

# Kent Academic Repository

## Full text document (pdf)

### Citation for published version

Bachman, Sebastian, Dunmore, Christopher J., Skinner, Matthew M., Pahr, Dieter H. and Synek, Alexander (2022) A computational framework for canonical holistic morphometric analysis of trabecular bone. Scientific Reports . ISSN 2045-2322.

### DOI

<https://doi.org/10.1038/s41598-022-09063-6>

### Link to record in KAR

<https://kar.kent.ac.uk/93840/>

### Document Version

Publisher pdf

#### Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

#### Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

#### Enquiries

For any further enquiries regarding the licence status of this document, please contact:

[researchsupport@kent.ac.uk](mailto:researchsupport@kent.ac.uk)

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>



OPEN

## A computational framework for canonical holistic morphometric analysis of trabecular bone

Sebastian Bachmann<sup>1✉</sup>, Christopher J. Dunmore<sup>2</sup>, Matthew M. Skinner<sup>2,3</sup>, Dieter H. Pahr<sup>1,4,5</sup> & Alexander Synek<sup>1,5</sup>

Bone is a remarkable, living tissue that functionally adapts to external loading. Therefore, bone shape and internal structure carry information relevant to many disciplines, including medicine, forensic science, and anthropology. However, morphometric comparisons of homologous regions across different individuals or groups are still challenging. In this study, two methods were combined to quantify such differences: (1) Holistic morphometric analysis (HMA) was used to quantify morphometric values in each bone, (2) which could then be mapped to a volumetric mesh of a canonical bone created by a statistical free-form deformation model (SDM). Required parameters for this canonical holistic morphometric analysis (cHMA) method were identified and the robustness of the method was evaluated. The robustness studies showed that the SDM converged after one to two iterations, had only a marginal bias towards the chosen starting image, and could handle large shape differences seen in bones of different species. Case studies were performed on metacarpal bones and proximal femora of different primate species to confirm prior study results. The differences between species could be visualised and statistically analysed in both case studies. cHMA provides a framework for performing quantitative comparisons of different morphometric quantities across individuals or groups. These comparisons facilitate investigation of the relationship between spatial morphometric variations and function or pathology, or both.

Living bone can functionally adapt<sup>1</sup> to loads it experiences and thus reflects, to some extent, the behaviour of individuals over their lifetime. This functional adaptation, caused by modelling and remodelling of bone<sup>2</sup>, has been experimentally demonstrated in external shape change<sup>1,3</sup>, the thickening of cortical bone<sup>1,4</sup>, enthesal shape change on the cortical bone surface<sup>5,6</sup>, and change in the architecture of trabecular bone<sup>7,8</sup>, or a combination of these phenomena. Unlike cortical bone, trabecular architecture is usually completely internal, concentrated in the epiphyses of long bones, and far more porous at the mesoscale<sup>9</sup>. As a result of these and many other differences, trabecular bone reacts to experienced loads differently than cortical bone, and thus records different information about these loads. Inference of function from trabecular bone architecture is complicated by genetic, ontogenetic<sup>10</sup>, and systemic<sup>11</sup> factors that also influence bone form. However, numerous studies have found trabecular architectures that reflect assumed habitually and significantly loaded joint postures<sup>12–24</sup>, regardless of whether this load is gravitational, muscular, or a combination of both<sup>25,26</sup>. Thus, analysis of trabecular architecture provides additional information that can be combined with studies of internal and external bone morphology, to provide a more holistic understanding of how bone reflects behaviour. We can use the bone morphology of recent species, with observable behaviours, to infer behaviour patterns of extinct species<sup>27</sup>. Therefore, methods to evaluate the trabecular morphologies of different species and compare them, offer an important tool to help answer evolutionary biological or anthropological questions.

Many different protocols exist to quantify the inner, trabecular structure of bone. These range from quantitative computed tomography (QCT) with a coarse resolution in the range of half a millimetre to high resolution techniques such as micro-computed tomography ( $\mu$ CT) with resolutions in the micrometre range. Using QCT it is possible to get estimates of bone-mineral density or bone volume fraction (BTV), however other morphometric quantities can only be assessed at a high resolution, which is able to depict the trabecular structure in

<sup>1</sup>Institute of Lightweight Design and Structural Biomechanics, TU Wien, Vienna, Austria. <sup>2</sup>School of Anthropology and Conservation, Skeletal Biology Research Centre, University of Kent, Canterbury, UK. <sup>3</sup>Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. <sup>4</sup>Department of Anatomy and Biomechanics, Division Biomechanics, Karl Landsteiner University of Health Sciences, Krems, Austria. <sup>5</sup>These authors contributed equally: Dieter H. Pahr and Alexander Synek. ✉email: bachmann@ilsb.tuwien.ac.at

detail.  $\mu$ CT allows for the imaging of the complete trabecular structure, and a variety of CT-based morphometric quantities exist to describe the bony network<sup>28</sup>. Two parameters, namely BTVV and local anisotropy, measured via a fabric tensor, are of special interest. BTVV alone accounts for approximately 87% and both values together for 97% of the trabecular stiffness, which is a proxy for the mechanical properties of the bone<sup>29</sup>.

Morphometric quantities can be measured for the whole trabecular volume, which yields a single value which can be compared between samples. However, the trabecular structure can be very heterogeneous—for example at the proximal femur—and thus it is not possible to gather precise differences using summary statistics on this multi-modal distribution. Therefore, a common approach is the use of regions of interest (ROI), where a subsection of the trabecular volume is selected for the investigation. One or multiple, spherical, cubic or arbitrarily shaped ROIs can be used. However, the placement of the ROI inside the bone at a homologous, i.e., anatomically equivalent, location is crucial when comparing different individuals and can be challenging in species with disparate morphology<sup>30</sup>. To ameliorate such problems, multiple ROIs<sup>31,32</sup> or sectors<sup>33–36</sup>, or the combination of both<sup>37</sup> can be used.

Another method that avoids many of these issues, is holistic morphometric analysis (HMA)<sup>12,38</sup>. HMA can map the morphometric quantities continuously over the whole trabecular volume onto volume meshes, which can then be visualised and compared qualitatively between different samples without ROI selection. It was successfully applied to a variety of bones, including carpals, metacarpals and phalanges<sup>12–16,36</sup>, proximal femora<sup>17–19</sup>, distal femora<sup>20</sup>, distal tibiae<sup>17</sup>, proximal humeri<sup>17</sup>, distal radii<sup>21</sup>, and first metatarsals<sup>22</sup>. However, HMA is not yet able to compare the site-specific morphometry, below the sector level, for different individuals quantitatively in homologous regions<sup>23,24,27</sup>.

The HMA method has also been successfully applied in combination with other methods. For example a sector-based analysis<sup>23,36</sup> can be employed, but, as with other sector methods, requires a-priori geometric division of the trabecular structure. Trabecular mapping<sup>39</sup> offers another way, where morphometric quantities are measured below the cortical bone and mapped onto the periosteal surface. A similar approach was applied to map HMA morphometric quantities to the endosteal surface of metacarpals<sup>15,16</sup>. However, both of these approaches are necessarily limited to analyse the edge of the trabecular network. Volumetric sliding semi-landmarks<sup>40</sup> and coherent point drift<sup>41</sup> allow for volumetric quantitative comparisons between samples using HMA<sup>24</sup>.

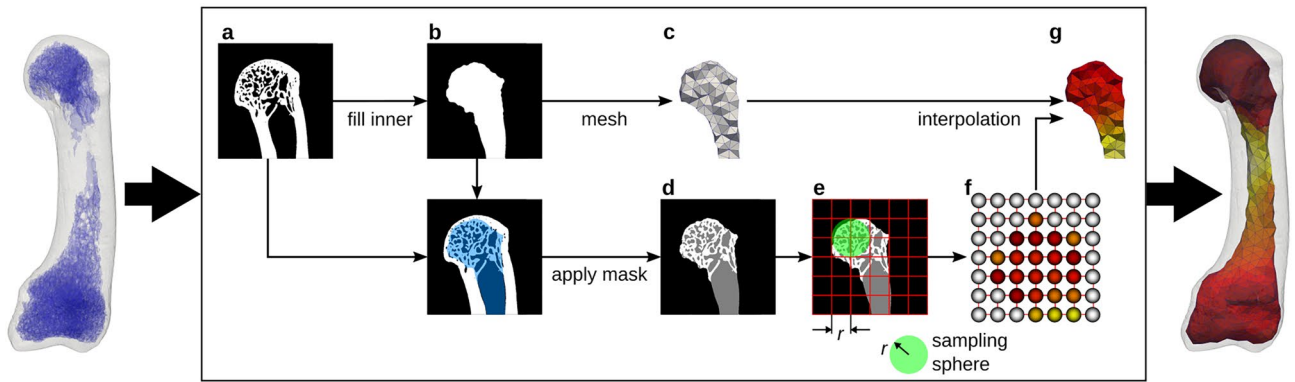
The principle in these methods is to find a canonical representation, either a surface or volume, on which all samples can be mapped, measured and finally compared. A volumetric canonical bone can also be created using statistical deformation models (SDM)<sup>42</sup>. The individual bones can then be registered onto the canonical bone and isotopological meshes can be created using mesh-morphing<sup>43</sup>. The advantage of the SDM approach is that it is landmark-free, and thus requires no extra annotation of the data. However, currently SDM methods have only been applied to QCT data, whereas  $\mu$ CT data is required in order to evaluate CT-based morphometric quantities. Furthermore, SDM approaches have, so far, only been used to create canonical models in a single species, which often have less inter-subject variation in bone shape.

