Full citation: Powell KA, Wiet GJ, Hittle B, Oswald GI, Keith JP, Stredney D, Andersen SAW. Atlas-based segmentation of cochlear microstructures in cone beam CT. Int J Comput Assist Radiol Surg. 2021 Mar;16(3):363-373.
DOI: 10.1007/s11548-020-02304-x

Title: Atlas-Based Segmentation of Cochlear MicroStructures in Cone Beam CT

Authors: Kimerly A. Powell Ph.D.[1], Gregory J. Wiet M.D., FACS, FAAP[2], Brad Hittle B.S.[1], Jason P. Keith[3], Grace I. Oswald[3], Don Stredney M.A.[4], Steven Arild Wuyts Andersen M.D., Ph.D.[2,5]

Affiliations:

[1] Department of Biomedical Informatics, The Ohio State University, Columbus, Ohio

[2] Department of Otolaryngology, The Ohio State University and Nationwide Children's Hospital, Columbus, Ohio

[3] Dept. of Biomedical Engineering, The Ohio State University, Columbus, Ohio

[4] Interface Laboratory, The Ohio State University, Columbus, Ohio

[5] Dept. of Otorhinolaryngology–Head & Neck Surgery, Rigshospitalet, Copenhagen, Denmark

Corresponding Author:
Kimerly A. Powell
Department of Biomedical Informatics
The Ohio State University
Columbus, Ohio 43210
Email – kimerly.powell@osumc.edu
ORCID ID - 0000-0003-0111-1047

Acknowledgements:
This research was supported by NIDCD/NIH 1R01-DC011321 and by funding from the Ohio State University College of Medicine Office of Research Bridge Funding Program and Nationwide Children's Hospital – Columbus OH. Dr. Andersen was funded by the Independent Research Fund Denmark (8026-00003B).

Abstract

Purpose: To develop an automated segmentation approach for cochlear microstructures [scala tympani (ST), scala vestibuli (SV), modiolus (Mod), mid-modiolus (Mid-Mod), and round window membrane (RW)] in clinical cone beam computed tomography (CBCT) images of the temporal bone for use in surgical simulation software and for preoperative surgical evaluation.

Methods: This approach was developed using the publicly available OpenEar (OE) Library that includes temporal bone specimens with spatially registered CBCT and 3D micro-slicing images. Five of these datasets were spatially aligned to our internal OSU atlas. An atlas of cochlear microstructures was created from one of the OE datasets. An affine registration of this atlas to the remaining OE CBCT images was used for automatically segmenting the cochlear microstructures. Quantitative metrics and visual review were used for validating the automatic segmentations.

Results: The average DICE metrics were 0.77 and 0.74 for the ST and SV, respectively. The average Hausdorff distance (AVG HD) was 0.11 mm and 0.12 mm for both scalae. The mean distance between the centroids for the round window was 0.32 mm, and the mean AVG HD was 0.09 mm. The mean distance and angular rotation between the mid-modiolar axes were 0.11 mm and 9.8 degrees, respectively. Visually, the segmented structures were accurate and similar to that manually traced by an expert observer.

Conclusions: An atlas-based approach using 3D micro-slicing data and affine spatial registration in the cochlear region was successful in segmenting cochlear microstructures of temporal bone anatomy for use in simulation software and potentially for pre-surgical planning and rehearsal.

Keywords: atlas-based segmentation; image registration; surgical simulation; pre-surgical planning, cochlea anatomy

Introduction

The temporal bone (TB) comprises the lateral and lower part of the skull and contains among other structures the inner ear including the cochlea that is responsible for the final transduction of sound into nerve signals sent to the brain. First, the sound waves pass through the external ear canal and reach the tympanic membrane. The vibration of the tympanic membrane is amplified through the ossicular chain before reaching the oval window membrane, where the fluid of the inner ear is put into motion, ultimately affecting the hair cells in the cochlea responsible for generating the nerve impulse. The snail shell-like shape of the cochlea serves to discern the frequency and amplitude of the sound signal and typically revolves with 2¾ turns around the modiolus—the conically shaped center housing the spiral ganglion and cochlear nerve. Microstructurally, the cochlea is compartmentalized into three separate fluid-filled spaces: the scala tympani, scala vestibuli, and scala media. The round window membrane at the other end of the system is displaced in opposite phase to the pressure applied through the oval window membrane.

In many cases of profound sensorineural hearing loss, especially congenital, temporal bone surgery with a cochlear implant (CI) for direct electrical stimulation of the cochlear nerve fibers is relevant. The surgery requires drilling of the temporal bone (mastoidectomy with facial recess approach) to gain access to the cochlea and the insertion of the CI electrode array through the round window membrane or through a cochleostomy (drilling of an artificial window) along the scala tympani. Next, the electrode array is slowly inserted and follows the curvature of the cochlea. Positioning of the electrode array is critical to obtain good patient hearing outcomes. Optimally, the electrode array is placed entirely in the scala tympani as any displacement into the scala vestibuli or tear of the partitioning membranes potentially causes reduced hearing outcomes for the patient post-operatively, accounting for some of the variability in patient hearing outcomes post implantation [1, 2].

This highlights the major role of the intracochlear anatomy and temporal bone microstructures in surgical management. However, the scala tympani and vestibuli are not visible on typical clinical computed tomography (CT) or cone-beam CT (CBCT) acquired pre-operatively. Micro-CT, histology, or micro-slicing provides the contours of the scala tympani, scala vestibuli, modiolus,

and round window membrane (Figure 1). Neither are clinically viable options due to radiation and/or the need to extract the temporal bone. Consequently, there is a need for other approaches to estimate these structures in clinical scans. This would have multiple important uses such as pre-surgical planning of for example, the surgical approach for electrode insertion, selection of electrode array length, pre-surgical training and rehearsal in a virtual reality (VR) simulation environment to optimize placement, and post-operative evaluation of electrode placement with possible implications for electrode programming and prediction of clinical outcomes.

We have previously reported on our experience with segmenting otic and surface-based temporal bone structures [3, 4] using atlas-based segmentation for use in our surgical simulation software [5, 6] and we now present a method for segmenting cochlear microstructures not readily observed in clinical CT images of the temporal bone.

Methods

The overall goal of this study was to develop an automated approach for segmenting the scala tympani (ST), scala vestibuli (SV), modiolus (Mod), mid-modiolus axis (Mid-Mod), and round window membrane (RW) in clinical CBCT images of temporal bone. The approach was based on images that were obtained from the publicly available OpenEar (OE) Library *ex vivo* specimens of temporal bones [7]. An atlas-based segmentation was developed that was based on the affine registration of one of 3D micro-slicing data from the OE library to the CBCT data. A flowchart of the workflow is presented in Figure 2.

Study Sample: A total of five datasets (DELTA, EPSILON, ETA, GAMMA, and THETA) from the OE Library were used to develop the atlas-based segmentation approach. One of the 3D micro-slicing data (ZETA) was of insufficient quality to accurately segment the cochlear microstructures for validation. All of these temporal bones were human adult specimens and did not have any otologic pathology.

