Practicable Assessment of Cochlear Size and Shape from Clinical CT Images

A. H. Gee, Y. Zhao, G. M. Treece and M. L. Bance

Technical Report ENG TR.004
ISSN 2633-6839
27 July 2020

Cambridge University Engineering Department
Trumpington Street
Cambridge CB2 1PZ
England

Corresponding e-mail: ahg@eng.cam.ac.uk

Abstract

There is considerable interpersonal variation in the size and shape of the human cochlea, with evident consequences for cochlear implantation. The ability to characterize a specific cochlea, from pre-operative computed tomography (CT) images, would allow the clinician to personalize the choice of electrode, surgical approach and post-operative programming. In this study, we present a fast, practicable and freely available method for estimating cochlear size and shape from clinical CT. The approach taken is to fit a template surface to the CT data, using either a statistical shape model (SSM) or a locally affine deformation (LAD). After fitting, we measure cochlear length, mean lumen area and a novel measure of basal turn non-planarity, which is expected to correlate with the risk of insertion trauma. Gold-standard measurements from a convenience sample of 18 micro-CT scans are compared with the same quantities estimated from low resolution, noisy, pseudo-clinical data synthesized from the same micro-CT scans. The results show that the LAD method outperforms the SSM method, with an expected error of around 11% of the gold-standard sample range for non-planarity and cochlear length. Results for lumen area are significantly worse, reflecting the difficulty of detecting the cochlear inner wall in clinical CT.

1 Introduction

The size and shape of an individual human cochlea are of profound interest when planning cochlear implant surgery. Knowing the size of the cochlea would enable the clinician to make an informed choice of electrode array and insertion depth, one aim being to preserve any residual, low frequency natural hearing, which can be beneficial in difficult listening conditions (Dhanasingh and Jolly, 2017; Sheffield et al., 2015). Over-insertion runs the risk of destroying residual hearing, while under-insertion may result in insufficient coverage of the frequency components required for good "electric hearing" through the implant (Lenarz et al., 2019). Additionally, since the cochlea is tonotopically organized, if the electrodes and their assigned frequencies can align with the natural tonotopic arrangement, speech recognition may be enhanced (Canfarotta et al., 2020).

Beyond cochlear size, a more nuanced knowledge of cochlear shape would allow the clinician to be forewarned of potentially traumatic insertion contact with the basillar membrane and the lateral wall. The immunological and fibrosis events that follow such trauma may damage the neural structures that respond to electrical stimulation. The nature, force and site of insertion contact will be determined by the shape of the cochlea, particularly in places where the array has to bend (Avci et al., 2017; De Seta et al., 2017; Meng et al., 2016; Nguyen et al., 2012). Unfortunately, this degree of patient-specific planning is difficult to achieve given the relatively low resolution of current, pre-operative computed tomography (CT) imaging. The human cochlea has outer dimensions of approximately 10 mm $\times$ 8 mm $\times$ 4 mm (Kjer et al., 2016), which corresponds to only 33 $\times$ 27 $\times$ 13 voxel widths in typical 0.3 mm clinical CT imaging.

High resolution morphometry of the inner ear is achievable through the use of cadaveric temporal bone specimens, which are either subject to micro-CT imaging with isotropic voxel dimensions of around 20 µm (Iyaniwura et al., 2018), or used to produce exquisite corrosion casts of the ductal structures which are then photographed and measured (Erixon et al., 2009). The emphasis of many of these studies is on measurements that are of relevance to cochlear implantation, these falling naturally into three broad categories: measurements of the cochlea's overall size (total coiling angle, diameter, length, cross-sectional ductal area) (Biedron et al., 2009; Escudé et al., 2006; Iyaniwura et al., 2018; Singla et al., 2015); measurements of the cochlea's "vertical" trajectory in the direction of the modiolar axis (Avci et al., 2014; Meng et al., 2016; Pietsch et al., 2017); and local measurements at the round window that constrain the initial insertion and bending angles of the electrode, when not performing a cochleostomy (Rask-Andersen et al., 2012; Tang et al., 2018).

Efforts to measure cochlear morphology from low-resolution, clinical CT images fall into two broad categories. Two-dimensional methods attempt to infer cochlear size and shape from a set of discrete measurements taken in specific planes (Rivas et al., 2017; Würfel et al., 2014). However, Koch et al. (2017) cast doubt on the accuracy of such measurements and suggest that full, three-dimensional analysis is preferable. Falling into this second category is the work of Noble et al. (2011), who built a statistical shape model (SSM) of the cochlea using six micro-CT scans of cadaveric temporal bones, and then assessed how well the SSM could be fitted to low-resolution, clinical scans of five of the same bones. Evaluation metrics were based on Dice similarity coefficients and surface errors, so it is unclear how well this method can estimate surgically

relevant parameters. Kjer et al. (2018) developed a similar approach using a statistical deformation model, reporting measurement accuracy and precision for cochlear length, width and height in addition to surface errors, but with no consideration of vertical trajectories. van der Jagt et al. (2017) describe an automatic, three-dimensional tracing method that was used to estimate inner and outer wall radii, duct diameter and vertical trajectory in low-resolution CT scans of 242 patients. Significant variation was observed in the cohort, but there was no validation against gold-standard measurements. Iyaniwura et al. (2018) describe a method to fit a grayscale cochlear atlas to low-resolution, clinical CT data using sequential landmark, affine and B-spline registration. Evaluation was performed using 20 specimens scanned at micro-CT and clinical CT resolutions. Gold-standard, micro-CT "A-values", which correlate well with cochlear duct length (Escudé et al., 2006), were compared with A-values derived from the fitted atlas and also A-values estimated by experts on the clinical CT images. There was no consideration of vertical trajectories.

In this work, we describe a fast, simple and freely available method to fit surface models of the otic capsule to CT data. The fitting may be directed by a statistical model, in the spirit of Noble et al. (2011) and Kjer et al. (2018), or constrained only by a smoothness criterion, in the spirit of Iyaniwura et al. (2018). We compare the performance of the two approaches, with specific reference to three-dimensional, surgically relevant measurements like those considered by van der Jagt et al. (2017). Particular emphasis is placed on an improved metric for characterizing the different vertical trajectories first described by Avci et al. (2014). Validation is by way of pseudo-clinical CT data synthesized from the original, gold-standard micro-CT images.

2 Materials and Methods

Temporal bone specimens and micro-CT scanning

A convenience sample of 18 human temporal bones was provided by the Department of Physiology, Development and Neuroscience at the University of Cambridge. The donors had provided consent before decease for the use of their bodies for anatomical research, in compliance with the UK Human Tissue Act 2004. The specimens were scanned using a Nikon Metrology XT H 225 ST micro-CT scanner (Nikon Metrology NV, Leuven, Belgium) at 125 kV, 120 µA, 1080 projections, 2 frames per projection and 1 s exposure time. Reconstruction was at an isotropic voxel resolution of around 25 µm, apart from specimen #18, which was a larger bone section reconstructed at 61 µm. The bones were a mixture of left and right sides and were all unimplanted apart from specimen #17, which was implanted before scanning (as part of a different study) and whose scans therefore suffered significant beam-hardening artefacts.

Construction of the template and statistical shape models

Figure 1, steps 1 and 2, show how the 18 micro-CT scans were segmented and the otic capsules represented as triangulated surface meshes. A template surface was then constructed as follows. One of the 18 specimens was selected, by eye, as being the most "average". The chosen mesh was registered to all 18 specimens, and the mean deformation was calculated and applied at each vertex, producing a mean otic capsule surface. This surface was re-triangulated to a reasonable resolution (11145 vertices), the resulting mesh serving as the template for all remaining experiments in this paper. Steps 3 and 4 show how the template mesh was then registered to each specimen using the sliding semilandmark algorithm, originally developed for planar morphometry (Bookstein, 1991, 1997) and subsequently extended to surfaces (Gunz et al., 2005). Segmentation and mesh construction were performed using Stradview[1], while surface registration was carried out in wxRegSurf[2], both of which are available for free download.

Following registration, the $n = 18$ sets of deformed template vertex coordinates were standardized for location, orientation and scale using Procrustes analysis (Goodall, 1991). This involves translating each specimen to a common origin, scaling to unit centroid size, and then rotating to minimize the sum of the squared distances between the vertices of each specimen and the undeformed template mesh. We then rescaled each

[1] `mi.eng.cam.ac.uk/Main/StradView`

[2] `mi.eng.cam.ac.uk/~ahg/wxRegSurf`

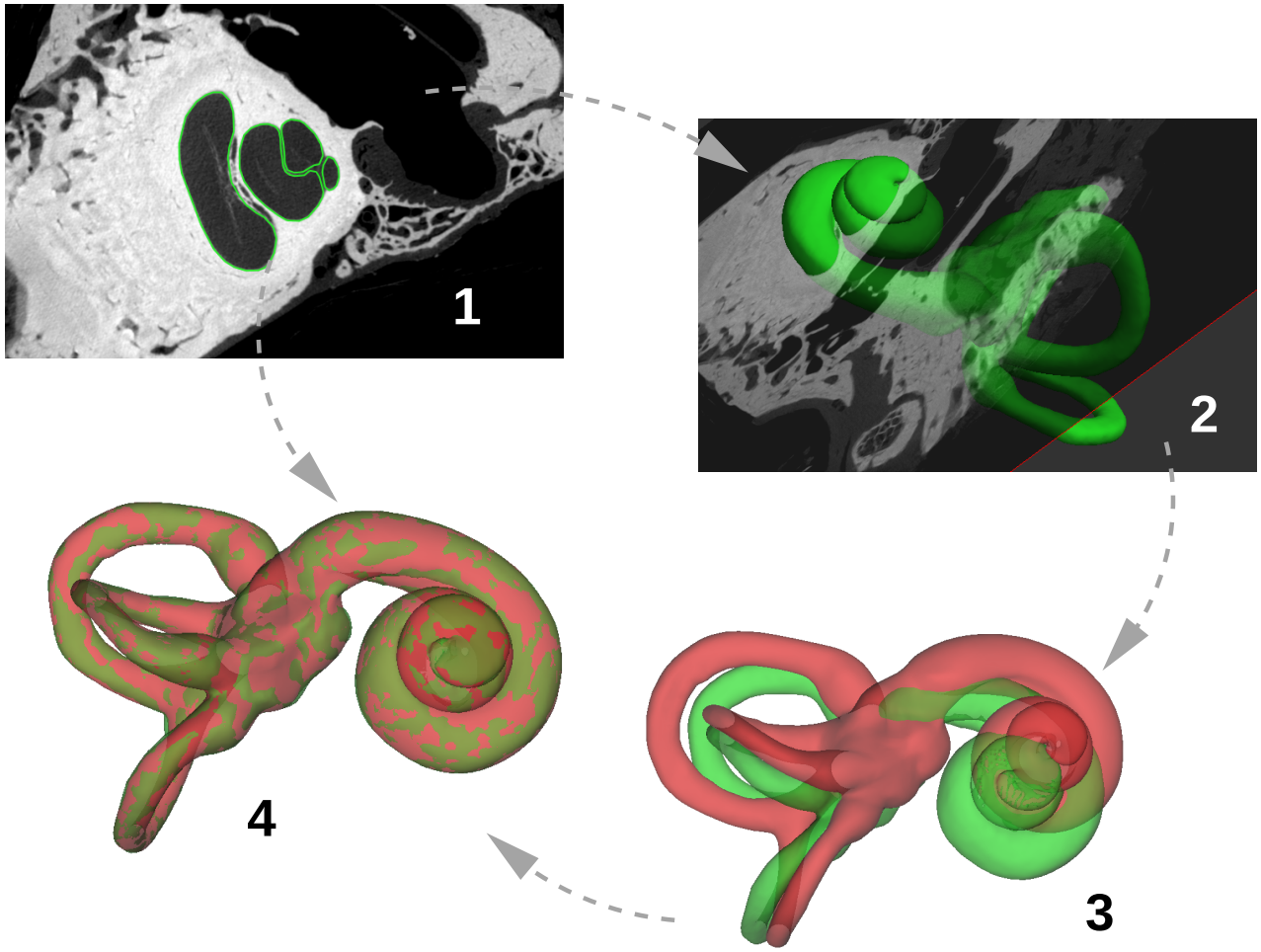
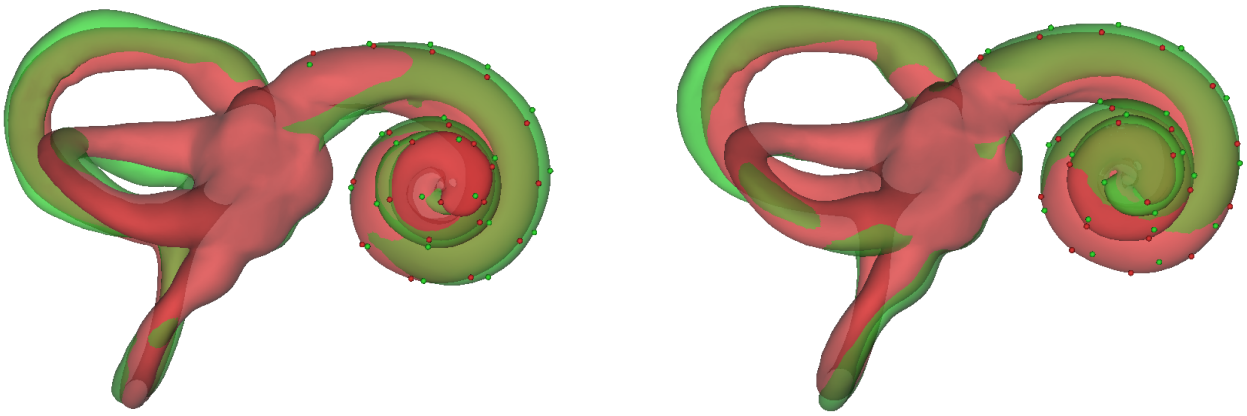