The aim of this work was to develop a workflow for quantitative volumetric comparison of bone morphometry at homologous sites directly using  $\mu$ CT images and representing bones of different species, without the need for landmarking nor definition of ROIs or subregions. The workflow consists of the canonical bone creation using SDM, registration of individual bones, mesh morphing and finally the application of HMA and statistical analysis. We refer to this workflow as canonical holistic morphometric analysis (cHMA). The objectives of this study were: (1) to identify cHMA method parameters suitable for the analysis of  $\mu$ CT images and comparison of different species, (2) to investigate the robustness of the cHMA method, and (3) to test the ability of cHMA to replicate previously published morphometric comparisons of metacarpals and femora of different primate taxa in two case studies.

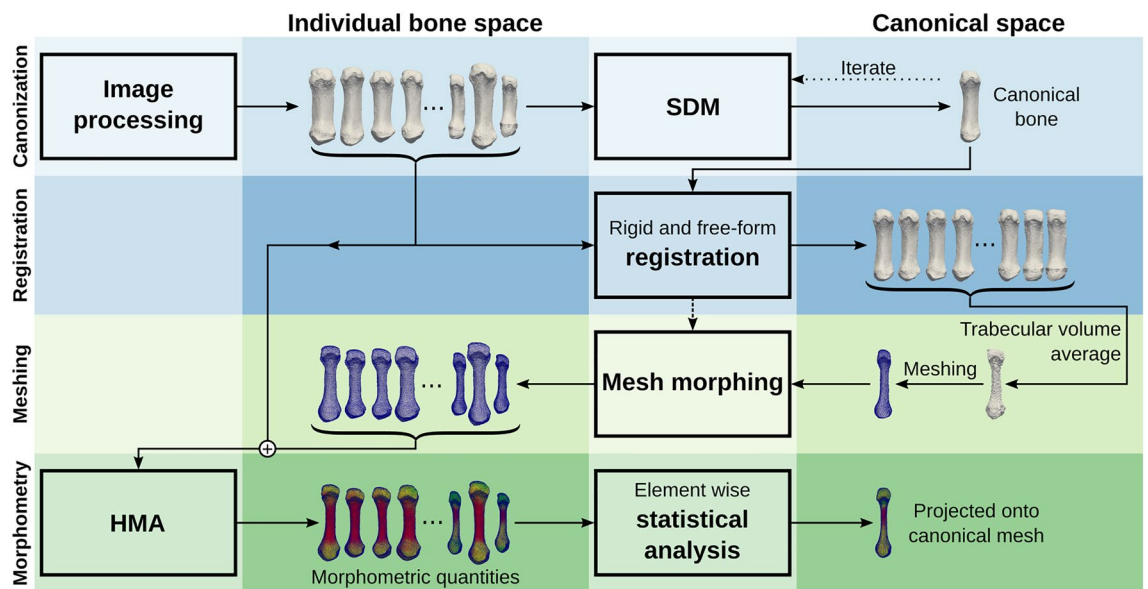
## Materials and methods

**HMA method.** The HMA method is the basis for the new cHMA workflow and is described in detail by Gross et al.<sup>38</sup> and Tsegai et al.<sup>12</sup>. Therefore only a brief outline shall be given here; the workflow is shown graphically in Fig. 1. HMA can map a wide variety of morphometric quantities of bones<sup>28</sup> from both trabecular and cortical volumes onto finite element meshes. Binary segmented  $\mu$ CT images, depicting the microstructure of the bone, are used as the input for HMA. The first step of the process is to separate the trabecular volume from the whole volume using the “fill” method<sup>44</sup>. This method creates two image masks for the bone: a cortical and a trabecular mask. Separate masks are required to measure morphometric quantities unique to either the cortical or the trabecular bone. Then a volumetric mesh is created for the cortical and the trabecular volume, respectively. In the next step a regular grid, the so called background grid, is created and morphometric quantities are evaluated in a spherical ROI centred at each grid vertex. The sphere has the same radius as the spacing of the grid, thus the sampling spheres overlap. Typically, 2.5 mm grid spacing and 5 mm sphere diameter are used due to the underlying homogenisation theory<sup>45,46</sup>. In the last step, the measured values at the background grid vertices are interpolated tri-linearly at the centroids of the mesh elements.

**The cHMA method.** *Outline.* The cHMA method extends the established HMA workflow and combines it with several other methods. A graphical summary can be seen in Fig. 2. It works by creating a canonical reference onto which different HMA results can be mapped and thus directly compared. The first stage of the workflow is to create a canonical bone image by using a statistical deformation model (SDM). In the next stage, all individual bone images are registered onto the canonical bone image, yielding transformations that map each bone volume onto the canonical bone image. The trabecular volume within all individual bone images is transformed into the canonical space, averaged and meshed using a tetrahedral mesher, which is then referred to as the canonical mesh. Isotopological meshes are created by morphing the canonical mesh back into each in-



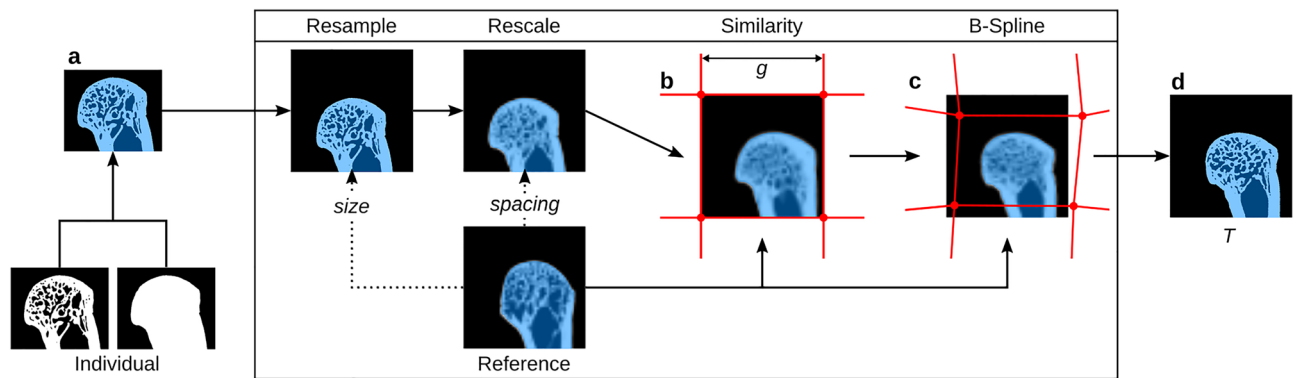
**Figure 1.** Mapping of morphometric quantities of trabecular bone onto finite-element meshes using the HMA method. The schematic overview shows the required image processing steps for a 2D slice, although the principle is similar in 3D. Original segmented  $\mu$ CT image (a). Trabecular mask (b). Mesh generated from trabecular mask (c). Masked trabecular volume (d). Evaluation of morphometric quantity using background grid with grid distance  $r$  and sampling sphere with radius  $r$  (e). Evaluated quantity at the background grid vertices (spheres not to scale) (f). Linear interpolated quantity on the mesh (g).



**Figure 2.** Graphical summary. Each block describes a separate step in the workflow, which runs from top to bottom.

dividual bone space. The HMA method can then be applied using the original bone images and these individual meshes, which possess homologous elements across the sample. In a further step, the mapped morphometric quantities can be statistically compared at each element across the samples. If required, results of HMA or the statistical analysis can be mapped back onto the canonical mesh and evaluated further.

*Statistical deformation model.* The SDM is created similar to Steiner et al.<sup>47</sup> but the process shall be repeated here briefly. The method operates on masked images of bones, where one image (the mask) describes the label of each voxel (bone or background) while the other image contains the actual grey-value. These can be the same images as required for the HMA workflow. All images are registered onto a randomly chosen reference image using a similarity transform. This type of transform allows only for translation, rotation and isotropic scaling. All similarity transformations are averaged by taking the arithmetic mean of the translation and scaling as well as averaging the rotation by quaternion averaging<sup>48</sup>. The centre of rotation is fixed in all transformations. The inverse of the averaged transformation is then applied to the reference image, to scale, rotate and translate the reference image into an average position. In the next step all images are registered onto the aligned reference image. First, a similarity transform is again used, then a free-form deformation is applied using a cubic B-spline transformation. Such a B-spline transformation is parametrized using a so called control point grid, which is deformed to create the overall image transformation. When all images are registered, the transformed images



**Figure 3.** Schematic overview of a single registration run in 2D, which is in principle similar to the real registration in 3D. Original binary segmented  $\mu$ CT image and mask are combined into masked image (a). This image is then resampled to the same size and rescaled to match the voxel spacing of the reference. Image after similarity transform using the reference image and undeformed B-spline control point grid with grid spacing  $g$  (b). Only grid nodes in the vicinity of the image domain are shown. Image after B-spline transformation with deformed control point grid (c). Transformed original image using the transformation  $T$  gained from the two registration steps (d).

and the B-spline control point grid displacements are averaged. The resulting average B-spline transformation is then inverted and applied to the averaged image, yielding the new reference image. The steps of registering and averaging can be iterated several times, until the model has converged. Diminishingly qualitative changes in the canonical bone shape indicate the convergence of the model, while the surface distance between two consecutive iterations can be used for quantification. The last reference image is then used as the canonical bone image for the further steps.

