Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Semi-automated sex estimation from cochlear shape using clinical CT imaging and OTOPLAN segmentation

Priya VISHNUMURTHY
IUSTI, Aix Marseille Univ, CNRS
Mary DAVAL
Hospital Fondation Adolphe de Rothschild
Luke HELPARD
MED-EL Corporation
Ismael OUZZINE
Aix Marseille Univ, APHM, CNRS, IUSTI, La Conception University Hospital
Ralph HADDAD
Aix Marseille Univ, APHM, CNRS, IUSTI, La Conception University Hospital
Nikola IVANOVIC
MED-EL Medical Electronics
Vincent JOURDES
MED-EL FRANCE
Justin MICHEL
Aix Marseille Univ, APHM, CNRS, IUSTI, La Conception University Hospital
Stéphane GARGULA
stephane.gargula@ap-hm.fr

Aix Marseille Univ, APHM, CNRS, IUSTI, La Conception University Hospital

Research Article

Keywords: Sex estimation, Cochlea, Bony labyrinth, Forensic anthropology, Computed tomography, Machine learning

Posted Date: July 7th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10093723/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Competing interest reported. The MED-EL-affiliated authors contributed expertise on cochlear segmentation and on the OTOPLAN export format. Study conception, data analysis, interpretation of results, and the decision to submit for publication were conducted independently by the academic authors. MED-EL had no role in the analysis or in the decision to publish.

Title: **Semi-automated sex estimation from cochlear shape using clinical CT imaging and OTOPLAN segmentation**

Running title: Semi-automated cochlear sex estimation

Article type: Research

Authors:

Priya VISHNUMURTHY[1], Mary DAVAL MD[2], Luke HELPARD PhD[3], Ismael OUZZINE MD[4], Ralph HADDAD MD[4], Nikola IVANOVIC PhD[5], Vincent JOURDES[6], Justin MICHEL MD PhD[4], Stéphane GARGULA MD[4]

1: IUSTI, Aix Marseille Univ, CNRS, Marseille, France

2 : Department of Otolaryngology, Hospital Fondation Adolphe de Rothschild, 29 Rue Manin, 75019, Paris, France

3 : MED-EL Corporation, Canada, 123 Edward Street, Suite 908, M5G 1E2 Toronto, CA

4 : Aix Marseille Univ, APHM, CNRS, IUSTI, La Conception University Hospital, ENT-HNS Department, Marseille, France

5: MED-EL Medical Electronics, Fuerstenweg 77a, 6020 Innsbruck, Austria

6: MED-EL FRANCE, 400 avenue Roumanille, Bât. 6 - CS 70062, 06902 Sophia-Antipolis, France

Corresponding Author: GARGULA Stéphane, stephane.gargula@gmail.com, 06 88 28 51 26, 147 Boulevard Baille, 13005, Marseille, France

ORCID: *0000-0002-5034-4210*

Semi-automated sex estimation from cochlear shape using clinical CT imaging and OTOPLAN segmentation

Abstract

Background. Sex estimation is fundamental in forensic anthropology and bioarchaeology, yet classical methods rely on skeletal elements that are often absent or poorly preserved. The bony labyrinth, enclosed within the petrous bone, is a promising alternative owing to its early maturation and taphonomic resistance, and cochlear shape carries a sex-specific signal. However, most approaches depend on manual segmentation, micro-CT, or complex shape-analysis frameworks that limit routine applicability**.**

Objective. To evaluate the feasibility of a semi-automated method for sex estimation from cochlear shape using standard clinical CT and automated OTOPLAN segmentation.

Materials and methods. We analyzed 139 cochleae from 87 adults using clinical CT scans (slice thickness ≤ 0.6 mm). The three-dimensional lateral wall path was extracted with OTOPLAN; after normalization and resampling, curvature- and torsion-based descriptors and scalar cochlear metrics were computed to train a supervised classifier. Preprocessing, feature, and hyperparameter choices were jointly optimized by Bayesian optimization on the training set.

Results. On an independent holdout set (n = 29), the XGBoost model reached an area under the ROC curve (AUC-ROC) of 0.77 (95% CI 0.55-0.95), with 79% overall and balanced accuracy. Cochlear diameter was the most influential feature, followed by the torsion-to-curvature (τ/κ) ratio across several spiral segments.

Conclusion. Sex-related cochlear shape information is recoverable from routine clinical CT using a semi-automated pipeline. Although accuracy remains moderate, the approach is non-

destructive, scalable, and clinically compatible, offering a potential adjunct for forensic anthropology and bioarchaeology when conventional methods are unavailable**.**

Keywords: Sex estimation; Cochlea; Bony labyrinth; Forensic anthropology; Computed tomography; Machine learning

INTRODUCTION

Sex estimation is a fundamental component of biological profile reconstruction in forensic anthropology and bioarchaeology, as it conditions subsequent estimations of age at death, stature, ancestry proxies, and ultimately individual identification[1, 2]. When skeletal remains are complete, the pelvis and skull provide high diagnostic accuracy. However, in many medico-legal and archaeological contexts, these skeletal elements are frequently absent, fragmented, or altered by taphonomic processes, fire, or trauma, limiting the applicability of classical sex estimation methods[3–5]. These limitations are particularly pronounced in commingled assemblages and in subadult remains, where sexually dimorphic traits are weakly expressed and unreliable[6, 7].

The petrous part of the temporal bone has emerged as a structure of particular interest in such contexts due to its exceptional density and resistance to post-mortem degradation[5, 8]. Within the petrous bone, the bony labyrinth of the inner ear is especially attractive for biological profiling because it reaches an adult-like morphology early in development and is largely unaffected by postnatal growth[9, 10]. These properties have motivated increasing interest in labyrinth-based approaches to sex estimation using computed tomography (CT).

Early studies focused primarily on linear and angular measurements of the semicircular canals and cochlea. Osipov et al. (2013) demonstrated that multivariate combinations of labyrinthine measurements could achieve sex classification accuracies of up to approximately 80–85% in adult samples[11]. Boucherie et al. (2021) extended these findings to both adult and subadult individuals, reporting moderate to good performance while emphasizing population specificity and the need for external validation[12]. More recently, practical approaches based on routine CT and cone-beam CT (CBCT) imaging have shown that sex estimation from the bony labyrinth can be achieved using simple, reproducible linear metrics compatible with clinical

and forensic workflows[13]. Together, these studies support the feasibility of inner ear–based sex estimation in situations where only dense petrous structures are preserved.

A major conceptual advance was proposed by Braga et al. (2019), who demonstrated that sexual dimorphism of the cochlea can be captured through three-dimensional shape analysis of the cochlear spiral[14]. By modeling the cochlea as a three-dimensional curve and applying advanced geometric morphometric techniques, they reported high classification accuracies exceeding 90% and showed that sex-related cochlear shape differences are present from early postnatal life. While these results highlight the strong discriminatory potential of cochlear shape, the methodological complexity of this approach-including high-resolution imaging, manual or semi-manual segmentation, and sophisticated shape-space modeling-currently limits its applicability in routine forensic and archaeological contexts.

In parallel, advances in clinical otology have led to the widespread adoption of automated or semi-automated cochlear segmentation tools for cochlear implant planning. Software platforms such as OTOPLAN enable reliable extraction of cochlear geometry from standard high-resolution clinical CT scans and offer a scalable, reproducible alternative to manual segmentation workflows. These developments raise the question of whether cochlear shape–based sex estimation can be translated into a semi-automated, clinically realistic pipeline without sacrificing robustness.

The objective of the present study was therefore to provide a proof of concept for sex estimation based on cochlear shape using routine clinical CT imaging and OTOPLAN-based segmentation. By combining standardized cochlear curve extraction, curvature- and torsion-derived descriptors, and supervised machine learning, we aimed to assess whether sex-related cochlear shape information remains detectable under conditions compatible with forensic anthropology and bioarchaeology practice.

MATERIALS AND METHODS

General workflow

The analytical pipeline was designed to evaluate whether cochlear shape, as extracted from routine clinical CT data, carries sufficient information for sex classification (**Fig.**). Raw cochlear lateral wall paths and scalar metrics exported from OTOPLAN were first normalized and submitted to quality control, then partitioned into a training set and an independent holdout set. A supervised classification model was trained and optimized exclusively on the training set, and its performance evaluated once on the holdout set to provide an unbiased estimate of generalization to unseen individuals.

