



OPEN Optimizing spatial normalization of multisubject inner ear MRI: comparison of different geometry-preserving co-registration approaches

Johannes Gerb^{1,3}✉, Valerie Kirsch^{1,2,3}, Emilie Kierig^{1,3}, Thomas Brandt^{1,2}, Marianne Dieterich^{1,2,3,4} & Rainer Boegle^{1,2}

Spatial normalization of multisubject inner ear imaging data is challenging, due to both substantial intraindividual differences and the small size of the organ compared to other intracranial structures. Automatic whole brain co-registration to standard space can only roughly co-align the peripheral vestibular endorgan, and complementary manual registration is highly time-consuming. Here, we compared the accuracy of four geometry-maintaining co-registration methods (one semi-manual method and three automatic methods). High-resolution structural T2-MRI of 153 inner ears from patients and healthy participants were co-registered to an inner-ear atlas. The semi-manual method used a three-point landmark-based approach (3P), two automatic methods were based on unassisted standard algorithms (Advanced Normalization Tools (ANTs), Elastix (EL)), while the fourth automatic method utilized a volumetrically dilated, atlas-based mask (thick inner ear, TIE) for probabilistic inner ear masking. Registration accuracy was evaluated by neurologists blinded to the respective registration paradigm, and the resulting median volumes were quantified using colocalization analyses. The mask-aided automatic approach showed the best ratings, followed by the semi-manual three-point landmark-based registration (mean ratings (lower: better) TIE 2.21 ± 1.15 ; 3P 2.58 ± 0.61 ; EL 3.42 ± 1.06 ; ANTs 3.49 ± 1.26). The semi-manual method had the lowest rate of insufficient registrations, followed by TIE (3P: 3.70%; TIE: 8.28%; EL: 22.66%; ANTs: 27.02%). TIE showed the highest colocalization metrics with the atlas. Only TIE and 3P allowed for sufficient semicircular canal visualization in method-wise average volumes. Overall, geometry-preserving spatial normalization of multisubject inner ear imaging data is possible and could allow groupwise examinations of the bony labyrinth or temporal bone morphology in the future.

Keywords Endolymphatic hydrops, Vertigo, Peripheral vestibular system, Inner ear, Spatial normalization, Atlas

Inner ear anatomy is known to be highly variable^{1,2}. Already in healthy participants, diameter and orientation of the semicircular canals (SCC) vary widely, sometimes even intraindividually^{1,3}. Disorders can involve SCC, the vestibular aqueduct (VA) or cochlea^{4,5}. No clear correlation between this anatomical variance of the bony labyrinth and peripheral vestibular function has been reported thus far, although an influence of temporal bone anatomy on caloric nystagmus intensity has been postulated⁶, and a relevant impact on, e.g., therapeutical effectiveness of positioning maneuvers in benign paroxysmal positional vertigo (BPPV) seems probable⁷. Systematic analysis, however, has been hindered by the lack of standardization in inner ear imaging data analysis.

Multiple approaches of automatic inner ear segmentation, based on MRI^{8,9} or CT scans¹⁰, have been proposed. Typically, they show good to very good segmentation performance (as evaluated by, e.g., the Sørensen–Dice

¹German Center for Vertigo and Balance Disorders (DSGZ), LMU University Hospital, Ludwig-Maximilians-University (LMU) Munich, Marchioninistraße 15, 81377 Munich, Germany. ²Graduate School of Systemic Neuroscience, LMU Munich, Munich, Germany. ³Department of Neurology, LMU University Hospital, LMU Munich, Munich, Germany. ⁴Munich Cluster for Systems Neurology (SyNergy), Munich, Germany. ✉email: Johannes.Gerb@med.uni-muenchen.de

coefficient¹¹) and work with little to no user input. Importantly, these methods only work on individual datasets, and do not allow for groupwise investigations. Co-registration of multisubject or multisubject/multimodal data to a common space, i.e., a shared coordinate system, is usually not performed. This need for standardized spatial coordinates, which are crucial to allow for groupwise image-based analysis, has been recently recognized by Copson et al.¹². The potential implications for vestibular research become even more evident when considering recent advances in Menière's disease (MD) research where petrous bone anatomy has been shown to correlate with disease progression¹³. Furthermore, while delayed inner ear imaging for endolymphatic hydrops (ELH) quantification has been used for many years^{14,15}, only recently has the focus shifted to extra-labyrinthine structures like the endolymphatic duct^{16–18} or the endolymphatic sac¹⁹. Omitting petrous bone tissue from MRI analysis pipelines (as is commonly done in automatic segmentation) might subsequently miss relevant information which could aid in differential diagnosis^{16,20}.

Multiple approaches aiming at fast and accurate rigid co-registration of inner ear imaging datasets have been proposed^{21,22}, often with a focus on cochlear structures in order to facilitate cochlear implant (CI) surgery^{23,24}. However, no consensus between different laboratories exists on the optimal preprocessing procedures.

In other fields of neuroimaging, the definition of common imaging coordinate systems (e.g., by the Montreal Neurological Institute (MNI)^{25,26} or in Talairach space^{27,28}) have been standard for decades and paved the way for exact colocalization of anatomical structures with, e.g., functional imaging data or cytoarchitectonic atlases, leading to increased understanding of physiological and disease-related brain function^{29,30}. This step enables group analyses detached from individual data variability and allowed the development of techniques such as voxel-based morphometry (VBM³¹) and the creation of generalized human connectomes, e.g., in the Human Connectome Project (HCP³²). For all of these data analysis pipelines with exact colocalization are crucial^{33,34}.

Recently, a freely-available atlas for neurotological research has been introduced by Ahmadi et al.³⁵, which can be used as a standard coordinate system. With the end-goal of developing VBM-similar techniques for the field of neurotology, reliable methods of coregistration of multisubject data are required. Given the anatomical variance of the structures involved, whole brain coregistration, however, does not result in usable shared coordinates for detailed inner ear anatomy (Fig. 1). Furthermore, automatic registration tools for whole-brain images or cropped petrous bone images methodically often result in misaligned inner ear structures for a simple reason. Since automatic coregistration algorithms work by (1) randomly selecting a given number of voxels, (2) calculating a metric for image similarity (e.g., Dice coefficient), and then (3) performing 2D or 3D transformation operations to optimize this metric until a predefined cutoff value for convergence is reached, small structures tend to be underrepresented. Inherently, voxels from larger structures in a given dataset will statistically be selected more commonly, leading to better alignment of large structures, but potential misalignment of smaller structures. This becomes even more problematic if anatomical structures show large variance, as the inner ear does. These challenges can be mitigated by introducing masks of relevant image areas. This approach has been previously demonstrated for brainstem coregistration³⁶.

