

**METHOD**

# Comparability of skeletal fibulae surfaces generated by different source scanning (dual-energy CT scan vs. high resolution laser scanning) and 3D geometric morphometric validation

Annalisa Pietrobelli<sup>1</sup>  | Rita Sorrentino<sup>1,2</sup>  | Veronica Notariale<sup>3</sup> | Stefano Durante<sup>3,4</sup> | Stefano Benazzi<sup>2,5</sup> | Damiano Marchi<sup>6,7</sup> | Maria Giovanna Belcastro<sup>1</sup>

<sup>1</sup>Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy

<sup>2</sup>Department of Cultural Heritage, University of Bologna, Ravenna, Italy

<sup>3</sup>IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy

<sup>4</sup>IRCCS Azienda Ospedaliero-Universitaria di Bologna Policlinico S. Orsola, Bologna, Italy

<sup>5</sup>Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

<sup>6</sup>Department of Biology, University of Pisa, Pisa, Italy

<sup>7</sup>Centre for the Exploration of the Deep Human Journey, University of the Witwatersrand, Johannesburg, South Africa

**Correspondence**

Annalisa Pietrobelli and Rita Sorrentino, Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy.

Email: [annalisa.pietrobelli2@unibo.it](mailto:annalisa.pietrobelli2@unibo.it); [rita.sorrentino2@unibo.it](mailto:rita.sorrentino2@unibo.it)

**Funding information**

Istituto Ortopedico Rizzoli

**Abstract**

This work aims to test accuracy and comparability of 3D models of human skeletal fibulae generated by clinical CT and laser scanner virtual acquisitions. Mesh topology, segmentation and smoothing protocols were tested to assess variation among meshes generated with different scanning methods and procedures, and to evaluate meshes-interchangeability in 3D geometric morphometric analysis. A sample of 13 left human fibulae were scanned separately with Revolution Discovery CT dual energy (0.625 mm resolution) and ARTEC Space Spider 3D structured light laser scanner (0.1 mm resolution). Different segmentation methods, including half-maximum height (HMH) and MIA-clustering protocols, were compared to their high-resolution standard generated with laser-scanner by calculating topological surface deviations. Different smoothing algorithms were also evaluated, such as Laplacian and Taubin smoothing. A total of 142 semilandmarks were used to capture the shape of both proximal and distal fibular epiphyses. After Generalized Procrustes superimposition, the Procrustes coordinates of the proximal and distal fibular epiphyses were used separately to assess variation due to scanning methods and the operator error. Smoothing algorithms at low iteration do not provide significant variation among reconstructions, but segmentation protocol may influence final mesh quality (0.09–0.24 mm). Mean deviation among CT-generated meshes that were segmented with MIA-clustering protocol, and laser scanner-generated ones, is optimal (0.42 mm, ranging 0.35–0.56 mm). Principal component analysis reveals that homologous samples scanned with the two methods cluster together for both the proximal and distal fibular epiphyses. Similarly, Procrustes ANOVA reveals no shape differences between scanning methods and replicates, and only 1.38–1.43% of shape variation is due to scanning device. Topological similarities

Stefano Benazzi, Damiano Marchi and Maria Giovanna Belcastro have equal contribution.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Journal of Anatomy* published by John Wiley & Sons Ltd on behalf of Anatomical Society.

support the comparability of CT- and laser scanner-generated meshes and validate its simultaneous use in shape analysis with potential clinical relevance. We precautionarily suggest that dedicated trials should be performed in each study when merging different data sources prior to analyses.

#### KEYWORDS

3D geometric morphometrics comparability, dual-energy CT scans, high-resolution surface scanning, human fibula

## 1 | INTRODUCTION

Several digitizing techniques, such as laser scanning, digital three-dimensional (3D) photogrammetry, computed tomography (CT) and micro-CT scanning, are widely implemented in morphological analyses, to evaluate bone shape and metric features. Such approaches consist of effective, repeatable, non-invasive diagnostic tools that allow the digitized object to be thoroughly examined without being physically manipulated, both in their external and internal geometry (Mantini & Ripani, 2009; Profico et al., 2019; Uldin, 2017; Weber & Bookstein, 2011; Weber et al., 2001; Weber, 2014, 2015).

Often, the digitized 3D surface meshes obtained by CT and laser scanner are interchangeably adopted together and merged in morphological analyses (e.g., Frelat et al., 2017; Mounier & Lahr, 2019; Sorrentino et al., 2021; Sorrentino, Carlson, et al., 2020). Several studies have compared surface meshes obtained through 3D surface scanning with meshes obtained through medical CT scanning (Adams et al., 2015; Brzobohatá et al., 2012; Choi et al., 2002; DeVries et al., 2008; Fahrni et al., 2017; Fourie et al., 2011; Lalone et al., 2015; Ramme et al., 2009; Waltenberger et al., 2021) or micro-CT scanning (Teeter et al., 2012) finding consistency among methods in measurement accuracy. Indeed, Adams et al. (2015) attested 0.4 mm mean surface deviations among surface meshes of hominin fossils digitized with a laser scanner, clinical CT and micro-CT scans. More recently, Waltenberger et al. (2021) compared surface meshes of human pelvises generated utilizing different digitizing methodologies (CT scans, laser scans, photogrammetry, MicroScribe digitizer) through a topological and landmark-based geometric morphometric (GM) approach. In their work, the average deviation among the surface meshes spanned between 100–200  $\mu\text{m}$ , and the GM results attested the comparability among 3D meshes acquired with CT scanning, 3D structured light scanning and photogrammetry (Waltenberger et al., 2021).

While most studies used the skull to investigate differences among scanning devices (e.g., Brzobohatá et al., 2012; Fahrni et al., 2017; Pinsky et al., 2006), others examined articular surfaces of long bones, revealing less than 0.4 mm average deviation among acquired elbow meshes (Lalone et al., 2015). Similar values were obtained when assessing mesh accuracy of the bones of the forearm (Oka et al., 2009). Considering the human femur, topological analyses on CT- and laser scanner-generated meshes found a deviation spanning

0.71–0.79 mm (Gelaude et al., 2008; Soodmand et al., 2018; Stephen et al., 2021). Additionally, Stephen et al. (2021) analyzed also the tibia and found comparable deviation of CT vs. laser scanner generated meshes (less than 0.71 mm for tibiae).

Besides the scanning devices, other factors such as the choice of segmentation protocols may heavily influence the final outcome of mesh reconstructions, with potential impact on morphometric studies. Previous studies have tested the effect of bone segmentation procedures on mesh accuracy and comparability in morphometric analyses, including visual -based segmentation protocol, canny edge detection and other automated algorithms (Toro-Ibacache, 2013; Gunz et al., 2012; Fourie et al., 2012; Engelbrecht et al., 2013; Rathnayaka et al., 2011; Ito, 2019). In particular, Ito (2019) found that single thresholding is often not suitable for a complex object with heterogeneous gray-value distributions and had statistically significant effect on shape and size variation.

Furthermore, the choice of smoothing procedures and the number of the iterations used have also shown possible effect on capturing anatomical information of virtually acquired bones, with potential impact on morphometric analyses (Profico et al., 2016; Veneziano et al., 2018).

The present work will explore the impact of different scanning devices, segmentation protocols and smoothing procedures on the reconstruction of surface models to test their comparability in morphometric analyses, with the aim of providing new data on the comparison of digitizing methodologies of long bones. Specifically, we test the accuracy and comparability of CT- and laser scanner-generated meshes in a sample of 13 modern human fibulae, considering the whole bone surface but also focusing on the extremities. The fibula has so far not been included in previous mesh accuracy assessments, despite its digitization in several biomechanical studies, alongside the tibia (e.g., Marchi, 2005, 2007, 2015b; Marchi, 2015a; Marchi et al., 2019, 2022). In addition, fibular 3D meshes are routinely utilized in reconstructive surgery, in the assessment of tibiofibular syndesmotomies performed after ankle fractures (e.g., Ebinger et al., 2013; Souleiman et al., 2021) and surgical planning of mandibular reconstructions from fibular free flaps (e.g., Damecourt et al., 2020; Ni et al., 2021; Ren et al., 2018).

We also aim to provide data on the whole virtual acquisition pipeline, by testing the impact of segmentation protocols (i.e., half-maximum height and MIA-clustering protocols), smoothing procedures

(i.e., Laplacian and Taubin smoothing) and 3D landmark configuration repeatability in inter-method comparisons. The results will add to our understanding of the use of 3D surface models obtained through different scanning techniques and procedures in 3D morphometric studies, with potential application in corrective surgical planning.

## 2 | MATERIALS AND METHODS

### 2.1 | The sample

The analyzed sample includes 13 left fibulae belonging to individuals from the Human Identified Skeletal Collection of Sassari, housed at the University of Bologna (Italy) and dating to the 19th–20th century. The fibulae were selected for their general good state of preservation, except for a case (SS23) intentionally selected for the presence of a small areas of cortical damage, to test the surface reconstruction performance in the case of a partially damaged bone. The human sample includes 8 males and 5 females spanning 20–45-year-old.

### 2.2 | Data acquisition, segmentation protocols, and smoothing procedures

#### 2.2.1 | Laser scanner acquisition

The sample of fibulae was scanned with ARTEC Space Spider 3D (Luxembourg), housed at the BONES Lab in the Department of Cultural Heritage of the University of Bologna. The ARTEC Space Spider 3D is a mobile and structured light laser scanner based on blue light technology, which uses trigonometric triangulation to calculate the distance among the points on the surface of the object and creates its relative point cloud. This machine provides a point-accuracy of 0.05 mm and mesh resolution of 0.1 mm. Moreover, it acquires both geometry and texture data (1.3 megapixel, 24-bits per pixel). The acquisition took place in two recording sessions for each specimen. The first scan was performed by holding the device approximately perpendicular to the surface of the fibular diaphysis, with one of the extremities that was in turn placed vertically into a polystyrene support with a cavity in the center that accommodates the epiphysis, and by rotating the turning table. The non-supported epiphysis and part of the diaphysis were then acquired and subsequently a second scan was performed by inverting the fibular extremities. The integrated scanner software (Artec Studio 9) was utilized to merge the two sets of acquisitions, which were first cleaned and then manually aligned. These roughly aligned scans were then globally registered, with outlying points removed, and subsequently, a single polygonal 3D mesh is created with the Sharp Fusion algorithm, finally generating the 3D surface model that was then saved in .stl (Little Endian) format, following the procedures that are presented by the manufacturer user documentation.