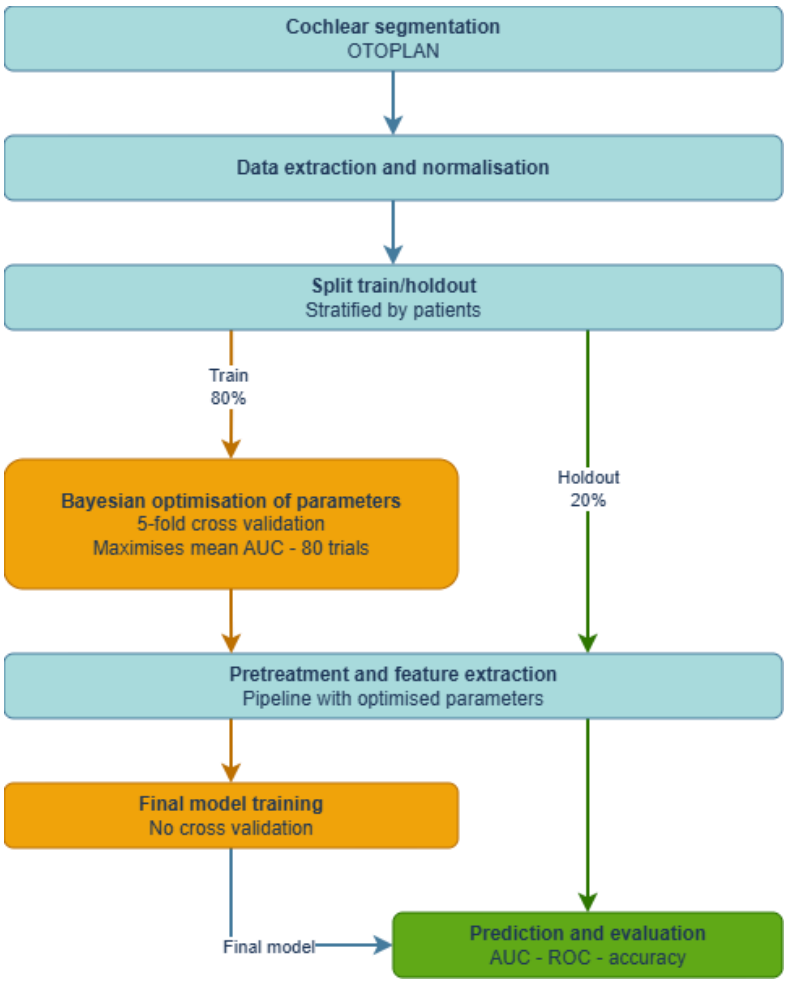


Fig. 1 *General workflow. Note : in yellow the step conducted on the train set, in green the steps conducted on the holdout set, in blue the steps conducted on the whole sample*

Study design and imaging data

This retrospective study was conducted on a series of high-resolution computed tomography (CT) scans acquired in a clinical setting for routine otologic indications. Only adult individuals were included to limit potential confounding related to developmental changes, although the cochlea is known to reach an adult-like morphology early in life. A total of 162 cochleae from 97 adult individuals of known biological sex were collected. Exclusion criteria comprised congenital inner ear malformations, advanced otosclerosis, post-surgical changes involving the cochlea, or significant imaging artifacts[15].

Human Ethics and Consent to Participate declarations

This retrospective study was approved by the Rothschild Foundation Hospital review board (IRB 00012801, study number CE_20230328_5_SGA) and conducted in accordance with the Declaration of Helsinki. Given the retrospective use of fully anonymized imaging data, the requirement for informed consent was waived. Clinical trial number: not applicable.

Cochlear segmentation and data extraction

Cochlear segmentation was performed using OTOPLAN (CASCINATION AG [Bern, Switzerland], in cooperation with MED-EL [Innsbruck, Austria]), a commercially available software platform developed for cochlear implant planning that enables automated segmentation of the cochlea from standard clinical CT data with minimal operator interaction[16–18]. CT scans with slice thickness exceeding 0.6 mm or in which the cochlea was absent from the scan field were excluded, as these datasets could not be processed by the software.

For each ear, OTOPLAN exported two types of data: a three-dimensional representation of the cochlear lateral wall path, consisting of 400 landmarks (x, y, z coordinates) regularly distributed along the cochlear duct, and a set of six scalar cochlear metrics: the cochlear diameter, the

cochlear width, the cochlear height, the estimated cochlear duct length (CDL), the estimated CDL from 0 (CDL_0) and the number of turns.

Data preprocessing and quality control

The extracted cochlear lateral wall paths were pre-processed using a standardized pipeline implemented in Python to ensure geometric comparability across individuals. Left cochlear curves were mirrored by negating the x-coordinate to ensure a uniform direction of rotation across all cochleae. Each curve was then translated to its centroid and uniformly scaled by its total norm, removing inter-individual differences in absolute position and size and focusing the analysis on shape.

Quality control was performed using automated outlier detection based on principal component analysis of the geometric features. Cochleae with a z-score exceeding 4.0 on the second principal component were excluded as this component was found to capture residual noise and segmentation artifacts. Scalar metrics with missing values were imputed using the global median computed on the training set, as described in the following section.

Split of the data

To obtain an unbiased estimate of classification performance, the dataset was partitioned into a training set (80%) and an independent holdout test set (20%) prior to any model development. The training set was used exclusively for model development and hyperparameter optimization, while the holdout set was reserved for final evaluation and accessed only once, after all modeling decisions had been finalized.

Intra-patient concordance between left and right cochleae was evaluated in patients with both ears using Lin's concordance correlation coefficient (CCC), applied separately to scalar cochlear metrics and landmarks of the lateral wall path (Appendix I). Local differences were identified along specific segments of the spiral, suggesting that both cochleae carry

complementary morphological information. However, overall concordance was good to excellent meaning that assigning them to different partitions would artificially inflate performance estimates by allowing the model to partially recognize individuals rather than generalizing to truly unseen cases. Therefore, the split was performed at the patient level using stratified sampling, ensuring a balanced male-to-female ratio across both sets and that both ears of a given patient were always assigned to the same partition to prevent any data leakage.

Curvature and torsion analysis

The cochlea was modelled as a three-dimensional parametric curve and its geometric properties characterized using curvature (κ) and torsion (τ), two intrinsic descriptors derived from differential geometry that are invariant to rigid transformations and scaling. Curvature quantifies the local bending of the curve, whereas torsion describes its three-dimensional twisting out of a plane. Both were computed numerically using finite-difference approximations of the derivatives with respect to arc length.

The normalized cochlear curve was divided into 19 equal-length segments from base to apex (Appendix II). Within each segment, three descriptors were computed for κ, the mean, standard deviation, and arc-length integral, four descriptors were computed for τ, the mean, standard deviation, arc-length integral, and absolute arc-length integral and two descriptors were computed for the τ/κ ratio: the mean and standard deviation. To assess whether these descriptors differed between male and female cochleae, segment-level values were compared between sexes using Mann-Whitney U tests, with p-values adjusted for multiple comparisons using the Benjamini-Hochberg False Discovery Rate correction (BH - FDR) (corrected p-value < 0.05).

Scalar cochlear metrics analysis

To assess whether the scalar cochlear metrics provided by OTOPLAN differed between male and female cochleae, each of the six variables was compared between sexes using Mann-

Whitney U tests. Effect sizes were estimated using the r statistic derived from the Mann-Whitney U statistic. P-values were adjusted for multiple comparisons using the Benjamini-Hochberg false discovery rate correction (corrected p-value < 0.05). Descriptive statistics (mean ± standard deviation) were computed for each variable by sex.

Pre-treatment pipeline and classification model optimization

The cochlear curve preprocessing pipeline and the classification model were jointly optimized using Bayesian optimization (Optuna framework, TPE sampler, 80 trials), treating preprocessing choices as free parameters explored simultaneously with model hyperparameters to ensure that the entire pipeline is tuned toward the classification objective. Three categories of parameters were explored: curve preprocessing parameters, geometric feature construction parameters, and the choice of classification algorithm with its associated hyperparameters. Optimization was performed exclusively on the training set using five-fold cross-validated AUC-ROC as the objective criterion, with both ears of a given patient always assigned to the same fold to prevent data leakage. The full parameter search space and retained values are reported in Appendix II.