Figure 1: Constructing otic capsule models. (1) The micro-CT scans were segmented in Stradview by simple thresholding followed by manual tidying up of the contours. (2) Stradview was then used to construct triangulated surface meshes of each specimen. (3) Since each mesh has a different number of triangles and vertices, the next step is to align a common template mesh (red) with each specimen (green). This allows statistical analysis of the deformation at each of the template's vertices, and subsequent construction of an SSM. (4) The alignment involves translation, rotation, isotropic scaling and (if necessary) reflection, followed by a nonrigid thin-plate spline deformation. The deformation was computed in wxRegSurf using the sliding semilandmark algorithm.



(a) Surface sliding semilandmarks only (b) Homologous landmarks on the outer cochlear wall

Figure 2: Statistical shape models. (a) shows the fourth mode, ± 3 standard deviations, of a statistical shape model constructed using surface sliding semilandmarks only. Variations in cochlear coiling are partially represented by this mode, as a local extension of the duct at the apex. (b) shows the first mode, ± 1.35 standard deviations, of an SSM constructed using surface sliding semilandmarks and 20 homologous point landmarks (the green and red spheres) placed at regular intervals along the outer cochlear wall. Variations in cochlear coiling are represented by this mode as a global elongation of the duct from the round window to the apex (compare the tangential locations of the red and green spheres).

specimen's vertex coordinates by its centroid size, and used principal component analysis to build a point-based SSM from the resulting n sets of coordinates. Let $\mathbf{X}_i$ be the 33435-element vector formed by concatenating the coordinates of individual i, and let $\hat{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i$. Then the principal modes of shape variation are the $n-1$ eigenvectors $\mathbf{m}_i$ of the sample covariance matrix $\frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{X}_i - \hat{\mathbf{X}})(\mathbf{X}_i - \hat{\mathbf{X}})^T$ with corresponding non-zero eigenvalues. In SSM-based segmentation, the surfaces of new specimens are restricted to anatomically plausible shapes by representing the mesh as a linear combination of the shape modes

$$\mathbf{X} = \hat{\mathbf{X}} + \sum_{i=1}^{n-1} S_i \mathbf{m}_i \tag{1}$$

where S_i are referred to as *shape coefficients*.

The nature of an SSM is entirely dependent on the choice of surface registration algorithm used to align the template with each specimen. The sliding semilandmark algorithm of Gunz et al. (2005) is highly flexible in this regard. At one extreme, the template and specimen may be aligned using surface sliding semilandmarks only, the effect being to warp the template onto the specimen with minimum bending energy and no further constraints on the tangential alignment. This produces a low variance SSM in which, for example, extra coiling of the cochlea is represented by local deformation at the apex, as shown in Figure 2(a). However, it can be argued (Gunz et al., 2012; Kjer et al., 2016) that such models do not capture the true nature of the inter-subject variation, since they do not account for correspondence between distinguished points and curves on the template and each specimen.

The sliding semilandmark algorithm can accommodate such point and curve homologies. For example, Figure 2(b) shows the first mode of an alternative SSM in which 20 point landmarks were placed at regular intervals along the outer cochlear wall of the template and each individual, and forced into correspondence during registration. This results in a much higher variance SSM with significant deformation tangential to the surface, as is evident by the wrapping and unwrapping of the entire cochlea in Figure 2(b). While such models are essential for meaningful shape-based classification in taxonomic studies, in this work we simply wish to represent the surfaces of our specimens with as little variance as possible, since any extra, tangential variance could result in implausible surfaces when fitting to noisy data. We therefore prefer the model built using surface sliding semilandmarks only, as in Figure 2(a). This model is available for free download as part of the Stradview package.

Synthesis of pseudo-clinical CT data

The micro-CT data, and the otic capsules segmented from them, provide the gold-standard measurements for the experiments in this paper. For the clinical measurements, we synthesized pseudo-clinical CT images from the micro-CT data. We achieved this by downsampling the micro-CT until the desired clinical resolution was achieved. We then projected the downsampled data into the CT detector space, in effect recovering the sinogram, added Gaussian noise to the sinogram, and then backprojected the noisy data into the world space. All of this processing was performed using wxDicom[3].

We synthesized three different classes of pseudo-clinical data, which we shall refer to as standard multi-detector CT (MDCT) (isotropic voxel dimension 0.3 mm), poor MDCT (isotropic voxel dimension 0.45 mm) and next-generation cone beam CT (CBCT) (isotropic voxel dimension 0.15 mm): see Figure 3. The level of Gaussian noise was adjusted by trial and error until the images resembled reference images from the literature[4]. For example, the standard MDCT images in Figure 3(b) resemble the exemplar clinical images in Figure 1 (c,e,g) of Phillips et al. (2012), while the poor MDCT images in Figure 3(c) are noticeably worse. The next generation CBCT images in Figure 3(a) are superior to those currently found in clinical practice, but resemble the state-of-the-art research images in Zou et al. (2017).

Fitting the model to CT data

Figure 4 shows the process of fitting the otic capsule model to new CT data. The data in Figure 4 is pseudo-clinical CT data, though the method is equally applicable to micro-CT data. The model-fitting process is freely available as part of the Stradview package and is designed to be clinically practicable, in that it requires around one minute of expert interaction, followed by at most a couple of minutes of computation.

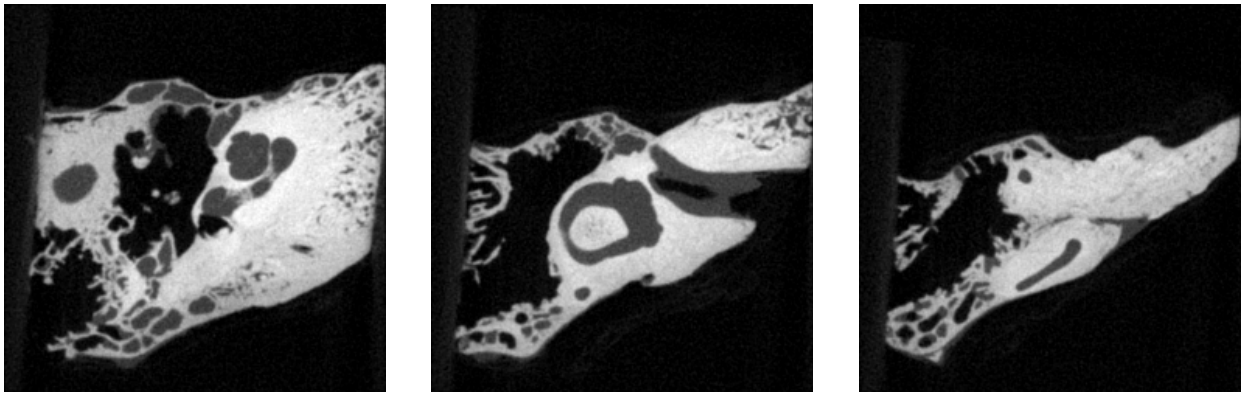
The first step is to position the template surface at approximately the correct location, by manually identifying three point landmarks in the data: the cochlear apex, the centre of the oval window and the posterior-anterior canal bifurcation (Figure 4, step 1). Stradview then computes the similarity transformation (rotation, translation and isotropic scaling) that best aligns these three points in the data with corresponding points predefined on the template mesh (Figure 4, step 2). The operator can optionally reflect the template in the plane of the three points, if the left-right fit was incorrect. The final manual interaction is to select an appropriate grayscale threshold to segment the boundary of the otic capsule (Figure 4, step 3). The thresholded contours define the point cloud (Figure 4, step 4) to which the model is now fitted automatically.

An initial, approximate alignment is computed using the iterative closest point (ICP) approach of Besl and McKay (1992). This approximate alignment is parameterized by a second similarity transformation (Figure 4, step 5). There follows a further iterative process to compute the additional, local displacement of each template vertex (Figure 4, step 6). Since the thresholded data is noisy (structures other than the otic capsule are captured, and some of the boundaries of the otic capsule, especially at the inner wall and the round and oval windows, are lost), this nonrigid registration must be regularized, to prevent overfitting of the model to the noise. Stradview offers two methods for regularized, nonrigid registration.