#### 2.2.2 | CT scan acquisition

All fibulae were also digitized through computed tomography (CT), utilizing a Revolution Discovery CT dual energy, with GSI Revolution and HD Revolution configurations, housed at Istituto Ortopedico Rizzoli (Bologna, Italy). The acquisition protocol chosen was Extremity GSI. Once the minimum Field of View (FOV) had been achieved to optimize the best resolution, an acquisition with polychromatic beam at 100 kV and 360 mA was performed with slice thickness and acquisition interval at 0.625 mm. A “Standard” reconstruction filter with WW 400 and WL 40 was selected. Then, two reconstructions with monochromatic beams were made. The first reconstruction is at 70 keV with “Detail” reconstruction filter, with superior detail of bone acquisition in respect to the “Bone” filter with WW 2000 and WL 350. The second reconstruction is at 40 keV, always using the “Detail” filter with WW 400 and WL 400. The obtained reconstructed DICOM (16-bit grayscale, signed, voxel size ranging from 0.39 to 0.507 × 0.625) images at 40 keV, chosen for subsequent analyses, were then processed with Avizo 9.2 (Thermo Fisher Scientific) for image segmentation.

#### 2.2.3 | Image segmentations

On the CT-generated meshes, we applied the half-maximum height (HMH) protocol outlined by Spoor et al. (1993), following the modified version detailed in Coleman and Colbert (2007), which included calculating the HMH for a row of pixels on the bone-to-air transition for 10 randomly selected slices, then averaged and applied to the whole stack. In addition, we also opted to visually segment the image stack utilizing a visual single-threshold based technique including voxels above –550, –600, –650 and –700 grayscale value intensity thresholds, to evaluate how possible differences of the chosen segmentation procedure influenced the 3D reconstruction overall quality, compared to the same surface data obtained by laser scanner. For both HMH and visual single-threshold based segmentations, thresholding was performed in Avizo 9.2 (Thermo Fisher Scientific, Waltham), determining grayscale intensity of selected voxels with either “histogram” or “magic wand” tools in the segmentation editor.

Moreover, we tested an additional segmentation protocol based on a machine-learning approach, MIA-clustering, which implements a clustering algorithm sorting the voxels of an image into clusters. This approach evaluates a global c-means clustering, and then separates the image into overlapping regions where more c-means iterations perform this sorting both globally and locally (Dunmore et al., 2018). For our sample, we tested different numbers of clusters (from 3 to 5) at a 2% threshold and selected for subsequent analysis the image segmented utilizing a grid size of 3 pixels and 4 clusters, which allowed us to better represent intensity inhomogeneity of the input CT-scans images.

Lastly, an isosurface (.stl in Little Endian format) was generated for each segmentation protocol.

## 2.2.4 | Smoothing procedures

Several smoothing procedures were tested on one specimen by calculating distances from the reconstructed CT-generated mesh and post-processed with different smoothing procedures to the homologous high-resolution laser scanner-generated mesh. Prior to smoothing testing, CT-generated meshes were segmented with the same segmentation protocol (HMH), chosen arbitrarily to reduce possible variation induced by different segmentation protocols and to control only for smoothing differences. The smoothing procedures tested are Laplacian smoothing with or without surface preservation and Taubin smoothing (Taubin, 1995) with different number iterations (0.5; 1; 3; 5), generated in Meshlab 2020.12 (Cignoni et al., 2008). Distances from CT-generated meshes smoothed with different procedures to the laser scanner-generated mesh were calculated, prior to alignment, following the procedures detailed below for all meshes comparisons.

## 2.3 | Surfaces comparisons

The comparison among 3D surface meshes was performed in R v. 4.0.3. (R Core Team, 2020). Both CT- and laser scanner-generated meshes were uploaded and aligned by their Principal Axes with the function `pcAlign()` from the package “Morpho” (Schlager, 2017), using the CT-generated mesh as target for the alignment of the laser scanner-generated mesh, with an optimization procedure that minimizes root mean square errors (RMSE) between reference and target mesh, by testing all possible axes alignments with a rigid iterative closest point (ICP) procedure. Both iterations and subsampled points were set at 200, to compute the optimization procedure.

Topological analysis of distances among vertices of the CT- and laser scanner-generated meshes was then implemented by the function `meshDist()` from the package “Morpho” (Schlager, 2017), with the aligned laser-scanner generated mesh as the reference and the CT-generated mesh as the target. Mesh vertices distances were then represented by a polychrome scale with values ranging from 2.50 to -2.50mm. Distances up to 0.20mm were depicted in green and considered not relevant. Descriptive statistics (mean, standard deviation, and maximum displacement of surface scan mesh surface from the CT mesh) of distance among the vertices of the triangles that form meshes, were also calculated. We also provide %variation among CT- and surface scanner-generated meshes, calculated by dividing mean variation by maximum diameter at fibular midshaft multiplied by 100 (Table S1).

## 2.4 | Geometric morphometric analysis

In addition to topological analysis, we assessed geometric distances between CT- and surface scanner-generated meshes by adopting a geometric morphometric approach based on fixed landmarks and curves and surface semilandmarks, specifically designed for evaluating the accuracy of the two meshes at epiphyseal areas. Procrustes distances between landmarked surfaces generated by laser scanner

and those generated by a CT-scanner were analysed. For each specimen, the laser scanner-generated meshes were compared to CT-scan generated meshes, which were segmented by MIA-clustering and HMH protocols separately. Additionally, a second digitization of the surface scanner-generated mesh was included in this analysis, to quantify possible intra-observer error due to template repeatability and not to due to scanning device.

A template of 142 landmarks and semilandmarks (Table 1 and Figure 1), modified from a previous version published by Marchi et al. (2022), was applied to targets (both CT- and laser scanner-generated models) utilizing Viewbox 4 software (dHal Software). Curve and surface semilandmarks were allowed to slide along the curves/surfaces to minimize the thin-plate spline (TPS) bending energy between the target and the template obtaining geometrically homologous semilandmarks (Slice, 2005; Gunz & Mitteroecker, 2013). The set of semilandmarks for the proximal epiphysis ( $N = 94$ ) and for the distal epiphysis ( $N = 48$ ) were separated in order to perform two separate Generalized Procrustes Analysis (GPA) using the R package “geomorph” version 3.0.7 (Adams et al., 2018). As a result, the semilandmark coordinates were superimposed with scale, position and orientation standardized, with semilandmarks being allowed to slide with each recursive update of the Procrustes consensus (Gunz et al., 2005; Mitteroecker & Gunz, 2009; Rohlf & Slice, 1990; Slice, 2005; Sorrentino et al., 2020,b).

Procrustes coordinates for each fibular epiphysis were then analyzed by Principal Component Analysis (PCA) to explore shape variations among specimens, differentiated by acquisition methodology, including for each individual the CT-generated meshes segmented with both HMH and MIA-clustering protocols, the surface scanner-generated mesh and a second landmark digitalization replica of the latter. Procrustes ANOVA was then adopted to test shape differences among scanning devices, utilizing Procrustes distances among specimens digitized with different scanning source, with a residual randomization procedure (RRPP = T, iterations = 1'000), with the R package “geomorph” version 3.0.7 (Adams et al., 2018). Initially, we performed a Procrustes ANOVA on shape distances among all specimens and its interaction with relative scanning devices in the model formula ( $\text{shape} \sim \text{device}$ ), considering CT-generated meshes segmented with HMH and MIA-clustering protocols as a group and both laser scanner replicates as another group (i.e., CT vs. LASER). Additionally, to exclude the influence of intra-observer error on shape distances due to template repeatability and not solely to surface acquisition and segmentation method, a second Procrustes ANOVA ( $\text{shape} \sim \text{method}$ ) was computed, this time separating CT-generated meshes segmented with HMH and MIA-clustering protocols and laser scanner replicates separately (i.e., CT-HMH vs. CT-MIA vs. LASER vs. LASER REPLICAS).

Furthermore, a Post-hoc test was calculated for the latter Procrustes ANOVA, to evaluate the interaction among CT-generated meshes segmented with HMH and MIA-clustering protocols and laser scanner replicates, with pairwise function from the RRPP package (Collyer et al., 2015; Collyer & Adams, 2018). This procedure controls for Type I errors in performing multiple tests, as the RRPP method performs the exact same random placement of residuals for every test statistic calculated, allowing us to