Classification and statistical analysis

The classification model was trained on the entire training set using the optimal parameter configuration and evaluated once on the holdout set. Primary outcome measures included overall classification accuracy, balanced accuracy, and AUC-ROC. The ROC curve was accompanied by a 95% confidence interval estimated by bootstrap resampling (1,000 iterations). The optimal classification threshold was identified using Youden's J index, which maximizes the sum of sensitivity and specificity. To assess the contribution of each feature to sex classification, permutation importance was computed on the training set: each feature was randomly shuffled in turn, and the resulting drop in AUC-ROC was recorded as a measure of its importance to the model. Model development and evaluation are reported in accordance with

the TRIPOD+AI statement for prediction models using machine learning [19]. This study corresponds to the development and internal validation of a prediction model on a single dataset partitioned into training and holdout sets. The sample was a convenience series of clinical CT scans meeting the inclusion criteria; no formal a priori sample-size calculation was performed, consistent with the exploratory, proof-of-concept nature of the study. Predictor extraction (lateral wall path and scalar metrics) was performed automatically by the segmentation and feature pipeline, without reference to the sex label. Calibration was not formally assessed; model performance is therefore reported in terms of discrimination and classification accuracy only.

RESULTS

Study sample

A total of 162 cochleae from 97 adult individuals were initially collected. Twenty-two cochleae were excluded due to slice thickness exceeding 0.6 mm or absence of the cochlea within the scan field, and one additional cochlea was excluded following PCA-based outlier detection (Appendix III). The final sample comprised 139 cochleae from 87 patients, partitioned into a training set of 110 cochleae and an independent holdout set of 29 cochleae (**Table 1**).

Table 1 : Distribution of cochleae and patients across the training and holdout sets, stratified by sex and ear side.

	Train	**Holdout**	**Total**
Patients (total)	69	18	87
Patients F	38	10	48
Patients M	31	8	39
Patients with one ear	28	7	35
Patients with two ears	41	11	52
Cochleae (total)	110	29	139
Cochleae F	62	15	77
Cochleae M	48	14	62

Curvature and torsion profiles

Visual inspection of the mean cochlear trajectories underlines overall morphological similarity between male and female cochleae across all three projections with moderate local variations (**Fig.**).

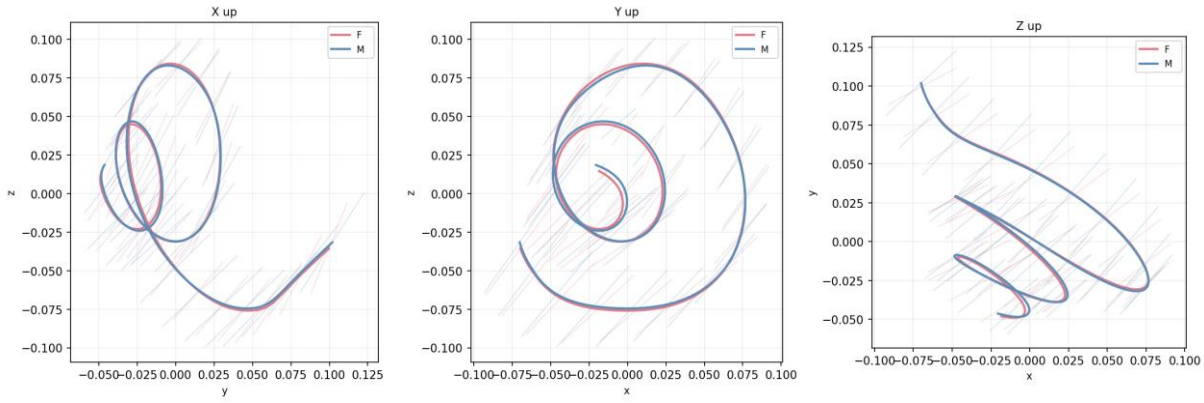


***Fig. 2** : Mean curve of male and female cochleae viewed from three viewpoints.*

Segment-level analysis identified ten statistically significant differences in κ, τ and τ/κ descriptors between male and female cochleae after Benjamini-Hochberg correction (**Table 2**, **Fig. 3**). However, all effect sizes were in the small-to-moderate range (r = 0.26 - 0.40), and significant differences were spatially scattered across the spiral without a dominant region or descriptor emerging as a reliable standalone discriminator.

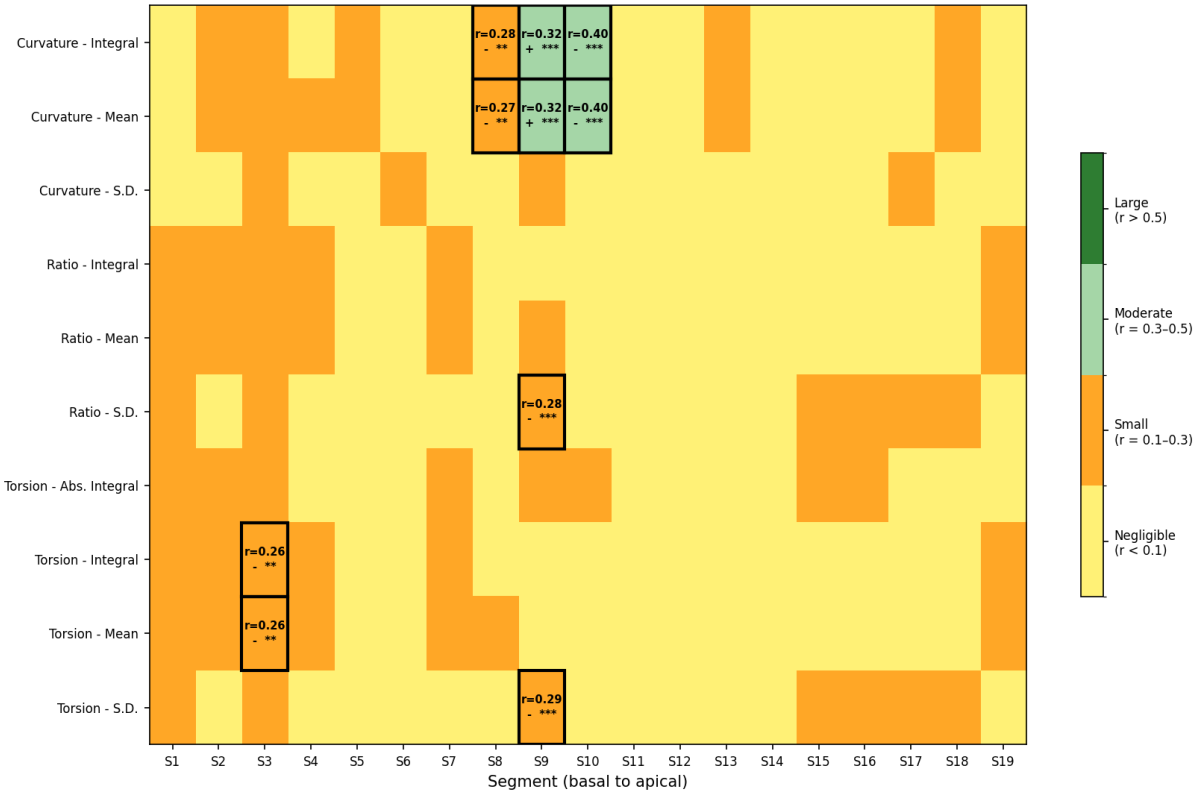


Fig. 3*: Sex differences of curvature and torsion metrics across the 19 segments (S1-S19). Note : Colors from yellow to green represent the effect size r ; black borders indicate segment with significant differences (p<0.05) ; * p < 0.05, ** p < 0.01, *** p < 0.001. The sign indicates the direction of the effect (+ : female > male; - : male > female)*

Table 2 : Statistically significant differences in curvature (κ), torsion (τ) and τ/κ ratio between male and female cochleae (Mann-Whitney U test, FDR-BH correction, FDR < 0.05). Notes: p: unadjusted p-value; p_fdr: Benjamini-Hochberg adjusted p-value.

Descriptor	Stat	Segment	F (mean ± SD)	M (mean ± SD)	r	p	p_fdr
κ	mean	S8	22.74 ± 2.80	24.21 ± 2.43	0.27	0.002	0.039
κ	integral	S8	1.03 ± 0.13	1.10 ± 0.12	0.28	0.001	0.032
κ	mean	S9	20.38 ± 2.81	18.63 ± 2.77	0.32	0.000	0.010
κ	integral	S9	1.02 ± 0.15	0.92 ± 0.15	0.32	0.000	0.009
κ	mean	S10	24.34 ± 3.12	26.80 ± 2.79	0.40	<0.001	<0.001
κ	integral	S10	1.10 ± 0.15	1.22 ± 0.13	0.40	<0.001	<0.001
τ	mean	S3	-3.90 ± 7.21	2.48 ± 12.60	0.26	0.002	0.048
τ	integral	S3	-0.28 ± 0.39	0.06 ± 0.66	0.26	0.003	0.049
τ	SD	S9	9.28 ± 7.71	14.47 ± 11.20	0.29	0.001	0.029
τ/κ	SD	S9	0.59 ± 0.62	1.04 ± 1.13	0.28	0.001	0.029

Scalar cochlear metrics

Four of the six OTOPLAN scalar metrics differed significantly between male and female cochleae after Benjamini-Hochberg correction (**Table 3**, **Fig. 4**), with male cochleae showing

consistently larger values. However, effect sizes remained small to moderate (r = 0.21-0.43), and distributions showed substantial overlap between sexes, indicating that no single scalar metric provides sufficient discriminative power for individual sex estimation.