Adaptations to the original workflow for QCT images are required for  $\mu$ CT images because the runtime of the registration is proportional to the image size and hence registration of  $\mu$ CT images can become impractical. Therefore, a resizing step was added to rescale the images to a resolution similar to that of QCT by increasing the voxel size by an isotropic factor. Rescaling the images before registration reduces the number of voxels and because the resulting transformations from the registration are smooth over the whole image domain, they can be applied to the original high-resolution images that accurately depict trabecular architecture as well<sup>49</sup>. The process of registration including the rescaling process is shown schematically in Fig. 3. However, the rescaling of the images also influences the registration quality, as with increasing scale factors, details are lost. Another parameter which influences the registration quality but also the overall runtime is the B-spline control grid spacing. A lower grid spacing can be used to register local features better at the expense of higher runtime. Therefore, optimal values for these two parameters have to be identified which allow for a feasible runtime with adequate registration quality.

Negative normalised cross correlation was used for both similarity and B-spline registration as cost-function for the optimisation. A conjugate gradient line search optimiser was used for the similarity and a L-BFGS-B optimiser<sup>50</sup> for the B-spline registration. All other parameters were used as described by Steiner et al.<sup>47</sup>.

**Mesh morphing.** After the canonical bone image is created, each image is registered onto the canonical bone image. This yields transformations that are used to morph a canonical mesh, generated on the canonical bone, onto each individual bone. In this step, only the vertices of the meshes are moved in space, without altering their connectivity. Thus, isotopological meshes are created, that is, the topology of the meshes does not change. However, in this process the elements might get distorted, which can, in the extreme cases, lead to degraded elements. The degradation of the elements can be checked using various mesh metrics, which detect different types of degradation<sup>51</sup>. Because isotopological meshes are a requirement for the cHMA workflow, common strategies to repair degraded elements after morphing, such as local or global remeshing, can not be applied. Therefore, excessive mesh degradation has to be avoided in the first place. Hence, the registration procedure should yield transformations, which add minimal mesh distortion during the morphing step.

**Parameter identification.** The registration method can be controlled by a set of parameters, which influence certain steps during the registration process. The cHMA workflow should allow analyses of high resolution ( $\mu$ CT) scans in a reasonable amount of time, capture sufficient detail in the anatomical structure, but should avoid overfitting during the registration and distortions of the mesh.

To identify parameters suitable for morphometric analysis of  $\mu$ CT images from bones of different species, a sample of thirty  $\mu$ CT images of first metacarpals was used from a previous study<sup>15</sup>. The set contained ten samples each of *Homo sapiens*, *Pan paniscus*, and *Pan troglodytes verus*, respectively. The images had different voxel sizes ranging from 0.023 to 0.031 mm due to the usage of different scanners. Therefore, all images were resampled to an isotropic voxel size of 0.03 mm and a total image size of  $674 \times 681 \times 1779$  voxels during the registration procedure. All images were already segmented and masked from the previous study<sup>15</sup>. Images from the left hand were mirrored to match the right side. A rendering of all bones can be found in the Supplementary Material S1.

Two main parameters were varied in order to identify a suitable parameter set: (1) Image rescaling factor and (2) B-spline control grid spacing. The previously presented model by Steiner et al.<sup>47</sup>, developed for the proximal femur, used a grid spacing of  $21 \times 15 \times 37$  mm with an isotropic voxel size of 0.6 mm. Therefore, two different rescaling factors of 10 and 20 were tested, resulting in a voxel size of 0.3 mm and 0.6 mm respectively and the grid spacing was varied with 4, 6, 10, 20 and 30 mm distance. As the grid spacing is not the same in the three cardinal directions, the given spacings resulted in a grid of  $5 \times 5 \times 13$ ,  $3 \times 3 \times 9$ ,  $2 \times 2 \times 5$ ,  $1 \times 1 \times 3$ , and  $1 \times 1 \times 2$  control points. A total of 10 different parameter runs were evaluated.

The full SDM workflow using two iterations, registration and mesh morphing was applied for each parameter run. Tetrahedral meshes were created using a characteristic edge length of 1 mm for all elements. The result was judged based on both image registration metrics and mesh distortion metrics as well as visual inspections of the registered images and morphed meshes. The overall runtime of the SDM workflow was measured for each parameter run. Image metrics, namely Dice coefficient, mean surface distance (MSD) and Hausdorff distance, were calculated between the transformed images and the canonical bone<sup>52</sup>. Mesh metrics, namely tet-collapse<sup>53</sup> and volume-skew<sup>54</sup> were calculated for the canonical and individual meshes. Briefly, tet-collapse measures the ratio of height and face area and is zero for a fully collapse tetrahedron and one for an optimal tetrahedron. Volume-skew measures the deviation in volume from an equilateral tetrahedron and is one for a degenerated tetrahedron and zero if the tetrahedron is equilateral. The final parameter set was selected by choosing the parameter set with the highest image metrics but lowest mesh degradation.

**Robustness.** To evaluate robustness of the cHMA method, three different tests were performed: (1) the convergence behaviour of the SDM was investigated, (2) the influence of the starting images was evaluated and (3) a tenfold cross-validation was performed. All tests were performed using the chosen parameter set from the parameter identification and the metacarpal bone sample as described above.

In order to identify the number of iterations needed in the SDM, a model was created using a total of ten iterations, saving the results of intermediate iterations. The intermediate images were compared using the same image metrics as used for the parameter identification using the image of the final iteration as reference.

The bias towards the starting image was tested by generating a canonical bone with two iterations but choosing a different random starting image for each run. The canonical image of the parameter identification served as reference and image metrics were calculated using nine additional starting configurations.

Finally, a tenfold cross-validation was performed, i.e., leaving out three samples per run for the canonical bone creation. The same starting image was used, except for the run where this image was excluded, to get comparable results without the influence of the starting image. Again, two iterations were used in the SDM and the final canonical bone was compared using the image metrics to the canonical bone of the parameter identification.

**Case studies.** To verify and showcase the applicability of the cHMA method, two case studies were performed on different bones. While the workflow can be applied using any morphometric quantity, only bone volume fraction was analysed for these case studies. Results were compared with previous studies to confirm that the same patterns can be observed with the new workflow.

*First metacarpals.* The first metacarpal data set, already described in the sections above, was used to replicate findings from a previous study on the same data set<sup>15</sup>. Specifically, Dunmore et al.<sup>15</sup> reported higher relative BVTV (rBVTV) in the radiopalmar aspects of the head and base in *Homo* compared to the great apes. These results were based on qualitative interpretation of the volumetric distribution of trabecular morphometry, while quantitative assessment was limited to subcortical bone below the joint surfaces.

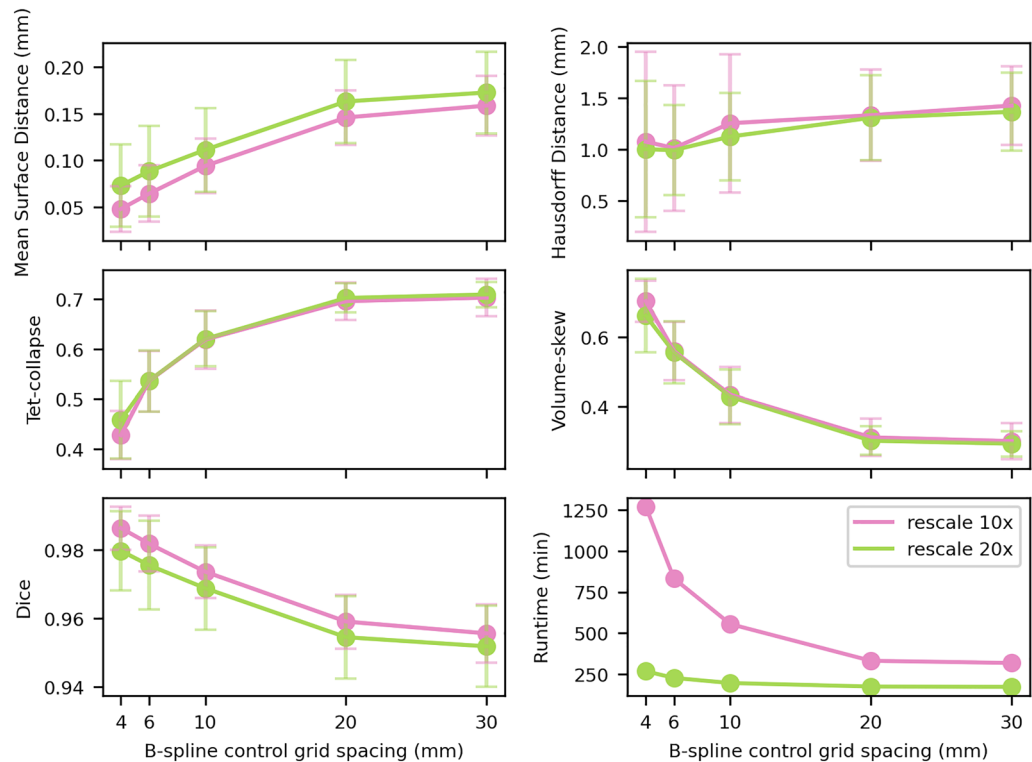
To test if cHMA can replicate and quantify these results for the entire volume, the workflow was run with the parameters as identified following the above sections. The canonical bone was created using two iterations in the SDM and the mesh was created using a characteristic element side length of 1 mm and morphed onto all bones. A background grid of 2.5 mm was used in the HMA process, with a sampling sphere diameter of 5 mm. BVTV was evaluated for all samples with HMA and rBVTV was calculated by normalising each sample's BVTV by its mean BVTV. This step was necessary to compensate for systematic differences between taxa and allowed for the analysis of bone volume distribution while controlling for magnitude<sup>16,23</sup>.