To fully understand potential disease-specific differences in peripheral vestibular disorders outside of overall inner ear anatomy variance, precise, inner ear-centered coregistration to a common space could therefore enable new analysis methods. Additionally, spatial normalization can aid in existing analyses, e.g., in semiquantitative ELH grading, by allowing exact slice selection as specified in the respective grading system (e.g., midmodiolar level for most grading conventions^{37–39}). Importantly, the registration algorithm needs to be adapted to the specific task since all methods expect rigid paradigms (which only allow translation and rotation around the x-, y- and z-axis, but forbid other image alteration) inherently alter the dataset's geometry. Paradigms with higher degrees of freedom (e.g., allowing for shearing or scaling) or non-linear deformations might result in higher coregistration overlap scores, but ultimately omit geometrical information and render measurements of geometrical petrous bone anatomical features impossible.

In this methodological investigation, we will therefore compare the accuracy on pre-cropped inner ear MRI T2-dataset coregistration from two automatic registration tools with a mask-aided automatic registration and a simple manual registration using anatomical landmarks. The inner ear datasets stem from participants who underwent MRI for diagnosis or exclusion of ELH and the data has been analyzed in previous studies^{15,40,41}.

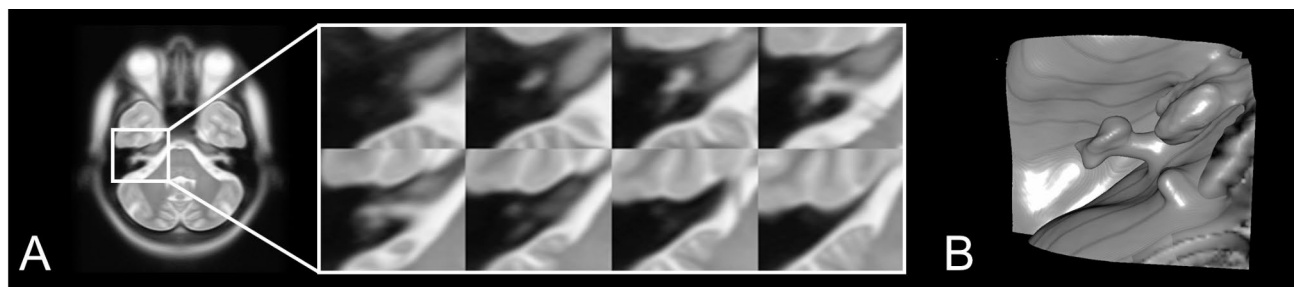


Fig. 1. Visualization of the inner ear currently available in the MNI152²⁰, based on whole-brain coregistration from 152 healthy adults. While the resulting atlas has been used in countless neuroimaging studies as a standard coordinate system, the inner ear anatomy is not depicted in a sufficient resolution. **(A)** Overview of the MNI high resolution T2 atlas and zoom into the inner ear region. **(B)** 3D-rendering of the inner-ear anatomy available in the MNI152.

Here, those datasets are exclusively used for a methodological investigation in the coregistration performance of T2 high-resolution inner ear MRI. While our primary goal was to determine which method results in the best alignment, our secondary goal was to show how even subpar inner ear registration into a common space still constitutes a significant step towards better data comparability and should be considered when performing inner ear imaging research.

Methods

Setting and institutional review board approval

All data was acquired at the interdisciplinary German Center for Vertigo and Balance Disorders (DSGZ) and the Department of Neurology of LMU University Hospital, Munich, between 2016 and 2019. The data protection clearance and Institutional Review Board of the Ludwig-Maximilians-University, Munich, Germany approved the study (no. 641-15). All participants provided informed oral and written consent before inclusion in the study. The study was conducted in accordance with the Declaration of Helsinki and its later amendments.

Study population

Inner ear datasets from patients with various neurotological disorders who underwent delayed, contrast-enhanced inner ear MRI (iMRI) for exclusion or verification of ELH, were included. Additionally, inner ear datasets from voluntary participants without any history of vestibular disorders were included. These participants underwent contrast MRI for clinical reasons and had previously agreed to perform a second inner ear MRI four hours after the clinical MRI for hydrops evaluation. The inclusion criteria were age between 18 and 85 years and normal kidney function. The exclusion criteria were other neurological or psychiatric disorders, as well as any MR-related contraindications⁴², poor image quality, or missing MR sequences.

Delayed intravenous gadolinium-enhanced MRI of the inner ear

Data acquisition

Four hours after intravenous injection of a standard dose (0.1–0.2 mmol/kg body weight, i.e., 0.1–0.1 mmol/kg body weight) of gadobutrol (Gadovist®, Bayer, Leverkusen, Germany), MR imaging (MRI) data were acquired in a whole-body 3 Tesla MRI scanner (Magnetom Skyra, Siemens Healthcare, Erlangen, Germany) with a 20-channel head coil. In this study, only the high-resolution, T2-weighted, spin-echo three-dimensional sampling perfection with application-optimized contrasts by using different flip angle evolutions (3D-SPACE) sequence of the temporal bones was used (TE 133 ms, TR 1000 ms, FA 100°, FOV 192 × 192 mm², 56 slices, base resolution 384, averages 4, acceleration factor of 2 using GRAPPA algorithm, 0.5 mm slice thickness, 0.5 × 0.5 × 0.5 mm³ spatial resolution).

Data processing

We used ImageJ/FIJI⁴³ to select and rescale (two-fold in x-, y- and z-direction, interpolation: bicubic) a rectangular ROI containing both left and right inner ear from the SPACE dataset. Left inner ear datasets were mirrored with respect to the sagittal plane to match the orientation of the right inner ear. All registrations were performed in 3D-Slicer⁴⁴ and the respective implementations of ANTs (Advanced Normalization Tools)^{45,46} and Elastix (EL)^{47,48} were used for registration. Both ANTs and EL are commonly used co-registration toolboxes⁴⁹. For all datasets, the T2 atlas from Ahmadi et al.³⁵ was used as a fixed volume (Fig. 2).

For the ANTs-registration, the following parameters were used: Stages: Rigid; initial transform: geometric image center; output interpolation: linear, computation precision: float. For EL, we used the preset “generic (rigid) all”. The TIE approach was performed with the Elastix plug-in using a volumetrically dilated (10px) inner ear mask (mask created in ImageJ/FIJI) defining the region of interest (ROI). The landmark-based approach utilized the “Fiducial Registration”-module to create a transformation matrix which was then hardened using the “Resample image (BRAIN)”-module (interpolation: linear). The landmarks were (1) the outmost (most lateral) part of the posterior SCC, (2) the upmost (most superior) part of the superior SCC under the arcuate eminence and (3) the apex cochleae. All landmarks were selected manually by one of the authors (JG). Landmark selection took roughly 15–20 s per inner ear.