OpenEar Library Specimen Preparation: The complete description of the OpenEar Library specimen preparation is found in the paper by Sieber *et al.* [7]. The general steps of their preparation were as follows: CBCT images of the temporal bone specimens were acquired at a

voxel size of 0.25 mm using a 3D Accuitomo 170 Digital CBCT scanner (J. Morita Tokyo Mfg. Corp., Japan). After CBCT imaging, the specimens were fixed in 4% Formol, dehydrated in a series of ethanol exchanges, stained with 0.1% Acid Fuchsin, and embedded and cured in epoxy resin. The embedded specimens were then re-imaged using CBCT at a voxel size of 0.125 mm. Micro-slicing was performed by sequential grinding and microscopic imaging of the specimen block face at 20x magnification. The 2D images were spatially registered to the high-resolution embedded CBCT scan. The final resolution of the registered 3D micro-slice images was an in-plane pixel size of 0.05 mm and a slice thickness of 0.15 mm.

OpenEar Library Atlas Preparation: For this study, the unembedded CBCT images were spatially aligned to our internal OSU atlas [3, 4] rescaled to 0.08 mm. The reason for setting the OSU atlas to this scale is to match the clinically acquired CBCT data at our institution. This scale can be changed based on the scale of the data being segmented. The OE CBCT images were rigid body registered to the OSU atlas using a region that included the otic capsule (Figure 3a and b). Elastix 4.7 [8-10] was used for the registration with the following parameters: A multi-resolution pyramid with 4 levels using Advanced Mattes Mutual Information, adaptive stochastic gradient descent optimizer, maximum number of iterations 500, and a 3rd order B-spline interpolator. The OE 3D micro-sliced images were then rigid body registered to the OE CBCT registered data using the same rigid body registration previously described. This registration was performed for the 5 datasets selected from the OE library. One of the OE datasets (THETA) was used as the OE atlas for subsequent validation in the other OE datasets.

The cochlear microstructures (scala tympani, scala vestibuli, modiolus, mid-modiolus, and round window membrane) were manually outlined by an expert (SA) using the OE spatially registered micro-slicing images and ITK-SNAP 3.6 [11] using a digital Wacom Cintiq Pro 13 pen display (Wacom Co., Ltd, Japan). This took 6-8 hours per data set.

Atlas Segmentation: The 3D micro-slicing image for THETA aligned to the OSU atlas was affine registered to the four other spatially aligned OE CBCT images using a cropped region that included only the cochlea (12 x 12 x 9.6 mm) (Figure 3c) and the same registration approach previously described. An affine registration was chosen primarily due the limited sample size

available for evaluating shape variations of the cochlea. It allows for proportional changes in shape due to anatomical variation between cochlea without the need for prior shape information that is required for the deformable registration. The cochlear microstructures manually outlined by the expert were spatially transformed to match the registration. Further 'cleaning' of the transformed structures were as follows: Any region of the transformed ST and SV that were outside the cochlea (i.e. in bone) were removed from the segmentation using a globally segmented bone image and Otsu multi-level threshold with three levels (bone = level 3) [12]. The atlas-transformed RW was translated to match an automatically segmented outline of the round window opening in the CBCT image. This round window opening was determined automatically by identifying the narrowest corresponding points on the outline of the cochlea in this region (Figure 4).

Validation: The results from the automated segmentations were then quantitatively and visually evaluated. For the quantitative evaluation, four quantitative metrics, the DICE, average Hausdorff distance (AVG HD), maximum Hausdorff distance (MAX HD), and volume similarity (VS), were calculated for the ST, SV, and the Mod. The DICE and VS were measured for the segmented volumes using Taha *et al.'s* quantitative evaluation tool [13]. The DICE coefficient is a measure of the amount of overlap between two segmented structures. The VS measures how close the volumes are to one another. The AVG HD and MAX HD were measured using the surface outlines of the manually and automatically segmented structures in MeshLab v1.3.4 [14]. The AVG HD, MAX HD, and the distance between the centroids of the automated and manual segmentation were measured for the RW. The difference in the angle (*dr*) and the distance (*dt*) was calculated between the automated and manual Mid-Mod axes using the line equations presented in Demarcy et al. [15],

$$d_r(L_i, L_j) = \left|\sin^{-1}\left(\left\|z_i \times z_j\right\|\right)\right| \epsilon [0, \pi]$$

$$d_t(L_i, L_j) = \left|\frac{z_i \times z_j}{\|z_i \times z_j\|} \cdot (p_j - p_i)\right| \geq 0$$

where the Mid-Mod axes are represented as lines (*Li*) and (*Lj*) and $L = \{p + sz | s \in R\}$, and *z* is a unit vector.

Visual verification was performed by comparing the automatically segmented structures with the manually segmented structures in a 2D and 3D viewer. No manual adjustments were made to the automated segmentations after visual verification.

The automated segmentation approach was further evaluated using a sample of 11 preoperative temporal bone CBCTs scans obtained from adult patients who had undergone CI surgery at the Ohio State University within three years prior to the study. The scans were all performed on a clinical CBCT scanner (3D Accuitomo 170, J. Morita Tokyo Mfg. Corp., Japan) at a voxel size of 0.08x0.08x0.08 mm. The cochlea, round window, and end points of the Mid-Mod axis were manually outlined on the clinical CBCTs by an expert in cochlear anatomy (SA). The scala tympani and vestibuli from the automated atlas-based segmentation were combined for comparison to the manually traced cochlea in the clinical CBCTs. The three cochlea microstructures were validated using the same metrics as described for the OE datasets.

Results

The results of the quantitative analysis are presented in Table 1. The mean DICE metric for the ST and SV was 0.77 and 0.74, respectively. The mean AVG HD for the ST and SV was 0.11 and 0.12 mm, respectively. The mean DICE metric for the Mod was 0.59, however the VS was 0.95 and AVG HD was 0.17 mm. The mean distance between the centroids for the RW was 0.32 mm and the mean AVG HD was 0.09 mm. The mean distance and angular rotation between the Mid-Mod axes were 0.11 mm and 9.8 degrees, respectively.

An example of the automated segmentation based on the atlas masks is presented in Figure 5. The red outlines in this image represent the manually segmented structures and the green outlines represent the atlas-based segmented structures. Overall, the segmentation of the structures appears consistent between the manually and automatically segmented images with slight differences observed for the scalae at the apex and slight shifts in the boundaries of the modiolus. The best segmentation was observed in ETA and the worst was observed in EPSILON. The increased errors observed with EPSILON were mainly due to processing artifacts in the micro-slicing images, that made manual segmentation of the cochlear microstructures difficult, as opposed to errors associated with the automated segmentation approach.

Three dimensional plots of HD distance for the automatically and manually segmented structures are presented in Figure 6. The left column of this figure illustrates the HD for the ST, the middle column HD for the SV, and the right column HD for the Mod. The rows from top to bottom are DELTA, EPSILON, ETA, and GAMMA. The mean MAX HD distance between manual and automated was 0.62, 0.60, and 0.33 mm for ST, SV, and Mod, respectively. However, the mean AVG HD distance between manual and automated segmentation was 0.12, 0.11, and 0.17 mm for ST, SV, and Mod, respectively. The greatest variation in HD in segmentation of the scalae were observed in the apical region, on the inside of the first half of the basal turn, and at the interface between the scalae in the 1-1.5 turn. These HDs were on the order of 0.30 - 0.35 mm. The greatest variation in HD on the Mod was also observed on the first half of the basal curve and were also on the order of 0.30-0.35 mm.