Element-wise statistical tests were performed on the resulting rBVTV values mapped onto the isotopological meshes. Pairwise independent samples t-tests were used, followed by permutation tests for family-wise error rate (FWER) control<sup>55</sup>, yielding a new t-value threshold for significance. The significance level was set to  $\alpha = 0.05$ . Moreover, principal component analysis (PCA) was calculated on the mesh elements to identify common and distinct rBVTV patterns among species. All values are given as mean  $\pm$  standard deviation, if not denoted otherwise.

*Proximal femur.* A second case study was performed in order to check whether the results of a study on a larger bone, not part of the parameter identification procedure, can also be replicated. Thirty proximal femora, cut below the lesser trochanter, from a previous study were used<sup>18,19</sup>. The set contained ten samples of *Homo sapiens*, *Gorilla gorilla*, and *Pan troglodytes*, respectively. All images were already segmented and masked similar to the first metacarpals. A rendering of all bones can be found in the Supplementary Material S2.

Georgiou et al.<sup>18</sup> found similar BVTV patterns for *Pan* and *Gorilla* while *Homo* showed a different pattern, lacking an anterior concentration in the femoral head. Moreover, the species specific pillar structures inside the femoral head, likely stemming from postural differences between the species, should be investigated further<sup>19</sup>. These findings were so far only described qualitatively.

Overall, the same settings as for the metacarpal sample set were used for both SDM and HMA. Only the image rescaling and mesh element size were adapted to reduce the runtime as a result of the larger bone size. Original



**Figure 4.** Average image and mesh metrics as well as runtime for the parameter identification studies. Each plot shows a single metric, split up for the two tested rescaling factors. The bars show the standard deviation. Runtimes are given for the canonical bone creation using 16 CPUs in parallel.

voxel sizes ranged from 0.05 to 0.07 mm and were all resampled to an isotropic voxel size of 0.06 mm, leading to an image size of  $1247 \times 1836 \times 1916$  voxels during the registration. The B-spline grid setting resulted in  $4 \times 5 \times 6$  grid points over the image domain and 10× rescaling resulted in a voxel size of 0.6 mm during registration. Meshes for HMA were created using a characteristic edge length of 3 mm to account for the larger size of the bones compared to the metacarpals<sup>38</sup>. Statistical analysis was similar to the metacarpal case study. However, in addition to the full bone analysis, the head region was cropped and analysed separately to be consistent with Georgiou et al.<sup>18,19</sup>. This included thresholding the rBTVV distribution of the femoral head at the 80-percentile to qualitatively identify the presence of the previously observed pillar structures.

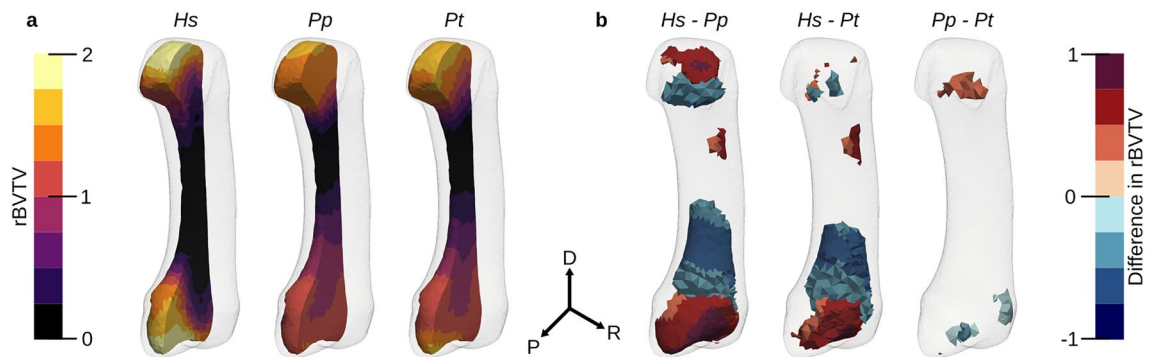
**Hardware and software.** The whole registration and SDM framework was implemented in SimpleITK 2.0<sup>56</sup> using Python 3.7 (Python Software Foundation, <https://www.python.org>). All other steps were performed directly in Medtool 4.5 (Dr. Pahr Ingenieure e.U., <http://www.medtool.at>). Meshing was done using CGAL 4.11<sup>57</sup> in Medtool. Visualisations were created in ParaView 5.9 (Kitware, <https://www.paraview.org>). Statistical analysis was done with SciPy 1.2.3<sup>58</sup> in Medtool.

All registration regarding the first metacarpal were run on a Dual Intel(R) Xeon(R) Gold 6144 @ 3.50 GHz using 16 cores in parallel. All registration regarding the proximal femur were run on a Dual Intel(R) Xeon(R) E5-2697 v3 @ 2.60 GHz using 28 cores in parallel.

## Results

**Parameter identification.** All ten parameter runs produced canonical bones and all bones were successfully registered onto it. Image and mesh metrics were calculated for each parameter run and the runtime for canonical bone creation and registration was tracked (Fig. 4). For larger B-spline grid spacings and higher rescaling factors, the runtime was lower than for small grid spacings and lower rescaling factors. The fastest run took 171 min, while the slowest run took 1270 min. Image metrics got better for lower grid spacings and lower rescaling factors, while mesh metrics got worse on average.

The mesh metrics showed a steep deterioration for grid spacings smaller than 20 mm, while the image metrics changed more gradually over the full range of tested spacings. As a result of this drastic change in mesh metrics, a 20 mm grid spacing was considered as the smallest possible grid spacing, without large mesh distortions. Single meshes had areas of distorted elements, apparent in the visual inspection, for grid spacings 10 mm and lower, while no distortions could be observed at 20 mm spacing or higher. The runtimes for 20 mm grid spacing were 330 min for 10× rescaling and 173 min for 20× rescaling. Despite longer runtimes, 10× rescaling was chosen as it had considerably better MSD on average (− 10%) with similar Hausdorff distance (+ 1.9%).



**Figure 5.** Average rBTVT and thresholded elements after FWER control. Average rBTVT per species cut along the sagittal plane (a). From left to right: *Hs* *Homo sapiens*, *Pp* *Pan paniscus*, *Pt* *Pan troglodytes*. Statistically significant groupwise differences, thresholded at critical level for  $\alpha = 0.05$  after FWER-control (b). Color gives the difference in rBTVT per species. For each pairwise comparison, a positive value (red) indicates that the first species has higher rBTVT while a negative value (blue) indicates that the second species has a higher rBTVT. *D* distal, *P* palmar, *R* radial.

**Robustness.** Using the identified parameters, a canonical bone was created using up to ten iterations. Qualitatively, no substantial change was visible for two or more iterations, suggesting that the model had converged (see Supplementary Fig. S1). Rigid body motion of the bone was observed, which resulted in a higher surface distance between two successive iterations than anticipated. The MSD between the final and penultimate iteration was  $0.027 \text{ mm} \pm 0.026 \text{ mm}$  with a Dice coefficient of 0.99. Two iterations were used for all further studies, as there was no visible change except for minor rigid body motion.

Qualitative inspection of the canonical bone with a different start image showed no considerable changes in bone shape but differences in bone size, translation and rotation (see Supplementary Fig. S3). The canonical bone size, ranged from 39 to 41.3 mm with a mean of 40.6 mm in the longitudinal axis. These size differences were also apparent in the MSD, which was 0.43 mm on average and ranged from 0.17 to 0.67 mm. Dice coefficient was 0.87 on average, ranging from 0.83 to 0.95.

All except one canonical bone in the cross-validation study were qualitatively similar in shape. This one canonical bone used a different starting image and thus resulted in a different sized bone, which is in line with the start image bias described above. MSD was 0.09 mm on average with values ranging from 0.04 to 0.34 mm and Dice coefficient was 0.97 on average, ranging from 0.91 to 0.99.

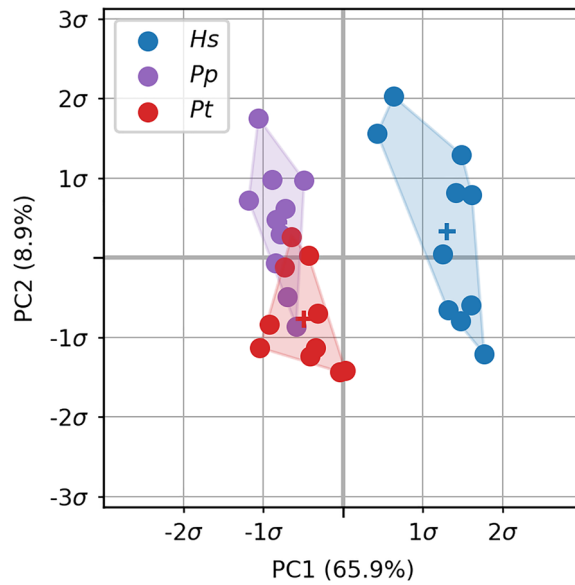
**Case studies.** *First metacarpals.* All images were successfully registered onto the canonical bone. Dice coefficient was  $0.96 \pm 0.008$ , Hausdorff distance  $1.33 \text{ mm} \pm 0.443 \text{ mm}$  and MSD  $0.146 \text{ mm} \pm 0.03 \text{ mm}$ . The canonical mesh consisted of 10880 linear tetrahedral elements. Volume-skew for the canonical mesh was 0.49  $\pm$  0.216 and tet-collapse  $0.59 \pm 0.153$ . After mesh morphing, the average volume-skew was 0.52 on average, ranging from 0.49 to 0.62 and tet-collapse was 0.56 on average, ranging from 0.5 to 0.58.