Subjective grading of coregistration accuracy

All resulting volumes were then given randomized filenames and rated by three of the authors (EK, RB, JG) blinded to the respective registration method used. Method identification was not possible from image characteristics alone. For registration comparison, we used the “Compare Volumes”-module from 3D-Slicer with the atlas as background and the respective registrations as foreground. All raters had multiple years of experience in inner ear imaging and rated each registration on a 6-point Likert scale (1 = best possible alignment, 6 = no alignment/non-convergence). This 6-point grading system is identical to the one being used in German schools and all raters were highly familiar with it. In unclear cases, intermittent half-steps were allowed (e.g., a grade of 2.5 was permitted when the coregistration was subjectively more accurate than other cases graded as 3, but not as accurate as cases graded as 2). Raters were previously instructed to evaluate inner ear geometry rather than, e.g., brainstem coregistration results.

Group averages were created using ImageJ/FIJI⁴³ by importing all method-specific volumes as a 4D-hyperstack and then using a slice-wise z-projection for median image intensity. In order to determine image similarities, we used the “Colocalization-Threshold”-plugin in ImageJ (thresholding via Costes auto-threshold⁵⁰, correlation with Spearman's Rho and Manders split⁵¹). To remove false-positive colocalization metrics from background noise or, e.g., brainstem tissue, a dilated inner ear mask was used to define the region of interest (ROI).

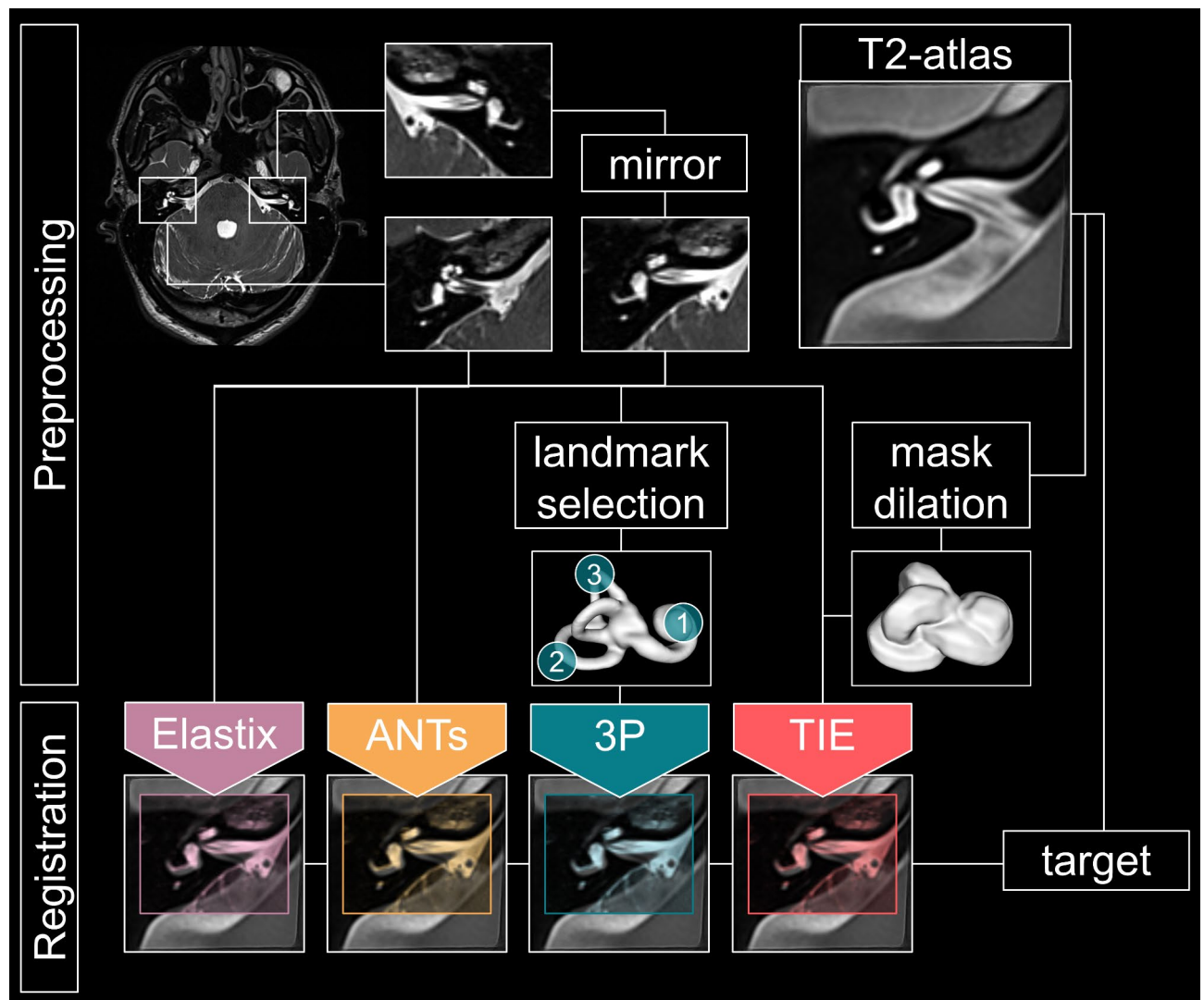


Fig. 2. Overview of the preprocessing, registration, and analysis steps for an individual dataset. After selecting both inner ears from the raw T2-SPACE, the left inner ear is mirrored along the sagittal plane to match the orientation of the T2-atlas. For every inner ear crop, two unaided automatic registrations (Elastix, ANTs), one semimanual registration (3P; using three landmarks: (1) apex cochlea, (2) lateral pSCC, (3) superior sSCC), and one mask-aided approach (TIE) were performed (target volume: atlas). While the landmark selection for the 3P-approach was done for every individual dataset, the atlas-derived dilated mask was only created once and used for all subsequent registrations. Registration accuracy was then rated by three experienced neurologists blinded to the method used. Furthermore, statistical colocalization analysis on the masked registration was performed (not depicted). *ANTs* advanced normalization tools, *3P* three point alignment, *TIE* thick inner ear mask coregistration.

Statistical analysis

We used JASP (version 0.18.3, <https://jasp-stats.org/>) to calculate mean method-wise scores per rater and performed ANOVA-testing to compare the average scores per registration method (post-hoc comparisons Bonferroni corrected). For dichotomic variables, Welch's t-test was used. Inter-rater-agreement was assessed using Spearman's rho due to the potentially different baseline ratings per rater which would have falsely influenced e.g., Cohen's kappa⁵². Additionally, Kendall's tau was calculated in order to be more sensitive regarding tied ranks in the 6-point rating scale.