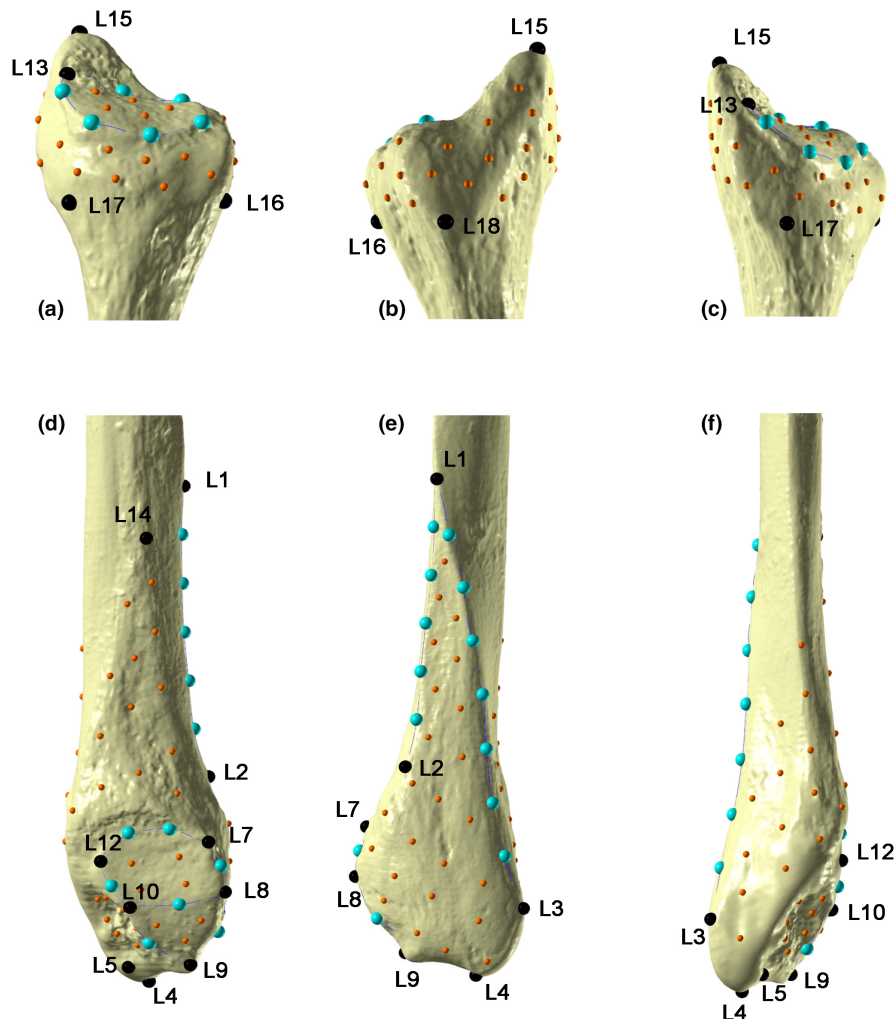
TABLE 1 Fibular landmarks and (semi)landmarks identification, definition, and number

Landmarks	Definition	
L1	Point where the <i>fibular anterolateral border</i> divides into two ridges: the proximal apex of the <i>subcutaneous triangular surface</i> (STS)	
L2	Most medial point of the medial border of the STS	
L3	Most lateral point of the lateral border of the STS	
L4	Most distal point of the <i>lateral malleolus</i> in anterior view	
L5	Most distal point of the posterior border of the <i>malleolar fossa</i>	
L7	Most anterior point on the anterior border of the <i>proximal fibular-talar articular facet</i> (PAF)	
L8	Point between the anterior border of PAF and the <i>anterior border of distal fibular-talar articular facet</i> (DAF)	
L9	Most distal point of DAF	
L10	Most proximal point on the posterior border of DAF	
L12	Most posterior point of the proximal border of PAF	
L13	Most proximal point of <i>proximal tibio-fibular articular facet</i>	
L14	Most proximal point of <i>interosseous tibio-fibular ligament</i> (ILA) insertion	
L15	Most proximal point on styloid process of fibular head in medial view	
L16	Most antero-proximal point on anterior border in medial view (above fibular neck)	
L17	Most postero-proximal point on posteromedial border in medial view (above fibular neck)	
L18	Most postero-proximal point on posterior border in lateral view (above fibular neck)	
Curves	Definition	Number of semi-landmarks
C_1->2	Medial border of the <i>subcutaneous triangular surface</i> (STS)	5
C_1->3	Posterior border of the STS	7
C_7->8	Anterior border of the <i>proximal fibular-talar articular facet</i> (PAF)	1
C_8->9	Anterior border of the <i>distal fibular-talar articular facet</i> (DAF)	1
C_9->10	Posterior border of the DAF	1
C_10->12	Posterior border of the PAF	1
C_8->10	Border between the PAF and DAF	1
C_13->13	Outline of <i>proximal tibio-fibular articular facet</i>	6
C_7->12	Proximal border of the PAF	2
Surfaces	Definition	Number of semi-landmarks
SSML_malleolar fossa	Surface of the <i>malleolar fossa</i> , attachment site of the <i>transverse tibiofibular</i> and <i>posterior talofibular ligaments</i> .	7
SSML_ILA	Attachment surface of <i>interosseous tibio-fibular ligament</i> and part of <i>interosseous membrane</i> (ILA)	13
SSML_fibular groove	Groove for tendons of <i>m. peroneus longus</i> and <i>m. tertius</i> and attachment site of <i>posterior tibiofibular ligament</i> .	13
SSML_STS	<i>Subcutaneous triangular surface</i> (STS)	24
SSML_FiTal1Ar	<i>Proximal fibular-talar articular facet</i> (PAF)	4
SSML_FiTal2Ar	<i>Distal fibular-talar articular facet</i> (DAF)	3
SSML_head	<i>Proximal tibio-fibular articular surface</i>	5
SSML_prox_ep	Surface of <i>proximal epiphysis</i>	32

consider these pairwise comparisons as separate inferences from the same test (Collyer et al., 2015). The replicate design allowed us to assess whether variation due to scanning methods could be proxy of operator error (i.e., intra-observer error), therefore denouncing similar (possibly negligible) approximation. Finally, size

comparisons among CT-generated meshes segmented with HMH and MIA-clustering protocols and the two laser scanner replicates were calculated using the centroid size, which is the square root of the summed squared distances between each semilandmark and the centroid of the semilandmark configuration (Slice, 2005).





**FIGURE 1** Landmarks (black), curve (light blue) and surface semilandmarks (orange) digitized on a left fibula (a, medial view, proximal; b, lateral view, proximal; c, postero-medial view, proximal; d, medial view, distal; e, lateral view, distal; f, posterior view, distal). See [Table 1](#) for a detailed description of the anatomical landmarks.

Centroid size distributions were evaluated visually by box-plots and differences among CT-generated meshes segmented with HMH and MIA-clustering protocols and laser scanner replicates were quantified by one-way ANOVA.

Landmarks and semilandmarks coordinates and sample list used to describe the specimens of the study are available (DOI: [10.5281/zenodo.6425379](https://doi.org/10.5281/zenodo.6425379)).

### 3 | RESULTS

#### 3.1 | Smoothing procedures

Consistency among smoothing procedures was revealed, but better results were offered when Laplacian smoothing with surface preserve was applied ([Table 2](#); [Figure S1](#)). Therefore, for further analysis we opted to follow the application of a Laplacian smoothing with surface preserve at 0.5 iteration, minimizing post-processing procedures, to all CT-generated meshes.

#### 3.2 | Topological and visual meshes comparison

Descriptive statistics for topological distances among every CT-generated meshes segmented at different grayscale threshold values and with HMH and MIA-clustering protocols, and laser scanner-generated meshes are presented in [Table 3](#). Topological differences among the CT-generated meshes with different segmentations and the laser scanner-generated meshes are presented in [Figure 2](#). Comprehensively, the mean deviation between CT-generated meshes and laser scanner-generated meshes, considering all individuals at different segmentation protocols, vary between 0.35 mm and 0.70 mm, while maximum deviation values vary between 1.23–4.16 mm ([Table 3](#)).

For each individual, some degree of variation is seen among all its CT-based meshes generated from different segmentations. Individually, the range of variation of the deviation between CT-generated meshes and surface scanner-generated meshes measured for different segmentation protocols was between 0.09–0.24 mm, and for maximum deviation values was between 0.16–2.54 mm ([Table 3](#)). While regions of variations (i.e., location on bone surface)

**TABLE 2** Descriptive statistics (mean, Standard deviation, Minimum and maximum values) of distances among vertexes of CT-generated mesh of SS163, segmented with HMH protocol, and subjected to different smoothing protocols, and its respective laser scanner-generated mesh. Smoothing protocols evaluated presence/absence of surface preserve filter and a different number of iterations (0.5, 1, 3, 5)

	Mean (mm)	Sd (mm)	Max (mm)	Min (mm)
Laplacian smoothing with SP				
0.5	0.51	0.27	1.71	0
1	0.51	0.28	1.79	0
3	0.51	0.29	1.75	0
5	0.51	0.32	1.76	0
Range	0.00	0.05	0.08	0.00
Laplacian smoothing				
0.5	0.52	0.29	1.73	0
1	0.52	0.31	1.97	0
3	0.55	0.25	1.97	0
5	0.52	0.27	2.1	0
Range	0.03	0.06	0.37	0.00
Taubin smoothing				
0.5	0.51	0.28	1.80	0
1	0.52	0.28	1.77	0
3	0.53	0.28	1.78	0
5	0.53	0.31	1.94	0
Range	0.02	0.03	0.17	0
No smoothing	0.51	0.27	1.71	0

Abbreviation: SP, surface preserve.

within each individual accounted by different segmentation protocols are quite subtle, both CT-generated meshes segmented with HMH and MIA-clustering protocols seem to better represent a larger portion of the diaphyses and epiphyseal areas, compared to other segmentation (-550, -600, -650 and -700 grayscale value intensity thresholds) (Figure 2, Figures S2–S12).

For specimen SS23 (Figure 3), taken singularly as an example due to the presence of a small cortical defect, we observe a more evident difference in regional distribution of variations accounted by segmentation protocol, located on the antero-medial portion of the fibular head, the styloid process, the proximal tibio-fibular articular facet, the distal half of the interosseous border and the medial surface, the area of attachment of the interosseous ligament and the malleolar fossa. The small area of cortical loss at proximal epiphysis exceeded the range of mesh deviations in the first two segmentations and is thus mostly colored in white (Figure 3a–c), but is captured by the last three protocols (Figure 3d–f).

Indeed, for all specimens the CT-generated mesh is optimized when the MIA-clustering segmentation protocol is adopted, followed by HMH (Table 2). Deviation between CT-generated meshes segmented with MIA-clustering protocol and laser scanner-generated

ones ranges between 0.35 and 0.56 mm of mean, while maximum deviation values are within a range of 1.27 and 2.99 mm. Overall, the average of the mean deviation among CT-generated meshes that were segmented with MIA-clustering protocol and laser scanner-generated ones is 0.42 mm, while the average of maximum deviations is 1.78 mm.