Relative BTVT (rBTVT) was evaluated for all samples and group-wise comparisons between the species were done element wise. Mean rBTVT distributions for each species and thresholded canonical meshes for the three pairwise comparisons can be seen in Fig. 5. *Pan paniscus* and *Pan troglodytes* had similar rBTVT distributions, which can be seen in the low number of significant elements after family-wise error rate (FWER) control. However, both species significantly differed in comparison to *Homo*, which showed higher densities in the radio-palmar region in the base. A smaller area of significant difference in the radio-palmar aspect of the head can be seen in comparison to *Pan paniscus*. These results are in line with the previous study, however there are also differences found which were not visible in the previous study due to the limited focus on trabecular distribution immediately below the subchondral region. *Homo* shows a region of lower density in the proximal shaft region above the base, in comparison to *Pan*. There is also a smaller region of higher density in *Homo* in the disto-radial region at the shaft. In the PCA plot (Fig. 6) of the first two principal components of rBTVT, *Homo* separates well from *Pan* via the first component. The first component accounts for 65.9% of the total variance, and the second component for 8.9%.

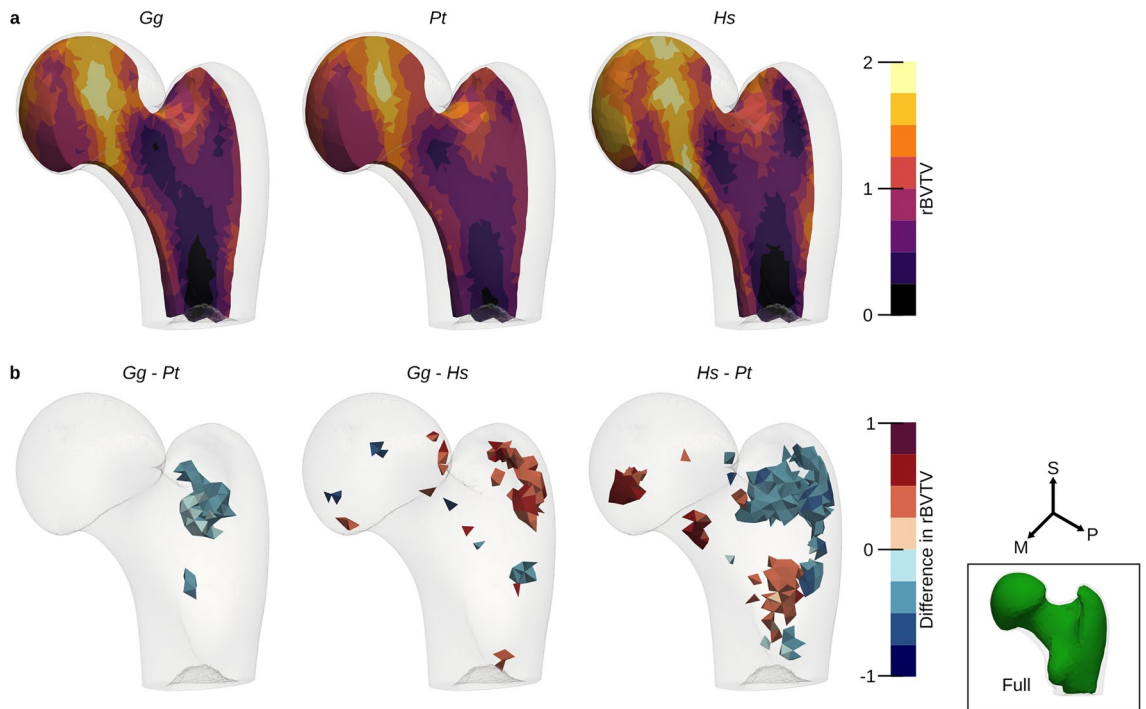
*Proximal femur.* The canonical bone was successfully created and all femora were registered onto the canonical bone. The canonical mesh contained 13075 linear tetrahedral elements. Volume-skew for the canonical mesh was  $0.25 \pm 0.147$  and tet-collapse  $0.74 \pm 0.12$ . After the mesh-morphing, volume-skew was 0.46 on average, ranging from 0.35 to 0.56 and tet-collapse was 0.6 on average, ranging from 0.53 to 0.67. Dice coefficient of the images after registration was  $0.98 \pm 0.005$  and Hausdorff distance was  $6.48 \text{ mm} \pm 4.022 \text{ mm}$ , while the MSD was  $0.246 \text{ mm} \pm 0.11 \text{ mm}$ .

The same HMA workflow was applied as for the first metacarpal bone. No clear pattern was visible in the univariate statistics for the full proximal femur (Fig. 7) after FWER control. Therefore, the elements in the head region were analysed separately. The head region, which consisted of 5140 elements, was extracted from the full mesh and rBTVT was calculated for this region separately. Less than 70 elements showed significant differences after FWER control in each pairwise comparison. The 80-percentile threshold of the rBTVT distribution of the



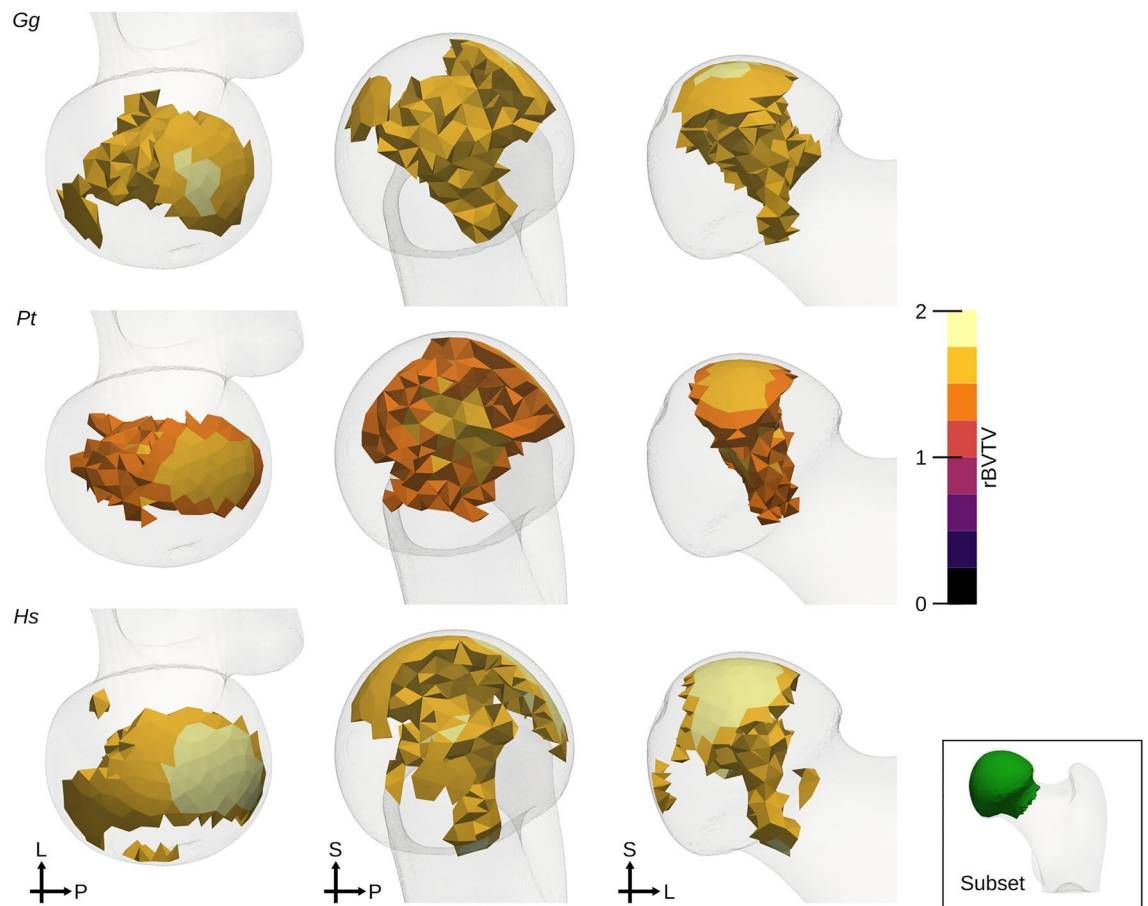


**Figure 6.** PCA plot of the first metacarpals. The plot shows the first two components in the principal component space for rBVTV. *Hs* *Homo sapiens*, *Pp* *Pan paniscus*, *Pt* *Pan troglodytes*.