Results

107 inner ear datasets from 54 individual patients with various neurotological disorders (mean age 51.3 ± 16.1 years, 29 females) and 46 inner ear datasets from 23 participants without any history of vestibular disorders (mean age 40.0 ± 17.2 years, 13 females) were included, resulting in a total of 153 inner ear datasets. Demographic characteristics of the included inner ears were comparable when divided by patient sex (mean age female inner

ears ($n=84$): 47.6 ± 18.3 years, mean age male inner ears ($n=69$): 48.3 ± 16.0 years; Welch's t -test of groupwise mean age not significant: t 0.55, p 0.59).

The average ratings (with lower scores representing higher subjective accuracy) were 2.21 ± 1.15 for the mask-aided TIE approach, 2.58 ± 0.61 for the semimanual three-point approach, 3.42 ± 1.06 for EL, and 3.49 ± 1.26 for ANTs (Fig. 3). ANOVA testing showed a highly significant effect of registration method on the mean rating (F 3,608 = 55.73, $p < 0.001$), with post-hoc comparison revealing clear differences between the unassisted automatic methods (ANTs, EL) and both the mask-aided and landmark-based approach (ANTs vs. TIE: mean difference 1.28, Cohen's d 1.22, p Bonf < 0.001***; ANTs vs. 3P: mean difference 0.91, Cohen's d 0.87, p Bonf < 0.001***; EL vs. TIE: mean difference 1.21, Cohen's d 1.16, p Bonf < 0.001***; EL vs. 3P: mean difference 0.85, Cohen's d 0.81, p Bonf < 0.001***). Between TIE and 3P, TIE showed significantly better ratings (mean difference 0.37, Cohen's d 0.35, p Bonf 0.01*). No significant difference was found between ANTs and EL (mean difference - 0.07, Cohen's d - 0.06, p Bonf 1.00, all p values adjusted for comparing a family of 4).

Due to the subjective nature of the rating system, the mean scores given by the three raters diverged around individual baselines (EK: 2.09 ± 1.36 , RB: 3.64 ± 1.31 , JG: 3.05 ± 1.27). However, all raters independently rated the TIE approach the best, followed by 3P (mean score TIE: 1.49 ± 1.31 (EK), 2.78 ± 1.20 (RB), 2.31 ± 1.28 (JG); mean score 3P: 1.70 ± 0.72 (EK), 3.35 ± 0.95 (RB), 2.64 ± 0.77 (JG); mean score EL: 2.48 ± 1.32 (EK), 4.27 ± 1.17 (RB), 3.43 ± 1.17 (JG); mean score ANTs: 2.69 ± 1.54 (EK), 4.16 ± 1.31 (RB), 3.64 ± 1.35 (JG); Fig. 3). The overall inter-rater agreement was good, as determined by a correlation analysis using Spearman's rho (EK/JG: rho 0.65, $p < 0.001$; EK/RB: rho 0.69, $p < 0.001$; RB/JG rho 0.71, $p < 0.001$) and Kendall's tau (EK/JG: tau 0.58, $p < 0.001$; EK/RB: tau 0.59, $p < 0.001$; RB/JG tau 0.60, $p < 0.001$).

The number of total insufficient registration outcomes (5 and 6 on the 6-point Likert scale, from all three independent raters combined, i.e., from all 612 co-registrations) were $n = 17$ for the semimanual approach, $n = 38$ for TIE, $n = 104$ for EL, and $n = 124$ for ANTs. Raters diverged in their overall number of registrations rated as insufficient (JG $n = 91$, EK $n = 59$, RB $n = 133$). However, both in grouped analysis as well as in rater-wise analysis,

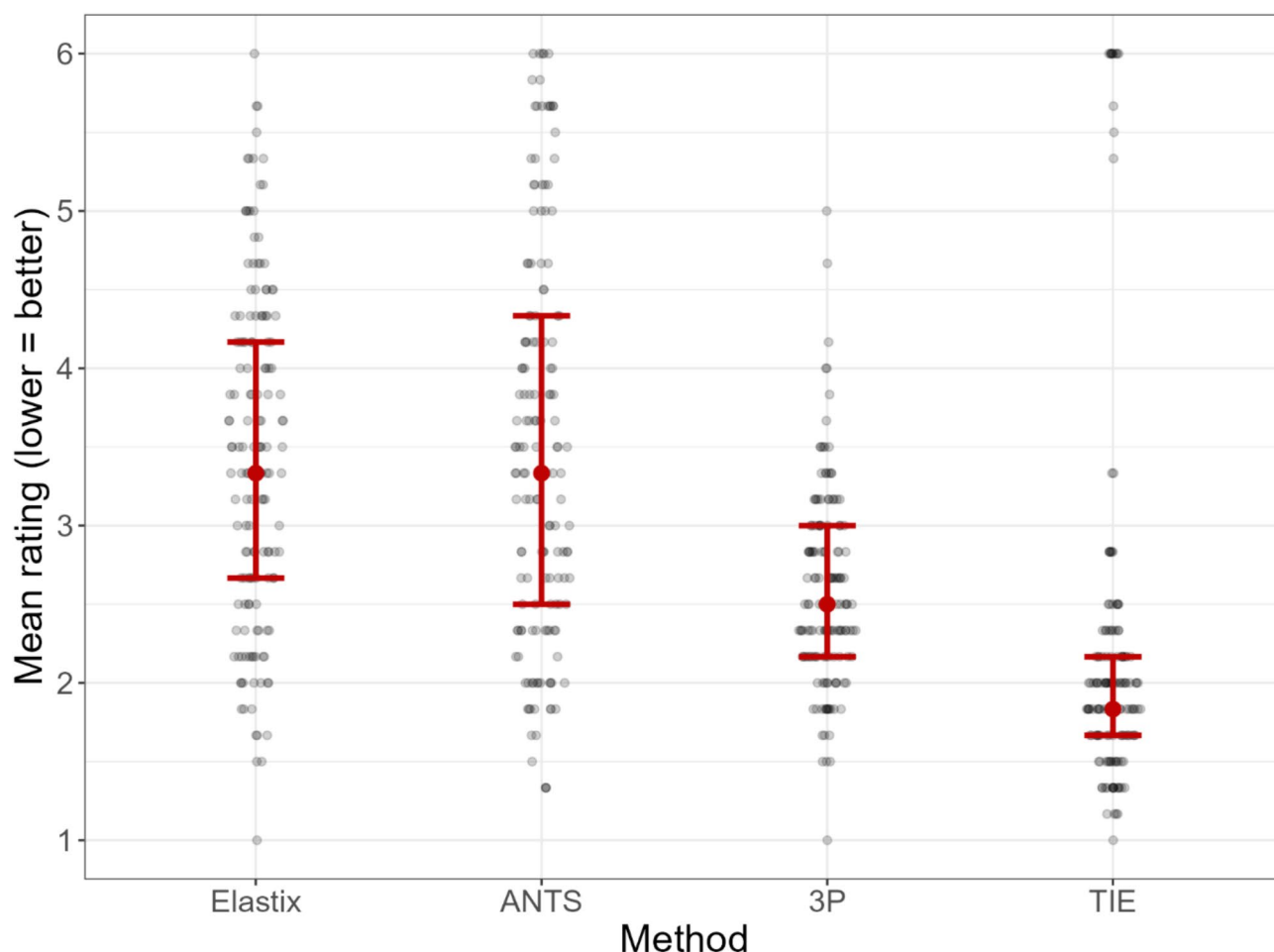


Fig. 3. Average ratings of three independent raters, blinded to the method (1 = best possible grade, 6 = worst possible grade; lower average scores equal to better performance). On average, TIE showed the best ratings; however, in some cases, non-convergence occurred. 3P results in slightly less accurate coregistration with almost no cases of non-convergence. Both automatic methods (ANTs, Elastix) showed significantly worse scores. ANTs advanced normalization tools, 3P three point alignment, TIE thick inner ear mask coregistration.