The results of the quantitative analysis for the clinical CBCTs are presented in Table 2. The mean DICE metric, AVG HD, and VS for the cochlea were 0.87, 0.11, and 0.98, respectively. The mean distance between the centroids for the RW was 0.24 mm and the mean AVG HD was 0.12 mm. The mean distance and angular rotation between the Mid-Mod axes were 0.11 mm and 8.5o, respectively. An example of the automated segmentation based on the OE atlas masks is presented in Figure 7. The red and green outlines in this image represent the automatically segmented ST and SV, respectively. The blue outline is the manually segmented cochlea. The outer surface of the combined ST and SV is closely aligned with the manually segmented cochlea as shown in a three-dimensional plot of the HD distance for one of the clinical datasets presented in Figure 8. The greatest variation in HD of the cochlea was observed at the first half of the basal turn near the Mod and was on the order of 0.30–0.35 mm.

The total time for automated segmentation of the cochlear microstructures was 108 secs, with 78 seconds for rigid body registration of the OE data to OSU atlas space, 21 seconds for the affine registration of the OE micro-slice atlas to the CBCT images, and 9 secs for ‘cleaning’ the transformed structures.

Discussion

Our overall goal was to develop an automated segmentation approach that could accurately map cochlear microstructures for use in our simulation and pre-surgical planning software. Visual inspection of the segmentations and the quantitative metrics for our method validate that the atlas-based affine registration of 3D micro-slicing data to CBCT data is good for this purpose. For example, the mean HD for ST and SV were 0.11 and 0.12, respectively. This was similar to what Kjer *et.al.* [16] observed using a statistical deformation model and B-spline parameters, and was lower than what Nobel *et al.* [17] observed (0.20 mm) using a statistical shape model. Large variations in HD, in the order of 0.3-0.35 mm, were observed for the scalae in the apical region beyond the second turn of the cochlea. This region is less critical in evaluating CI position within the cochlea since CI electrodes do not cover this region within the cochlea. Large variations in HD were also observed at the interface between the scalae in the middle turns of the cochlea. However, the basilar membrane is on the order of 0.23-0.28 mm in this region [18] which is similar to the variation we observed in this region. The last region that we observed large variation in HD was on the inside of the first half of the basal turn. The CI electrode stimulates this region of the cochlea and consequently, accuracy in the calculation of the distance of the CI electrode contacts to the spiral ganglion (modiolar distance used as a proxy) is more important for this region for example if this information is to be used for image-guided programming [19]. The larger variations in HD for the Mod were consistent with that observed in the modiolar region of the scalae and is consistent with variation observed in the adjacent structures.

The center of the RW and the Mid-Mod are important landmarks for defining an internal cochlear coordinate system [20] used in evaluating individual electrode locations with respect to auditory nerve endings (spiral ganglion cells). The Mid-Mod constitutes the Z-axis in this coordinate system. Demarcy *et al.* [15] developed an automated method for defining the Mid-Mod based on the centerline of the cochlea spiral. They used the mean distance and angular rotation between two lines defined in 3D space for evaluating the errors in predicting this axis automatically. They compared their automated method with manual estimations made by four experts and calculated that the inter-observer variability in manually identifying the Mid-Mod was 10.0^{o} for the rotation angle, and 0.14 mm for the distance between the axes. We observed a mean difference in Mid-Mod angular rotation between our automated and manual method of 9.8^{o}, and distance between the axes of 0.11 mm. These are withinthe variability of the manual

estimations made by the expert observers reviewed by Demarcy. The observed Mid-Mod rotation in our analysis is similar to that in Demarcy's automated method based on the centerline of the cochlea, however, the distance between the axes is slightly better using our method. The center of the round window is used as the zero-reference angle for this coordinate system. We observed a mean difference in distance between the centroids of the RW determined automatically and manually of 0.32 mm and a mean AVG HD of the RW surface of 0.09 mm.

The application of our approach to a clinical CBCT dataset resulted in validation metrics similar to those observed for the OE datasets. However, the DICE metric for the cochlea compared to the combined scalae was significantly better at 0.87. This may be due in part to the larger size and more regular shape of the structure being tested (i.e., whole cochlea vs. individual scala). Although there is no ground truth available for the individual scala in these datasets, we observed a uniform scaling of these structures from the base to the apex within the cochlea. Previous studies indicate that there is a significant variation in the shape of these microstructures between different specimens, however, the overall change in diameter, height, and cross-sectional area within a single cochlea decreases fairly uniformly from the base to apex and this pattern has been shown to be well conserved and rather uniformly scaled between cochlea [21].

The results of our study suggest that affine registration may be sufficient for cochlea microstructure segmentation in clinical CBCT images, however, additional testing with a larger number of samples is needed to confirm this. Two factors that may have contributed to our results are (1) that our method is based on volumetric registration of images as opposed to using extracted structures or surfaces of structures, and (2) we used CBCT images resampled to an isotropic voxel size of 0.08 mm as opposed to standard clinical CT images with a typical voxel size on the order of 0.2–0.3 mm. We evaluated whether the affine registration approach differed from a rigid body registration and observed that the quantitative validation metrics were similar for both registrations. However, a visual review of the HD for the surface of both scala indicated that the affine registration was consistently better than the rigid body for identifying the outside edge in the first half of the basal turn. The improvement in this region was 0.25 mm or more.

The primary advantage of our proposed method is that it is fully automated and fast (< 2 min), making it clinically useful for pre-surgical planning and post-implant evaluation. It also does not require an accurately segmented cochlea prior to fitting a model to the data which can be difficult in clinically acquired images where patient motion and image artifacts can be significant.

Since there are few landmarks available within the clinically acquired CBCT images to segment some of these microstructures directly, we must rely upon the information obtained from high-resolution ex vivo imaging methods such as micro-MRI [22], micro-CT [15–17], or microscopy [7, 23] imaging to predict them in clinical images. Although histology and micro-slicing images can provide high-resolution information about these structures, preparation processing artifacts can negatively impact manual outlining of microstructures from these images. Additionally, there is inherent variability in the manual segmentation of images used for the validation of these approaches [24]. This variability has largely been ignored in the evaluation of cochlear microstructure in previous studies due to the significant amount of time required to manually trace the structures for validation. Nonetheless these errors exist and can be quite significant (i.e., in the order of 10–20% error for measurements in volume).