Indeed, taking as reference specimen SS163, considered representative of the variations seen in all specimens (Figure 4) for both HMH and MIA-clustering protocols there are areas with displacements from 0.2 to -0.2 mm, visible in green, in most of the posterior surface fibular diaphysis, while anterior and medial surfaces are slightly more displaced than the former, staying however within maximum 0.50 mm distance. The lowest divergence between CT and laser scanners is observed at the level of epiphyses, at the proximal tibio-fibular articular facet and the distal fibular-talar articular facet, whereas some degree of displacement is found around the fibular head, and at the superior border of the proximal fibular-talar articular facet.

### 3.3 | 3D GM analysis for meshes comparison and intra-observer error

The PCA plots of the shape coordinates for the proximal and distal epiphysis show that the CT-generated specimens segmented with both HMH and MIA-clustering protocols either overlapped or clustered in the very proximity of the homologous individual digitized with laser scanner and its replica (Figure 5; Figures S13 and S14). Results are corroborated by the Procrustes ANOVA for both distal and proximal epiphysis (Table 4) comparing shape variation due to scanning device considering CT-generated meshes (pooled meshes segmented with HMH and MIA protocols) and laser scanner (original and replicates combined together), with no significant differences found. Considering both proximal and distal epiphyses, only 1.38%–1.43% of variance ( $Rsq$ ) is explained by differences in scanning device, while 98.57%–98.62% of variance is explained by residuals, suggesting that biological variations may account for the large proportion of this variance, independently from the scanning devices.

Moreover, the PCA plots for the proximal and distal epiphysis show that the replicates (digitized on the laser scanner meshes) maintain similar variances and plot closely to their homologous meshes coming from the first digitization on laser scanner (Figure 5). Indeed, when Procrustes ANOVA is computed considering CT-generated meshes segmented with HMH and MIA-clustering protocols and laser scanner replicates separately (Table 5), again no significant difference among all configurations emerged ( $p > 0.05$  for both epiphyses). The Procrustes distance between LASER vs. LASER REPLICA is actually lower than distances among all other comparisons (Table 6), ultimately denoting low intra-observer error in the application of the template. This result likely indicates that (1) the intra-observer error is negligible, (2) and that the inter-method comparisons are not influenced by repeating the template on the same individual (intra-observer error).

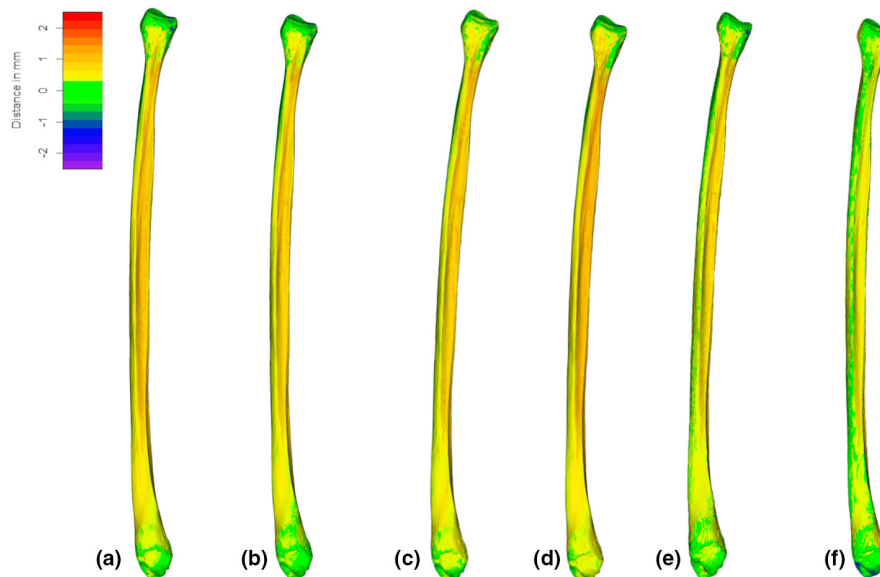
**TABLE 3** Descriptive statistics (mean, standard deviation, and maximum values) of distances calculated among CT- and laser scanner-generated meshes (vertexes) for the 13 specimens analysed. %var is calculated as (mean distance/max diameter at midshaft\*100). Maximum midshaft diameters at midshaft for each specimen are presented in Table S1. For each specimen, segmentations at four different grayscale values are considered (-550, -600, -650, and -700; for SS163: -500, -550, -600, -650). For each specimen, total individual range variation in deviations considering all segmentation protocols is indicated at the bottom of each specimen's section. The total average of deviations considering every segmentation with half maximum height (HMH) and MIA-clustering protocols for all specimen, are indicated in the last rows. Values are in mm

	Mean (mm)	Sd (mm)	Max (mm)	%		Mean (mm)	Sd (mm)	Max (mm)	%
<b>SS4</b>					<b>SS95</b>				
-550	0.59	0.25	1.39	3.44	-550	0.48	0.32	3.33	3.45
-600	0.62	0.25	1.49	3.61	-600	0.52	0.33	3.52	3.74
-650	0.68	0.26	1.51	3.96	-650	0.56	0.37	3.81	4.03
-700	0.68	0.26	1.52	3.96	-700	0.6	0.38	3.7	4.32
MIA	0.56	0.2	1.57	3.26	MIA	0.4	0.31	2.99	2.88
HMH	0.56	0.22	1.5	3.26	HMH	0.46	0.31	3.04	3.31
range	0.12	0.06	0.18	0.7	range	0.2	0.07	0.82	1.41
<b>SS14</b>					<b>SS105</b>				
-550	0.48	0.3	4.06	3.38	-550	0.63	0.27	1.83	3.57
-600	0.51	0.27	3.3	3.59	-600	0.64	0.26	1.72	3.63
-650	0.54	0.29	2.57	3.81	-650	0.68	0.28	1.8	3.85
-700	0.59	0.29	2.46	4.16	-700	0.7	0.27	1.88	3.97
MIA	0.39	0.24	1.94	2.75	MIA	0.51	0.33	2.18	2.89
HMH	0.46	0.32	4.16	3.24	HMH	0.61	0.28	1.75	3.46
range	0.2	0.08	2.22	1.41	range	0.19	0.07	0.46	1.08
<b>SS23</b>					<b>SS115</b>				
-550	0.54	0.55	3.76	3.3	-550	0.55	0.2	1.23	3.44
-600	0.5	0.42	3.34	3.06	-600	0.58	0.21	1.33	3.63
-650	0.5	0.38	3.03	3.06	-650	0.6	0.2	1.25	3.75
-700	0.5	0.36	3.6	3.06	-700	0.63	0.2	1.55	3.94
MIA	0.45	0.29	2.26	2.75	MIA	0.4	0.16	1.27	2.5
HMH	0.47	0.41	3.62	2.87	HMH	0.54	0.22	1.39	3.38
range	0.09	0.26	1.5	0.55	range	0.23	0.06	0.32	1.62
<b>SS26</b>					<b>SS163</b>				
-550	0.59	0.27	2.04	3.5	-550	0.53	0.31	1.7	3.7
-600	0.61	0.27	1.88	3.61	-600	0.54	0.31	1.54	3.77
-650	0.64	0.26	2.01	3.79	-650	0.58	0.27	1.48	4.05
-700	0.66	0.27	2.01	3.91	-700	0.66	0.36	1.73	4.61
MIA	0.56	0.26	2.03	3.32	MIA	0.42	0.21	1.48	2.94
HMH	0.58	0.26	2.04	3.44	HMH	0.51	0.27	1.71	3.56
range	0.10	0.01	0.16	0.59	range	0.24	0.10	0.23	1.68
<b>SS28</b>					<b>SS215</b>				
-550	0.46	0.28	3.03	4.14	-550	0.58	0.22	1.66	4.19
-600	0.49	0.25	2.19	4.41	-600	0.61	0.24	1.92	4.41
-650	0.47	0.25	2.24	4.23	-650	0.63	0.27	2.11	4.56
-700	0.54	0.27	1.49	4.86	-700	0.67	0.33	2.9	4.84
MIA	0.39	0.23	1.63	3.51	MIA	0.45	0.22	1.86	3.25
HMH	0.45	0.26	2.93	4.05	HMH	0.57	0.26	2	4.12
range	0.15	0.05	1.54	1.35	range	0.22	0.11	1.24	1.55



TABLE 3 (Continued)

	Mean (mm)	Sd (mm)	Max (mm)	%		Mean (mm)	Sd (mm)	Max (mm)	%
SS48					SS224				
-550	0.44	0.22	3.32	3.06	-550	0.45	0.26	2.44	2.86
-600	0.45	0.2	2.99	3.13	-600	0.49	0.28	2.86	3.12
-650	0.48	0.2	2.75	3.34	-650	0.53	0.28	2.79	3.37
-700	0.54	0.23	2.37	3.76	-700	0.58	0.28	2.71	3.69
MIA	0.36	0.17	2.12	2.5	MIA	0.41	0.24	1.99	2.61
HMH	0.41	0.23	3.39	2.85	HMH	0.46	0.24	2.16	2.92
range	0.18	0.06	1.27	1.25	range	0.17	0.04	0.87	1.08
Average MIA	0.42	0.22	1.78	2.8	SS234				
Average HMH	0.45	0.27	2.29	3.02	-550	0.37	0.27	3.81	2.55
					-600	0.38	0.24	3.29	2.62
					-650	0.42	0.25	2.79	2.89
					-700	0.46	0.27	1.47	3.17
					MIA	0.35	0.21	1.27	2.41
					HMH	0.36	0.53	2.54	2.48
					range	0.11	0.32	2.54	0.76