the semimanual approach (3P) had the lowest rate of non-convergence, followed by TIE (3P: 3.70%; TIE: 8.28%; EL: 22.66%; ANTs: 27.02%).

No effect of participant sex on the overall registration accuracy was found (Welch's t-test: n.s.). ANOVA-testing with Bonferroni-corrected post-hoc comparison additionally ruled out method-specific sex effects (TIE mean sex difference: 0.08, pBonf 1.00; 3P mean sex difference: 0.09, pBonf 1.00; EL mean sex difference: 0.13, pBonf 1.00; ANTs mean sex difference: 0.34, pBonf 1.00).

Visual analysis of the group-medians showed the highest overlap with the atlas in the TIE-approach and the 3P-method (Fig. 4). The SCCs could not be clearly separated from background noise in group-medians from the unaided registration algorithms (EL, ANTs; Fig. 4).

The colocalization analysis between median method-specific volumes and the source atlas showed the highest correlation scores (Spearman's rho) for the TIE-approach (TIE: 0.931; 3P: 0.780, EL 0.761, ANTs 0.777). The thresholded Mander-Splits (automatic thresholding into two channels) were the highest for the TIE-approach (TIE Channel 1: 0.9796, Channel 2: 0.9530; 3P Channel 1: 0.9926, Channel 2: 0.9170; ANTs Channel 1: 0.8730, Channel 2: 0.7730; EL Channel 1: 0.8092, Channel 2: 0.7546, cp. Fig. 5). A simple linear regression model showed the most ideal fit (i.e., an estimated slope approximating 1) for the TIE approach, followed by 3p (TIE: $y = 0.829x + 0.0$; 3p: $y = 0.777x + 0.1$; ANTs: $y = 0.581x + 0.2$; EL: $y = 0.575x - 0.6$).

Discussion

In this methodological investigation, it was shown that despite anatomical variance, automated and fast spatial normalization of multi-subject inner ear imaging datasets to a common space is feasible. From the automatic algorithms, the mask-aided approach using an atlas-based dilated inner ear mask (TIE) significantly outperformed the unaided methods investigated in this study and showed the highest scores overall, both in subjective rating scores and in quantitative colocalization metrics. As a slower alternative, semimanual landmark based inner ear registration using only three fiducials (3P) showed the second-highest ratings, and had the lowest rate of non-convergence. All methods investigated resulted in clearer defined group-wise inner ears than whole-brain coregistrations (Figs. 1 and 4), and would allow basic group-wise atlas-based morphometrical analyses. However, finer inner ear structures such as the SCCs were only discernible in TIE or 3P.

Notably, the subjective rating scores showed no sex differences, indicating an unbiased performance. Especially when considering known sex differences in the vestibular system^{53–55}, methodical impartialness is a necessary prerequisite for the analysis of potential sex-specific differences in inner ear anatomy after establishing a stable multi-subject coregistration method.

The comparative evaluation of the current study has some limitations. Firstly, the subjective rating scale to determine the registration performances is intrinsically subjective, despite raters being blinded to the method used. As expected, this resulted in different average scores per rater, with some raters giving better grades overall than others. However, the ranked inter-rater agreement was very high (Spearman's rank correlation, Kendall's Tau), and all raters agreed on the superior performance of the TIE approach, followed by 3P, as could be shown in the subjective ratings. While a correlation of subjective and objective coregistration scores would be desirable, conventional colocalization metrics such as the Dice-coefficient suffer from the same methodological drawbacks that make the development of inner-ear centered registration approaches necessary in the first place. Since larger structures tend to be overrepresented in Dice-score calculation, inadequate inner ear registrations can result in higher Dice-scores than adequate ones, if e.g., brainstem or petrous bone tissue coalignment is higher. For this reason, we utilized in the current study colocalization quantification metrics which do not suffer from these limitations. These unbiased metrics (automated thresholding and voxelwise correlation) confirmed the subjective scores, with the TIE-approach and the semimanual method showing the highest performance.

A second potential limitation could be that the landmark selections for all 3P datasets were done by a single author (JG). Here, further research is required to quantify intra-observer and inter-observer variability during landmark selection, and their potential effects on the final registration.

This study did not include major congenital inner ear malformations. All inner ears were from patients and healthy participants with overall normal inner ear anatomy. For such patients, a modified spatial template might be necessary, and the 3P-approach might not be applicable.

All patients included in this study underwent standardized MR imaging with predefined sequences, acquired in a single center on the same scanner unit. However, it should be noted that all datasets were acquired as part of the clinical routine, rather than as part of an elaborate, extensive neuroimaging study: the scanner unit used in this study is a widely available model (Magnetom Skyra, Siemens Healthcare, Erlangen, Germany), and the employed MR sequence (3D-SPACE) constitutes a rather standard inner ear imaging sequence. All employed coregistration methods can be applied to multisite MR data as well, e.g., in multicenter studies, although we did not investigate multisite-harmonization performance in the current study. In principle, however, one major advantage of the semimanual 3P-technique is its generalizability since the three landmarks used in this study should be easily identifiable in high-resolution MR or CT sequences, independent of vendor or sequence characteristics. Future developments could include automatic landmark selection, for example, using deep learning⁵⁶.