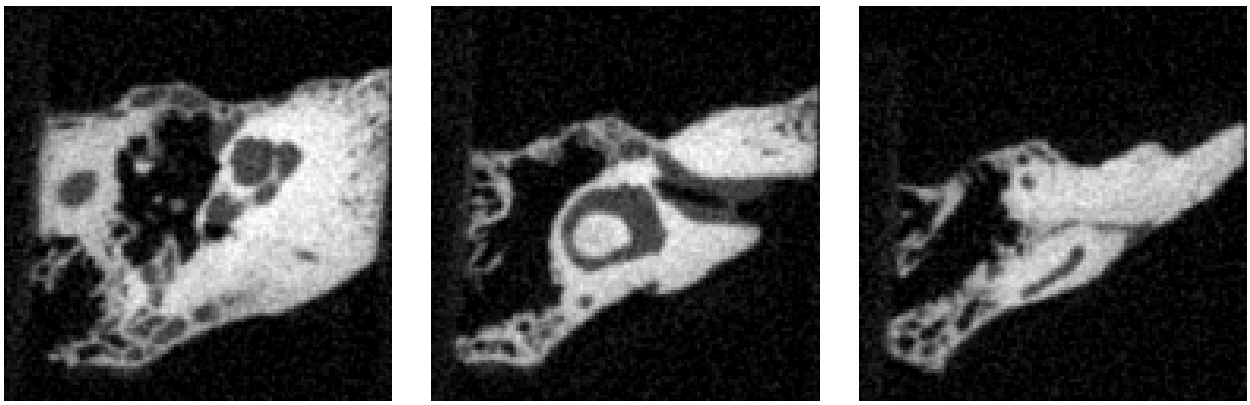
The first is the locally affine registration algorithm of Feldmar and Ayache (1996). Associated with each vertex k of the template is a set of neighbouring vertices N_k, where each member of N_k lies within a distance d of vertex k. At iteration i, every vertex on the template is paired with the closest point in the cloud. Then, for each vertex k on the template, the rigid transformation $\boldsymbol{R}_{k,i}$ is found that minimizes the sum of the squared distances between the transformed vertices in N_k and their partners in the point cloud. The local displacement of vertex k is then set to a proximity-weighted average of all the rigid transformations $\boldsymbol{R}_{k,i}$ within N_k. At iteration $i+1$, the closest neighbours and consequent rigid transformations $\boldsymbol{R}_{k,i+1}$ are recomputed, and so on, until convergence. d is the algorithm's only parameter, its effect being to control the amount of allowable deformation. Smaller values of d permit more deformation and closer alignment of the template to the point

[3] `mi.eng.cam.ac.uk/Main/GMT_wxDicom`

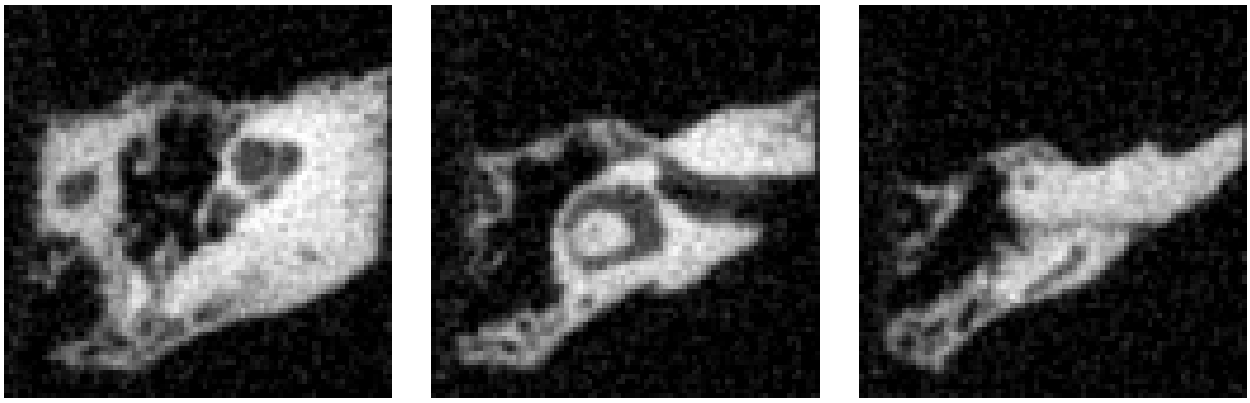
[4] There is no convenient way to quantify the noise levels other than relative to each other. Expressed arbitrarily as the value of wxDicom's *Detection Noise* slider, the levels were 60 dB for the next-generation CBCT data, 70 dB for the standard MDCT data and 75 dB for the poor MDCT data.



(a) Next generation pseudo-clinical CBCT (0.15 mm voxel dimension, 60 dB added noise).



(b) Standard pseudo-clinical MDCT (0.3 mm voxel dimension, 70 dB added noise).



(c) Poor pseudo-clinical MDCT (0.45 mm voxel dimension, 75 dB added noise).

Figure 3: Pseudo-clinical data. The images in (b) are of comparable quality to those found in current clinical practice, see for example Figure 1 (c,e,g) of Phillips et al. (2012).

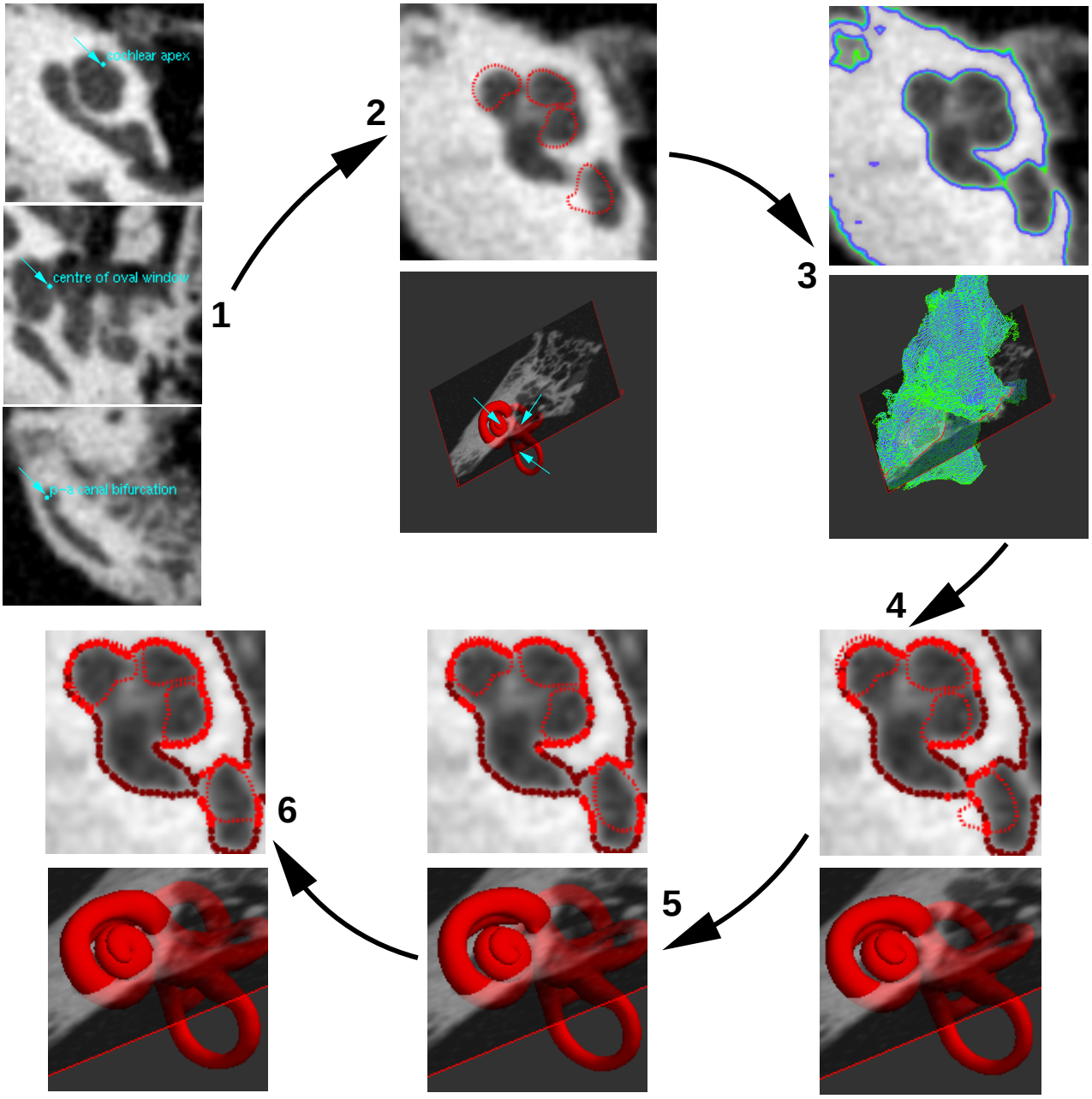


Figure 4: Fitting a model to CT data. (1) The operator scrolls through the axial CT images and places landmarks at the cochlear apex, the centre of the oval window and the posterior-anterior canal bifurcation (at the tip of the crus commune). (2) These landmarks are matched with corresponding, pre-defined landmarks on the model. An initial, approximate alignment is provided by the similarity transformation (rigid body plus uniform scaling) that minimizes the sum of the squared distances between the three pairs of landmarks. (3) The operator selects a suitable segmentation threshold: this subfigure shows contours thresholded at grayscale 160 (green), 170 (cyan) and 180 (purple). (4) The contour vertices at the selected threshold (170 in this example) provide the point cloud to which the model is fitted. Points closest to a model vertex are displayed in bright red, other points in dark red. (5) Iterative closest point (ICP) registration of the model to the point cloud, with a similarity transformation. (6) Finally, the fit is refined using ICP registration with a nonrigid transformation, either a statistical shape model or a locally affine deformation.

cloud, while larger values of d favour smooth displacement fields over alignment accuracy. We shall refer to this algorithm using the acronym LAD (Locally Affine Deformation).

In Stradview's second method, the nonrigid deformation is governed by the SSM, with the template's vertices constrained according to Equation 1. The registration again proceeds within an ICP framework. Each of the template's vertices is paired with the closest point in the cloud. Then, the SSM shape coefficients S_i are found that minimize the sum of the squared distances between the deformed template vertices and their partners in the point cloud. At iteration $i + 1$, the closest neighbours and consequent shape coefficients S_i are recomputed, and so on, until convergence. This algorithm is parameter-free, since we use all the available SSM modes in Equation 1.

Clinically relevant shape and size measurements

Figure 5 summarises the three measurements we make on the cochlear surfaces, to compare the similarity of the meshes fitted to pseudo-clinical CT data with their gold-standard counterparts. All three measurements are made on the first 270° of the basal turn[5]. Limiting the measurements to this angular range avoids any influence of the middle and apical turns, which are of little relevance to most current implant designs (Dhanasingh and Jolly, 2017; Pelliccia et al., 2014). The 270° range covers the most common sites of insertion trauma (Avci et al., 2017; De Seta et al., 2017; Meng et al., 2016; Nguyen et al., 2012).

Mean lumen area is measured on gross cross-sections of the cochlea, with no attempt to isolate the scala tympani. The remaining two measurements are made on the curve that delineates the cochlear outer wall. The first, which we shall refer to as "reach" (and is equivalent to common measures of cochlear length, including the "A-value" of Escudé et al. (2006)), is the distance from the round window to the furthest point on the curve. It characterizes the cochlea's size, focusing on the clinically relevant basal turn. The second concerns the cochlea's vertical trajectory, in which a down-then-up "rollercoaster" profile was identified by Avci et al. (2014) as a potential risk factor for insertion trauma. However, Demarcy et al. (2017) observed that vertical trajectories are sensitive to the definition of "vertical", which is normally taken to be the modiolar axis (Avci et al., 2014). We further explore this point in Figure 6, which shows two vertical trajectories of the *same* cochlea, with the vertical axis defined by the modiolar axis in (a) and the normal to the basal plane in (b). We note not only the previously observed sensitivity to the vertical direction, but also that the rollercoaster profile in (a) does not necessarily imply a challenging insertion.