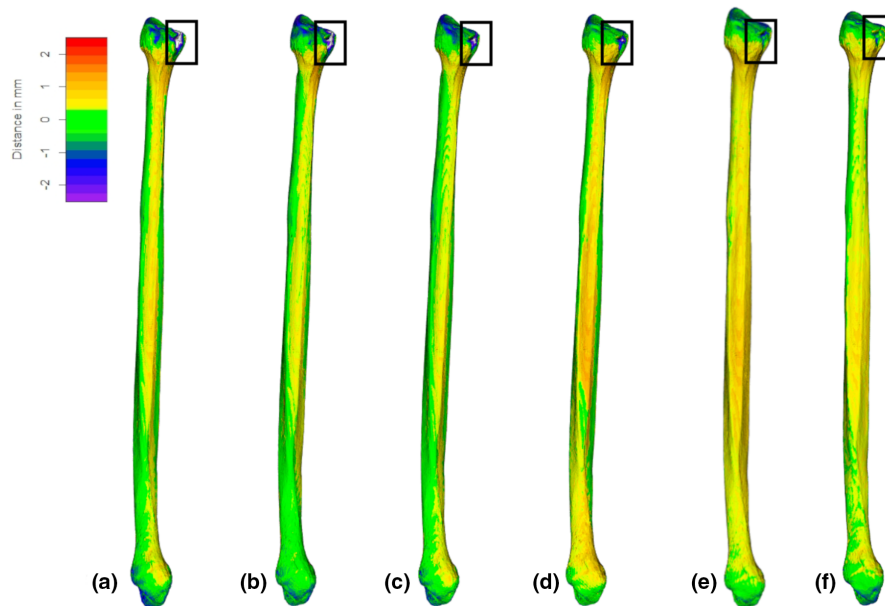


**FIGURE 2** Computed tomography (CT)-generated meshes for every segmentation protocol for a representative specimen (SS163), with relative distances (-2.5 to 2.5 mm, from violet to red; in white, areas where deviations exceed this range), from their laser-scanner generated equivalent. (a), CT segmented at -550 grayscale intensity; (b), CT segmented at -600 grayscale intensity; (c), CT segmented at -650 grayscale intensity; (d), CT segmented at -700 grayscale intensity; and e, CT segmented with half maximum height protocol (HMH) and f, MIA-clustering protocol (MIA). Other specimens are represented in Figures S2-S14 in supplementary information.

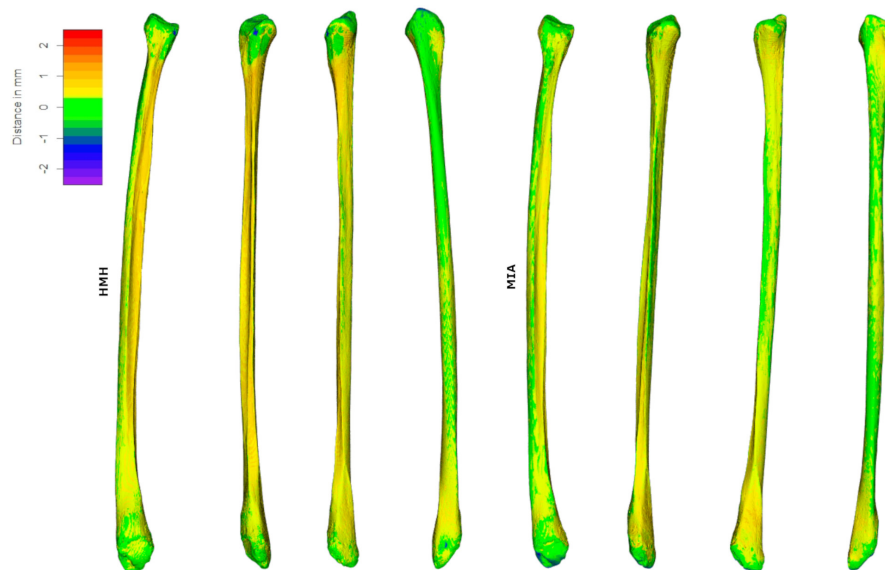
Centroid size differences among CT-generated meshes segmented with HMH and MIA-clustering protocols and laser scanner replicates are appreciable but statistically not significant for both epiphyses (Figure 6). Overall, CT-generated meshes either segmented with HMH and MIA-clustering protocols tend to slightly overestimate mesh size, but ANOVA test confirm that for both epiphyses comparisons are not significant ( $p > 0.05$ ).

#### 4 | DISCUSSION AND CONCLUSION

Both topological and geometric morphometric analysis demonstrated that CT scanning and laser scanning provide similar meshes of the fibula, fully comparable to one another, suggesting these acquisition devices could be adopted interchangeably in morphological analyses. Deviations among the meshes obtained



**FIGURE 3** Computed tomography-generated meshes on one specimen (SS23, left fibula) segmented with every segmentation protocol with relative distances (-2.5 to 2.5 mm, from violet to red; in white, deviations exceed this range), from its laser-scanner generated equivalent; (a) CT segmented at -550 grayscale intensity; (b) CT segmented at -600 grayscale intensity; (c) CT segmented at -650 grayscale intensity; (d) CT segmented at -700 grayscale intensity; and (e) CT segmented with half maximum height protocol (HMH) and (f) MIA-clustering protocol (MIA). Note the small area of cortical defect (black square).



**FIGURE 4** Computed tomography-generated meshes on one representative specimen (SS163, left fibula) segmented with half maximum height protocol (HMH) and f, MIA-clustering protocol (MIA), with relative distances (-2.5 to 2.5 mm, from violet to red; in white, deviations exceed this range), from its laser-scanner generated equivalent, seen in medial, anterior lateral, and posterior view (from left to right).

by using the two devices are comparable with those reported in previous studies (Brzobohatá et al., 2012; Fahrni et al., 2017; Lalone et al., 2015; Waltenberger et al., 2021). Our results agree with those obtained on long bones by Soodmand et al. (2018), adopting a similar acquisition protocol: they found a mean root mean square error below 0.84 mm between CT and laser-scanning meshes and reported consistent inter-observer agreement among

processing of different laboratories acquiring and elaborating the same model. Our data are also consistent with those of Stephen et al. (2021), who examined a sample of long bones, which despite some differences in the CT scanning protocols in relation to slice thickness (1 mm instead of our 0.65 mm), found similar mean deviation values between CT and laser-scanning meshes (tibia: 0.71–0.75 mm; femur: 0.69–0.80 mm), with no significant difference

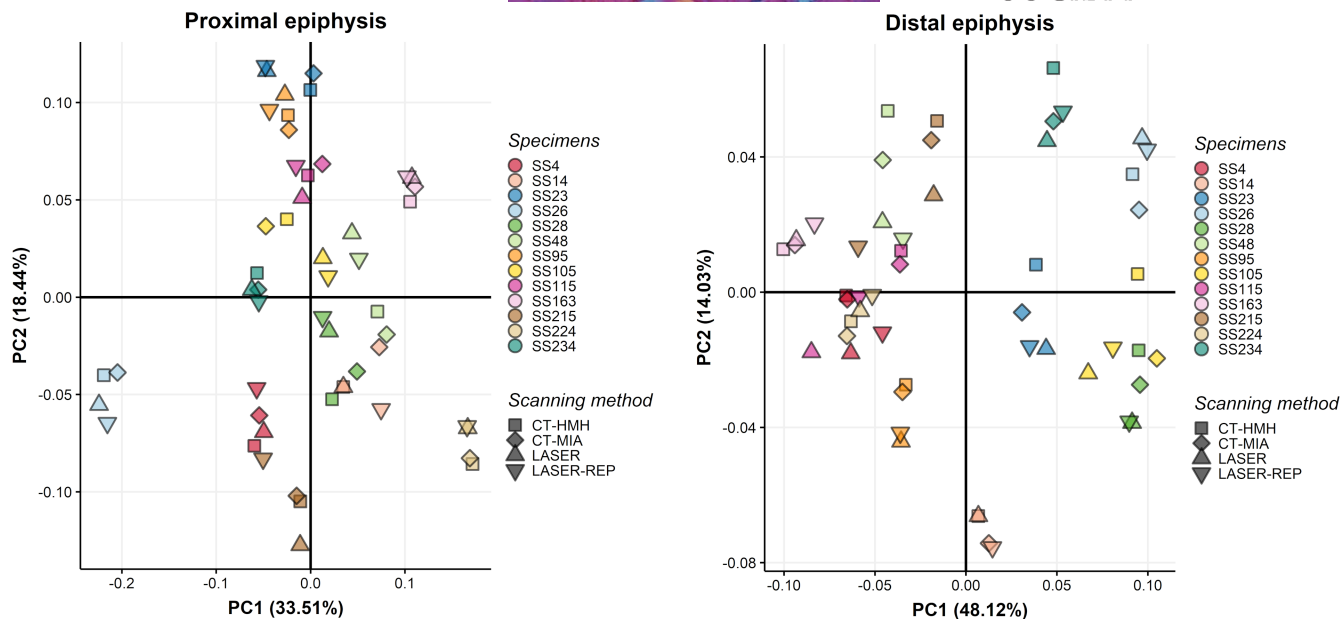


FIGURE 5 Principal component analysis plot considering landmark configurations applied to computed tomography (CT)-generated segmented with HMH (square), and with MIA-clustering (rhombus), and laser scanner-generated individuals and replicates (triangles) for the distal and the proximal epiphysis. PC1: Principal component 1; PC2: Principal component 2.

