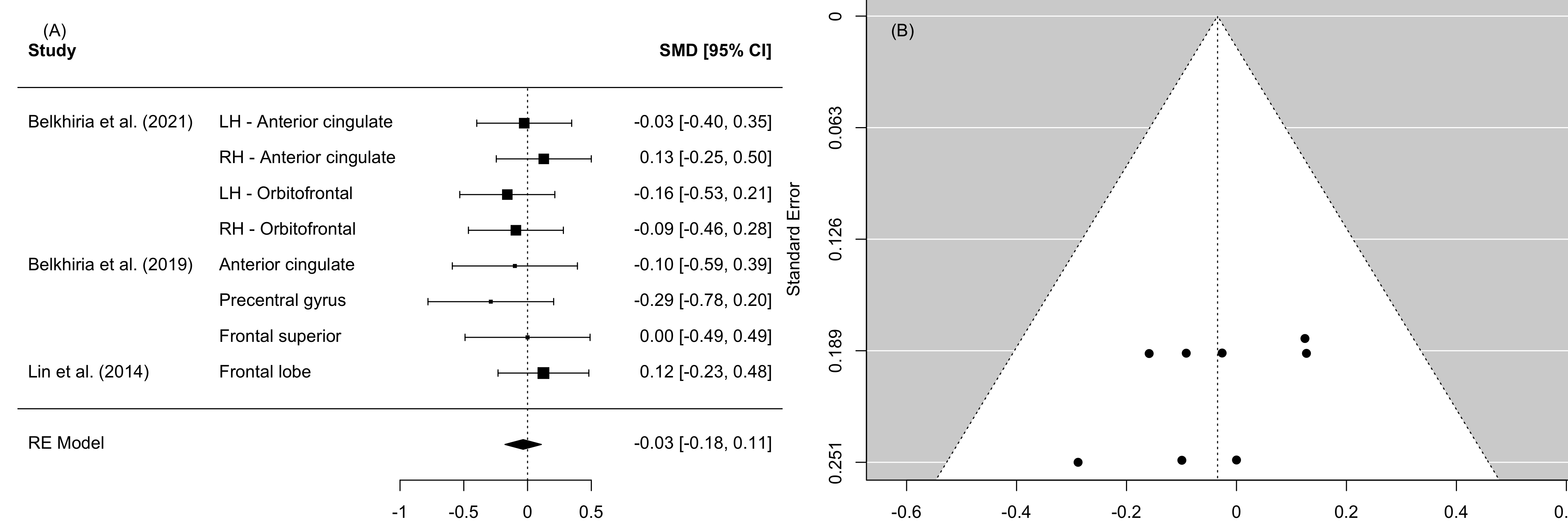


Figure 4. Forest plot (A) and funnel plot (B) of the frontal lobe analysis. Note: In the forest plot (A) negative values on the x-axis indicate grey matter atrophy.