We therefore propose an alternative way to characterize the vertical trajectory. We define the "basal plane" as the best fit plane to the first 270° of the outer wall contour. Vertical trajectories are measured along the normal to this plane, as in Figure 6(b), thus avoiding any sensitivity to the less germane anatomy of the middle and apical turns. Defining the vertical direction in this way is consistent with the consensus approach to cochlear coordinate systems (Verbist et al., 2010). Having established a reliable vertical trajectory, Figure 5 illustrates how we summarise the "non-planarity" of the basal turn as the mean absolute distance between the first 270° of the outer wall contour and the basal plane. The hypothesis is that cochleas with lower non-planarity are less susceptible to insertion trauma than those with higher non-planarity.

3 Experiments, results and discussion

The otic capsule model was fitted to each of the 54 pseudo-clinical scans (18 specimens, 3 different resolutions) at grayscale thresholds of 160, 170 and 180. The threshold of 170 was observed to produce visually appropriate segmentations in most cases, with ± 10 re-runs to assess sensitivity. For each data set at each threshold, the model was fitted three times: using the "full" SSM, trained using all 18 micro-CT data sets; using a "leave-one-out" SSM, trained using 17 of the micro-CT data sets, but not the specimen on which it was being evaluated; and using the LAD method, with a fixed parameter $d = 5$ mm.

[5] Since we do not detect the modiolus in this study, the 270° angle is not measured in the usual polar coordinate system defined by the modiolus (origin) and the round window (0°). Instead, we consider the angle through which the tangent to the outer wall contour has turned with respect to its initial trajectory at the round window. 270° in this paper's notation corresponds to somewhat more than 270° in round window/modiolar polar coordinates.

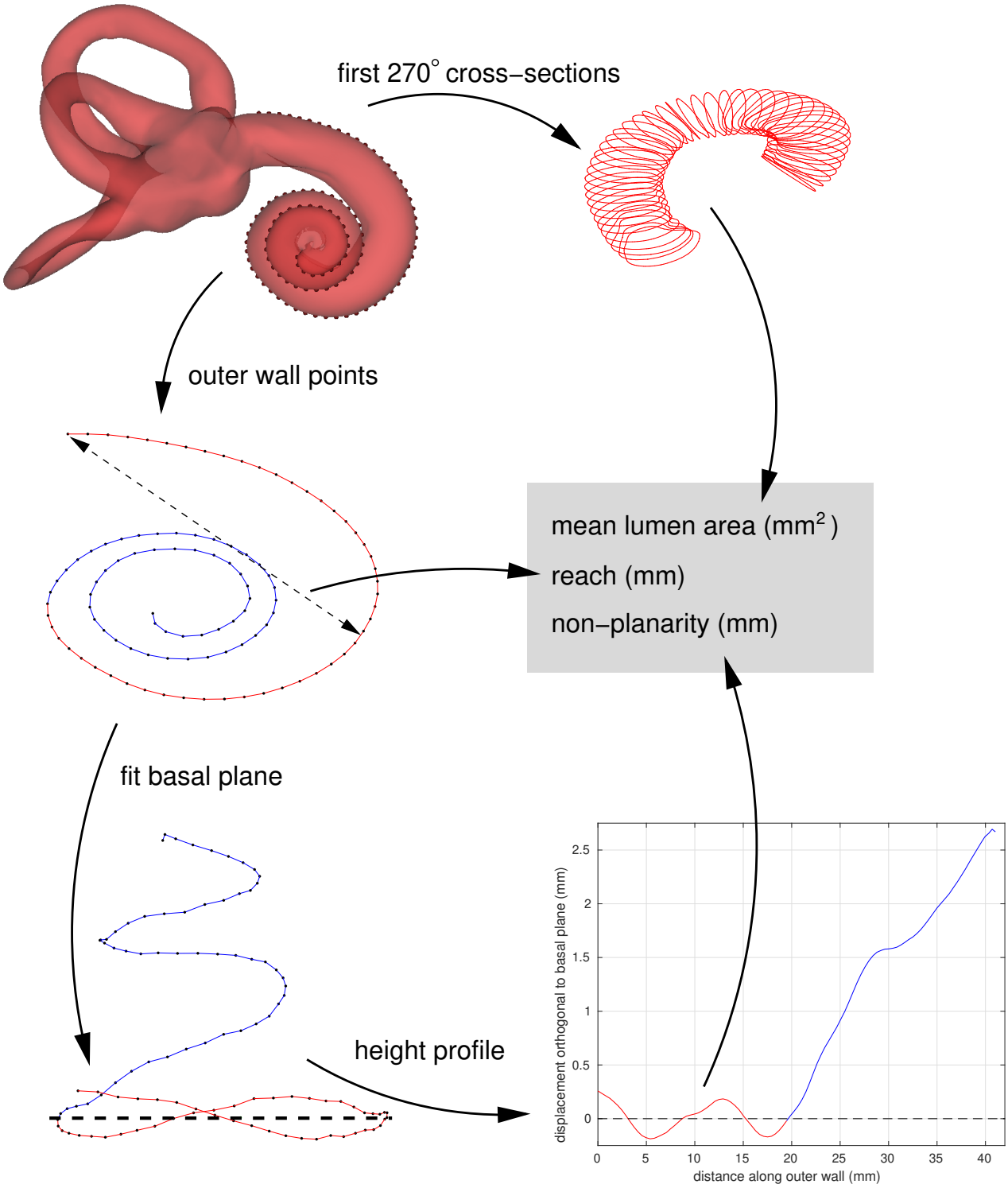
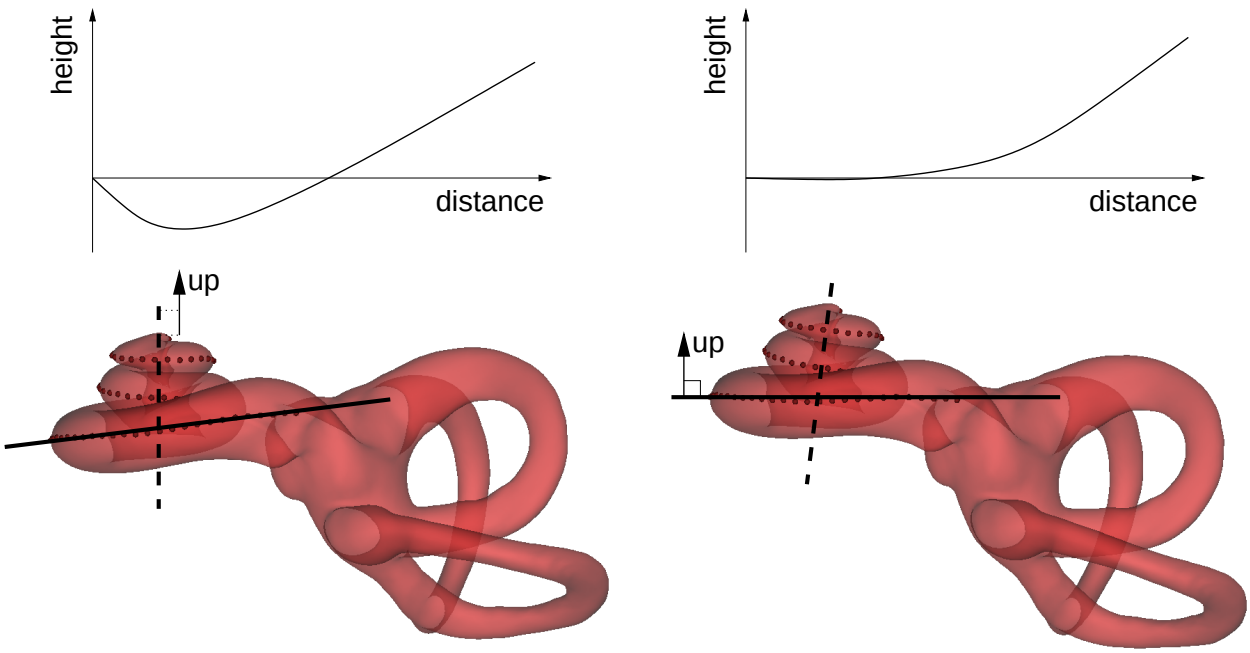


Figure 5: Clinically relevant measurements. In evaluating the success or otherwise of the model fit, we consider the cochlea's non-planarity, reach and lumen area. These measurements are of direct clinical utility, unlike more conventional and generic measures of surface misalignment, such as average vertex error. The three measurements are made on the first 270° of the basal turn (red), ignoring the rest of the spiral (blue). The basal plane (black dotted line) is defined as the best fit plane to the first 270° of the outer wall contour. It is a coincidence that the contour passes through the basal plane at around 270° in this example.



(a) Defined by modiolar axis

(b) Defined by basal plane

Figure 6: Modiolar and basal coordinates. Previous studies have defined cochlear vertical trajectories with reference to the modiolar axis (dashed line, a). However, estimation of this axis is not straightforward and is influenced by the middle and apical turns, which are of little relevance to most current implant designs (Dhanasingh and Jolly, 2017; Pelliccia et al., 2014). Furthermore, "rollercoaster" height profiles that go down and then up may nevertheless correspond to planar insertion trajectories that present little risk of trauma. An alternative is to replace the modiolar axis with the normal to the best fit plane through the first 270° of the basal turn (solid line, b). In this coordinate system, the "height" axis corresponds to deviation from the best fit plane, and it is clear at which point the insertion trajectory becomes nonplanar and potentially traumatic to the cochlear structures.

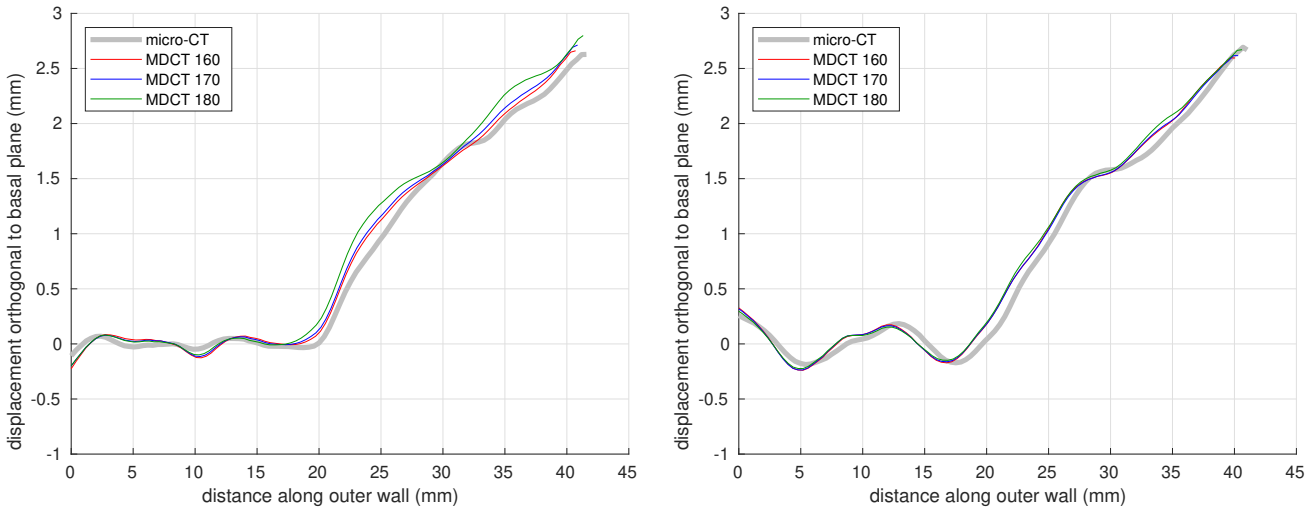


Figure 7: Vertical trajectories of the least (left) and most (right) nonplanar cochleas. The gold-standard micro-CT profiles are displayed in gray. The other profiles are derived from LAD model fits to the standard pseudo-clinical MDCT data, thresholded at 160 (red), 170 (blue) and 180 (green).