TABLE 4 Results of Procrustes ANOVA for distal and proximal epiphyses, considering computed tomography (CT)-generated meshes segmented with HMH and MIA-clustering protocols together and both laser scanner replicates together, comparing shape variation due to scanning device (i.e., CT vs. LASER)

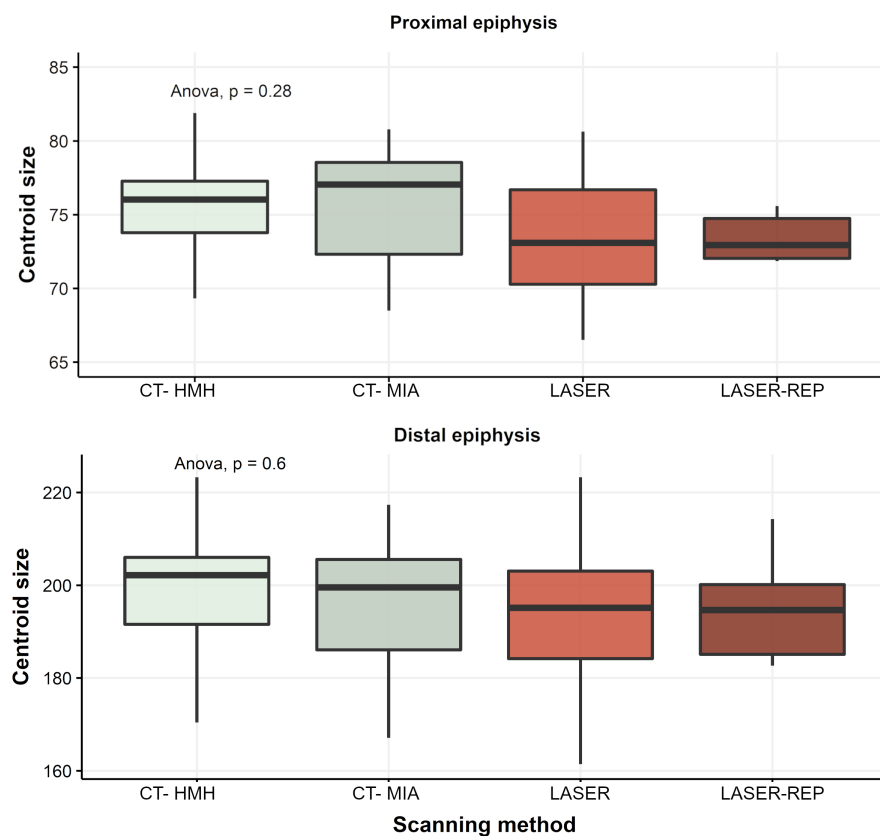
Procrustes ANOVA	Df	SS	MS	Rsq	F	p value
Distal epiphysis						
Device	1	0.00605	0.0060516	0.01378	0.6986	ns
Residuals	50	0.43313	0.0086626	0.98622		
Total	51	0.43918				
Proximal epiphysis						
Device	1	0.01786	0.017857	0.0143	0.7253	ns
Residuals	50	1.23096	0.024619	0.9857		
Total	51	1.24881				

TABLE 5 Results of Procrustes ANOVA for distal and proximal epiphyses, considering computed tomography (CT)-generated meshes segmented with HMH and MIA-clustering protocols and laser scanner replicates separately (i.e., CT-HMH vs. CT-MIA vs. LASER vs. LASER REPLICAS)

Procrustes ANOVA	Df	SS	MS	Rsq	F	p value
Distal epiphysis						
Method	3	0.00807	0.0026917	0.01839	0.2997	ns
Residuals	48	0.43111	0.0089814	0.98161		
Total	51	0.43918				
Proximal epiphysis						
Method	3	0.02117	0.0070581	0.01696	0.276	ns
Residuals	48	1.22764	0.0255758	0.98304		
Total	51	1.24881				

**TABLE 6** Post-hoc Procrustes distances pairwise comparisons among computed tomography (CT)-generated meshes segmented with HMH and MIA-clustering protocols and two sets of laser scanner replicates considered separately. D = Procrustes distance; UCL = upper confidence limit

	<i>d</i>	UCL (95%)	Z	<i>p</i> value	Global <i>p</i> value
<b>Distal epiphysis</b>					
CT-HMH: CT-MIA	0.016164350	0.05585739	-2.608923	ns	ns
CT-HMH: Laser1	0.024059392	0.05637040	-1.038789	ns	
CT-HMH: Laser2	0.024081671	0.05640215	-1.042411	ns	
CT-MIA: Laser1	0.022243496	0.05902505	-1.353978	ns	
CT-MIA: Laser2	0.022798490	0.05591536	-1.278960	ns	
Laser1: Laser2	0.007071018	0.05844946	-6.105541	ns	
<b>Proximal epiphysis</b>					
CT-HMH: CT-MIA	0.01782615	0.08976858	-4.264950	ns	ns
CT-HMH: Laser1	0.03426822	0.08563265	-1.721336	ns	
CT-HMH: Laser2	0.03956952	0.08771015	-1.360360	ns	
CT-MIA: Laser1	0.03783650	0.08875301	-1.517139	ns	
CT-MIA: Laser2	0.04281501	0.08964866	-1.073916	ns	
Laser1: Laser2	0.01387727	0.08989073	-4.294620	ns	



**FIGURE 6** Centroid size comparisons among landmark configurations applied to computed tomography (CT)-generated segmented with HMH, with MIA-clustering (in gray), and laser scanner-generated individuals and replicates (in red), displayed for both proximal and distal epiphysis and results of one-way ANOVA.

between surface accuracy and methodology. Our results on the fibula fall within this accuracy range, spanning an average of 0.35–0.56 mm when CT-obtained meshes segmented with MIA-clustering protocol are considered.

The comparison of our results to those of previous studies on comparability of surface reconstructions (Brzobohatá et al., 2012; Choi et al., 2002; DeVries et al., 2008; Fahrni et al., 2017; Ramme et al., 2009) should account for differences in protocols of acquiring

and elaborating volumetric data, and the variety of the examined anatomic components (skull, phalanges). For instance, Brzobohatá et al. (2012), Choi et al. (2002), DeVries et al. (2008), Ramme et al. (2009) utilized a slice thickness of 1.0mm, within their specific image acquisition parameters on medical CT scans. DeVries et al. (2008) revealed that the overall mean difference between the manually segmented CT models of hand phalanges and their laser surface scan was 0.20mm, with minimal effect of different smoothing procedures on surface representations, also coherent with our results (Table 1; Figure S1).

Concerning the smoothing procedures, our results highlight only minimal variation related to procedure choice (Table 2) in agreement with a previous work on smoothing procedures that have showed that smoothing algorithm affected mesh topology differently, despite retaining measurement accuracy in most reconstructions (Veneziano et al., 2018). While Veneziano et al. (2018) found that the Taubin algorithm ensured avoiding information loss on the surface of the mesh, we found a lower mean distance between the CT-generated mesh smoothed with Laplacian algorithm with surface preserve. It is important to notice, though, that in our testing the best results were obtained when the iterations were minimal.

Regarding the segmentation protocols, our results indicate that different chosen single-thresholding value influences topological differences of CT scan and laser scanner meshes, suggesting the necessity of a homogeneous segmentation protocol. Previous works (Engelbrecht et al., 2013; Fourie et al., 2012; Rathnayaka et al., 2011) found that, despite differences in accuracy of the 3D model in relation the segmentation methods, even the least accurate of the analyzed segmentation protocols (the single-threshold technique) produced comparable 3D reconstruction, minimally deviating from their high-resolution gold standard (0.18–0.24mm). This is consistent with our study, where the adoption of a single-threshold segmentation protocol provided average deviations mostly below the maximum resolution of the CT scan of 0.625mm in slice thickness, even in cases of small cortical defects, despite with a less accurate performance. On the other hand, MIA-clustering and HMH protocols offered non-relevant deviations (–0.02–0.2mm) at epiphyses and in larger portions of the diaphysis (Figure 3), compared to meshes generated with other segmentations based on single grayscale values, possibly reflecting shortcomings of single grayscale value thresholding in detecting thinner cortical feature on the bone surface. Indeed, Ito (2019) found that single grayscale value thresholding, including HMH protocol, might induce excessive erosion in 3D reconstructions, especially in low-density areas, suggesting caution when combining segmentation protocols within a dataset for detailed GM analysis (e.g., asymmetry), since inter-method error may be greater than shape variations explained by asymmetry. In fact, single thresholding segmentations may induce errors due to intensity inhomogeneity of CT data, and therefore may not capture fine bone surface features (Scherf & Tilgner, 2009). Adopting a segmentation protocol such as MIA-clustering, based on a machine-learning approach that adopts a global and local fuzzy c-means clustering, here implemented on medical CT images, yielded better results

than single-thresholding segmentations, providing a practical, inexpensive solution to possible low resolution and grayscale intensity shortcomings (Dunmore et al., 2018). This segmentation approach could also be useful in clinical applications, for the 3D evaluation of the tibio-fibular syndesmotoc space, as it is a timesaving technique (Ebinger et al., 2013). Such variations may also have implications for GM analysis, as landmarks situated on sharp peaks may be sensitive to volume-averaging effect, causing blurring in bone regions with low density or in adjacent bones with adjacent soft tissue (Wang et al., 2006).

In our results, we found that our landmark and semilandmark configurations mostly overlap in homologous individuals scanned with CT and laser scanner (Figure 5) and only 1.38%–1.43% of total variance in the dataset is explained by difference in scanning device, with no statistically significant differences in morphology between the two methods. The results of our PCA analyses (Figure 5) are comparable to those of Marcy et al. (2018) and Waltenberger et al. (2021), even when both sets of scanning digitization are added, with each individual, either CT- or laser scanned, clustering together. However, while Waltenberger et al. (2021) found no significant difference in their Procrustes ANOVA in accordance with our own results, Marcy et al. (2018), who compared laser scanner and micro-CT derived mesh reconstructions of mice skulls, found that the percentage of variance explained by asymmetry (directional and fluctuating) was lower than the percentage of variance explained by scanning device. According to the authors, this would suggest that analyses of asymmetry with a combination of scanning methods scans may be subject to systematic error. In both ours and Waltenberger et al. (2021) cases, it is possible that the relatively big size of the fibula in contrast to the diminutive size of mice skulls may have prevented this effect, therefore supporting the comparability of scanning methodology through geometric morphometric methods and the combination of different datasets in GM analysis of human bone elements. Moreover, as in our sample 98.57%–98.62% of variance is explained by factors other than scanning methodology, it is possible to infer that such shape variations may account for biological diversity within our sample.

Centroid size comparisons for both epiphyses revealed that, while CT-generated meshes are larger than laser scanner-generated meshes, their size is comparable. The larger centroid size of CT-generated meshes also confirms the observations of Table 3, with mean distances between CT-generated meshes and laser scanner-generated meshes indicating all positive values, suggesting that the latter meshes possess indeed smaller surface and areas than the formers.

In our analysis, inter-method comparisons showed good reproducibility of landmark digitization, with lower Procrustes distance between each set of CT and laser scanner replicates than inter-methods, suggesting low intra-observer error. This is consistent with Fruciano et al. (2017), who found that when combining landmark data from multiple devices and digitized by multiple operators and test for the presence of bias, a larger amount of variance was explained by the operator compared to the device. The authors

however also found that, similar to Robinson and Terhune (2017), inter-method and inter-observer errors may be greater than small intraspecific or closely related species shape variations. Indeed, the authors advocate for testing accuracy trials of different methods and comparability of observers prior to combination of different data source.