In Figure 9, we illustrate how the segmented cochlear microstructures can be used to visualize the placement of CI electrodes in 3D for a cochlear implant patient. These images were created using our OE atlas-based segmentation approach applied to a patient after CI implant surgery at our institution. CBCT images (0.08 mm isotropic voxels) were acquired pre- and post-implant and spatially aligned to the OSU atlas as described previously [3, 4]. The post-implant image was then rigid body registered to the pre-implant image and the electrodes were segmented in the post-implant image using a global threshold. The OE atlas was affine registered to the pre-implant image, and the cochlear microstructures were segmented. The cochlear microstructures and electrodes were combined into a single-color overlay and merged with the pre-implant image for interactive 3D volume visualization. In Figure 9a, you can readily observe how the CI electrode (yellow) enters through the round window membrane (orange) and wraps around the modiolus (magenta) inside the scala tympani (red). Future work will focus on using pre-implant images in our surgical simulator [5, 6] for pre-surgical planning to optimize implant placement

and post-implant images for visualizing final outcomes and for image-guided programming of the CI as described by Noble et al. [19].

A major limitation of this study is the small sample size used for validating the model. However, these types of studies (embedding, micro-slicing, and processing) are laborious and costly, and we are grateful to our colleagues for making the OpenEar data available to the community.

Compliance with Ethical Standards

Disclosure of potential conflicts of interest: This study was funded by NIDCD/NIH (1R01-DC011321), The Ohio State University College of Medicine Office of Research Bridge Funding Program and Nationwide Children's Hospital, Columbus, OH.

Ethical approval: All procedures performed in this study involving human participants were in accordance with the ethical standards of the Ohio State University Institutional Review Board and have been performed in accordance with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent: For this type of study formal consent was not required.

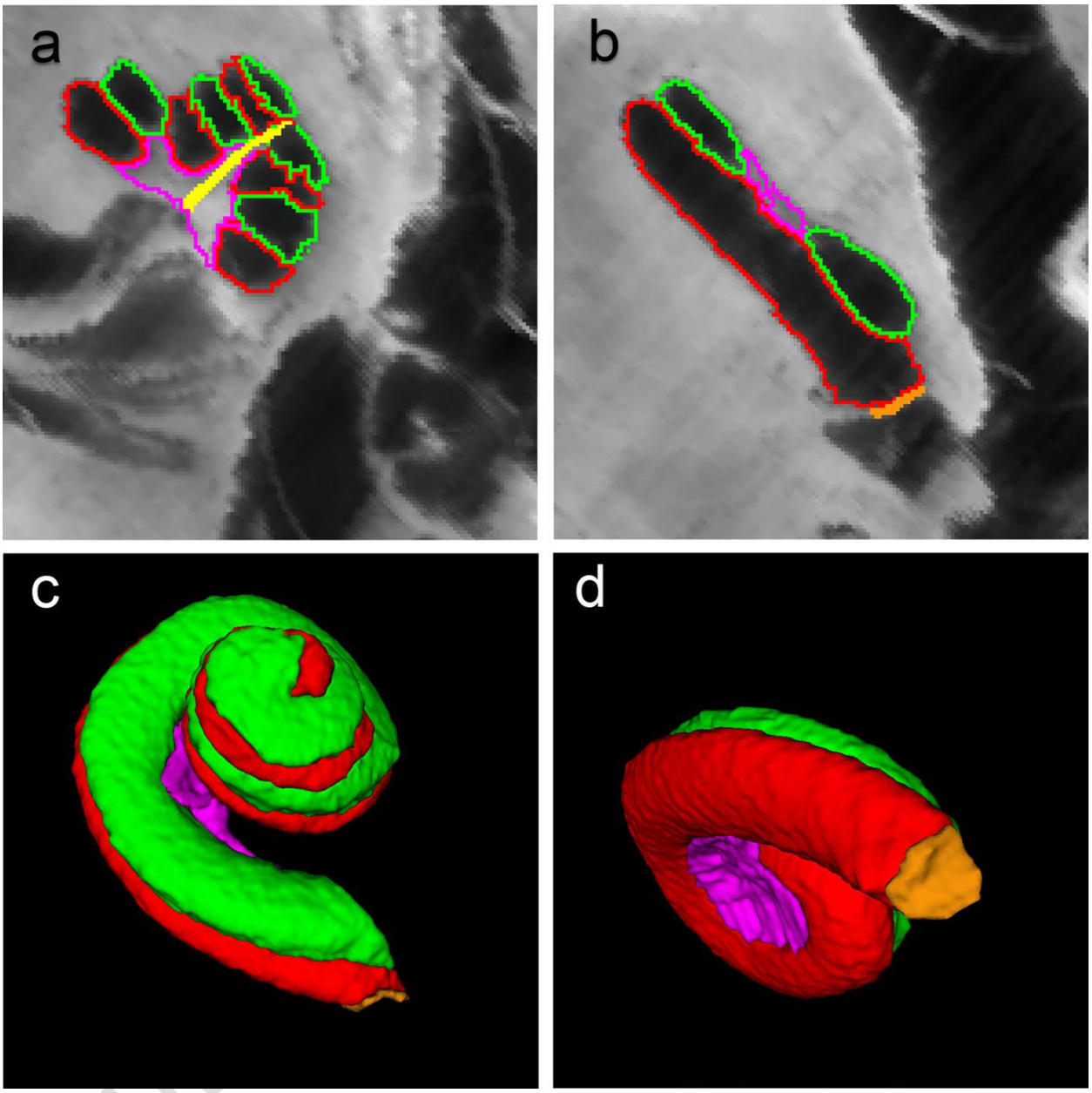
References

1. Holden LK, Finley CC, Firszt JB, Holden TA, Brenner C, Potts LG, Gotter BD, Vanderhoof SS, Mispagel K, Heydebrand G, Skinner MW (2013) Factors affecting open-set word recognition in adults with cochlear implants. *Ear Hear.* 34(3)342-60. https://doi.org/10.1097/AUD.0b013e3182741aa7
2. Chakravorti S, Noble JH, Gifford RH, Dawant BM, O'Connell BP, Wang J, Labadie RF (2019) Further Evidence of the Relationship Between Cochlear Implant Electrode Positioning and Hearing Outcomes. Otol Neurotol, **40**(5): p. 617-624. https://doi.org/10.1097/MAO.0000000000002204
3. Powell KA, Kashikar T, Hittle B, Stredney D, Kerwin T, Wiet GJ (2019) Atlas-based segmentation of temporal bone surface structures. *Int J Comput Assist Radiol Surg* 14(8):1267-1273. https://doi.org/10.1007/s11548-019-01978-2
4. Powell KA, Liang T, Hittle B, Stredney D, Kerwin T, Wiet GJ (2017) 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
5. Wiet G, Stredney D, Sessanna D, Bryan JA, Welling DB, Schmalbrock P (2002) Virtual temporal bone dissection: An interactive surgical simulator. Otolaryngol Head Neck Surg. 127(1):79-83. https://doi.org/10.1067/mhn.2002.126588.
6. Wiet GJ, Stredney D, Kerwin T, Hittle B, Fernandez SA, Abdel-Rasoul M, Welling DB (2012) Virtual temporal bone dissection system: OSU virtual temporal bone system: development and testing. Laryngoscope. 122 Suppl 1:S1-12. https://doi.org/10.1002/lary.22499
7. Sieber D, Erfurt P, John S, Santos GRD, Schurzig D, Sørenson MS, Lenarz T (2019) The OpenEar library of 3D models of the human temporal bone based on computed tomography and micro-slicing. *Sci Data*. 6:180297. https://doi.org/10.1038/sdata.2018.297
8. Klein S, Staring M, Murphy K, Viergever MA, Pluim JPW (2010) Elastix: a toolbox for intensity based medical image registration. *IEEE Transactions on Medical Imaging* 29:196 - 205. https://doi.org/10.1109/TMI.2009.2035616