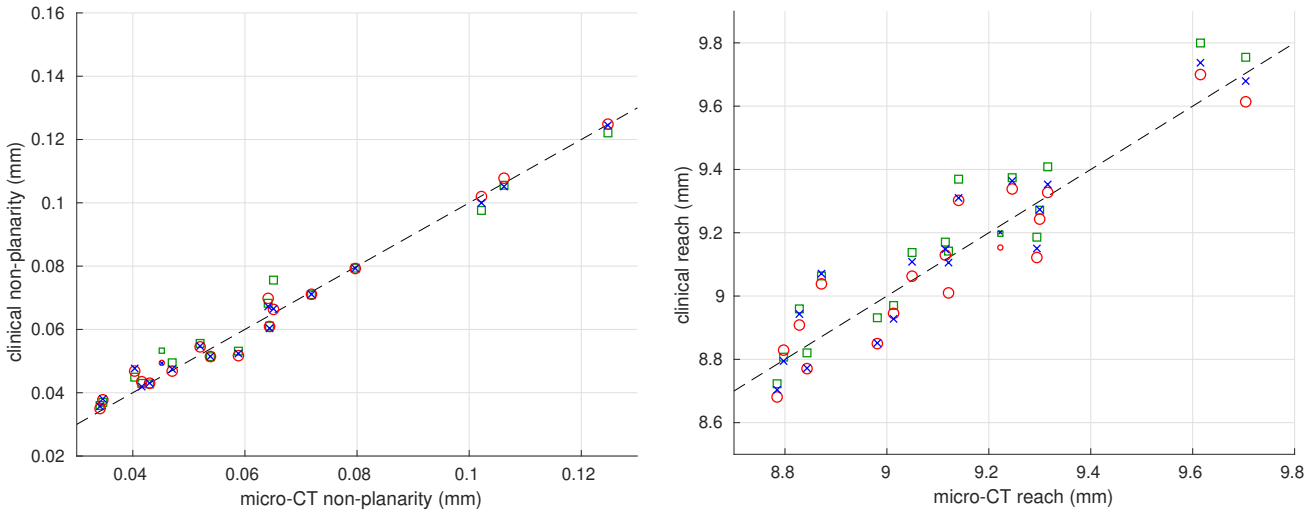
		full model			leave-one-out model			LAD		
		160	170	180	160	170	180	160	170	180
CBCT	non-planarity	2.47	2.46	3.35	15.7	16.0	14.8	12.1	10.8	8.63
	reach	9.36	9.16	9.57	26.7	24.7	23.6	13.0	11.2	10.9
MDCT	non-planarity	3.04	3.67	4.52	15.1	14.2	15.5	11.7	10.6	8.3
	reach	13.6	11.2	10.4	26.3	24.1	26.2	13.7	11.5	12.8
poor MDCT	non-planarity	4.58	4.61	6.27	16.8	16.0	14.7	17.6	17.4	23.6
	reach	12.9	9.46	16.9	27.5	28.9	35.4	19.0	17.7	18.1

Table 1: Summary results for the 17 specimens without an implanted electrode. Average measurement errors are expressed as $100 \times$ mean(absolute error) / gold-standard sample range.

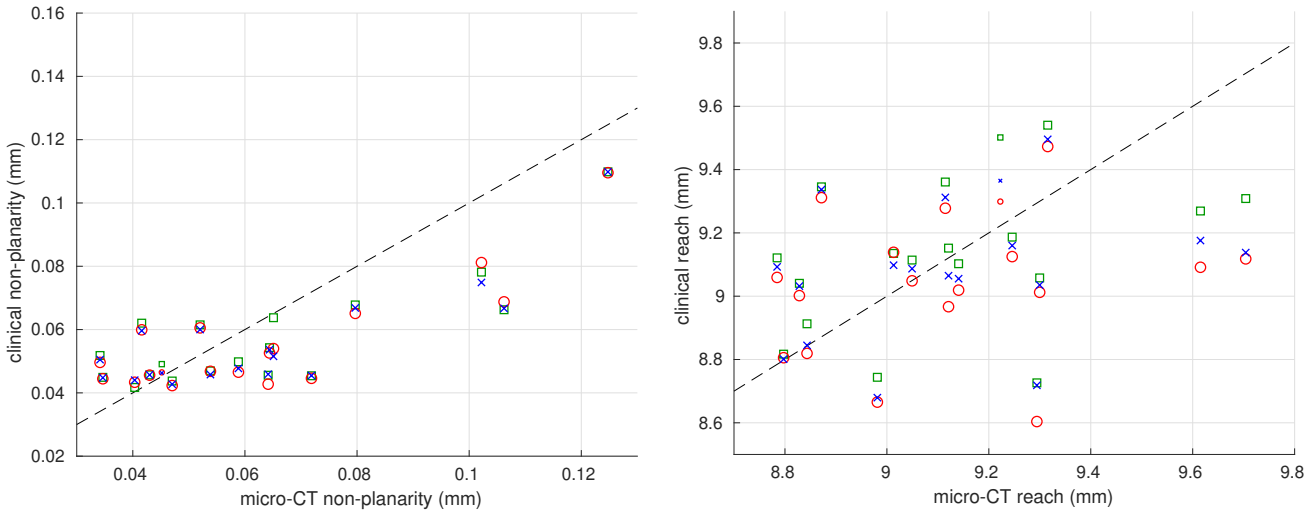
The full SSM provides an upper bound on SSM performance, with an effectively perfect model and fitting compromised only by the image resolution and detector noise. In contrast, the leave-one-out results are indicative of expected performance on unseen specimens with a model trained using only 17 exemplars. Since the LAD method does not require training, the results presented here are expected to generalise to new specimens without gross malformations. The parameter $d = 5\,\text{mm}$ was chosen since it produced visually plausible nonrigid deformations in all cases, without over-fitting to noise.

Figure 7 illustrates the LAD method's ability to recover the vertical trajectories of the least and most nonplanar cochleas, when applied to the standard pseudo-clinical MDCT data. Sensitivity to the grayscale segmentation threshold appears to be reasonable. Figures 8, 9 and 10 show the full set of non-planarity and reach results for next-generation CBCT, standard MDCT and poor MDCT data respectively, with the gold-standard measurements on the x-axis and the measurements derived from the pseudo-clinical images on the y-axis. The results show the expected degradation in performance with lower resolution, more noisy data. Also as expected, the full model performs significantly better than the leave-one-out model. The specimen with the implanted electrode is identified by half-sized markers and is a frequent outlier, since the scans suffered from beam-hardening artefacts that corrupted the thresholded point cloud.

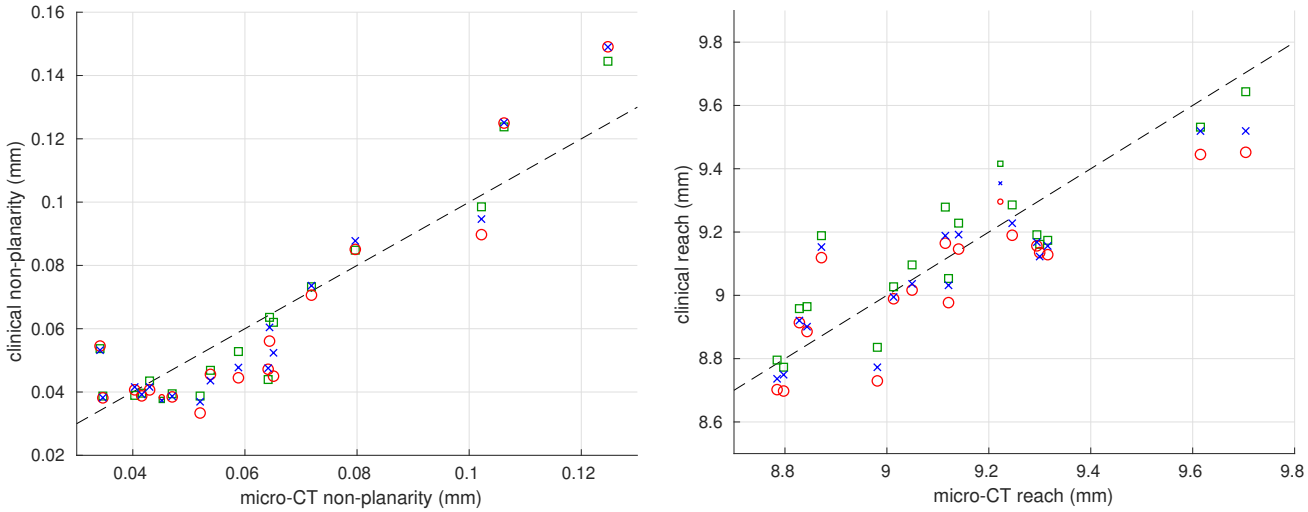
The non-planarity and reach results are summarised in Table 1, omitting the specimen with the implanted electrode. The tabulated numbers are the mean absolute error expressed as a percentage of the gold-standard sample range. Thus, for example, when applied to standard MDCT data, the LAD method produces measurements with an expected error of around 11% of the gold-standard range (maximum minus minimum) of the measured quantity. For comparison, the corresponding errors for the *unfitted* model (i.e. measured directly on