In conclusion, we validated the comparability and combination of 3D reconstruction of human fibulae, generated by dual-energy CT scan and two sets of 3D laser scanner replicates, confirming substantial topological similarity of the meshes reconstructed with the two methods and corroborating their interchangeability in GM studies, as no significant inter-method and intra-observer difference is found in shape analysis, with potential applications in clinical context (Souleiman et al., 2021). A possible limitation of the present study is the use of laser scanner-generated meshes as the gold standard comparison from which distances of CT-generated meshes are calculated, assuming their closeness to the real bone surface as being more accurate. In our analyses, in fact, we did not evaluate this proximity to the real specimens, but on the basis of previous evaluations utilizing the same laser scanner employed here (Adams et al., 2015), and given that the 3D accuracy stated by the manufacturer is to 50 µm, we could indeed infer that the laser scanner surfaces are the most accurate approximations. We suggest, in accordance with previous studies (Marcy et al., 2018; Robinson & Terhune, 2017; Waltenberger et al., 2021), that prior to formal analysis, coherence of both CT and laser scanning protocols, including assessment of scanning settings and resolutions, segmentations and reconstruction post-processing, should be thoroughly evaluated, in case of merging different data sources, possibly with dedicated trials.

#### AUTHORS' CONTRIBUTIONS

AP conceptualization, investigation, formal analysis, validation, writing-original draft, methodology, writing-review, and editing; RS validation, writing-original draft, methodology, writing-review, and editing; VN data acquisition, formal analysis, validation; SD conceptualization, supervision of the research, methodology, data acquisition, validation, software; DM, SB, MGB conceptualization, supervision of the research, methodology, project administration, software, writing-review and editing.

#### ACKNOWLEDGMENTS

The authors would like to thank the Radiology Unit at Istituto Ortopedico Rizzoli for their kind support and cooperation during data acquisition, and for allowing the use of their facility for this research. We thank Dr. Carla Figus for her advice on segmentations. We are also grateful to Madison Smith for language proof-reading. Lastly, we would also like to express our gratitude to the reviewers of the manuscript. Open Access Funding provided by Università degli Studi di Bologna within the CRUI-CARE Agreement.

#### CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

#### DATA AVAILABILITY STATEMENT

All data acquired and analyzed in this work is available upon request addressed to the corresponding author. Landmarks coordinates and sample lists for geometric morphometric analysis are available (DOI: [10.5281/zenodo.6425379](https://doi.org/10.5281/zenodo.6425379)).

#### ORCID

Annalisa Pietrobelli  <https://orcid.org/0000-0002-2558-1118>

Rita Sorrentino  <https://orcid.org/0000-0002-6529-7250>

#### REFERENCES

- Adams, D. C., Collyer, M. L. & Kaliontzopoulou, A. (2018) *Geomorph: Software for geometric morphometric analyses. R package version 3.0.7* (Version 3.2.0).
- Adams, J.W., Olah, A., McCurry, M.R. & Potze, S. (2015) Surface model and tomographic archive of fossil primate and other mammal holotype and paratype specimens of the Ditsong National Museum of Natural History, Pretoria, South Africa. *PLoS One*, 10(10), 1–14. <https://doi.org/10.1371/journal.pone.0139800>
- Brzobohatá, H., Prokop, J., Horák, M., Jancárek, A. & Velemínská, J. (2012) Accuracy and benefits of 3D bone surface modelling: a comparison of two methods of surface data acquisition reconstructed by laser scanning and computed tomography outputs. *Collegium Antropologicum*, 36(3), 801–806.
- Choi, J.Y., Choi, J.H., Kim, N.K., Kim, Y., Lee, J.K., Kim, M.K. et al. (2002) Analysis of errors in medical rapid prototyping models. *International Journal of Oral and Maxillofacial Surgery*, 31(1), 23–32. <https://doi.org/10.1054/ijom.2000.0135>
- Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F. & Ranzuglia, G. (2008) MeshLab: an open-source mesh processing tool. In: *Eurographics Italian Chapter Conference*. Abstract retrieved from <https://diglib.org/handle/10.2312/LocalChapterEvents.ItalChap.ItalianChapConf2008.129-136>
- Coleman, M.N. & Colbert, M.W. (2007) Technical note: CT thresholding protocols for taking measurements on three - dimensional models. *American Journal of Physical Anthropology*, 133(1), 723–725.
- Collyer, M.L. & Adams, D.C. (2018) RRPP: an r package for fitting linear models to high-dimensional data using residual randomization. *Methods in Ecology and Evolution*, 9, 1772–1779. <https://doi.org/10.1111/2041-210X.13029>
- Collyer, M.L., Sekora, D.J. & Adams, D.C. (2015) A method for analysis of phenotypic change for phenotypes described by high-dimensional data. *Heredity*, 115(4), 357–365. <https://doi.org/10.1038/hdy.2014.75>
- Damecourt, A., Nieto, N., Galmiche, S., Garrel, R. & de Boutray, M. (2020) In-house 3D treatment planning for mandibular reconstruction by free fibula flap in cancer: our technique. *European Annals of Otorhinolaryngology, Head and Neck Diseases*, 137(6), 501–505. <https://doi.org/10.1016/j.anorl.2020.02.002>
- DeVries, N.A., Gassman, E.E., Kallemeyn, N.A., Shivanna, K.H., Magnotta, V.A. & Grosland, N.M. (2008) Validation of phalanx bone three-dimensional surface segmentation from computed tomography images using laser scanning. *Skeletal Radiology*, 37(1), 35–42. <https://doi.org/10.1007/s00256-007-0386-3>
- Dunmore, C.J., Wollny, G. & Skinner, M.M. (2018) MIA-clustering: a novel method for segmentation of paleontological material. *PeerJ*, 6, e4374.
- Ebinger, T., Goetz, J., Dolan, L. & Phisitkul, P. (2013) 3D model analysis of existing CT syndesmosis measurements. *The Iowa Orthopaedic Journal*, 33, 40–46.
- Engelbrecht, W.P., Fourie, Z., Damstra, J., Gerrits, P.O. & Ren, Y. (2013) The influence of the segmentation process on 3D measurements from cone beam computed tomography-derived surface