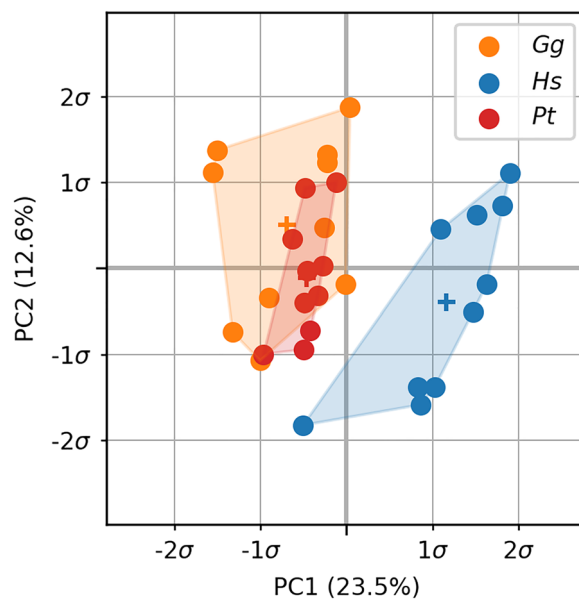


**Figure 7.** Mean rBVTV distribution (a) and statistical analysis after FWER control of the proximal femur (b). From left to right: *Gg* *Gorilla gorilla*, *Pt* *Pan troglodytes*, *Hs* *Homo sapiens*. S superior, M medial, P posterior.

mean mesh of each species showed a pillar structure inside the femoral head (Fig. 8). This pillar structure could be found in all three species and distally bifurcated into two pillars ending at the posterior and anterior aspects articular surface. The posterior pillar was present in all three species and extended anteriorly at the trabecular surface in only *Homo*. *Gorilla* and *Pan* exhibited an anterior pillar deep in the trabecular structure, which reached the trabecular surface in only *Gorilla*. The PCA plot (Fig. 9) for rBVTV in the head region showed a good separation of *Homo* versus *Gorilla* and *Pan*. However, the first principal component only accounted for 23.5% and the second for 12.6%.



**Figure 8.** Thresholded femoral heads. rBVTV after thresholding at 80-percentile of each distribution in the femoral head only, shown here in three different views per species. From top to bottom: *Gg Gorilla gorilla*, *Pt Pan troglodytes*, *Hs Homo sapiens*. L lateral, S superior, P posterior.



**Figure 9.** PCA plot for the femoral head. The plot shows the first two components in the principal component space for rBVTV. *Gg Gorilla gorilla*, *Hs Homo sapiens*, *Pt Pan troglodytes*.

## Discussion

The cHMA workflow shows the applicability of statistical deformation models to  $\mu$ CT data, using a lower resolution image for the registration and the original, high-resolution image for the morphometric analysis with HMA. The method was able to replicate results from previous studies and also showed new, previously undetected differences between species. Furthermore, it was possible to confirm previously reported qualitative differences in a quantitative way. The canonical bone creation was robust and the registration produced images with high Dice coefficients and low MSD on average for both studied bones.

A similar method was recently presented by DeMars et al.<sup>24</sup>. The method is similar to the approach shown here, as HMA is also used to evaluate morphometric quantities. However, the meshes are created on the individual bones first and then correspondence is established using a coherent point drift algorithm. That means that meshes are created for all bones separately, while the here presented workflow uses isotopological meshes without the need of interpolation of values onto the canonical point cloud. Another similar workflow was presented by Taghizadeh et al.<sup>59</sup>. A canonical bone is constructed and meshed, however the individual meshes are then used to build a statistical model of shape, BTV and anisotropy rather than doing element-wise statistical analysis. Another similar approach was used by Marangalou et al.<sup>60</sup>. However, in their method, a mesh-morphing approach with landmarks was used and not an SDM. All these three models were so far only applied to human bones. In contrast to other existing methods, the cHMA method was tested on bones of different species and provides a high level of automation. For instance, no landmarks need to be defined to run the model and  $\mu$ CT data can directly be fed into the workflow.

Despite the high level of automation, the cHMA workflow still requires a number of parameters to be set manually. The parameter identification procedure highlighted the need to find a balance between capturing anatomical details and maintaining high quality meshes. Here, the parameters were identified by subjective criteria, i.e., minimal distortion in the meshes, maximal overlap of images, and visual inspection of the resulting meshes. A more objective measure would be beneficial in the future, however due to the lack of a gold-standard, it was not possible to select the parameters based on absolute measures. Except for the mesh size and voxel size during registration, the parameters identified from metacarpals could also be used for femoral bones. Even though the effective voxel size during registration was larger in the femora study than for the metacarpals, to reduce the runtime, the registration performance was still acceptable.

Although the method was generally robust, a marginal start image bias was observed that resulted in different location, rotation and scale of the canonical bone. This bias can be explained by the use of the similarity transform. The similarity transformation did not always produce images which would overlap in a true anatomical homologous position. For example, a rotation might also minimise the metric in the same way as scaling would do. To enforce a better initial alignment of the bones, anatomical landmarks could be used, similar to shape models. However, landmarking requires extra manual work and introduces new errors into the model. Another option would be the use of a different metric or tweaking the optimiser to find a better starting position.

The image registration worked in all cases, yielding high Dice coefficients above 95% on average and low MSD, less than the size of a single voxel in the rescaled versions (0.3 mm for the first metacarpals and 0.6 mm for the proximal femur). Hausdorff distances were slightly higher than expected, especially in the femur, with four individuals reaching distances above 10 mm. In particular, high surface distances were seen in the distal region of the femoral shaft. As the femora images were cropped at the shaft, it is likely that the high surface distance stem from those regions, similar to other studies where cut sections of bones were used<sup>61</sup>. However, these parts of the proximal femur are of less interest for trabecular analysis, as they contain almost no trabecular bone. Therefore, it is possible to remove parts of the bone from the analysis, by removing parts of the canonical bone before mesh generation or cropping the mesh afterwards.

Mesh morphing resulted in only minor deterioration of the mesh quality metrics, but individual elements in the final morphed meshes were still strongly distorted. This might be caused by the initial tetrahedral quality, which already showed some distorted elements. However, only a maximum of 0.12% of the elements failed for volume-skew metric (volume-skew larger than 0.9), and only up to 0.06% of the elements failed for aspect ratio, a metric commonly used in finite element modelling, for which an acceptable region is lower than 5% of all elements<sup>62</sup>. Furthermore, mesh quality is not a pressing issue here, as the models are not used for finite element analysis (FEA). When using mesh-morphing to create FE meshes, the initial mesh quality should be good, as deformation usually leads to distorted elements<sup>63</sup>. For further studies the canonical mesh quality should therefore be improved, to produce better morphed meshes. Other automatic meshing tools might yield better results and existing meshes could be re-meshed for better quality. However, even with a perfect canonical mesh, there is the chance that morphed meshes deteriorate, which can only be controlled for by using an optimiser procedure which incorporates information about mesh distortion, i.e., some sort of regularisation.

It was possible to run the full workflow on both metacarpal as well as femoral bones using the identified parameters. In both case studies, known differences between species could be found, and new, previously unreported differences in the trabecular bone could be observed. These novel significant differences were located in the trabecular bone, farther away from the cortical surface, which was previously not analysed. For the metacarpal bones, the spatial extent of differences at the base between *Homo* and *Pan* could be observed, including the density differences in the medullary cavity. Density patterns for the femoral head were consistent with previous results<sup>18,19</sup>. However, there was no significant difference in the full proximal femur or the femoral head using univariate statistics and the pillar structure was not as prominent in the averaged images as it was shown to be in individuals. Only multivariate statistics, using PCA, on the cropped femoral head demonstrated the previously seen difference between the species. These findings suggest that the pillar structure might be more variable between individuals and thus is not found in the averaged images. Hence, element-wise comparisons are not able to pick up the signal because the variance for this small sample set is too high.

This study has some limitations. Only a small subset of all available parameters was tested in a parameter study. All other parameters were used from literature. Additional parameter tuning could be done individually for each investigated bone to further improve the accuracy of the cHMA workflow. There was no regularisation method added in the B-spline registration, leading to distortions on the resulting meshes if the grid spacing was too small. The large B-spline grid spacing, used here, already regularised the deformation field to some extent, but other methods are available to gain more control over the B-spline, for example the imposition of bounds on the control grid displacements or control of the smoothness or rigidity of the resulting deformation field<sup>64</sup>. The creation of the canonical bone can be time-consuming, especially if lower B-spline grid spacing is used. Methods to further improve runtimes include using a B-spline pseudo inversion<sup>65</sup> or iterative shape averaging<sup>66</sup> for canonical bone creation. The sample set in these studies was relatively small, with only ten individuals per species. Furthermore, there was no differentiation regarding sex and age for both case studies, thus effects which are caused by ageing or due to sexual dimorphism might also explain the variability within the species. Therefore it should be repeated with a larger sample set and potentially also further separation of sexes and age groups. Additionally, all bones used in this study were not damaged or pathological, and high-quality CT-scans of them were available. Caution is warranted when the trabecular structure of potentially taphonomically altered bones is analysed in general<sup>67,68</sup>, and these cases could require adaptations to the cHMA method. FWER control was done using permutation tests, while random field theory (RFT) offers a faster, parametric approach<sup>69,70</sup>. However, RFT implementation for irregular tetrahedral meshes is challenging, and the requirements for the use of RFT might not be met by the data. Moreover, RFT currently does not support the use of non-parametric tests.

The here presented cHMA method shows an extension to the existing HMA method to quantitatively compare the trabecular morphology of bones of different species using a canonical bone and isotopological meshes. It works independent of the size of bones, is robust with respect to the choice of samples and starting image and the canonical bone creation converges quickly. The method can be applied directly to existing  $\mu$ CT data sets and only requires segmented and masked images; it does not require landmarks. While the method was currently only tested on very well conserved bones of recent species, it shall be used for fossils as well in the future, where taphonomic damages pose an extra challenge in the registration and trabecular analysis. Furthermore, the cHMA workflow shall be applied to other bones, using more morphometric quantities and more samples to investigate also subtle differences and the local influence of morphometric quantities other than relative bone volume fraction.