On a technical note, we tested the 3D-Slicer implementations of the automatic registration tools which are not as actively developed as the standalone versions. While software updates cannot solve the inherent pitfalls of automatic, unaided inner ear image coregistration, code improvements might still result in notable performance increases in the standalone versions compared to the 3D-Slicer implementations. Importantly, however, the registration inaccuracies observed in the automatic registration tools do not stem from faulty algorithms or their wrongful implementation. The software worked just as intended, correctly optimizing the registration process which inherently (due to the aforementioned anatomical and methodological pitfalls) resulted in misaligned inner ears. This is one of the reasons why no hyper-parameter optimization of the automatic algorithms was

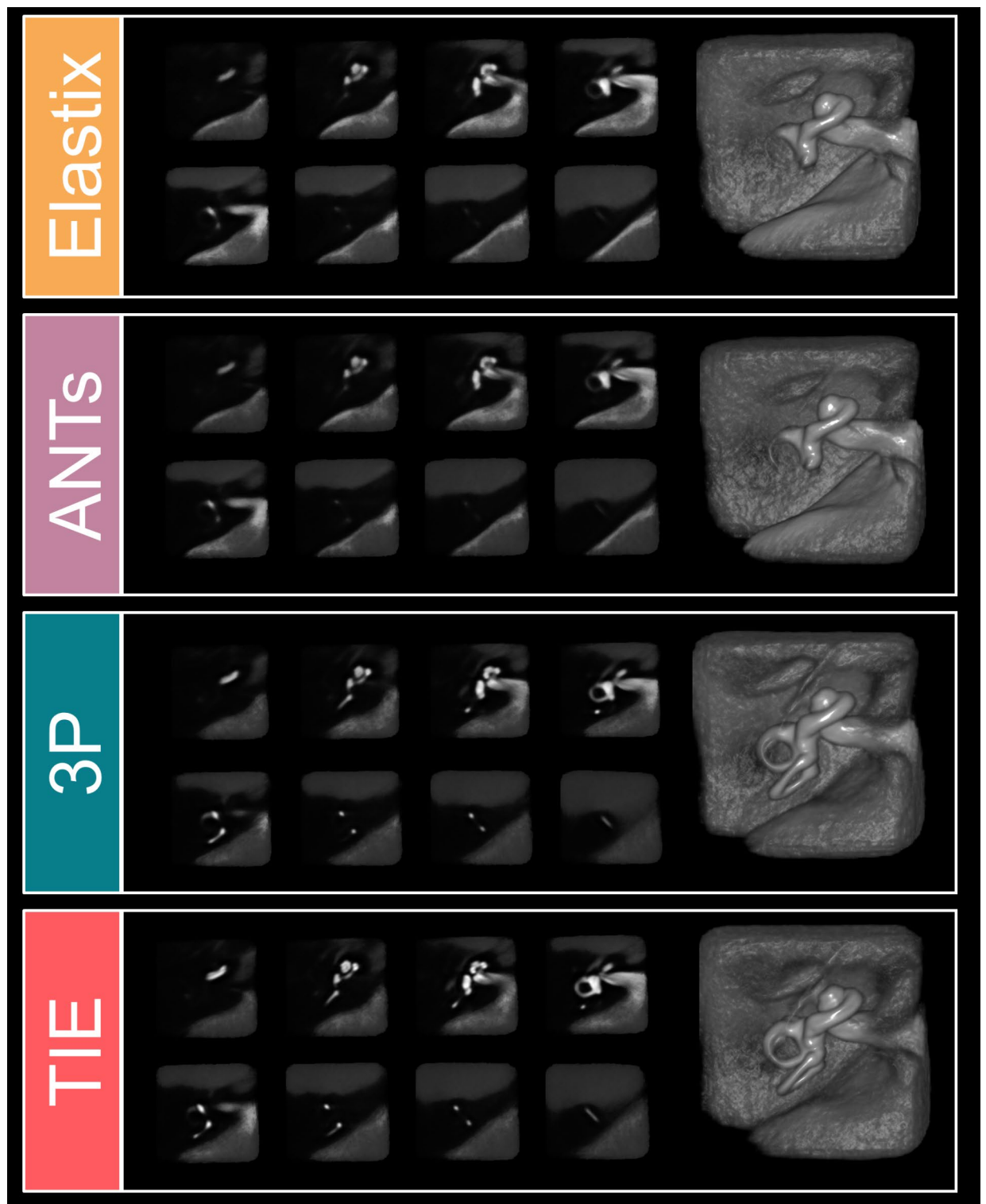


Fig. 4. Slicewise comparisons of the 4 examined registration methods and 3D-renderings of the resulting median volumes from 153 human inner ears (from top to bottom: automatic Elastix registration, automatic ANTs registration, semimanual three-point registration (3P), automatic mask-aided TIE-approach). Finer inner ear details such as the semicircular canals are only discernible in the 3P and TIE methods.

performed; while parameter tuning might potentially result in slightly higher accuracy, these inherent pitfalls would not be avoided.

It should be noted that the aim of this study was the analysis of rigid registrations that preserve petrous bone anatomy and geometry. Depending on one's research question, non-linear registration of a presegmented

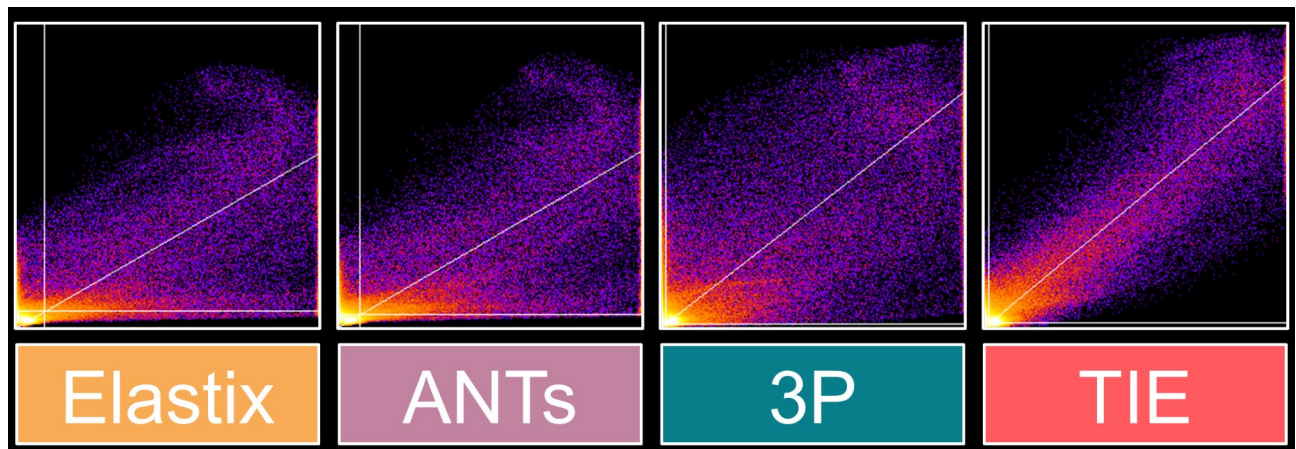


Fig. 5. Colocalization scatterplots of each registration method with the target atlas (from left to right: Elastix, ANTs, 3P, TIE). In order to exclude background noise, an inner ear mask was used as an ROI (diagonal line: linear regression fit of colocalizations; horizontal/vertical lines: Costes auto thresholds for channel separation). While exact colocalization (i.e., all data points on the diagonal line) is not possible since the moving image and target image are inherently different, TIE results in the most accurate coregistration, i.e., data points located in close proximity to the diagonal line.

individual inner ear dataset onto an anatomically labeled inner atlas might suffice, despite inherently altering petrous bone geometry. For this, numerous registration toolboxes with great accuracy are available^{49,57}. However, when petrous bone anatomy needs to be analyzed without introducing geometrical modifications, rigid registration is necessary (e.g., with TIE and 3P). As noted earlier, this can only result in approximate colocalization in a multipatient dataset.