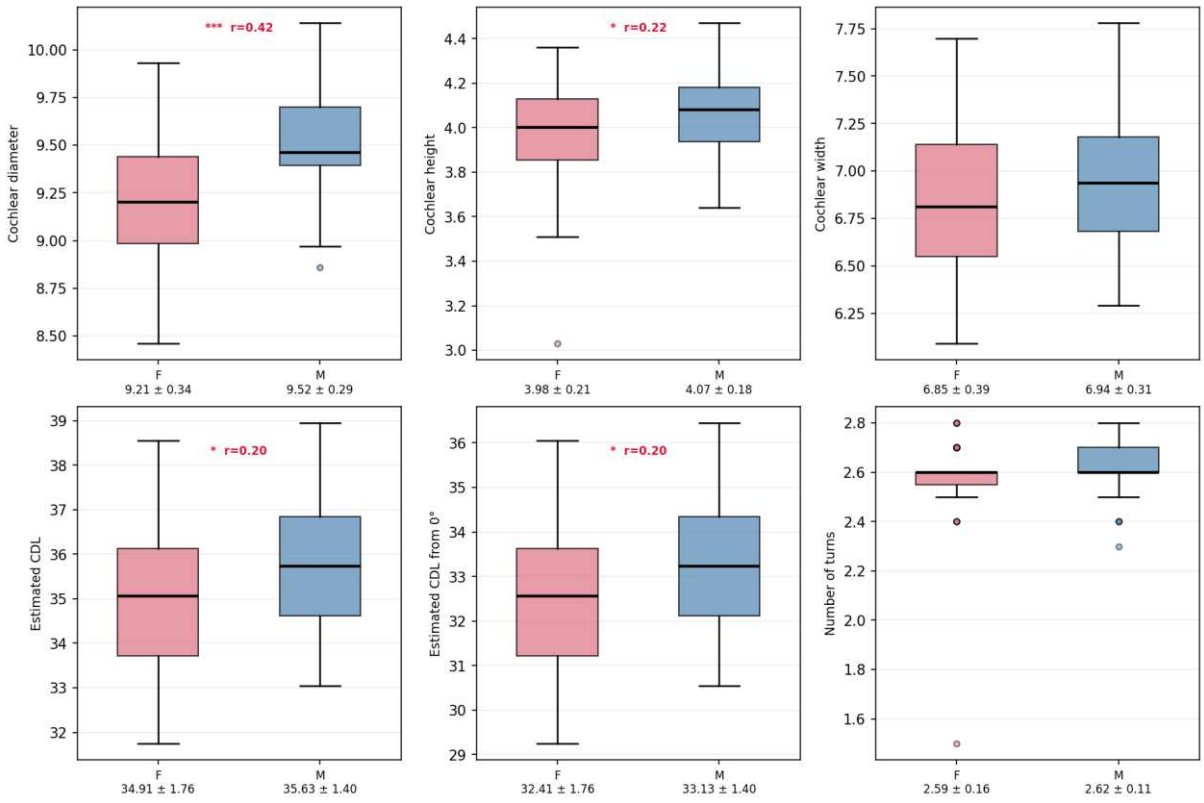


Fig. 4 *: Boxplots of the six scalar metrics in female (pink, n = 77) and male (blue, n = 62) cochleae. Values below each boxplot indicate mean ± standard deviation. Notes: Significant differences after Benjamini-Hochberg correction are annotated in red with the corresponding effect size r (Mann-Whitney U test). p_fdr: Benjamini-Hochberg adjusted p-value. * p_fdr < 0.05; ** p_fdr < 0.01; *** p_fdr < 0.001.*

Table 3 : Statistically significant differences in scalar cochlear metrics between male and female cochleae (Mann-Whitney U test, FDR-BH correction, FDR < 0.05). Notes: p: unadjusted p-value; p_fdr: Benjamini-Hochberg adjusted p-value.

Variable	**F (mean ± SD)**	**M (mean ± SD)**	**r**	**p**	**p_fdr**
Cochlear diameter	9.21 ± 0.34	9.52 ± 0.29	0.43	<0.001	<0.001
Cochlear height	3.98 ± 0.21	4.07 ± 0.18	0.22	0.009	0.022
Estimated CDL	34.91 ± 1.73	35.63 ± 1.40	0.21	0.014	0.022
Estimated CDL from 0	32.41 ± 1.73	33.13 ± 1.40	0.21	0.014	0.022

Model selection

To assess the respective contributions of scalar and geometric information, three configurations were evaluated, each independently optimized using the same Bayesian optimization procedure: scalar metrics alone, geometric descriptors alone, and their combination (**Fig. 5**). The scalar-only model showed marked overfitting, with an AUC during optimization of 0.82 collapsing to 0.59 when tested on the holdout set. The geometry-only model generalized more reliably (holdout AUC = 0.69), confirming that curvature- and torsion-based descriptors capture genuine sex-related shape variation. The combined model outperformed both configurations on all holdout metrics (AUC = 0.77, accuracy = 79%, balanced accuracy = 79%), supporting the complementarity of scalar and geometric information. The following results therefore describe this combined model exclusively. Given the limited size of the holdout set, the numerical differences between the three configurations fall within overlapping confidence intervals and should be regarded as indicative rather than definitive. The preference for the combined model rests on the convergence of cross-validation and holdout estimates and on the biological plausibility of combining size and shape information, rather than on a statistically significant margin.

Model optimization

Bayesian hyperparameter optimization identified an XGBoost classifier as the best-performing model, achieving a mean cross-validated AUC of 0.81 on the training set and selected combining scalar metrics with the τ/κ ratio descriptors only. Curvature and torsion metrics were not kept for the classification model. The optimal configuration is detailed in Appendix II.

Classification performance

On the independent holdout set (n = 29 cochleae), the XGBoost model achieved an AUC-ROC of 0.771 (95% CI: 0.549-0.953), an overall accuracy of 79.3% and a balanced accuracy of 79.1% (**Fig. 5 and 6**). These three metrics provide complementary information about model performance. The AUC-ROC reflects the model's ability to rank male cochleae above female cochleae regardless of the classification threshold, indicating good discriminative ability above chance (0.50) but below perfect separation (1.0). The wide confidence interval of the AUC reflects the limited size of the holdout set and should be interpreted with caution. The overall accuracy and balanced accuracy, nearly identical here given the nearly equal class sizes in the holdout set, indicate that approximately 8 out of 10 cochleae were correctly classified.

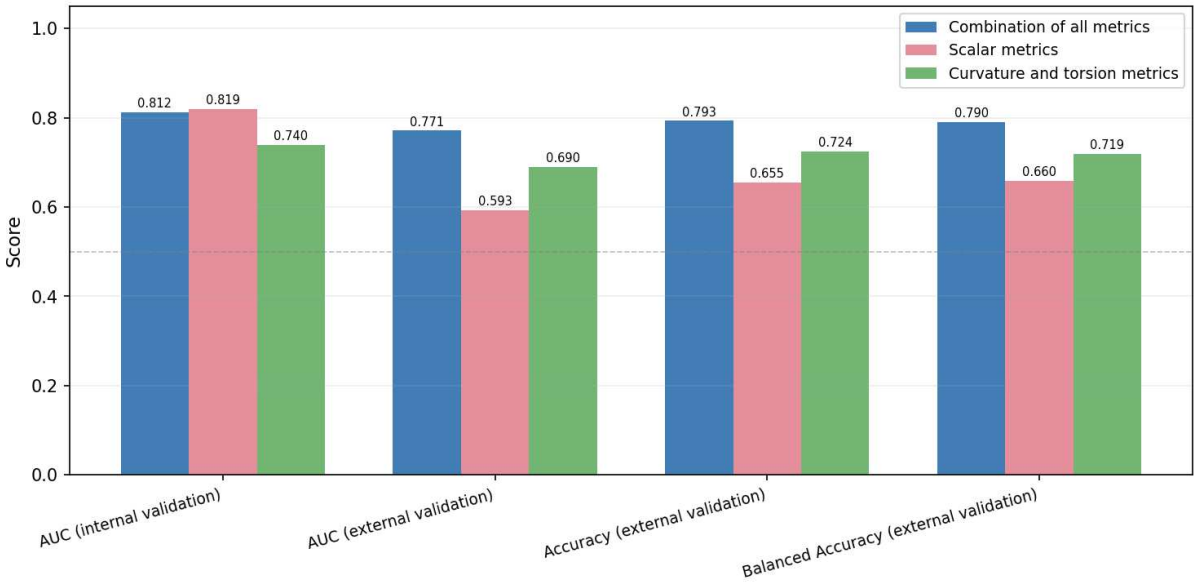


Fig. 5 *: Performance comparison of the three model configurations on cross-validation and holdout sets. Note : The dashed line indicates chance level (0.50). CV AUC: mean cross-validated AUC-ROC on the training set obtained during the model optimization phase; Holdout AUC: AUC-ROC on the independent holdout set; Holdout Acc.: overall accuracy; Holdout Bal. Acc.: balanced accuracy. In blue combination of geometric and scalar metrics; in pink, scalar metrics only; in green geometric metrics only*