Received: 8 October 2021; Accepted: 14 March 2022

Published online: 25 March 2022

## References

- Ruff, C., Holt, B. & Trinkaus, E. Who's afraid of the big bad Wolff?: "Wolff's law" and bone functional adaptation. *Am. J. Phys. Anthropol.* **129**, 484–498. <https://doi.org/10.1002/ajpa.20371> (2006).
- Seeman, E. Bone modeling and remodeling. *Crit. Rev. Eukaryot. Gene Expr.* **19**, 219–233. <https://doi.org/10.1615/critrevueukargeneexpr.v19.i3.40> (2009).
- Carlson, K. J. & Judex, S. Increased non-linear locomotion alters diaphyseal bone shape. *J. Exp. Biol.* **210**, 3117–3125. <https://doi.org/10.1242/jeb.006544> (2007).
- Byron, C. D., Herrel, A., Pauwels, E., Muynck, A. D. & Patel, B. A. Mouse hallux metatarsal cross-sectional geometry in a simulated fine branch niche. *J. Morphol.* **276**, 759–765. <https://doi.org/10.1002/jmor.20376> (2015).
- Turcotte, C. M., Rabey, K. N., Green, D. J. & McFarlin, S. C. Muscle attachment sites and behavioral reconstruction: An experimental test of muscle-bone structural response to habitual activity. *Am. J. Biol. Anthropol.* **177**, 63–82. <https://doi.org/10.1002/ajpa.24410> (2022).
- Karakostis, F. A., Jeffery, N. & Harvati, K. Experimental proof that multivariate patterns among muscle attachments (entheses) can reflect repetitive muscle use. *Sci. Rep.* **9**, 16577. <https://doi.org/10.1038/s41598-019-53021-8> (2019).
- Pontzer, H. *et al.* Trabecular bone in the bird knee responds with high sensitivity to changes in load orientation. *J. Exp. Biol.* **209**, 57–65. <https://doi.org/10.1242/jeb.01971> (2006).
- Barak, M. M., Lieberman, D. E. & Hublin, J.-J. A wolf in sheep's clothing: Trabecular bone adaptation in response to changes in joint loading orientation. *Bone* **49**, 1141–1151. <https://doi.org/10.1016/j.bone.2011.08.020> (2011).
- Currey, J. D. The structure and mechanics of bone. *J. Mater. Sci.* **47**, 41–54. <https://doi.org/10.1007/s10853-011-5914-9> (2011).
- Wallace, I. J., Demes, B. & Judex, S. Ontogenetic and genetic influences on bone's responsiveness to mechanical signals. In *Building Bones: Bone Formation and Development in Anthropology* (eds Percival, C. J. & Richtsmeier, J. T.) 233–253 (Cambridge University Press, 2017).
- Tsegai, Z. J., Skinner, M. M., Pahr, D. H., Hublin, J.-J. & Kivell, T. L. Systemic patterns of trabecular bone across the human and chimpanzee skeleton. *J. Anat.* **232**, 641–656. <https://doi.org/10.1111/joa.12776> (2018).
- Tsegai, Z. J. *et al.* Trabecular bone structure correlates with hand posture and use in hominoids. *PLoS ONE* **8**, e78781. <https://doi.org/10.1371/journal.pone.0078781> (2013).
- Skinner, M. M. *et al.* Human-like hand use in *Australopithecus africanus*. *Science* **347**, 395–399. <https://doi.org/10.1126/science.1261735> (2015).
- Stephens, N. B. *et al.* Trabecular architecture in the thumb of pan and homo: Implications for investigating hand use, loading, and hand preference in the fossil record. *Am. J. Phys. Anthropol.* **161**, 603–619. <https://doi.org/10.1002/ajpa.23061> (2016).
- Dunmore, C. J., Bardo, A., Skinner, M. M. & Kivell, T. L. Trabecular variation in the first metacarpal and manipulation in hominids. *Am. J. Phys. Anthropol.* **171**, 219–241. <https://doi.org/10.1002/ajpa.23974> (2019).
- Dunmore, C. J., Kivell, T. L., Bardo, A. & Skinner, M. M. Metacarpal trabecular bone varies with distinct hand-positions used in hominid locomotion. *J. Anat.* **235**, 45. <https://doi.org/10.1111/joa.12966> (2019).
- Tsegai, Z., Skinner, M., Pahr, D., Hublin, J.-J. & Kivell, T. Ontogeny and variability of trabecular bone in the chimpanzee humerus, femur and tibia. *Am. J. Phys. Anthropol.* **167**, 713–736. <https://doi.org/10.1002/ajpa.23696> (2018).
- Georgiou, L., Kivell, T. L., Pahr, D. H., Buck, L. T. & Skinner, M. M. Trabecular architecture of the great ape and human femoral head. *J. Anat.* **234**, 679–693. <https://doi.org/10.1111/joa.12957> (2019).
- Georgiou, L. *et al.* Evidence for habitual climbing in a pleistocene hominin in South Africa. *Proc. Natl. Acad. Sci.* **117**, 8416–8423. <https://doi.org/10.1073/pnas.1914481117> (2020).