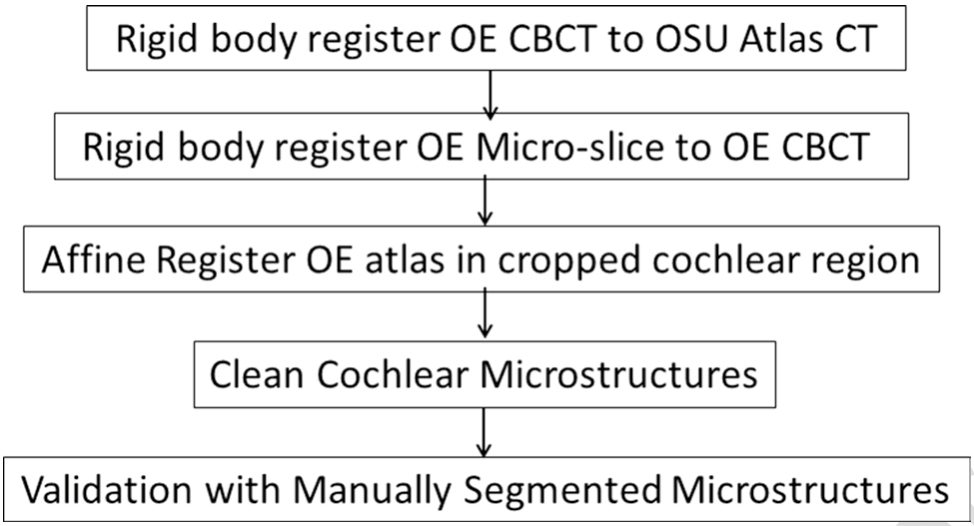
9. Shamonin DP, Bron EE, Lelieveldt BPF, Smits M, Klein S, Staring M (2014) Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease. *Frontiers in Neuroinformatics* 7:1-15. https://doi.org/10.3389/fninf.2013.00050
10. Klein S, Staring M. Elastix the manual, (2014) (http://elastix.isi.uu.nl/).
11. Yushkevich PA, Piven J, Cody Hazlett H, Gimpel Smith R, Ho S, Gee JC, Gerig G (2006) User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* 1;31(3):1116-28. https://doi.org/10.1016/j.neuroimage.2006.01.015
12. Liao PS, Chen TS, Chung PC (2001) A fast algorithm for multilevel thresholding. *Journal of Information Science and Engineering* 17:713-727.
13. Taha AA, Hanbury A (2015) Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Medical Imaging* 15(29):1-28. https://doi.org/10.1186/s12880-015-0068-x
14. Cignoni, C. Rocchini, R. Scopigno (1998) Metro: measuring error on simplified surfaces. *Computer Graphics Forum* 17 (2), 167-174.
15. Demarcy T, Vandersteen C, Guevara N, Raffaelli C, Gnansia D, Ayache N, Delingette H (2017) Automated analysis of human cochlea shape variability from segmented mCT images. *Computerized Medical Imaging and Graphics* 59: 1-12. https://doi.org /10.1016/j.compmedimag.2017.04.002
16. Kjer HM, Fagertun J, Wimmer W, Gerber N, Vera S, Barazzetti L, Mangado N, Ceresa M, Peilla W, Stark T, Stauber M, Reyes M, Weber S, Caversaccio M, Gonzalez Ballester, Paulsen RR (2018) Patient-specific estimation of detailed cochlear shape from clinical CT images. *International Journal of Computer Assisted Radiology and Surgery* 13:389–396. https://doi.org/10.1007/s11548-017-1701-7
17. Noble JH, Labadie RF, Majdani O, Dawant BM (2011) Automatic segmentation of Intra-Cochlear anatomy in conventional CT. I*EEE Trans Biomed Eng* 58(9):2625-2632. https://doi.org/10.1109/TBME.2011.2160262
18. Liu W, Atturo F, Aldaya R, Santi P, Cureoglu S, Obwegeser S, Gluekert R, Pfaller K, Schrott-Fischer A, Rask-Andersen H. (2015) Macromolecular organization and fine structure of the human basilar membrane - RELEVANCE for cochlear implantation. *Cell Tissue Res* 360:245–262. https://doi.org/ 10.1007/s00441-014-2098-z

19. Noble JH, Hedley-Williams AJ, Sunderhaus L, Dawant BM, Labadie RF, Camarata SM, Gifford RH (2016) Initial results with image-guided cochlear implant programming in children. Otol Neurotol 37(2):E63–E69. https://doi.org/10.1097/mao.0000000000000909
20. Verbist BM, Skinner MW, Cohen LT, Leake PA, James C, Boëx, Holden TA, Finley CC, Roland PS, Roland JT, Haller M, Patrick JF, Jolly CN, Faltys MA, Briaire JJ, Frijns JHM (2010) Consensus panel on a cochlear coordinate system applicable in histologic, physiologic, and radiologic studies of the human cochlea. Otol Neurotol. 31(5):722-730. https://doi.org/10.1097/MAO.0b013e3181d279e0
21. Avc E, Nauwelaers T, Lenarz T, Hamacher V, Kral A (2014) Variations in the microanatomy of the human cochlea. J Comp Neurol 522:3245–3261. https ://doi.org/10.1002/cne.23594
22. Lane JI, Witte RJ, Henson OW, Driscoll CLW, Camp J, Robb RA (2005) Imaging microscopy of the middle and inner ear. *Clin Ana* 18:409-415. https://doi.org/ 10.1002/ca.20152
23. Bellos C, Rigas G, Spiridon IF, Bibas A, Iliopoulou D, Böhnke F, Koutsouris D, Fotiadis DI (2014) Reconstruction of cochlea based on micro-CT and Histological images of the human ear. *Biomed Research International* 2014:1-7. https://doi.org/ 10.1155/2014/485783
24. Joskowicz L, Cohen D, Caplan N, Sosna J (2018) Inter-observer variability of manual contour delineation of structures in CT. Eur Radiol 29(3):1391–1399. https ://doi.org/10.1007/s00330-018-5695-5