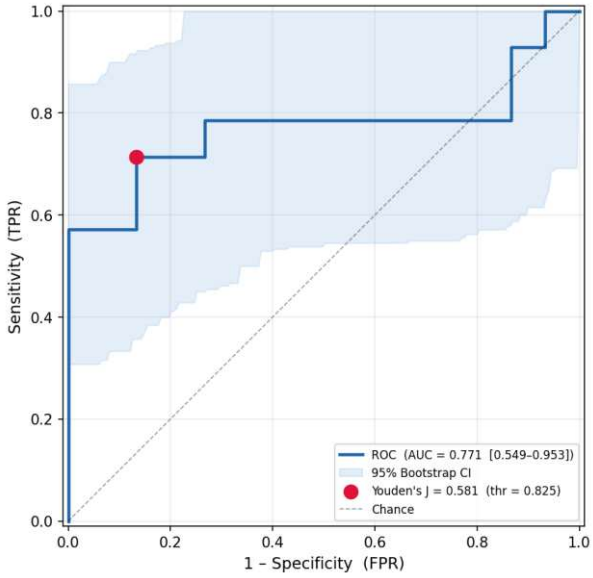


Fig. 1 *: ROC curve on the holdout set*

At the optimal classification threshold identified by Youden's J index (J = 0.58, threshold = 0.825), the model correctly classified 10 out of 14 male cochleae (sensitivity of 71%) and 13 out of 15 female cochleae (specificity of 87%) (**Fig. 16**). Male cochleae showed a broader and more dispersed distribution of predicted probabilities, suggesting greater morphological heterogeneity in this group, while female cochleae were clustered more consistently below the threshold (**Fig. 7**).

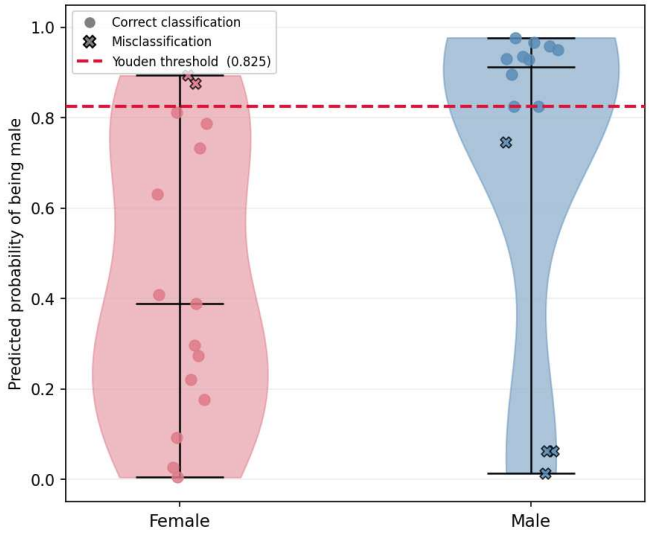


***Fig. 7** : Distribution of predicted probabilities by sex on the holdout set. Note : Each point represents one cochlea. The dashed red line indicates the optimal classification threshold identified by Youden's J index. The width of the violin at each score value reflects the density of cochleae receiving that predicted probability, with wider regions indicating a higher concentration of observations*

Feature importance

Permutation importance analysis identified cochlear diameter as the dominant feature (importance = 0.25 ± 0.05), followed at a distance by τ/κ ratio descriptors distributed across several segments of the spiral (**Fig. 8**). The large gap between cochlear diameter and all other features suggests that size-related information drives most of the classification signal, with geometric shape contributing a secondary but complementary role. This is consistent with the univariate analysis, where cochlear diameter showed the strongest sex difference ($r = 0.43$), and with the model selection results, where the scalar-only model collapsed on the holdout set despite a high cross-validated AUC, indicating that diameter alone, without geometric context, is insufficient for robust generalization.

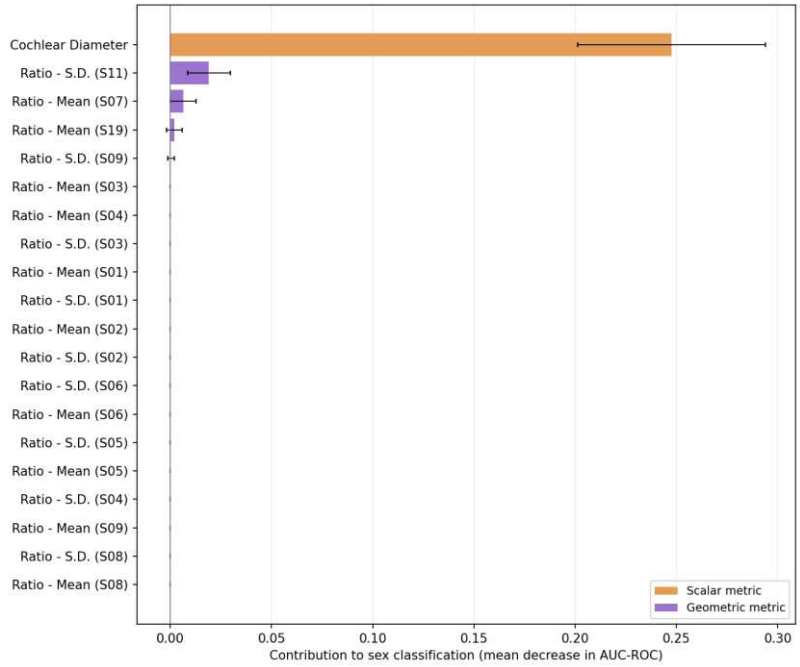


***Fig. 2** : Permutation feature importance on the training set. Note : S.D. : Standard Deviation; in bracket the number of segments*

DISCUSSION

Main results and interpretation

The present study provides a proof of concept that sex-related cochlear shape information can be extracted from routine clinical CT data using an automated segmentation workflow (OTOPLAN) and translated into a reproducible classification framework. An AUC-ROC of 0.77, an overall accuracy of 79%, and a balanced accuracy of 79% were achieved on an independent holdout set. Although moderate, this performance is consistent with several previous studies of bony labyrinth sexual dimorphism based on CT-derived morphometric features and demonstrates that cochlear shape differences remain detectable under clinically realistic imaging conditions[11, 12].

The feature importance analysis revealed a clear hierarchy: cochlear diameter dominated classification performance, followed at a distance by the τ/κ ratio distributed across several segments of the spiral. Cochlear diameter captures overall cochlear size, which is larger in

males consistent with general sexual size dimorphism. However, the finding that a scalar-only model showed marked overfitting, with a cross-validated AUC of 0.82 collapsing to 0.59 on the holdout set, demonstrates that size alone is insufficient for robust generalization: without geometric context, the model memorizes training patterns rather than learning a transferable discriminative rule.

The τ/κ ratio was the only geometric descriptor selected by Bayesian optimization. The fact that this ratio outperforms κ and τ as standalone descriptors suggests that cochlear sexual dimorphism is not simply reducible to differences in bending magnitude but involves subtle differences in the spatial organization of the spiral that the ratio captures more stably. Together, cochlear diameter and τ/κ appear to reflect two biologically distinct and complementary dimensions of cochlear sexual dimorphism, which jointly provide the most robust basis for classification.

Comparison with previous studies on inner ear sexual dimorphism

Previous studies based on linear and angular labyrinthine measurements have reported sex classification accuracies ranging from approximately 80 to 85% in adult and mixed-age samples [11–13]. In the present study, classification performance falls within the moderate range of these values, despite relying on a different feature space and a semi-automated segmentation pipeline.