Conclusions

Mask-aided inner ear coregistration, followed by semimanual landmark-based registration, outperformed unaided automatic coregistration methods, while all methods were substantially more accurate than simply a whole-brain coregistration. In general, spatial normalization should be integrated into inner ear imaging data analysis pipelines to aid in research comparability and to lay the foundation for more complex analyses of neurotological datasets. To allow other researchers to implement spatial normalization steps, the TIE method (which proved superior in this analysis) will be made freely available.

Data availability

All software used in this study is open source and available for free download. The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Received: 5 November 2024; Accepted: 17 February 2025

Published online: 21 February 2025

References

- Bradshaw, A. P. et al. A mathematical model of human semicircular canal geometry: A new basis for interpreting vestibular physiology. *JARO* **11**, 145–159 (2010).
- Cortés-Domínguez, I., Fernández-Seara, M. A., Pérez-Fernández, N. & Burguete, J. Systematic method for morphological reconstruction of the semicircular canals using a fully automatic skeletonization process. *Appl. Sci.* **9**, 4904 (2019).
- Johnson Chacko, L. et al. Analysis of vestibular labyrinthine geometry and variation in the human temporal bone. *Front. Neurosci.* **12**, 107 (2018).
- Lewis, M. A. et al. The spectrum of cochlear malformations in CHARGE syndrome and insights into the role of the CHD7 gene during embryogenesis of the inner ear. *Neuroradiology* **65**, 819–834 (2023).
- Roesch, S., Rasp, G., Sarikas, A. & Dossena, S. Genetic determinants of non-syndromic enlarged vestibular aqueduct: A review. *Audiol. Res.* **11**, 423–442 (2021).
- Patki, A. U., Ronen, O., Kaylie, D. M., Frank-Ito, D. O. & Piker, E. G. Anatomic variations in temporal bones affect the intensity of nystagmus during warm caloric irrigation. *Otol. Neurotol.* **37**, 1111–1116 (2016).
- von Brevern, M. et al. Benign paroxysmal positional vertigo: Diagnostic criteria. *VES* **25**, 105–117 (2015).
- Vaidyanathan, A. et al. Deep learning for the fully automated segmentation of the inner ear on MRI. *Sci. Rep.* **11**, 2885 (2021).
- Ahmadi, S.-A., Frei, J., Vivar, G., Dieterich, M. & Kirsch, V. IE-Vnet: Deep learning-based segmentation of the inner ear's total fluid space. *Front. Neurol.* **13**, 663200 (2022).
- Neves, C. A., Tran, E. D., Kessler, I. M. & Blevins, N. H. Fully automated preoperative segmentation of temporal bone structures from clinical CT scans. *Sci. Rep.* **11**, 116 (2021).
- Yao, A. D., Cheng, D. L., Pan, I. & Kitamura, F. Deep learning in neuroradiology: A systematic review of current algorithms and approaches for the new wave of imaging technology. *Radiol. Artif. Intell.* **2**, e190026 (2020).
- Copson, B. et al. Definition of a coordinate system for multi-modal images of the temporal bone and inner ear. *PLoS ONE* **19**(10), e0294828. <https://doi.org/10.1371/journal.pone.0294828> (2024).