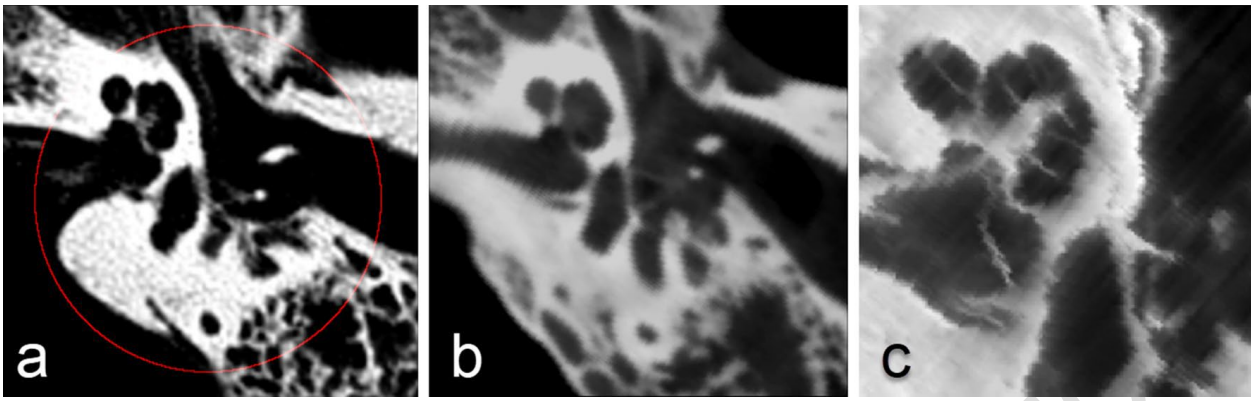
Figure Legends



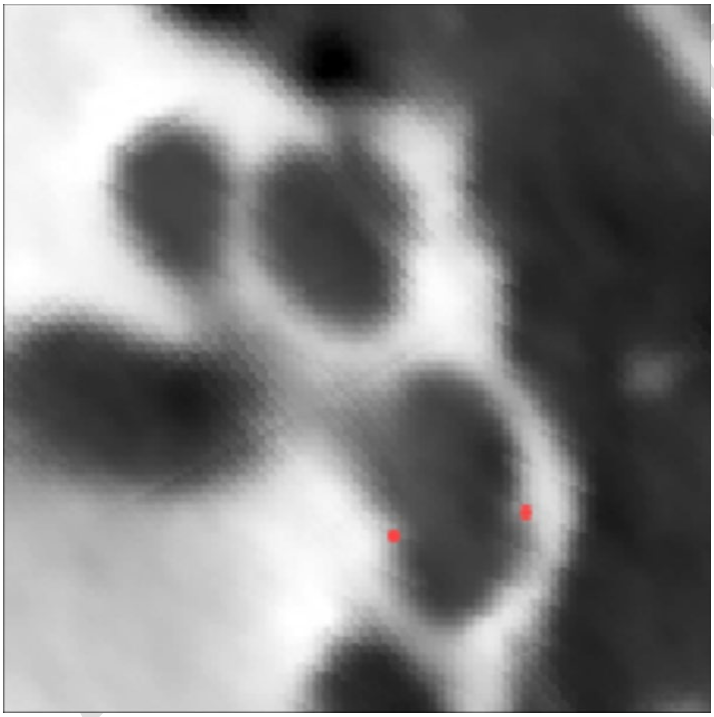
1. a. and b. 2D axial micro-slicing images of temporal bone with labeled structures: ST (red), SV (green), Mod (magenta), Mid-Mod (yellow), and RW (orange) **c. and d.** Segmented cochlear structures rendered in 3D.



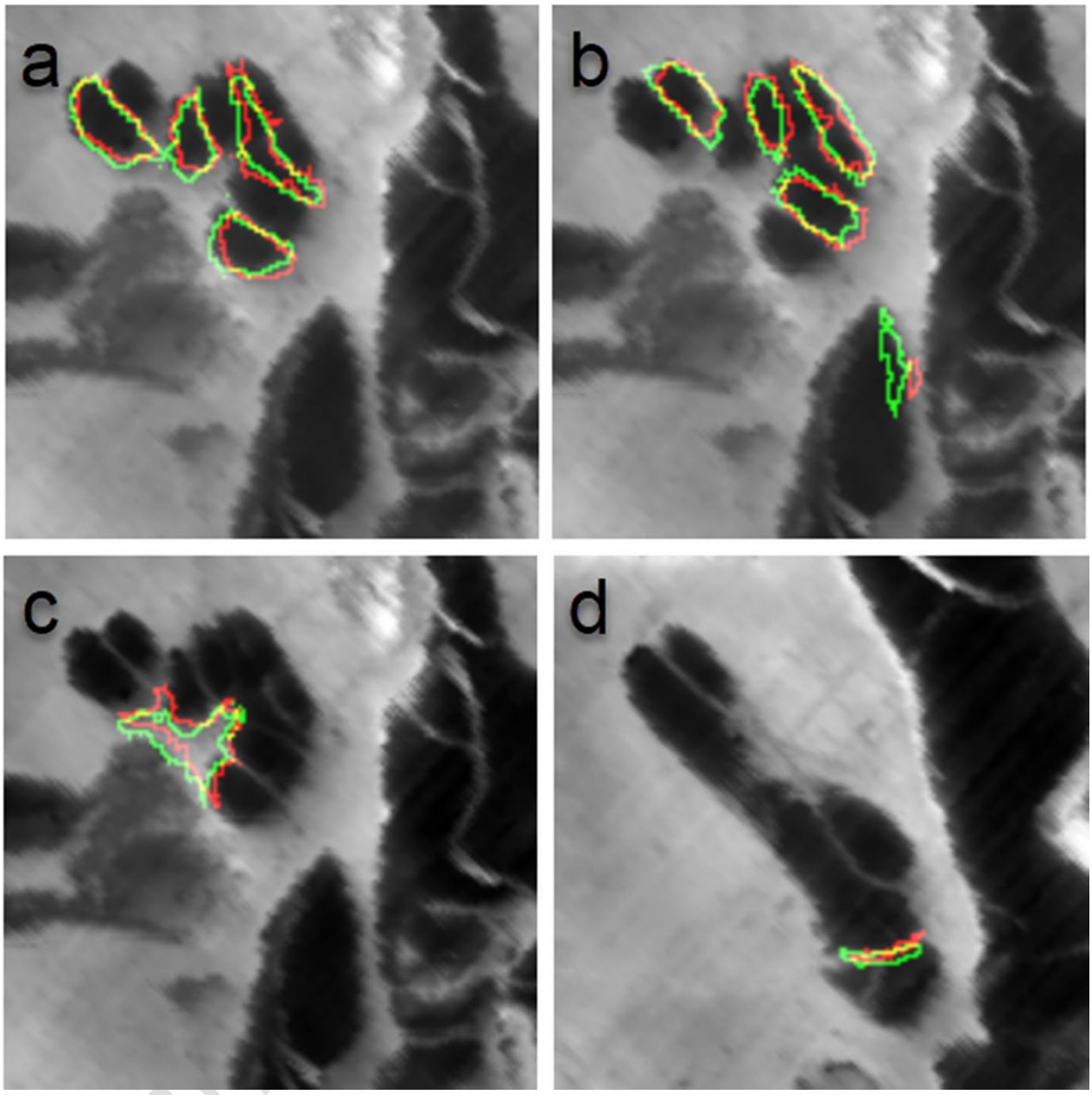
2. Atlas-based segmentation flowchart. The first two rigid body registrations are performed to orient the OpenEar (OE) datasets to the OSU atlas space. The affine registration of the OE atlas to the spatially oriented data is the for atlas base segmentation of the structures.