(a) Full SSM

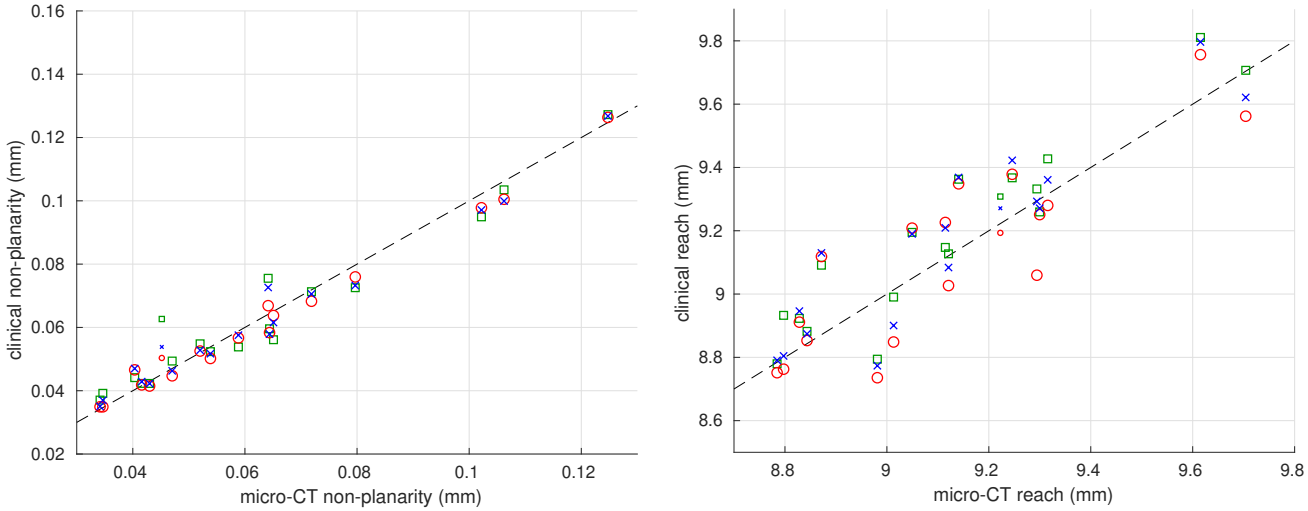


(b) Leave-one-out SSM

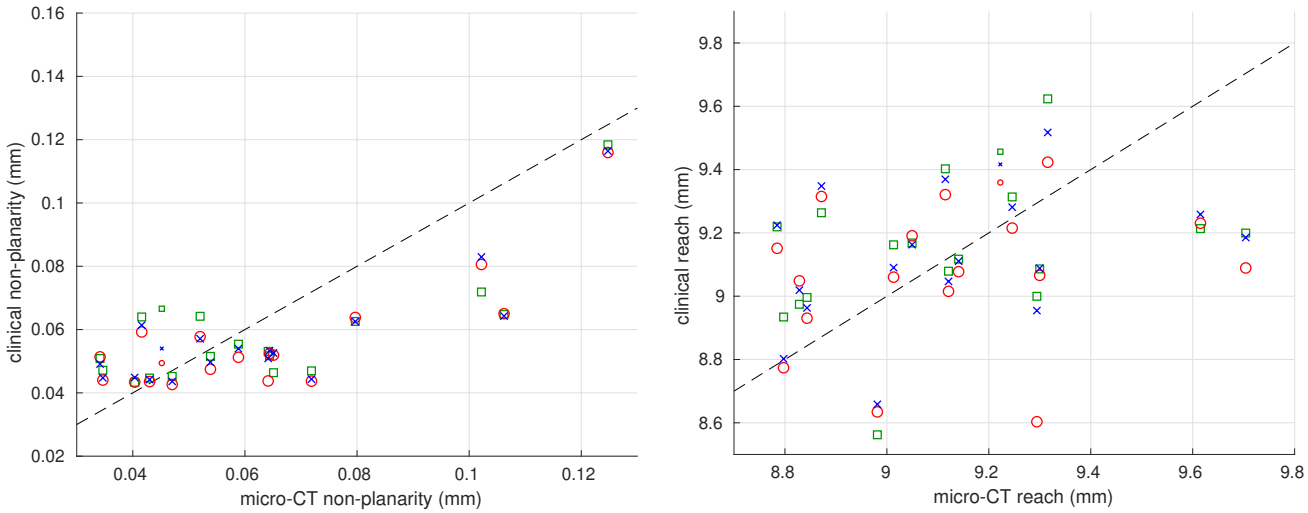


(c) LAD

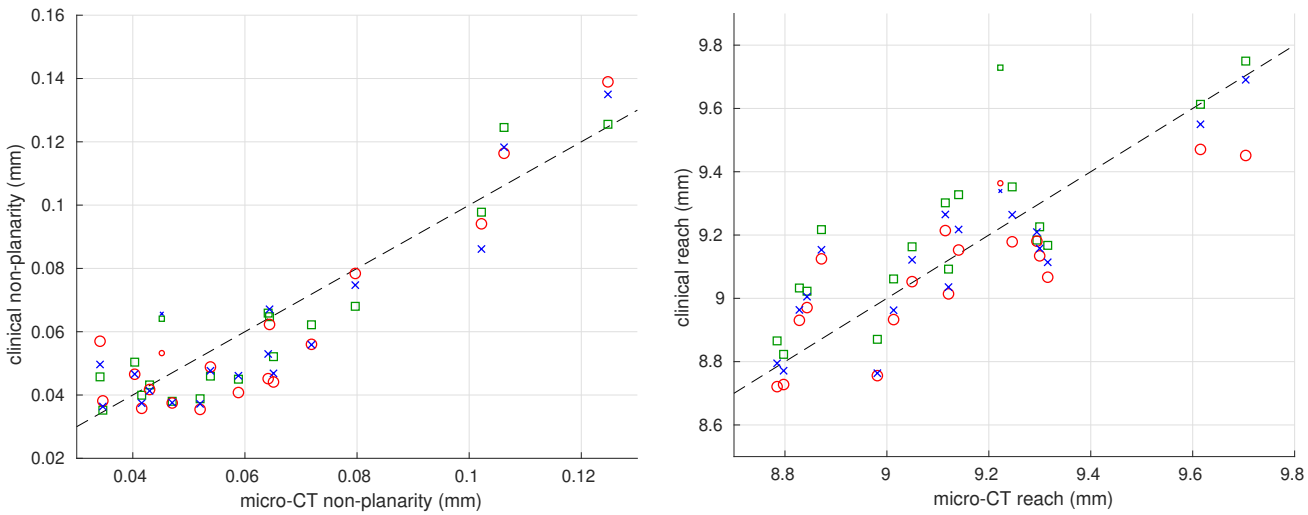
Figure 8: Cochlear non-planarity and reach estimated from next generation pseudo-clinical CBCT. Fully automatic processing with grayscale thresholds of 160 (red circles), 170 (blue crosses) and 180 (green squares). The half-size markers are for the specimen with the implanted electrode.



(a) Full SSM

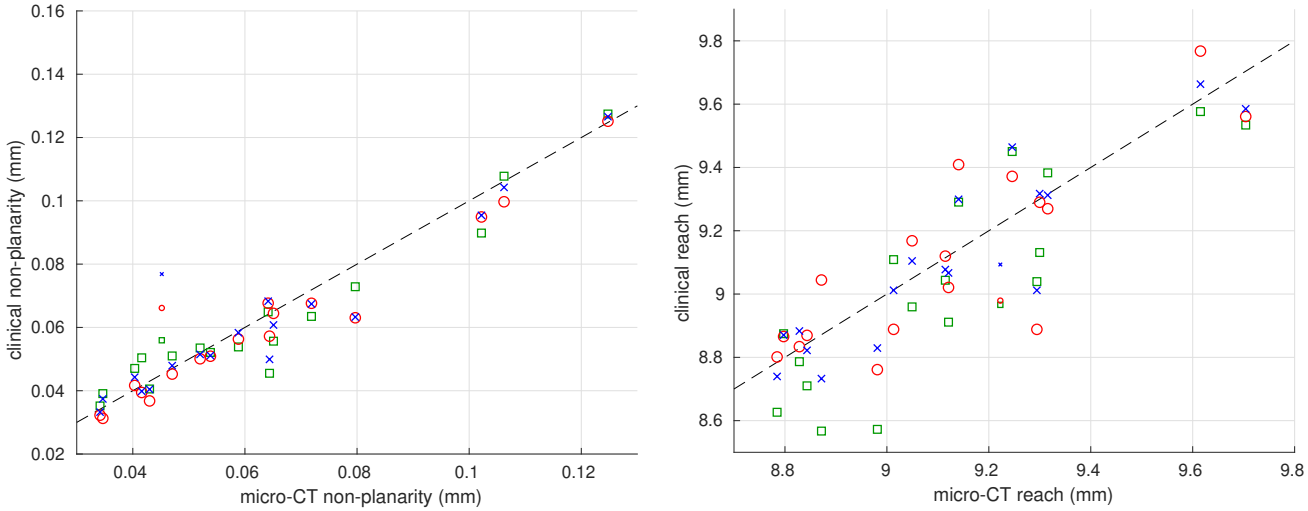


(b) Leave-one-out SSM

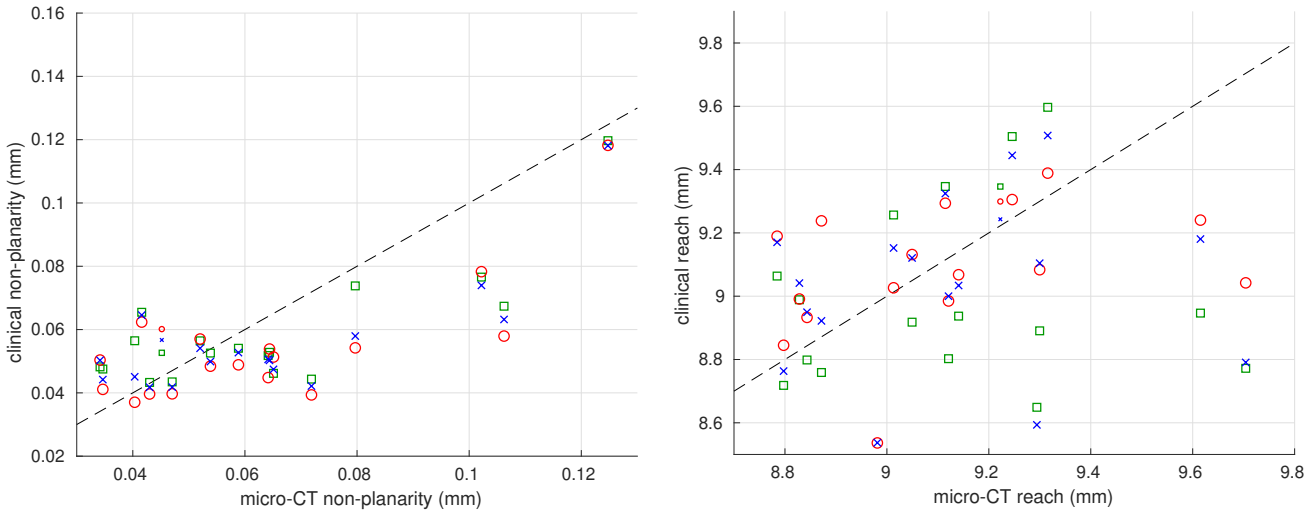


(c) LAD

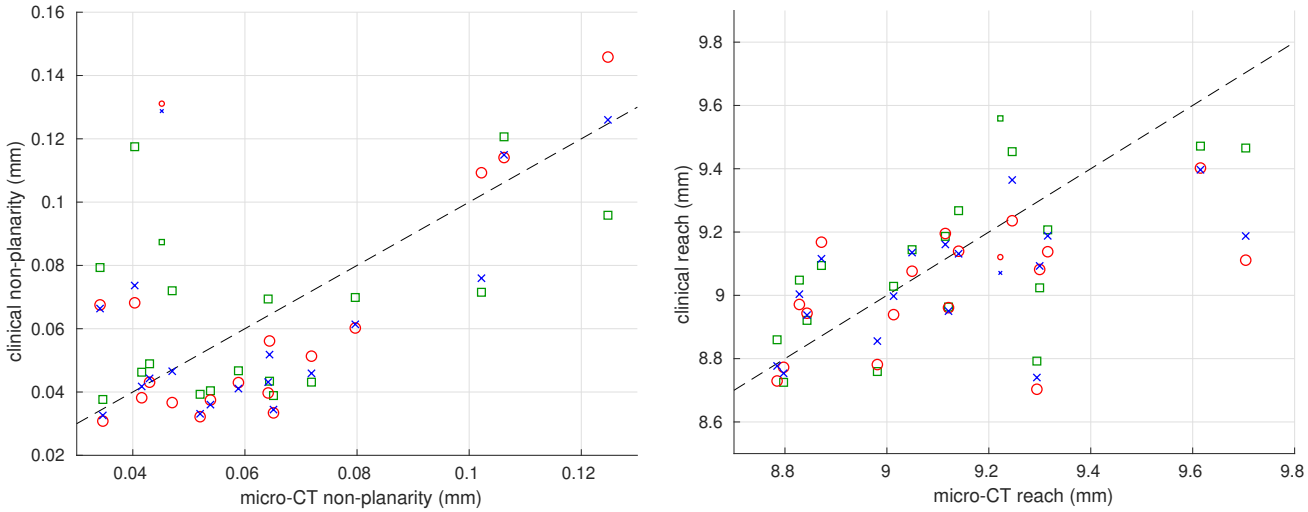
Figure 9: Cochlear non-planarity and reach estimated from standard pseudo-clinical MDCT. Fully automatic processing with grayscale thresholds of 160 (red circles), 170 (blue crosses) and 180 (green squares). The half-size markers are for the specimen with the implanted electrode.