13. Bächinger, D., Schuknecht, B., Dlugacz, J. & Eckhard, A. H. Radiological configuration of the vestibular aqueduct predicts bilateral progression in Meniere's disease. *Front. Neurol.* **12**, 674170. <https://doi.org/10.3389/fneur.2021.674170> (2021).
14. Naganawa, S. et al. Visualization of endolymphatic hydrops in Ménière's disease with single-dose intravenous gadolinium-based contrast media using heavily T(2)-weighted 3D-FLAIR. *Magn. Reson. Med. Sci.* **9**, 237–242 (2010).
15. Kirsch, V. et al. Imaging endolymphatic space of the inner ear in vestibular migraine. *J. Neurol. Neurosurg. Psychiatry* <https://doi.org/10.1136/jnnp-2024-334419> (2024).
16. Leng, Y. et al. Non-contrast MRI of inner ear detected differences of endolymphatic drainage system between vestibular migraine and unilateral Ménière's disease. *Front. Neurol.* **13**, 814518 (2022).
17. Li, H. et al. A synchrotron and micro-ct study of the human endolymphatic duct system: Is Meniere's disease caused by an acute endolymph backflow? *Front. Surg.* **8**, 662530 (2021).
18. Connor, S. et al. Assessing the optimal MRI descriptors to diagnose Ménière's disease and the added value of analysing the vestibular aqueduct. *Eur. Radiol.* <https://doi.org/10.1007/s00330-024-10587-w> (2024).
19. Möller, M. N., Kirkeby, S., Vikeså, J., Nielsen, F. C. & Caye-Thomasen, P. Expression of histamine receptors in the human endolymphatic sac: The molecular rationale for betahistine use in Menieres disease. *Eur. Arch. Otorhinolaryngol.* **273**, 1705–1710 (2016).
20. Lei, P., Leng, Y., Li, J., Zhou, R. & Liu, B. Anatomical variation of inner ear may be a predisposing factor for unilateral Ménière's disease rather than for ipsilateral delayed endolymphatic hydrops. *Eur. Radiol.* **32**, 3553–3564 (2022).
21. Bartling, S. H. et al. Registration and fusion of CT and MRI of the temporal bone. *J. Comput. Assist. Tomogr.* **29**(3), 305–310. <https://doi.org/10.1097/01.rct.0000160425.63884.5b> (2005).
22. Powell, K. A. et al. Atlas-based segmentation of temporal bone anatomy. *Int. J. Comput. Assist. Radiol. Surg.* **12**(11), 1937–1944. <https://doi.org/10.1007/s11548-017-1658-6> (2017).
23. Al-Dhamari, I. et al. Automatic intra-subject registration and fusion of multimodal cochlea 3D clinical images. *plos ONE* **17**(3), e0264449. <https://doi.org/10.1371/journal.pone.0264449> (2022).
24. Kakir, A., Labadie, R. F., Zuniga, M. G., Dawant, B. M. & Noble, J. H. Evaluation of rigid cochlear models for measuring cochlear implant electrode position. *Otol. Neurotol.* **37**(10), 1560–1564. <https://doi.org/10.1097/MAO.0000000000001245> (2016).
25. Fonov, V. et al. Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage* **54**, 313–327 (2011).
26. Evans, A. C., Collins, D. L. & Milner, B. An MRI-based stereotactic atlas from 250 young normal subjects. *J. Soc. Neurosci.* **18**, 408 (1992).
27. Chau, W. & McIntosh, A. R. The Talairach coordinate of a point in the MNI space: How to interpret it. *NeuroImage* **25**, 408–416 (2005).
28. Talairach, J. & Tournoux, P. *Co-planar Stereotaxic Atlas of the Human Brain: An Approach to Cerebral Imaging* (Thieme Classics, 1988).
29. Evans, A. C., Janke, A. L., Collins, D. L. & Baillet, S. Brain templates and atlases. *NeuroImage* **62**, 911–922 (2012).
30. Amunts, K. et al. BigBrain: An ultrahigh-resolution 3D human brain model. *Science* **340**, 1472–1475 (2013).
31. Ashburner, J. & Friston, K. J. Voxel-based morphometry—The methods. *NeuroImage* **11**, 805–821 (2000).
32. Glasser, M. F. et al. The human connectome project's neuroimaging approach. *Nat. Neurosci.* **19**, 1175–1187 (2016).
33. Bookstein, F. L. "Voxel-based morphometry" should not be used with imperfectly registered images. *NeuroImage* **14**, 1454–1462 (2001).
34. Mangin, J.-F. et al. Spatial normalization of brain images and beyond. *Med. Image Anal.* **33**, 127–133 (2016).
35. Ahmadi, S.-A., Raiser, T. M., Rühl, R. M., Flanagan, V. L. & Eulenburg, P. IE-Map: A novel in-vivo atlas and template of the human inner ear. *Sci. Rep.* **11**, 3293 (2021).
36. Napadow, V., Dhond, R., Kennedy, D., Hui, K. K. S. & Makris, N. Automated brainstem co-registration (ABC) for MRI. *NeuroImage* **32**, 1113–1119 (2006).
37. Nakashima, T. et al. Grading of endolymphatic hydrops using magnetic resonance imaging. *Acta Otolaryngol.* **129**, 5–8 (2009).
38. Baráth, K. et al. Detection and grading of endolymphatic hydrops in Ménière disease using MR imaging. *AJNR Am. J. Neuroradiol.* **35**, 1387–1392 (2014).
39. Boegle, R. et al. Intravenous delayed gadolinium-enhanced MR imaging of the endolymphatic space: A methodological comparative study. *Front. Neurol.* **12**, 647296 (2021).
40. Boegle, R., Kirsch, V., Gerb, J. & Dieterich, M. Modulatory effects of magnetic vestibular stimulation on resting-state networks can be explained by subject-specific orientation of inner-ear anatomy in the MR static magnetic field. *J. Neurol.* **267**, 91–103 (2020).
41. Gerb, J. et al. VOLT: a novel open-source pipeline for automatic segmentation of endolymphatic space in inner ear MRI. *J. Neurol.* **267**, 185–196 (2020).
42. Dill, T. Contraindications to magnetic resonance imaging: Non-invasive imaging. *Heart* **94**, 943–948 (2008).
43. Schindelin, J. et al. Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
44. Fedorov, A. et al. 3D slicer as an image computing platform for the quantitative imaging network. *Magn. Reson. Imaging* **30**, 1323–1341 (2012).
45. GitHub. *Git*Hub - *netstim/SlicerANTs: This Extension Implements ANTs Registration into Slicer* (2024). <https://github.com/netstim/SlicerANTs>. Accessed 15 Jan 2024.
46. Avants, B. B. et al. The Insight ToolKit image registration framework. *Front. Neuroinform.* **8**, 44 (2014).
47. Klein, S., Staring, M., Murphy, K., Viergever, M. A. & Pluijm, J. P. W. elastix: A toolbox for intensity-based medical image registration. *IEEE Trans. Med. Imaging* **29**, 196–205 (2010).
48. Shamonin, D. P. et al. Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease. *Front. Neuroinform.* **7**, 50 (2013).
49. Keszei, A. P., Berkels, B. & Deserno, T. M. Survey of non-rigid registration tools in medicine. *J. Digit. Imaging* **30**(1), 102–116. <https://doi.org/10.1007/s10278-016-9915-8> (2017).
50. Costes, S. V. et al. Automatic and quantitative measurement of protein–protein colocalization in live cells. *Biophys. J.* **86**, 3993–4003 (2004).
51. Manders, E. M. M., Verbeek, F. J. & Aten, J. A. Measurement of co-localization of objects in dual-colour confocal images. *J. Microsc.* **169**, 375–382 (1993).
52. Gisev, N., Bell, J. S. & Chen, T. F. Interrater agreement and interrater reliability: Key concepts, approaches, and applications. *Res. Soc. Adm. Pharm.* **9**, 330–338 (2013).
53. Gerb, J., Padovan, L., Lehrer, N., Brandt, T. & Dieterich, M. Differential effects of gender and age on dynamic subjective visual vertical. *Neurol. Res. Pract.* **5**, 37 (2023).
54. Smith, P. F., Agrawal, Y. & Darlington, C. L. Sexual dimorphism in vestibular function and dysfunction. *J. Neurophysiol.* **121**, 2379–2391 (2019).
55. Gerb, J. et al. Sex differences in caloric nystagmus intensity: Should reference values be updated? *Ann. N. Y. Acad. Sci.* <https://doi.org/10.1111/nyas.15310> (in press).
56. Stebani, J. et al. Towards fully automated inner ear analysis with deep-learning-based joint segmentation and landmark detection framework. *Sci. Rep.* **13**, 19057 (2023).
57. Reuter, M., Rosas, H. D. & Fischl, B. Highly accurate inverse consistent registration: A robust approach. *Neuroimage* **53**(4), 1181–1196. <https://doi.org/10.1016/j.neuroimage.2010.07.020> (2010).

Acknowledgements

The present study was supported by the Deutsche Stiftung Neurologie (DSN; projects 80766113 and 80766125) and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (Munich Cluster for Systems Neurology: EXC 2145 SyNergy) to MD, and partially funded by the "Verein zur Förderung von Wissenschaft und Forschung an der Medizinischen Fakultät der LMU" (Association for the Promotion of Science and Research at the LMU Medical Faculty), and the German Federal Ministry of Education and Research (BMBF) via the German Center for Vertigo and Balance Disorders (DSGZ, Grant No. 01 EO 0901).

Author contributions

JG: conception, design of the study, acquisition, analysis of the data, creating figures, drafting and revising the manuscript. VK: acquisition of the data, providing funding. EK: acquisition and analysis of the data. TB: revising the manuscript. MD: revising the manuscript, providing funding. RB: conception, analysis of the data, drafting and revising the manuscript. All authors contributed to the article and approved the submitted version.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025