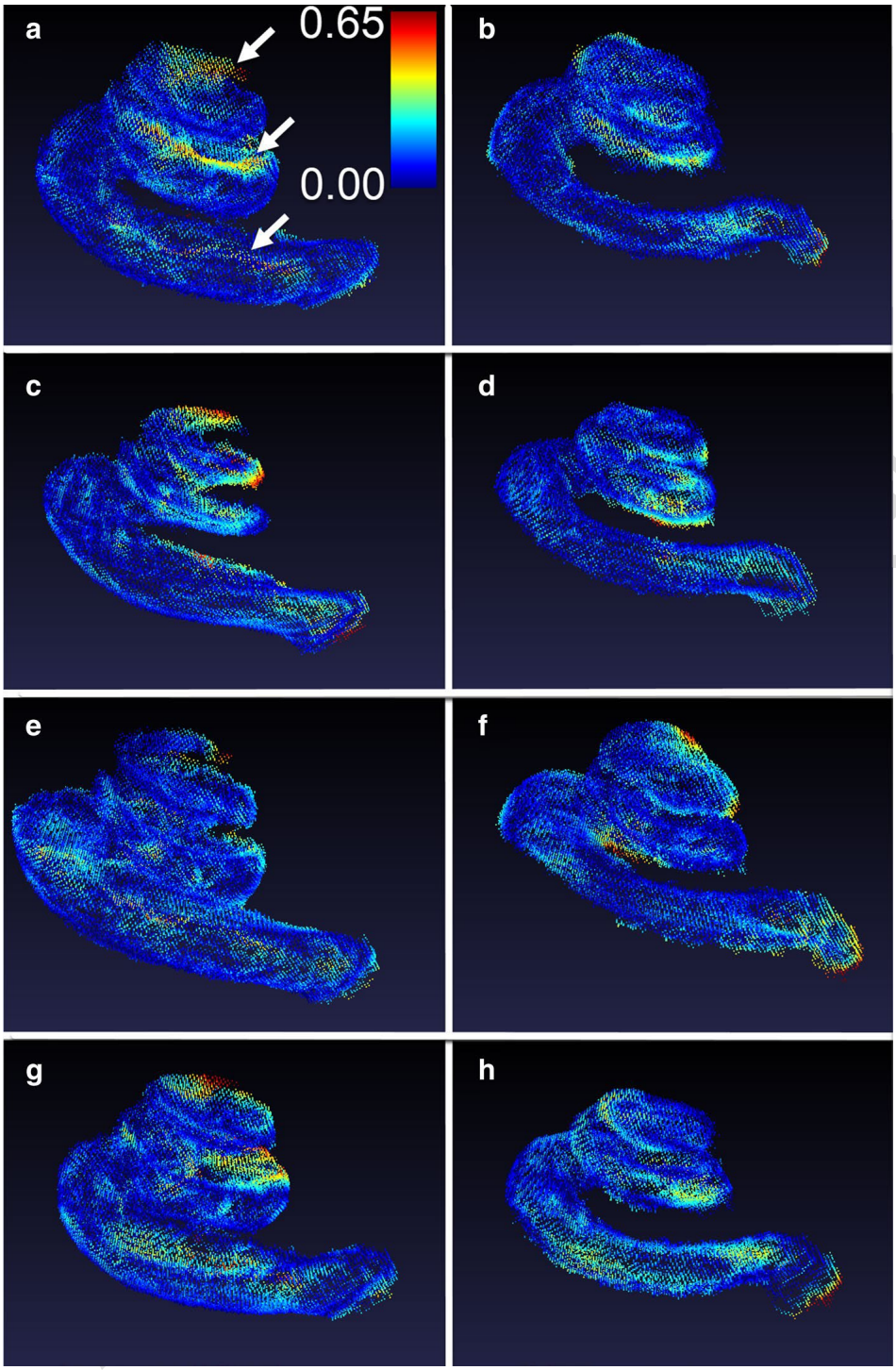
3. 2D axial slice of **a.** OSU atlas (red outline indicates region used for rigid body registration to OSU atlas), **b.** (OE) THETA CBCT spatially registered to OSU atlas, and **c.** cropped ROI of OE THETA micro-slice data spatially registered to OE THETA CBCT used for the affine registration to cropped spatially aligned OE CBCT images.



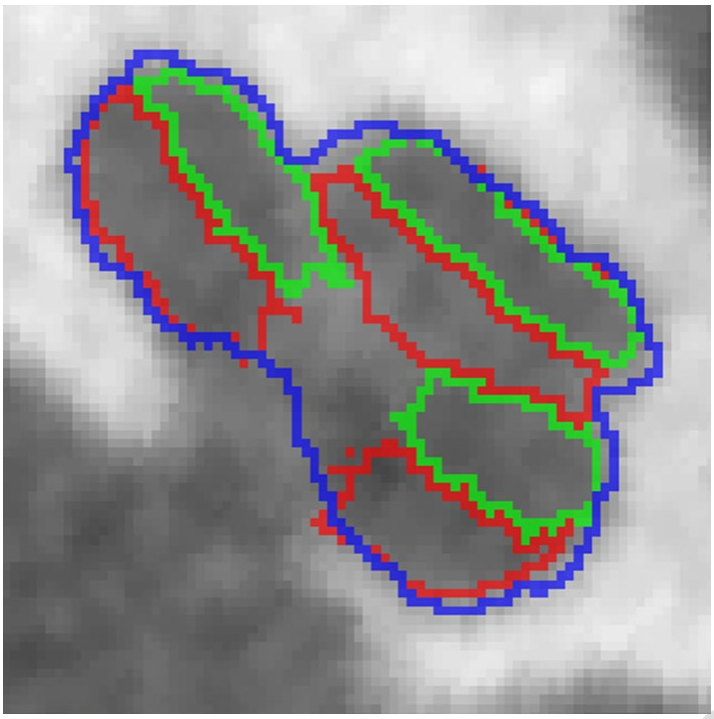
4. 2D axial CBCT image of cochlea illustrating automatically identified points (red) on round window opening.



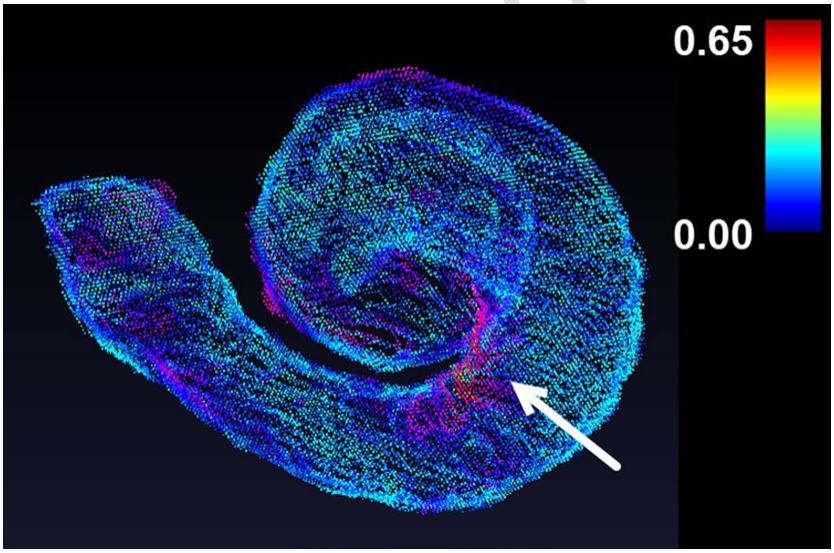
5. 2D axial micro-slice images with manually labeled structures (red) and automatically labeled structures (green). **a.** ST, **b.** SV, **c.** Mod, **d.** and RW.