When compared with the curve-based approach proposed by Braga et al. (2019), the more moderate performance observed here is likely attributable to substantial methodological differences. In particular, Braga et al. relied on high-resolution imaging, manual or semi-manual curve extraction, and advanced shape-space modeling, whereas the present work deliberately prioritizes automation, reproducibility, and compatibility with routine clinical CT imaging[14]. This trade-off between maximal discriminatory power and real-world applicability is central to the intended forensic and bio-archaeological positioning of the method.

Strengths of the approach

Several methodological features distinguish the present pipeline from prior work and support its practical applicability. First, cochlear segmentation relies on OTOPLAN, a validated, commercially available platform already deployed in clinical cochlear implant planning workflows, which ensures reproducibility without requiring dedicated imaging protocols or manual expertise. Second, the joint optimization of preprocessing and modeling parameters via Bayesian optimization removes the risk of arbitrary parameter choices inflating performance estimates. Third, the analysis is entirely non-destructive, preserving the specimen for subsequent genetic testing.

Finally, to facilitate the application of the proposed method to new cases, a standalone Python script is made available alongside this article. The script extracts the three-dimensional lateral wall coordinates and the six scalar cochlear metrics from the OTOPLAN output files, applies the same normalization, resampling, and feature extraction pipeline used during model training, and returns a sex prediction for each cochlea using the trained classifier. For each prediction, the script reports the internal probability assigned by the model to the male class, the distance of this score from the decision boundary (% Confidence), and a pipeline stability index derived from 200 bootstrap resampling iterations of the 400 landmarks, which reflects the sensitivity of the prediction to small variations in cochlear segmentation (% Coherence). Users should note that none of these indicators represents the probability that the classification is correct: a coherent or high-confidence prediction may still be incorrect.

Relevance to forensic identification in fragmented and degraded remains

The forensic relevance of cochlear shape–based sex estimation becomes particularly apparent when considering the high frequency of medico-legal cases involving incomplete or severely degraded skeletal remains. Retrospective studies from forensic anthropology units indicate that approximately **30–40% of cases** involve only a single bone or a very limited number of skeletal

elements, often as a consequence of advanced decomposition, scavenging, fire, or mechanical trauma[5, 20]. In such situations, classical sex estimation methods relying on the pelvis or cranial morphology are frequently inapplicable.

This limitation is further amplified in mass fatality incidents, including fires, explosions, and aviation disasters, where remains are often heavily fragmented, commingled, or thermally altered[21]. Well-documented disaster victim identification (DVI) operations have shown that a substantial proportion of victims may be represented only by small skeletal fragments, and that complete skeletal remains are the exception rather than the rule[21]. Under these extreme conditions, identification relies primarily on DNA analysis, and even then, a subset of victims may remain unidentified due to the absence of recoverable biological material.

Among the skeletal elements that tend to survive such conditions, the petrous part of the temporal bone occupies a unique position. Owing to its exceptional density, the petrous bone is highly resistant to taphonomic degradation, mechanical stress, and thermal alteration, and may survive even severe fires or cremation with minimal structural damage[22, 23]. Consequently, it is often the most frequently recovered cranial fragment in advanced decomposition or fire-related contexts. The bony labyrinth enclosed within the petrous portion therefore represents a rare and valuable source of morphological information when other skeletal indicators are no longer available.

Positioning cochlear shape analysis within identification workflows

The classification accuracy achieved in the present study does not support the use of cochlear shape analysis as a stand-alone method for sex estimation in individual forensic casework. However, forensic identification rarely relies on a single criterion. Instead, biological profiling is based on the convergence of multiple probabilistic indicators, each contributing complementary information[24, 25]. Within this framework, cochlear shape–based sex

estimation may serve as a valuable adjunct in cases where only dense skeletal elements such as the petrous bone are available.

Even a moderate gain in sex attribution probability may substantially narrow the pool of potential matches in missing persons databases, particularly when combined with age estimation, ancestry proxies, contextual information, or DNA analysis. Importantly, cochlear shape analysis can be performed non-destructively from CT or CBCT data, preserving the specimen for subsequent genetic testing, which remains the gold standard for identification in highly degraded cases[24].

Emerging imaging technologies and future perspectives

Beyond methodological refinements, ongoing advances in imaging technology are likely to substantially improve the performance and applicability of cochlear shape–based sex estimation in the near future. In particular, the increasing availability of high-resolution cone-beam CT (CBCT) systems capable of imaging dry bone specimens represents a major opportunity for forensic anthropology and bioarchaeology. Modern CBCT devices can achieve isotropic voxel sizes substantially smaller than those of conventional clinical CT, while being portable, relatively affordable, and deployable in field laboratories or forensic facilities[26]. Several studies have demonstrated the suitability of CBCT for detailed assessment of temporal bone and inner ear anatomy, suggesting that cochlear shape descriptors could be extracted with improved geometric fidelity compared with routine clinical CT[27].

In parallel, rapid developments in advanced CT technologies, including photon-counting CT, promise further improvements in spatial resolution, contrast-to-noise ratio, and artifact reduction[28]. Photon-counting detectors allow more precise characterization of fine bony structures at clinically acceptable radiation doses and have already demonstrated improved visualization of temporal bone anatomy compared with energy-integrating CT systems[29]. Enhanced delineation of the cochlear spiral and bony labyrinth would be expected to reduce

segmentation noise, stabilize curvature and torsion estimates, and ultimately improve classification performance.

Importantly, these technological advances align particularly well with the present semi-automated workflow, which is designed to leverage improvements in input data without requiring fundamental changes to the analytical framework. As imaging resolution and segmentation accuracy improve, the same curvature- and torsion-based descriptors may capture subtler sex-related shape differences that are currently obscured by resolution limits. In this context, the moderate classification accuracy observed in the present study should be interpreted not as an intrinsic limitation of cochlear shape–based sex estimation, but rather as a lower bound under contemporary clinical imaging conditions.

Limitations and concluding remarks

Several limitations of the present study should be acknowledged. First, the analysis was based on routine clinical CT scans acquired using heterogeneous scanners and acquisition protocols. While this heterogeneity reflects real-world forensic and clinical conditions and supports the robustness of the proposed pipeline, it may also limit optimal performance compared with standardized high-resolution imaging. Future studies using harmonized acquisition protocols or advanced imaging modalities may help refine performance estimates.

Second, cochlear segmentation relied on a semi-automated, commercially available software platform. Although such tools offer important advantages in terms of reproducibility and scalability, the internal segmentation parameters are not fully transparent or user adjustable. As a result, some degree of methodological dependence on proprietary algorithms remains unavoidable.

Third, the present dataset lacked detailed anthropological metadata beyond biological sex, precluding the evaluation of potential confounding factors such as body size, ancestry, or

population-specific variation. These factors have been shown to influence labyrinthine morphology and may partly account for inter-individual variability observed in the present study.

From a methodological perspective, cochlear shape was characterized using curvature- and torsion-based descriptors, which provide an interpretable but simplified representation of three-dimensional geometry. More sophisticated shape-space approaches may capture additional discriminatory information but would come at the cost of increased computational complexity and reduced accessibility.

The present study was conducted on in vivo imaging data, and the transferability of the proposed approach to dry bone specimens remains to be formally validated. Differences in imaging contrast, noise characteristics, and segmentation behavior between living and archaeological or forensic skeletal material may affect performance and should be addressed in future work using dedicated CBCT or high-resolution CT protocols on dry bone collections.

The relatively small size of the holdout set (n = 29 cochleae) represents an additional limitation. Although the holdout design provides an unbiased estimate of generalization, the wide bootstrap confidence interval observed for the AUC-ROC (95% CI: 0.55-0.95) reflects the instability inherent to performance estimation on small samples. Larger independent validation cohorts will be needed to obtain more precise and reliable estimates of classification performance. Model calibration was not evaluated. Given the small holdout set, calibration estimates would have been unstable, and the predicted probabilities should therefore be interpreted as relative scores rather than as calibrated probabilities of being male. Assessment of calibration in larger cohorts is an important direction for future work.

Finally, a mild sex imbalance was present in the dataset, with 77 female and 62 male cochleae. Although stratified sampling was used to preserve this ratio across the training and holdout sets, and balanced accuracy was reported alongside overall accuracy to account for this imbalance,

this distribution may have influenced certain classification metrics and should be considered when interpreting the results.