(a) Full SSM



(b) Leave-one-out SSM



(c) LAD

Figure 10: Cochlear non-planarity and reach estimated from poor pseudo-clinical MDCT. Fully automatic processing with grayscale thresholds of 160 (red circles), 170 (blue crosses) and 180 (green squares). The half-size markers are for the specimen with the implanted electrode.

	full model	leave-one-out model	LAD
CBCT	0.0891	0.144	0.117
MDCT	0.104	0.151	0.123
poor MDCT	0.120	0.161	0.138

Table 2: Average vertex errors (mm) for the 17 specimens without an implanted electrode, at a grayscale threshold of 170. The tabulated numbers are the average distances between vertices on the fitted model and the gold-standard mesh, after optimal rigid body alignment of the two meshes.

the template without fitting to the individual) are 24.0% for non-planarity and 22.9% for reach.

Table 2 quantifies the segmentation accuracy of the 17 non-implanted specimens thresholded at a grayscale value of 170. For comparison, the average vertex error for the unfitted model is 0.185 mm. While such results are difficult to interpret from a clinical perspective, they do allow cautious comparison with the work of Noble et al. (2011), who achieved average vertex errors of around 0.2 mm (fitted) and 0.27 mm (unfitted). This comparison does require considerable qualification though, since Noble et al. (2011) segmented the scala tympani, while the results in Table 2 are for the entire otic capsule. Kjer et al. (2018) reported mean surface errors of 0.11 mm for the cochlear scalae.

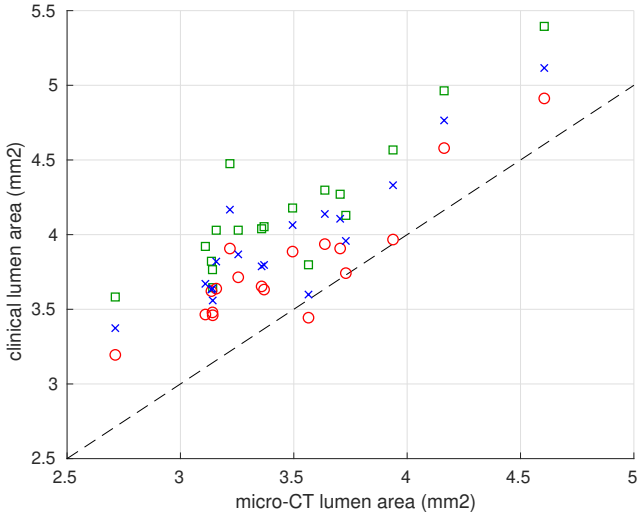
Figure 11 presents the results for the mean lumen area, just for the standard pseudo-clinical MDCT data. Performance is poor compared with non-planarity and reach. The principal difference between the two types of measurement is involvement of the cochlear inner wall in the lumen estimation, whereas non-planarity and reach require only the outer wall geometry. As Figure 11(e) shows, the clinical CT contrast at the inner wall is significantly worse than at the outer wall, resulting in poor segmentation generally and systematic overestimation of lumen area with the full SSM. Unfortunately, cochlear reach is not a good surrogate for lumen area, as demonstrated in Figure 11(f). Estimation of the ϕ angle between the plumb line of the round window and the tangent to the inner wall of the basal turn (Tang et al., 2018) — of interest since it constrains the initial insertion and bending angles of the electrode — produced similarly poor results, which we do not present here for reasons of concision. We conclude that anatomical measurements involving the cochlear inner wall are currently infeasible with this methodology.

On the basis of these investigations, we recommend the LAD approach as a practicable way to estimate clinically relevant anatomy of the human cochlea from standard clinical MDCT. The model fitting software is available for free download. Analysis of one cochlea requires around one minute of expert interaction followed by a couple of minutes of computation. The expert does need to exercise reasonable care when selecting the segmentation threshold: the one outlying result for the LAD approach, for non-planarity at a threshold of 180 with the poor MDCT data, was due to this threshold failing to capture part of the outer wall in many of the scans.

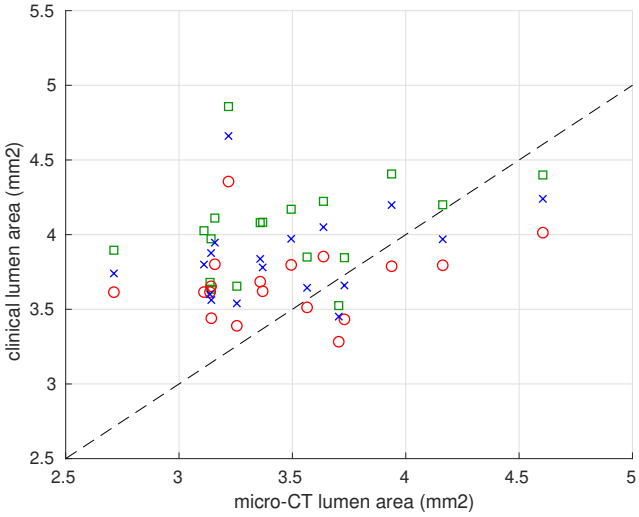
At the central threshold of 170, the LAD approach is able to estimate cochlear reach with a mean absolute error of 11.5% of the gold-standard sample range, or 1.16% $\pm$ 0.88% (mean $\pm$ one standard deviation) of the gold-standard values. This compares favourably with the method of Iyaniwura et al. (2018), where the absolute error in equivalent "A-value" estimates was 2.7% $\pm$ 2.1% of the gold-standard values. Kjer et al. (2018) reported signed errors for cochlear length down to 0.02 ± 0.2 mm: our equivalent values for reach are -0.00116 ± 0.134 mm.

At the same threshold of 170, cochlear non-planarity was estimated with an average absolute error of 10.6% of the gold-standard sample range. This is a novel metric that we suggest might correlate with the risk of insertion trauma, and may be more reliable than the "rollercoaster" classification of Avci et al. (2014), which is sensitive to estimation of the modiolar axis (Demarcy et al., 2017). While van der Jagt et al. (2017) demonstrated automatic estimation of cochlear vertical trajectories from clinical CT scans, to the best of our knowledge this is the first study to validate such measurements against micro-CT gold standards.

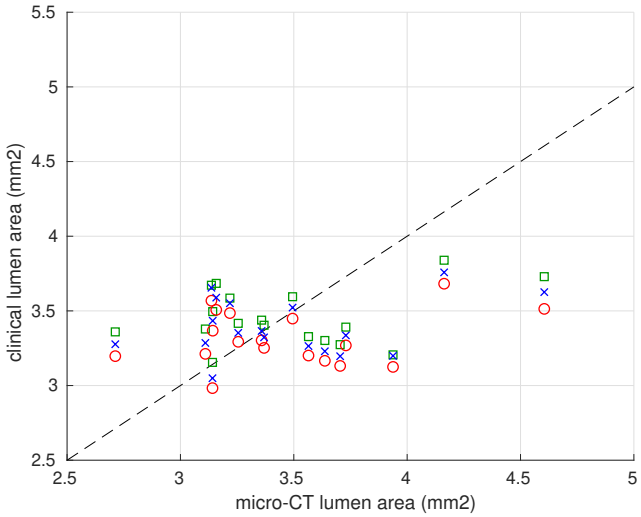
The principal limitation of this study is the use of pseudo-clinical data for the low resolution model fitting. Real MDCT scans of the temporal bones would arguably have provided a more sound basis for the work, but they were not available. That said, real MDCT data is no panacea: a dissected temporal bone imaged in a



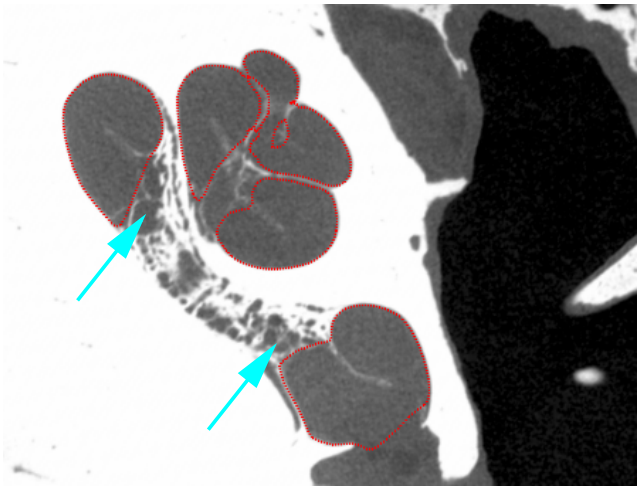
(a) Full SSM



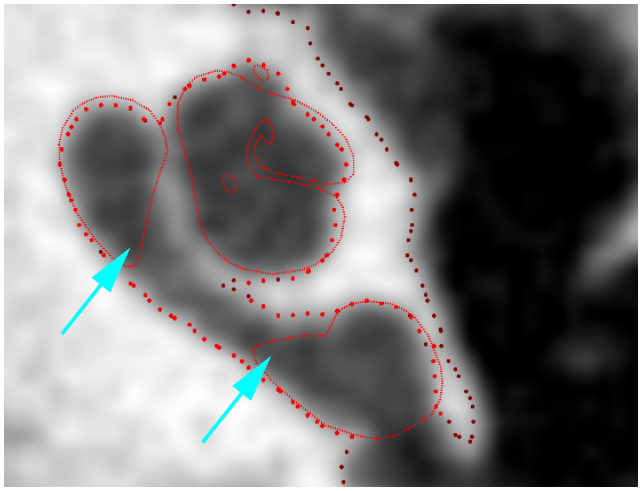
(b) Leave-one-out SSM



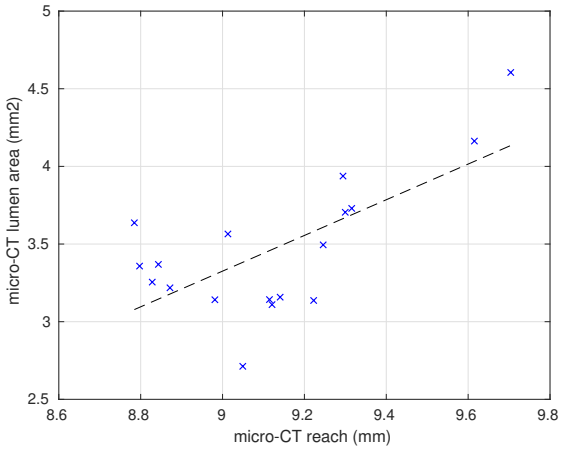
(c) LAD



(d) Micro-CT segmentation



(e) Pseudo-clinical MDCT fitted model



(f) Reach–area correlation ($\rho = 0.68$)

Figure 11: Cochlear mean lumen areas estimated from standard pseudo-clinical MDCT. (a)–(c) Fully automatic processing with grayscale thresholds of 160 (red circles), 170 (blue crosses) and 180 (green squares). (d) Micro-CT segmentation, correctly excluding the spiral ganglion at the inner wall. (e) Systematic overestimation in (a) can be attributed to the MDCT threshold including these inner-wall structures. (f) Lumen area is only weakly correlated with reach.

clinical MDCT scanner would not appear identical to the same bone scanned intact in a living human being. A reassuring indicator of the validity of the present study is the failure to estimate lumen area or any metric involving identification of the cochlear inner wall.