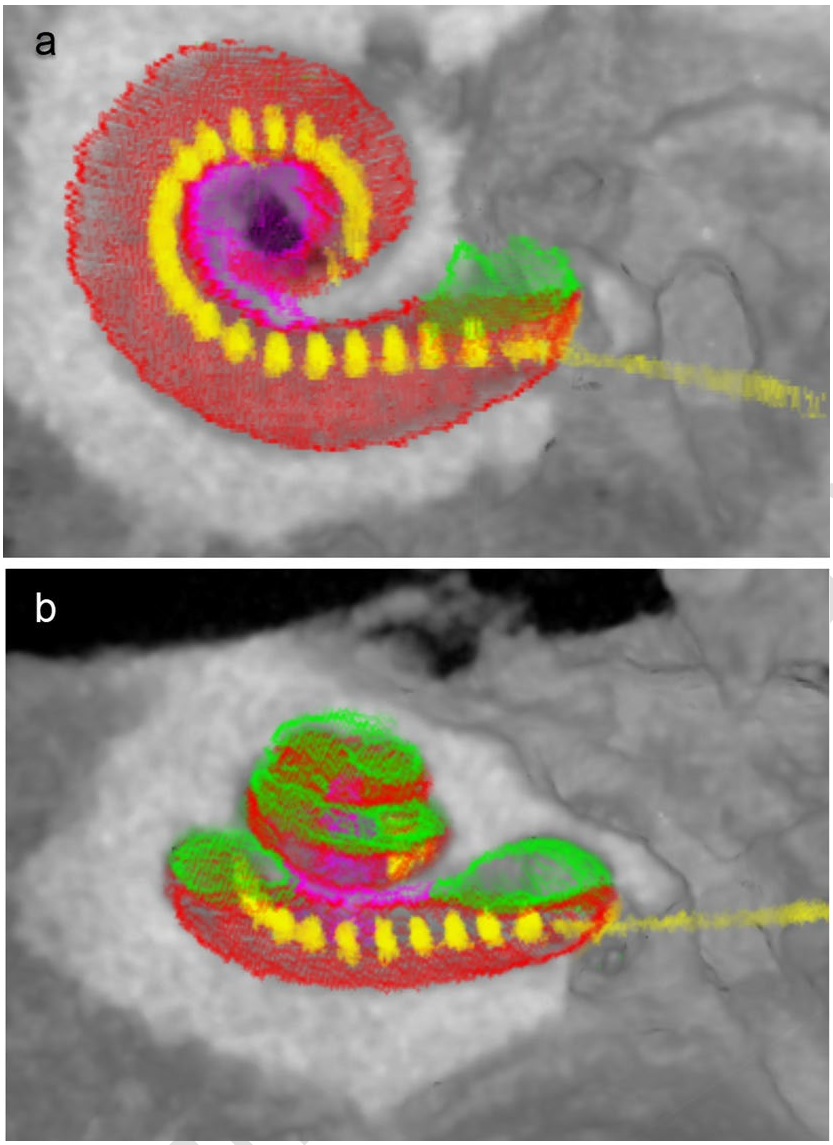
6. HD between automatically and manually segmented ST (**a** DELTA, **c** EPSILON, **e** ETA, **g** GAMMA) and SV (**b** DELTA, **d** PSILON, **f** ETA, **h** GAMMA). The white arrows in **a** represent the regions where we typically observed variations in HD in the range of 0.30–0.35 mm.



7. 2D axial clinical CBCT image of cochlea illustrating automatically identified scala (red = ST, green = SV) and manually identified cochlea (blue)



8. HD between automatically and manually segmented cochlea for one of the clinical CBCT images. The white arrow represents a region near the modiolus where we typically observed variations in HD in the range of 0.30–0.35 mm. This is mainly due to the limitation of manually tracing the cochlea in this region



9. 3D rendering of cochlear implant (yellow) in patient CBCT illustrating how the implant electrode enters through the round window (orange) and the electrodes wrap around the modiolus (magenta) and are placed within the ST (red) and do not breach the SV (green).

Table 1. Quantitative Validation Metrics - Mean (SD)

Structure	DICE	AVG HD (mm)	MAX HD (mm)	VS
Scala Tympani				
Delta	0.765	0.1	0.72	0.95
Epsilon	0.703	0.13	0.89	0.92
Eta	0.846	0.1	0.79	0.92
Gamma	0.759	0.12	0.73	0.89
Mean (s.d)	0.77 (0.06)	0.11 (0.01)	0.78 (0.08)	0.92 (0.03)
Scala Vestibuli				
Delta	0.69	0.11	0.65	0.97
Epsilon	0.73	0.12	0.65	0.97
Eta	0.80	0.11	0.85	0.91
Gamma	0.73	0.13	0.89	0.92
Mean (s.d)	0.74 (0.05)	0.12(0.01)	0.76(0.13)	0.95 (0.03)
Modiolus				
Delta	0.57	0.17	0.83	0.98
Epsilon	0.54	0.17	0.99	0.91
Eta	0.70	0.13	0.66	0.97
Gamma	0.56	0.2	1.39	0.92
Mean (s.d)	0.59 (0.07)	0.17(0.03)	0.97 (0.31)	0.95(0.04)
Round Window Membrane	**Distance between centroids (mm)**	**AVG HD (mm)**	**MAX HD (mm)**	
Delta	0.44	0.1	0.44	
Epsilon	0.41	0.1	0.56	
Eta	0.11	0.08	0.29	
Gamma	0.33	0.08	0.39	
Mean (s.d)	0.32 (0.15)	0.09 (0.01)	0.42(0.11)	
Mid Modiolus	**Distance between lines (mm)**	**Angular Rot (degrees)**		
Delta	0.15	5.9		
Epsilon	0.05	16.4		
Eta	0.01	6.5		
Gamma	0.25	10.2		
Mean (s.d)	0.11 (0.11)	9.8(4.8)		

Table 2. Quantitative Validation Metrics clinical CBCT- Mean (SD)

Structure	DICE	AVG HD (mm)	MAX HD (mm)	VS
Cochlea	0.87 (0.02)	0.11 (0.01)	0.63 (0.02)	0.98 (0.02)
	Distance between centroids	**AVG HD (mm)**	**MAX HD (mm)**	
RW membrane	0.24 (0.18)	0.12 (0.07)	0.28 (0.06)	
	Distance between lines	**Angular rotation (degrees)**		
Mid-Modiolus	0.11 (0.06)	8.46 (5.48)		