CONCLUSION

In conclusion, this study demonstrates that sex-related cochlear shape information can be extracted from routine clinical CT data using a semi-automated and reproducible workflow. While classification accuracy remains moderate, the method directly targets scenarios in which classical skeletal indicators are unavailable and only dense elements such as the petrous bone persist. Cochlear shape analysis should therefore be viewed not as a replacement for established forensic methods, but as a complementary tool that may contribute valuable probabilistic information in some of the most challenging identification contexts.

DECLARATIONS

Acknowledgments: This manuscript was edited for language clarity using a large language model (Claude AI, Anthropic) under full author supervision. The authors remain responsible for all content.

Data availability: The code supporting this study is openly archived on Zenodo in two deposits. The development and analysis pipeline used to train the models is available at https://doi.org/10.5281/zenodo.20542412 . The standalone tool that applies the model to new OTOPLAN segmentations, together with the trained model package for the geometric + scalar metrics configuration is available at https://doi.org/10.5281/zenodo.20542460. Each deposit includes a version-pinned requirement.txt and was run under Python 3.12. The trained model packages for the two other configurations (scalars only and geometry only) are available from the corresponding author on reasonable request.

Author contributions:. Priya Vishnumurthy: Methodology, Software, Formal analysis, Data curation, Visualization, Writing - review & editing. Mary Daval (ORCID: 0000-0002-9815-1602): Resources, Investigation, Data curation, Writing - review & editing. Luke Helpard: Software, Resources, Writing - review & editing. Ismael Ouzzine: Data curation, Resources. Ralph Haddad (ORCID: 0000-0002-1388-3848): Methodology, Supervision. Nikola Ivanovic: Software, Resources, Writing - review & editing. Vincent Jourdes: Resources, Writing - review & editing. Justin Michel (ORCID: 0000-0002-5275-909X): Conceptualization, Supervision, Writing - review & editing. Stéphane Gargula (ORCID: 0000-0002-5034-4210): Conceptualization, Methodology, Writing - original draft, Project administration, Supervision.

Ethics approval. This retrospective study was approved by the Rothschild Foundation Hospital review board (IRB 00012801, study number CE_20230328_5_SGA) and conducted in accordance with the Declaration of Helsinki. Given the retrospective use of fully anonymized imaging data, the requirement for informed consent was waived.

Clinical trial number: not applicable

Funding. This research received no specific grant from any funding agency.

Conflicts of interest. Luke HELPARD, Nikola IVANOVIC and Vincent JOURDES are employees of MED-EL (Innsbruck, Austria). All other authors declare no conflicts of interest. OTOPLAN is a commercial product of CASCINATION AG in cooperation with MED-EL.

The MED-EL-affiliated authors contributed expertise on cochlear segmentation and on the OTOPLAN export format. Study conception, data analysis, interpretation of results, and the decision to submit for publication were conducted independently by the academic authors. MED-EL had no role in the analysis or in the decision to publish.

REFERENCES

1. Durić M, Rakocević Z, Donić D (2005) The reliability of sex determination of skeletons from forensic context in the Balkans. Forensic Sci Int 147:159–164. https://doi.org/10.1016/j.forsciint.2004.09.111

2. İşcan MY (2013) The human skeleton in forensic medicine, Third edition. Charles C Thomas Publisher, LTD, Springfield, Illinois, U.S.A

3. Bruzek J (2002) A method for visual determination of sex, using the human hip bone. Am J Phys Anthropol 117:157–168. https://doi.org/10.1002/ajpa.10012

4. Peckmann T, Orr K, Meek S, Manolis S (2015) Sex determination from the talus in a contemporary Greek population using discriminant function analysis. Journal of Forensic and Legal Medicine 33:14–19. https://doi.org/10.1016/j.jflm.2015.03.011

5. Gaudio D, Fernandes DM, Schmidt R, et al (2019) Genome-Wide DNA from Degraded Petrous Bones and the Assessment of Sex and Probable Geographic Origins of Forensic Cases. Sci Rep 9:8226. https://doi.org/10.1038/s41598-019-44638-w

6. Schutkowski H (1993) Sex determination of infant and juvenile skeletons: I. Morphognostic features. Am J Phys Anthropol 90:199–205. https://doi.org/10.1002/ajpa.1330900206

7. Cardoso HFV, Saunders SR (2008) Two arch criteria of the ilium for sex determination of immature skeletal remains: A test of their accuracy and an assessment of intra- and inter-observer error. Forensic Science International 178:24–29. https://doi.org/10.1016/j.forsciint.2008.01.012

8. Humphrey LT (1998) Growth patterns in the modern human skeleton. American Journal of Physical Anthropology 105:57–72. https://doi.org/10.1002/(SICI)1096-8644(199801)105:1%253C57::AID-AJPA6%253E3.0.CO;2-A

9. Spoor F, Zonneveld F (1995) Morphometry of the primate bony labyrinth: a new method based on high-resolution computed tomography. J Anat 186:271–286

10. Spoor F, Zonneveld F (1998) Comparative review of the human bony labyrinth. Am J Phys Anthropol Suppl 27:211–251. https://doi.org/10.1002/(sici)1096-8644(1998)107:27+%253C211::aid-ajpa8%253E3.3.co;2-m

11. Osipov B, Harvati K, Nathena D, et al (2013) Sexual dimorphism of the bony labyrinth: A new age-independent method. American Journal of Physical Anthropology 151:290–301. https://doi.org/10.1002/ajpa.22279

12. Boucherie A, Polet C, Lefèvre P, Vercauteren M (2021) Sexing the bony labyrinth: A morphometric investigation in a subadult and adult Belgian identified sample. J Forensic Sci 66:808–820. https://doi.org/10.1111/1556-4029.14663

13. Gargula S, Michel J, Ayache D, et al (2024) Sex determination by linear measurements of the bony labyrinth in cone-beam and computed tomography: a practical method. Forensic Sci Med Pathol 20:310–314. https://doi.org/10.1007/s12024-023-00675-6

14. Braga J, Samir C, Risser L, et al (2019) Cochlear shape reveals that the human organ of hearing is sex-typed from birth. Sci Rep 9:10889. https://doi.org/10.1038/s41598-019-47433-9

15. Veillon F, Stierle JL, Dussaix J, et al (2006) [Otosclerosis imaging: matching clinical and imaging data]. J Radiol 87:1756–1764. https://doi.org/10.1016/s0221-0363(06)74157-9

16. Gatto A, Tofanelli M, Costariol L, et al (2023) Otological Planning Software-OTOPLAN: A Narrative Literature Review. Audiol Res 13:791–801. https://doi.org/10.3390/audiolres13050070

17. Távora-Vieira D, Voola M, Kuthubutheen J, et al (2023) Evaluation of the Performance of OTOPLAN-Based Cochlear Implant Electrode Array Selection: A Retrospective Study. J Pers Med 13:1276. https://doi.org/10.3390/jpm13081276

18. Paouris D, Kunzo S, Goljerová I (2023) Validation of Automatic Cochlear Measurements Using OTOPLAN® Software. J Pers Med 13:805. https://doi.org/10.3390/jpm13050805

19. Collins GS, Moons KGM, Dhiman P, et al (2024) TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 385:e078378. https://doi.org/10.1136/bmj-2023-078378

20. Indra L, Lösch S (2021) Forensic Anthropology Casework from Switzerland (Bern): Taphonomic Implications for the Future. Forensic Science International: Reports 4:100222. https://doi.org/10.1016/j.fsir.2021.100222

21. de Boer HH, Roberts J, Delabarde T, et al (2020) Disaster victim identification operations with fragmented, burnt, or commingled remains: experience-based recommendations. Forensic Sci Res 5:191–201. https://doi.org/10.1080/20961790.2020.1751385

22. Wahl J, Graw M (2001) Metric sex differentiation of the pars petrosa ossis temporalis. Int J Legal Med 114:215–223. https://doi.org/10.1007/s004140000167

23. Uhl A, Karakostis FA, Harvati K (2022) The Application of Bony Labyrinth Methods for Forensic Affinity Estimation. Biology (Basel) 11:1088. https://doi.org/10.3390/biology11071088

24. Nagare SP, Chaudhari RS, Birangane RS, Parkarwar PC (2018) Sex determination in forensic identification, a review. J Forensic Dent Sci 10:61–66. https://doi.org/10.4103/jfo.jfds_55_17

25. Wang X, Liu G, Wu Q, et al (2024) Sex estimation techniques based on skulls in forensic anthropology: A scoping review. PLoS One 19:e0311762. https://doi.org/10.1371/journal.pone.0311762

26. Scarfe WC, Farman AG (2008) What is cone-beam CT and how does it work? Dent Clin North Am 52:707–730, v. https://doi.org/10.1016/j.cden.2008.05.005

27. Debeaupte M, Hermann R, Pialat J-B, et al (2019) Cone beam versus multi-detector computed tomography for detecting hearing loss. Eur Arch Otorhinolaryngol 276:315–321. https://doi.org/10.1007/s00405-018-5214-y

28. Willemink MJ, Persson M, Pourmorteza A, et al (2018) Photon-counting CT: Technical Principles and Clinical Prospects. Radiology 289:293–312. https://doi.org/10.1148/radiol.2018172656

29. Hermans R, Boomgaert L, Cockmartin L, et al (2023) Photon-counting CT allows better visualization of temporal bone structures in comparison with current generation multi-detector CT. Insights Imaging 14:112. https://doi.org/10.1186/s13244-023-01467-w

APPENDIX I : Lin's CCC

Table 4 : Lin's CCC comparing right and left ears of a same patient. Note : These results can be interpreted as follow : <0 : poor correlation; 0-0.20 : slight; 0.21-0.40 : fair; 0.41-0.60 : moderate; 0.61 - 0.80 : substantial, 0.81-1 : almost perfect; 1 : perfect.