In comparison with the LAD method, the performance of the SSM approach was disappointing. Improved performance may be achievable through more sophisticated grayscale modelling (Kjer et al., 2018), or by using more training data, or by limiting the model to the cochlea alone (Kjer et al., 2018; Noble et al., 2011). That said, a benefit of including the canals is to leverage the posterior-anterior canal bifurcation as a readily identifiable landmark for initial model positioning. Notwithstanding potential improvements, SSM approaches suffer the disadvantage of requiring training on representative exemplars, with no guarantee of subsequent generalisation to new specimens, especially where there is malformation. Indeed, it was a failure to generalise that limited their efficacy in the present study, as evidenced by the relative performance of the full and leave-one-out models.

4 Conclusions

We have demonstrated a simple, rapid and freely available technique for estimating cochlear morphology from clinical MDCT scans. Average vertex errors are comparable with the state of the art, as are estimates of cochlear size. A further contribution of this study is an enhanced understanding of the cochlea's vertical trajectory, leading to a robust metric for characterizing the non-planarity of the basal turn. The non-planarity metric can be estimated from clinical scans with an average absolute error of 10.6% of the gold-standard sample range. The hope is that these morphologic estimates may one day assist in personalized implant selection and surgical planning, in the same way that similar methods have already been shown to improve implant programming (Noble et al., 2013).

Acknowledgments

The authors acknowledge the generosity of the deceased donors and thank Dr Cecilia Brassett and Dr Matthew Mason, of the Department of Physiology, Development and Neuroscience, University of Cambridge, for providing access to the cadaveric temporal bones and performing the micro-CT scans. The scans were funded by the Department of Physiology, Development and Neuroscience. Thanks are also due to Mrs Maria Wright, Mr Darren Broadhurst, Mr Roger Gray, Mr Matt O'Connor, Mr Prabhvir Singh Marway and Mr Will Ashley-Fenn for their expertise in preparing the temporal bones for scanning. MB acknowledges the support of the Evelyn Trust, the MRC Confidence in Concept Fund and the Cambridge Hearing Trust.

References

Avci, E., Nauwelaers, T., Hamacher, V., Kral, A., 2017. Three-dimensional force profile during cochlear implantation depends on individual geometry and insertion trauma. Ear and Hearing 38 (3).

Avci, E., Nauwelaers, T., Lenarz, T., Hamacher, V., Kral, A., 2014. Variations in microanatomy of the human cochlea. Journal of Comparative Neurology 522 (14), 3245–3261.

Besl, P. J., McKay, N. D., 1992. A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence 14 (2), 239–256.

Biedron, S., Westhofen, M., Ilgner, J., 2009. On the number of turns in human cochleae. Otology & Neurotology 30 (3), 414–417.

Bookstein, F. L., 1991. Morphometric Tools for Landmark Data: Geometry and Biology. Cambridge University Press.

Bookstein, F. L., 1997. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Medical Image Analysis 1 (3), 225–243.

Canfarotta, M. W., Dillon, M. T., Buss, E., Pillsbury, H. C., Brown, K. D., O'Connell, B. P., 2020. Frequency-to-place mismatch: Characterizing variability and the influence on speech perception outcomes in cochlear implant recipients [published online ahead of print March 19, 2020]. Ear and Hearing. doi: 10.1097/AUD.0000000000000864.

De Seta, D., Torres, R., Russo, F. Y., Ferrary, E., Kazmitcheff, G., Heymann, D., Amiaud, J., Sterkers, O., Bernardeschi, D., Nguyen, Y., 2017. Damage to inner ear structure during cochlear implantation: Correlation between insertion force and radio-histological findings in temporal bone specimens. Hearing Research 344, 90–97.

Demarcy, T., Vandersteen, C., Guevara, N., Raffaelli, C., Gnansia, D., Ayache, N., Delingette, H., 2017. Automated analysis of human cochlea shape variability from segmented μCT images. Computerized Medical Imaging and Graphics 59, 1–12.

Dhanasingh, A., Jolly, C., 2017. An overview of cochlear implant electrode array designs. Hearing Research 356, 93–103.

Erixon, E., Högstorp, H., Wadin, K., Rask-Andersen, H., 2009. Variational anatomy of the human cochlea: Implications for cochlear implantation. Otology & Neurotology 30 (1), 14–22.

Escudé, B., James, C., Deguine, O., Cochard, N., Eter, E., Fraysse, B., 2006. The size of the cochlea and predictions of insertion depth angles for cochlear implant electrodes. Audiology and Neurotology 11, 27–33.

Feldmar, J., Ayache, N., 1996. Rigid, affine and locally affine registration of free-form surfaces. International Journal of Computer Vision 18 (2), 99–119.

Goodall, C., 1991. Procrustes methods in the statistical analysis of shape. Journal of the Royal Statistical Society. Series B (Methodological) 53 (2), 285–339.

Gunz, P., Mitteroecker, P., Bookstein, F. L., 2005. Semilandmarks in three dimensions. In: Slice, D. E. (Ed.), Modern Morphometrics in Physical Anthropology. Springer, Boston, pp. 73–98.

Gunz, P., Ramsier, M., Kuhrig, M., Hublin, J.-J., Spoor, F., 2012. The mammalian bony labyrinth reconsidered, introducing a comprehensive geometric morphometric approach. Journal of Anatomy 220 (6), 529–543.

Iyaniwura, J. E., Elfarnawany, M., Ladak, H. M., Agrawal, S. K., 2018. An automated A-value measurement tool for accurate cochlear duct length estimation. Journal of Otolaryngology — Head & Neck Surgery 47, 5.

Kjer, H. M., Fagertun, J., Vera, S., Gil, D., Ballester, M. Á. G., Paulsen, R. R., 2016. Free-form image registration of human cochlear μCT data using skeleton similarity as anatomical prior. Pattern Recognition Letters 76, 76–82.

Kjer, H. M., Fagertun, J., Wimmer, W., Gerber, N., Vera, S., Barazzetti, L., Mangado, N., Ceresa, M., Piella, G., Stark, T., Stauber, M., Reyes, M., Weber, S., Caversaccio, M., Ballester, M. Á. G., Paulsen, R. R., 2018. Patient-specific estimation of detailed cochlear shape from clinical CT images. International Journal of Computer Assisted Radiology and Surgery 13, 389–396.

Koch, R. W., Ladak, H. M., Elfarnawany, M., Agrawal, S. K., 2017. Measuring cochlear duct length — a historical analysis of methods and results. Journal of Otolaryngology — Head & Neck Surgery 46, 19.

Lenarz, T., Timm, M. E., Salcher, R., Büchner, A., 2019. Individual hearing preservation cochlear implantation using the concept of partial insertion. Otology & Neurotology 40 (3).

Meng, J., Li, S., Zhang, F., Li, Q., Qin, Z., 2016. Cochlear size and shape variability and implications in cochlear implantation surgery. Otology & Neurotology 37 (9), 1307–1313.

Nguyen, Y., Miroir, M., Kazmitcheff, G., Sutter, J., Bensidhoum, M., Ferrary, E., Sterkers, O., Bozorg Grayeli, A., 2012. Cochlear implant insertion forces in microdissected human cochlea to evaluate a prototype array. Audiology and Neurotology 17 (5), 290–298.

Noble, J. H., Labadie, R. F., Gifford, R. H., Dawant, B. M., 2013. Image-guidance enables new methods for customizing cochlear implant stimulation strategies. IEEE Transactions on Neural Systems and Rehabilitation Engineering 21 (5), 820–829.

Noble, J. H., Labadie, R. F., Majdani, O., Dawant, B. M., 2011. Automatic segmentation of intracochlear anatomy in conventional CT. IEEE Transactions on Biomedical Engineering 58 (9), 2625–2632.

Pelliccia, P., Venail, F., Bonafé, A., Makeieff, M., Iannetti, G., Bartolomeo, M., Mondain, M., 2014. Cochlea size variability and implications in clinical practice. Acta otorhinolaryngologica Italica 34 (1), 42–49.

Phillips, G. S., LoGerfo, S. E., Richardson, M. L., Anzai, Y., 2012. Interactive web-based learning module on CT of the temporal bone: Anatomy and pathology. RadioGraphics 32 (3), E85–E105.

Pietsch, M., Aguirre Dávila, L., Erfurt, P., Avci, E., Lenarz, T., Kral, A., 2017. Spiral form of the human cochlea results from spatial constraints. Scientific Reports 7, 7500.

Rask-Andersen, H., Liu, W., Erixon, E., Kinnefors, A., Pfaller, K., Schrott-Fischer, A., Glueckert, R., 2012. Human cochlea: Anatomical characteristics and their relevance for cochlear implantation. The Anatomical Record 295 (11), 1791–1811.

Rivas, A., Cakir, A., Hunter, J. B., Labadie, R. F., Zuniga, M. G., Wanna, G. B., Dawant, B. M., Noble, J. H., 2017. Automatic cochlear duct length estimation for selection of cochlear implant electrode arrays. Otology & Neurotology 38 (3), 339–346.

Sheffield, S. W., Jahn, K., Gifford, R. H., 2015. Preserved acoustic hearing in cochlear implantation improves speech perception. Journal of the American Academy of Audiology 26 (2), 145–154.

Singla, A., Sahni, D., Gupta, A. K., Aggarwal, A., Gupta, T., 2015. Surgical anatomy of the basal turn of the human cochlea as pertaining to cochlear implantation. Otology & Neurotology 36 (2), 323–328.

Tang, J., Tang, X., Li, Z., Liu, Y., Tan, S., Li, H., Ke, R., Wang, Z., Gong, L., Tang, A., 2018. Anatomical variations of the human cochlea determined from micro-CT and high-resolution CT imaging and reconstruction. The Anatomical Record 301 (6), 1086–1095.

van der Jagt, A. M. A., Kalkman, R. K., Briaire, J. J., Verbist, B. M., Frijns, J. H. M., 2017. Variations in cochlear duct shape revealed on clinical CT images with an automatic tracing method. Scientific Reports 7, 17566.

Verbist, B. M., Skinner, M. W., Cohen, L. T., Leake, P. A., James, C., Boëx, C., Holden, T. A., Finley, C. C., Roland, P. S., Roland, J. T. J., Haller, M., Patrick, J. F., Jolly, C. N., Faltys, M. A., Briaire, J. J., Frijns, J. H. M., 2010. Consensus panel on a cochlear coordinate system applicable in histologic, physiologic, and radiologic studies of the human cochlea. Otology & Neurotology 31 (5).

Würfel, W., Lanfermann, H., Lenarz, T., Majdani, O., 2014. Cochlear length determination using cone beam computed tomography in a clinical setting. Hearing Research 316, 65–72.

Zou, J., Lähelmä, J., Arnisalo, A., Pyykkö, I., 2017. Clinically relevant human temporal bone measurements using novel high-resolution cone-beam CT. Journal of Otology 12 (1), 9–17.