- models. *Clinical Oral Investigations*, 17(8), 1919–1927. <https://doi.org/10.1007/s00784-012-0881-3>
- Fahrni, S., Campana, L., Dominguez, A., Uldin, T., Dedouit, F., Delémont, O. & Grabherr, S. (2017) CT-scan vs. 3D surface scanning of a skull: first considerations regarding reproducibility issues. *Forensic Sciences Research*, 2(2), 93–99. <https://doi.org/10.1080/20961790.2017.1334353>
- Fourie, Z., Damstra, J., Gerrits, P.O. & Ren, Y. (2011) Evaluation of anthropometric accuracy and reliability using different three-dimensional scanning systems. *Forensic Science International*, 207(1–3), 127–134. <https://doi.org/10.1016/j.forsciint.2010.09.018>
- Fourie, Z., Damstra, J., Schepers, R.H., Gerrits, P.O. & Ren, Y. (2012) Segmentation process significantly influences the accuracy of 3D surface models derived from cone beam computed tomography. *European Journal of Radiology*, 81(4), e524–e530. <https://doi.org/10.1016/j.ejrad.2011.06.001>
- Frelat, M.A., Shaw, C.N., Sukhdeo, S., Hublin, J.J., Benazzi, S. & Ryan, T.M. (2017) Evolution of the hominin knee and ankle. *Journal of Human Evolution*, 108, 147–160. <https://doi.org/10.1016/j.jhevol.2017.03.006>
- Fruciano, C., Celik, M.A., Butler, K., Dooley, T., Weisbecker, V. & Phillips, M.J. (2017) Sharing is caring? Measurement error and the issues arising from combining 3D morphometric datasets. *Ecology and Evolution*, 7, 7034–7046. <https://doi.org/10.1002/ece3.3256>
- Gelaude, F., Sloten, J.V. & Lauwers, B. (2008) Accuracy assessment of CT-based outer surface femur meshes. *Computer Aided Surgery*, 13(4), 188–199. <https://doi.org/10.3109/10929080802195783>
- Gunz, P. & Mitteroecker, P. (2013) Semilandmarks: a method for quantifying curves and surfaces. *Hystrix*, 24(1), 103–109. <https://doi.org/10.4404/hystrix-24.1-6292>
- Gunz, P., Mitteroecker, P. & Bookstein, F.L. (2005) Semilandmarks in three dimensions. In: Slice, D.E. (Ed.) *Modern morphometrics in physical anthropology*. New York, Springer: Science & Business Media, 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.
- Ito, T. (2019) Effects of different segmentation methods on geometric morphometric data collection from primate skulls. *Methods in Ecology and Evolution*, 10, 1972–1984. <https://doi.org/10.1111/2041-210X.13274>
- Lalone, E.A., Willing, R.T., Shannon, H.L., King, G.J. & Johnson, J.A. (2015) Accuracy assessment of 3D bone reconstructions using CT: an intro comparison. *Medical Engineering & Physics*, 37(8), 729–738. <https://doi.org/10.1016/j.medengphy.2015.04.010>
- Mantini, S. & Ripani, M. (2009) Modern morphometry: new perspectives in physical anthropology. *New Biotechnology*, 25(5), 325–330. <https://doi.org/10.1016/j.nbt.2009.03.009>
- Marchi, D. (2005) The cross-sectional geometry of the hand and foot bones of the Hominoidea and its relationship to locomotor behavior. *Journal of Human Evolution*, 49, 743–761.
- Marchi, D. (2007) Relative strength of the tibia and fibula and locomotor behavior in hominoids. *Journal of Human Evolution*, 53, 647–655.
- Marchi, D. (2015a) Using the morphology of the hominoid distal fibula to interpret arboreality in Australopithecus afarensis. *Journal of Human Evolution*, 85, 136–148.
- Marchi, D. (2015b) Variation in tibia and fibula diaphyseal strength and its relationship with arboreal and terrestrial locomotion: extending the investigation to non-hominoid primates. *Journal of Anthropological Sciences*, 93, 1–4.
- Marchi, D., Harper, C.M., Chirchir, H. & Ruff, C.B. (2019) Relative fibular strength and locomotor behavior in KNM-WT 15000 and OH 35. *Journal of Human Evolution*, 131, 48–60. <https://doi.org/10.1016/j.jhevol.2019.02.005>
- Marchi, D., Rimoldi, A., García-Martínez, D. & Bastir, M. (2022) Morphological correlates of distal fibular morphology with locomotion in great apes, humans, and Australopithecus afarensis. *American Journal of Biological Anthropology*, 178(2), 286–300. <https://doi.org/10.1002/ajpa.24507>
- Marcy, A.E., Fruciano, C., Phillips, M.J., Mardon, K. & Weisbecker, V. (2018) Low resolution scans can provide a sufficiently accurate, cost- and time-effective alternative to high resolution scans for 3D shape analyses. *PeerJ*, 6, e5032. <https://doi.org/10.7717/peerj.5032>
- Mitteroecker, P. & Gunz, P. (2009) Advances in geometric morphometrics. *Evolutionary Biology*, 36, 235–247. <https://doi.org/10.1007/s11692-009-9055-x>
- Mounier, A. & Lahr, M.M. (2019) Deciphering African late middle Pleistocene hominin diversity and the origin of our species. *Nature Communications*, 10(1), 1–13.
- Ni, Y., Zhang, X., Meng, Z., Li, Z., Li, S., Xu, Z.F. et al. (2021) Digital navigation and 3D model technology in mandibular reconstruction with fibular free flap: a comparative study. *Journal of Stomatology, Oral and Maxillofacial Surgery*, 122(4), e59–e64. <https://doi.org/10.1016/j.jormas.2020.11.002>
- Oka, K., Murase, T., Moritomo, H., Goto, A., Sugamoto, K. & Yoshikawa, H. (2009) Accuracy analysis of three-dimensional bone surface models of the forearm constructed from multidetector computed tomography data. *International Journal of Medical Robotics and Computer Assisted Surgery*, 5, 452–457. <https://doi.org/10.1002/rcs.277>
- Pinsky, H.M., Dyda, S., Pinsky, R.W., Misch, K.A. & Sarment, D.P. (2006) Accuracy of three-dimensional measurements using cone-beam CT. *Dento Maxillo Facial Radiology*, 35(6), 410–416. <https://doi.org/10.1259/dmfr/20987648>
- Profico, A., Bellucci, L., Buzi, C., Di Vincenzo, F., Micarelli, I., Strani, F. et al. (2019) Virtual anthropology and its application in cultural heritage studies. *Studies in Conservation*, 64(6), 323–336. <https://doi.org/10.1080/00393630.2018.1507705>
- Profico, A., Veneziano, A., Lanteri, A., Piras, P., Sansalone, G. & Manzi, G. (2016) Tuning geometric morphometrics: an r tool to reduce information loss caused by surface smoothing. *Methods in Ecology and Evolution*, 7, 1195–1200. <https://doi.org/10.1111/2041-210X.12576>
- R Core Team. (2020) *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ramme, A.J., DeVries, N., Kallemyr, N.A., Magnotta, V.A. & Grosland, N.M. (2009) Semi-automated phalanx bone segmentation using the expectation maximization algorithm. *Journal of Digital Imaging*, 22(5), 483–491. <https://doi.org/10.1007/s10278-008-9151-y>
- Rathnayaka, K., Sahama, T., Schuetz, M.A. & Schmutz, B. (2011) Effects of CT image segmentation methods on the accuracy of long bone 3D reconstructions. *Medical Engineering & Physics*, 33(2), 226–233. <https://doi.org/10.1016/j.medengphy.2010.10.002>
- Ren, W., Gao, L., Li, S., Chen, C., Li, F., Wang, Q. et al. (2018) Virtual planning and 3D printing modeling for mandibular reconstruction with fibula free flap. *Medicina Oral, Patología Oral y Cirugía Bucal*, 23(3), e359–e366. <https://doi.org/10.4317/medoral.22295>
- Robinson, C. & Terhune, C.E. (2017) Error in geometric morphometric data collection: combining data from multiple sources. *American Journal of Physical Anthropology*, 164(1), 62–75. <https://doi.org/10.1002/ajpa.23257>
- Rohlf, F.J. & Slice, D.E. (1990) Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology*, 39(1), 40–59. <https://doi.org/10.2307/2992207>
- Scherf, H. & Tilgner, R. (2009) A new high-resolution computed tomography (CT) segmentation method for trabecular bone architectural analysis. *American Journal of Physical Anthropology*, 140, 39–51. <https://doi.org/10.1002/ajpa.21033>

- Schlager, S. (2017) Morpho and Rvcg - shape analysis in R. In: Zheng, G., Li, S. & Szekely, G. (Eds.) *Statistical shape and deformation analysis*. London: Academic Press, pp. 217-256.
- Slice, D.E. (Ed.) (2005) *Modern morphometrics in physical anthropology*. New York, NY: Springer. <https://doi.org/10.1007/0-387-27614-9>
- Soodmand, E., Kluess, D., Varady, P.A., Cichon, R., Schwarze, M., Gehweiler, D. et al. (2018) Interlaboratory comparison of femur surface reconstruction from CT data compared to reference optical 3D scan. *Biomedical Engineering Online*, 17(1), 29. <https://doi.org/10.1186/s12938-018-0461-0>
- Sorrentino, R., Carlson, K., Bortolini, E., Minghetti, C., Feletti, F., Fiorenza, L. et al. (2020) Morphometric analysis of the hominin talus: evolutionary and functional implications. *Journal of Human Evolution*, 142, 102747.
- Sorrentino, R., Stephens, N.B., Carlson, K.J., Figus, C., Fiorenza, L., Frost, S. et al. (2020) The influence of mobility strategy on the modern human talus. *American Journal of Physical Anthropology*, 171(3), 456-469.
- Sorrentino, R., Stephens, N.B., Marchi, D., DeMars, L.J.D., Figus, C., Bortolini, E. et al. (2021) Unique foot posture in Neanderthals reflects their body mass and high mechanical stress. *Journal of Human Evolution*, 161, 103093.
- Souleiman, F., Heilemann, M., Hennings, R., Hennings, M., Klengel, A., Hepp, P. et al. (2021) A standardized approach for exact CT-based three-dimensional position analysis in the distal tibiofibular joint. *BMC Medical Imaging*, 21(1), 41. <https://doi.org/10.1186/s12880-021-00570-y>
- Spoor, C.F., Zonneveld, F.W. & Macho, G.A. (1993) Linear measurements of cortical bone and dental enamel by computed tomography: applications and problems. *American Journal of Physical Anthropology*, 91, 469-484. <https://doi.org/10.1002/ajpa.1330910405>
- Stephen, J.M., Calder, J.D., Williams, A. & El Daou, H. (2021) Comparative accuracy of lower limb bone geometry determined using MRI, CT, and direct bone 3D models. *Journal of Orthopaedic Research*, 39(9), 1870-1876. <https://doi.org/10.1002/jor.24923>
- Taubin, G. (1995) Curve and surface smoothing without shrinkage. In: *Proceedings of IEEE International Conference on Computer Vision*, 21-23 June 1995, Cambridge, MA, pp. 852-857.
- Teeter, M.G., Brophy, P., Naudie, D.D. & Holdsworth, D.W. (2012) Comparison of micro-computed tomography and laser scanning for reverse engineering orthopaedic component geometries. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 226(3), 263-267. <https://doi.org/10.1177/09544119111434674>
- Toro-Ibacache, M.V. (2013) *A finite element study of the human cranium: The impact of morphological variation on biting performance* (PhD Dissertation). York: University of Hull and University of York.
- Uldin, T. (2017) Virtual anthropology - a brief review of the literature and history of computed tomography. *Forensic Sciences Research*, 2(4), 165-173. <https://doi.org/10.1080/20961790.2017.1369621>
- Veneziano, A., Landi, F. & Profico, A. (2018) Surface smoothing, decimation, and their effects on 3D biological specimens. *American Journal of Physical Anthropology*, 166(2), 473-480. <https://doi.org/10.1002/ajpa.23431>
- Waltenberger, L., Rebay-Salisbury, K. & Mitteroecker, P. (2021) Three-dimensional surface scanning methods in osteology: a topographical and geometric morphometric comparison. *American Journal of Physical Anthropology*, 174, 846-858. <https://doi.org/10.1002/ajpa.24204>
- Wang, L.I., Greenspan, M. & Ellis, R. (2006) Validation of bone segmentation and improved 3-D registration using contour coherency in CT data. *IEEE Transactions on Medical Imaging*, 25(3), 324-334. <https://doi.org/10.1109/TMI.2005.863834>
- Weber, G.W. (2014) Another link between archaeology and anthropology: virtual anthropology. *Digital Applications in Archaeology and Cultural Heritage*, 1, 3-11. <https://doi.org/10.1016/j.daach.2013.04.001>
- Weber, G.W. (2015) Virtual anthropology. *American Journal of Physical Anthropology*, 156, 22-42. <https://doi.org/10.1002/ajpa.22658>
- Weber, G.W. & Bookstein, F.L. (2011) *Virtual anthropology: a guide to a new interdisciplinary field*. Wien: Springer-Verlag.
- Weber, G.W., Schäfer, K., Prossinger, H., Gunz, P., Mitteroecker, P. & Seidler, H. (2001) Virtual anthropology: the digital evolution in anthropological sciences. *Journal of Physiological Anthropology and Applied Human Science*, 20, 69-80.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Pietrobelli A, Sorrentino R, Notariale V, et al. Comparability of skeletal fibulae surfaces generated by different source scanning (dual-energy CT scan vs. high resolution laser scanning) and 3D geometric morphometric validation. *J. Anat.* (2022); 241: 667-682. Available from: <https://doi.org/10.1111/joa.13714>