	Estimate	Lower C.I.	Upper C.I.
Diameter	0.919	0.864	0.952
Height	0.831	0.731	0.896
Width	0.843	0.744	0.906
EstimatedCDL	0.890	0.816	0.935
EstimatedCDLFrom0	0.890	0.816	0.935
NumberOfTurns	0.413	0.211	0.580

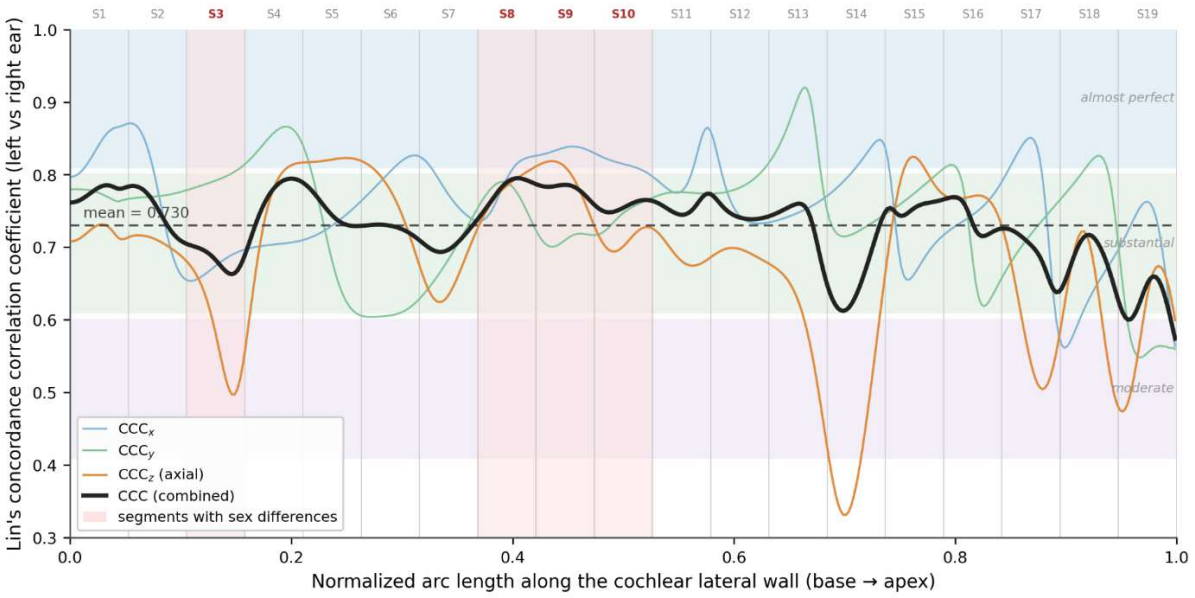


Fig. 3 *: Landmark-level concordance between the left and right cochleae along the lateral wall path. Note : Lin's concordance correlation coefficient (CCC) was computed at each of the 400 landmarks of the cochlear lateral wall path in the 52 patients with both ears available, after mirroring of the left cochleae. The thick black curve represents the combined CCC; the thin curves represent the coordinate-wise CCC for the x, y and z (axial) components. The x-axis corresponds to the normalized arc length along the spiral, from base (0) to apex (1). The*

horizontal dashed line indicates the mean combined CCC (0.730). Background horizontal bands correspond to Lin's interpretation thresholds (moderate: 0.41-0.60; substantial: 0.61-0.80; almost perfect: 0.81-1.00). Vertical lines delimit the 19 equal-length segments (S1-S19); segments showing significant sex differences in curvature or torsion descriptors (S3, S8, S9, S10; Table 2) are shaded in red. Concordance was consistently lowest along the axial (z) component, in line with the limited through-plane resolution of clinical CT

APPENDIX II : Pre-treatment pipeline and classification model optimization

Ten parameters were included in the optimization search space:

Curve preprocessing parameters. The first parameter controlled whether a basal truncation was applied (binary: yes/no). The basal segment of the extracted cochlear curve includes a short vertical extension that is not part of the cochlear spiral. When truncation was activated, the point of maximum turning angle within the first 40 points was identified and all preceding points were removed. The second parameter controlled the resampling density (100 to 400 points, step = 50). All curves were reparametrized by normalized arc length and resampled to a fixed number of equidistant points along the spiral, ensuring regional comparisons along homologous portions of the spiral. The third and fourth parameters governed the Savitzky-Golay smoothing filter (window size: 5 to 31 points; polynomial order: 2 to 4), controlling the trade-off between noise reduction and preservation of local geometric detail.

Feature construction parameters. The fifth parameter defined the number of equal-length segments into which the normalized spiral was divided (5 to 20 segments). The sixth, seventh and eighth parameters were binary flags independently controlling whether curvature (κ), torsion (τ), and the τ/κ ratio were included in the feature vector. Any combination excluding all three geometric descriptor groups was assigned an AUC of 0.5 to prevent Optuna from exploring configurations devoid of shape information. The ninth parameter controlled whether the scalar cochlear metrics were appended to the geometric feature vector, avoiding unnecessary inflation of the feature space with potentially redundant variables.

Model and hyperparameters. The tenth parameter selected the classification algorithm among five candidates: Gradient Boosting, Support Vector Machine, Random Forest, XGBoost, and LightGBM. Algorithm-specific hyperparameters were optimized conditionally on this choice.

Table 5 : Hyperparameter search space and optimal values identified by Bayesian optimization

Parameter	Tested range	Value retained
Truncation	Yes/No	No
Resampling number of landmarks	100-400, step = 50	200
Savitzky-Golay window length	5, 31, step=2	13
Savitzky-Golay polynomial order	2, 4, step = 1	2
Number of segments	5,20, step = 1	19
Use κ	Yes/No	No
Use τ	Yes/No	No
Use τ/κ	Yes/No	Yes
Use scalar metrics	Yes/No	Yes
Classification algorithm	GBM, SVM, RF, XGBoost, LightGBM	XGBoost

APPENDIX III : Outlier PCA

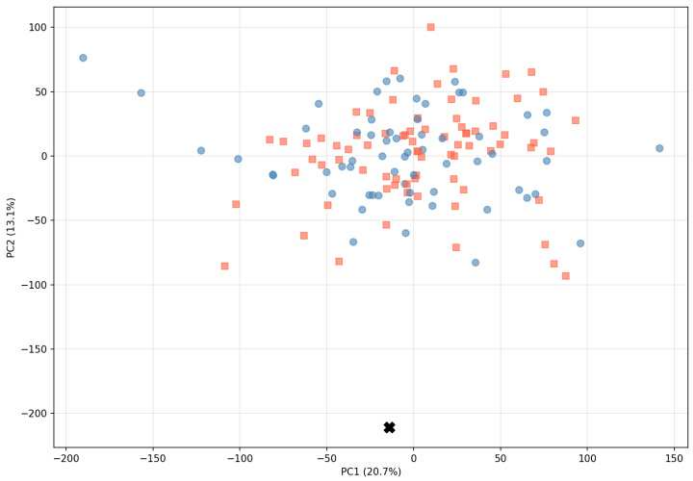


Fig. 4 *: Principal component analysis for outlier detection. Each point represents one cochlea projected onto the first two principal components (PC1 and PC2) computed from the full geometric feature set. PC1 and PC2 explain 20.7% and 13.1% of the total variance, respectively. Notes: Blue circle: male cochlea; Pink square: female cochlea. Cross: cochlea excluded as outlier (|z(PC2)| > 4.0)*