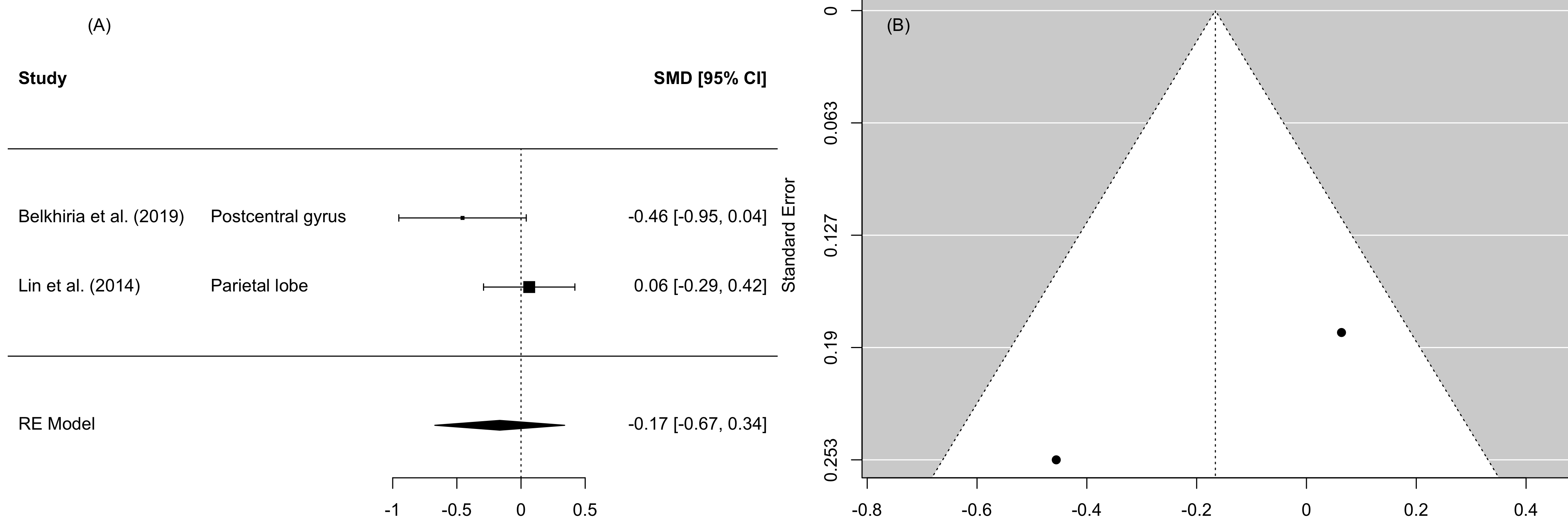


Figure 5. Forest plot (A) and funnel plot (B) of the parietal lobe analysis. Note: In the forest plot (A) negative values on the x-axis indicate grey matter atrophy.

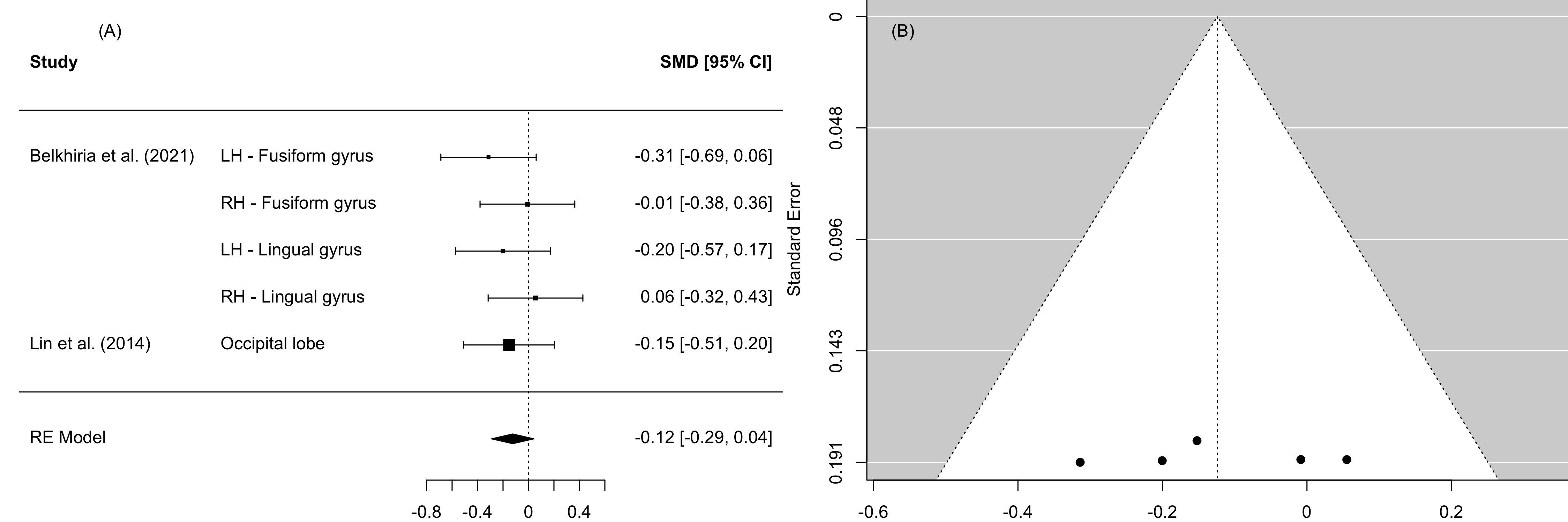


Figure 6. Forest plot (A) and funnel plot (B) of the occipital lobe analysis. Note: In the forest plot (A) negative values on the x-axis indicate grey matter atrophy.