20. Georgiou, L., Kivell, T., Pahr, D. & Skinner, M. Trabecular bone patterning in the hominoid distal femur. *PeerJ* **6**, e5156. <https://doi.org/10.7717/peerj.5156> (2018).
21. Colombo, A. *et al.* Trabecular analysis of the distal radial metaphysis during the acquisition of crawling and bipedal walking in childhood: A preliminary study. *Bull. et Mem. de la Soc. d'Anthropol. de Paris* **31**, 43–51. <https://doi.org/10.3166/bmsap-2018-0041> (2019).
22. Komza, K. & Skinner, M. First metatarsal trabecular bone structure in extant hominoids and swartkrans hominins. *J. Hum. Evol.* **131**, 1–21. <https://doi.org/10.1016/j.jhevol.2019.03.003> (2019).
23. Sukhdeo, S., Parsons, J., Niu, X. M. & Ryan, T. M. Trabecular bone structure in the distal femur of humans, apes, and baboons. *Anat. Rec.* **303**, 129–149. <https://doi.org/10.1002/ar.24050> (2018).
24. DeMars, L. J. D. *et al.* Using point clouds to investigate the relationship between trabecular bone phenotype and behavior: An example utilizing the human calcaneus. *Am. J. Hum. Biol.* **33**, e23468. <https://doi.org/10.1002/ajhb.23468> (2020).
25. Judex, S. & Carlson, K. J. Is bone's response to mechanical signals dominated by gravitational loading? *Med. Sci. Sports Exerc.* **41**, 2037–2043. <https://doi.org/10.1249/mss.0b013e3181a8c6e5> (2009).
26. Robling, A. G. Is bone's response to mechanical signals dominated by muscle forces? *Med. Sci. Sports Exerc.* **41**, 2044–2049. <https://doi.org/10.1249/mss.0b013e3181a8c702> (2009).
27. Kivell, T. L. A review of trabecular bone functional adaptation: What have we learned from trabecular analyses in extant hominoids and what can we apply to fossils? *J. Anat.* **228**, 569–594. <https://doi.org/10.1111/joa.12446> (2016).
28. Boussein, M. L. *et al.* Guidelines for assessment of bone microstructure in rodents using micro-computed tomography. *J. Bone Miner. Res.* **25**, 1468–1486. <https://doi.org/10.1002/jbmr.141> (2010).
29. Maquer, G., Musy, S. N., Wandel, J., Gross, T. & Zysset, P. K. Bone volume fraction and fabric anisotropy are better determinants of trabecular bone stiffness than other morphological variables. *J. Bone Miner. Res.* **30**, 1000–1008. <https://doi.org/10.1002/jbmr.2437> (2015).
30. Kivell, T. L., Skinner, M. M., Lazenby, R. & Hublin, J.-J. Methodological considerations for analyzing trabecular architecture: An example from the primate hand. *J. Anat.* **218**, 209–225. <https://doi.org/10.1111/j.1469-7580.2010.01314.x> (2011).
31. Griffin, N. L. *et al.* Comparative forefoot trabecular bone architecture in extant hominids. *J. Hum. Evol.* **59**, 202–213. <https://doi.org/10.1016/j.jhevol.2010.06.006> (2010).
32. Chirchir, H., Zeininger, A., Nakatsukasa, M., Ketcham, R. A. & Richmond, B. G. Does trabecular bone structure within the metacarpal heads of primates vary with hand posture? *C.R. Palevol.* **16**, 533–544. <https://doi.org/10.1016/j.crpv.2016.10.002> (2017).
33. Mueller, T. L. *et al.* Non-invasive bone competence analysis by high-resolution pqct: An in vitro reproducibility study on structural and mechanical properties at the human radius. *Bone* **44**, 364–371. <https://doi.org/10.1016/j.bone.2008.10.045> (2009).
34. Sode, M., Burghardt, A. J., Kazakia, G. J., Link, T. M. & Majumdar, S. Regional variations of gender-specific and age-related differences in trabecular bone structure of the distal radius and tibia. *Bone* **46**, 1652–1660. <https://doi.org/10.1016/j.bone.2010.02.021> (2010).
35. Rubinacci, A. *et al.* Comparative high-resolution pqct analysis of femoral neck indicates different bone mass distribution in osteoporosis and osteoarthritis. *Osteoporos. Int.* **23**, 1967–1975. <https://doi.org/10.1007/s00198-011-1795-7> (2011).
36. Stephens, N. B., Kivell, T. L., Pahr, D. H., Hublin, J.-J. & Skinner, M. M. Trabecular bone patterning across the human hand. *J. Hum. Evol.* **123**, 1–23. <https://doi.org/10.1016/j.jhevol.2018.05.004> (2018).
37. Du, J. *et al.* Characterising variability and regional correlations of microstructure and mechanical competence of human tibial trabecular bone: An in-vivo hr-pqct study. *Bone* **121**, 139–148. <https://doi.org/10.1016/j.bone.2019.01.013> (2019).
38. Gross, T., Kivell, T. L., Skinner, M. M., Nguyen, N. & Pahr, D. H. A ct-image-based framework for the holistic analysis of cortical and trabecular bone morphology. *Palaeontol. Electron.* **17**, 1–13 (2014).
39. Sylvester, A. D. & Terhune, C. E. Trabecular mapping: Leveraging geometric morphometrics for analyses of trabecular structure. *Am. J. Phys. Anthropol.* **163**, 553–569. <https://doi.org/10.1002/ajpa.23231> (2017).
40. Gunz, P. & Mitteroecker, P. Semilandmarks: A method for quantifying curves and surfaces. *Hystrix Ital. J. Mammal.* **24**, 103. <https://doi.org/10.4404/hystrix-24.1-6292> (2013).
41. Myronenko, A. & Song, X. Point set registration: Coherent point drift. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**, 2262–2275. <https://doi.org/10.1109/tpami.2010.46> (2010).
42. Rueckert, D., Frangi, A. F. & Schnabel, J. A. Automatic construction of 3d statistical deformation models using non-rigid registration. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, Vol. 77–84. [https://doi.org/10.1007/3-540-45468-3\\_10](https://doi.org/10.1007/3-540-45468-3_10) (Springer, 2001).
43. Bijar, A., Rohan, P.-Y., Perrier, P. & Payan, Y. Atlas-based automatic generation of subject-specific finite element tongue meshes. *Ann. Biomed. Eng.* **44**, 16–34. <https://doi.org/10.1007/s10439-015-1497-y> (2016).
44. Pahr, D. H. & Zysset, P. K. From high-resolution ct data to finite element models: Development of an integrated modular framework. *Comput. Methods Biomech. Biomed. Eng.* **12**, 45–57. <https://doi.org/10.1080/10255840802144105> (2009).
45. Pahr, D. H. & Zysset, P. K. Influence of boundary conditions on computed apparent elastic properties of cancellous bone. *Biomech. Model. Mechanobiol.* **7**, 463–476. <https://doi.org/10.1007/s10237-007-0109-7> (2007).
46. Pahr, D. H. & Zysset, P. K. A comparison of enhanced continuum fe with micro fe models of human vertebral bodies. *J. Biomech.* **42**, 455–462. <https://doi.org/10.1016/j.jbiomech.2008.11.028> (2009).
47. Steiner, L., Synek, A. & Pahr, D. H. Femoral strength can be predicted from 2D projections using a 3D statistical deformation and texture model with finite element analysis. *Med. Eng. Phys.* **93**, 72–82. <https://doi.org/10.1016/j.medengphys.2021.05.012> (2021).
48. Markley, F. L., Cheng, Y., Crassidis, J. L. & Oshman, Y. Averaging quaternions. *J. Guid. Control. Dyn.* **30**, 1193–1197. <https://doi.org/10.2514/1.28949> (2007).
49. Taghizadeh, E. *et al.* Biomechanical role of bone anisotropy estimated on clinical CT scans by image registration. *Ann. Biomed. Eng.* **44**, 2505–2517. <https://doi.org/10.1007/s10439-016-1551-4> (2016).
50. Zhu, C., Byrd, R. H., Lu, P. & Nocedal, J. L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. *ACM Trans. Math. Softw.* **23**, 550–560. <https://doi.org/10.1145/279232.279236> (1997).
51. Shewchuk, J. R. What is a good linear element? Interpolation, conditioning, and quality measures. In *Eleventh International Meshing Roundtable (Ithaca, New York)*, 115–126 (2002).
52. Taha, A. A. & Hanbury, A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Med. Imaging* **15**, 29. <https://doi.org/10.1186/s12880-015-0068-x> (2015).
53. MSC Software. *Patran 2012 Reference Manual Part 3: Finite Element Modeling* (2012).
54. Wang, B., Mei, G. & Xu, N. Method for generating high-quality tetrahedral meshes of geological models by utilizing CGAL. *MethodsX* **7**, 101061. <https://doi.org/10.1016/j.mex.2020.101061> (2020).
55. Nichols, T. & Hayasaka, S. Controlling the familywise error rate in functional neuroimaging: A comparative review. *Stat. Methods Med. Res.* **12**, 419–446. <https://doi.org/10.1191/0962280203sm341ra> (2003).
56. Lowekamp, B. C., Chen, D. T., Ibáñez, L. & Blezek, D. The design of SimpleITK. *Front. Neuroinform.* **7**, 45. <https://doi.org/10.3389/fninf.2013.00045> (2013).
57. The CGAL Project. *CGAL User and Reference Manual* 4th edn. (CGAL Editorial Board, 2017).
58. Virtanen, P. *et al.* SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nat. Methods* **17**, 261–272. <https://doi.org/10.1038/s41592-019-0686-2> (2020).

59. Taghizadeh, E., Chandran, V., Reyes, M., Zysset, P. & Büchler, P. Statistical analysis of the inter-individual variations of the bone shape, volume fraction and fabric and their correlations in the proximal femur. *Bone* **103**, 252–261. <https://doi.org/10.1016/j.bone.2017.07.012> (2017).
60. Marangalou, J. H., Ito, K., Taddei, F. & van Rietbergen, B. Inter-individual variability of bone density and morphology distribution in the proximal femur and t12 vertebra. *Bone* **60**, 213–220. <https://doi.org/10.1016/j.bone.2013.12.019> (2014).
61. Joshi, A. A., Leahy, R. M., Badawi, R. D. & Chaudhari, A. J. Registration-based morphometry for shape analysis of the bones of the human wrist. *IEEE Trans. Med. Imaging* **35**, 416–426. <https://doi.org/10.1109/TMI.2015.2476817> (2016).
62. Burkhart, T. A., Andrews, D. M. & Dunning, C. E. Finite element modeling mesh quality, energy balance and validation methods: A review with recommendations associated with the modeling of bone tissue. *J. Biomech.* **46**, 1477–1488. <https://doi.org/10.1016/j.jbiomech.2013.03.022> (2013).
63. Grassi, L. *et al.* Evaluation of the generality and accuracy of a new mesh morphing procedure for the human femur. *Med. Eng. Phys.* **33**, 112–120. <https://doi.org/10.1016/j.medengphy.2010.09.014> (2011).
64. Rueckert, D. & Aljabar, P. Non-rigid registration using free-form deformations. In *Handbook of Biomedical Imaging* (eds Paragios, N. *et al.*) 277–294 (Springer, 2015).
65. Yu, W., Tannast, M. & Zheng, G. Non-rigid free-form 2d–3d registration using a b-spline-based statistical deformation model. *Pattern Recogn.* **63**, 689–699. <https://doi.org/10.1016/j.patcog.2016.09.036> (2017).
66. Rohlfing, T., Brandt, R., Maurer, C. & Menzel, R. Bee brains, B-splines and computational democracy: Generating an average shape atlas. In *Proc. IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA 2001)*, 187–194. <https://doi.org/10.1109/MMBIA.2001.991733> (2001).
67. Spoor, F., Jeffery, N. & Zonneveld, F. *Development, Growth and Evolution: Implications for the Study of the Hominid Skeleton*, Chap. *Imaging Skeletal Growth and Evolution*. *Linnean Society Symposium Series*, 1 edn, 123–162 (Academic Press, 2000).
68. Bishop, P. J., Clemente, C. J., Hocknull, S. A., Barrett, R. S. & Lloyd, D. G. The effects of cracks on the quantification of the cancellous bone fabric tensor in fossil and archaeological specimens: A simulation study. *J. Anat.* **230**, 461–470. <https://doi.org/10.1111/joa.12569> (2016).
69. Worsley, K., Andermann, M., Koulis, T., MacDonald, D. & Evans, A. Detecting changes in nonisotropic images. *Hum. Brain Mapp.* **8**, 98–101 (1999).
70. Adler, R. J., Bartz, K., Kou, S. C. & Monod, A. *Estimating Thresholding Levels for Random Fields via Euler Characteristics*, Vol. 1704, 08562 (2017).

## Acknowledgements

This research has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 819960). The authors are grateful to: Jean-Jacques Hublin (Collège de France, Paris), Christophe Boesch and Uta Schwarz (MPI-EVA, Leipzig), Emmanuel Gilissen and Wim Welden (Royal Museum for Central Africa, Tervuren), Maria Teschler-Nicola and Ronald Muehl (Natural History Museum, Vienna), Jacopo Moggi-Cecchi and Silvia Bortoluzzi (University of Florence, Florence), and Birgit Großkopf (Johann-Friedrich-Blumenbach Institut für Zoologie und Anthropologie der Georg-August-Universität Göttingen, Göttingen), Inbal Livne (Powell-Cotton Museum, Birchington) and Nicole Edmison (Smithsonian Institution National Museum of Natural History NMNH, Washington D.C.), for access to specimens in their care. For technical assistance we would also like to thank Habiba Chirchir (Marshall University, Huntington), Patrick Schoenfeld and David Plotzki (MPI-EVA, Leipzig).

## Author contributions

S.B., C.J.D., M.S., A.S., and D.H.P. designed the study; Software was implemented by S.B.; Data analyses was performed by S.B. and C.J.D.; D.H.P. and A.S. supervised the work; all authors reviewed and approved the manuscript.

## Competing interests

D.H.P. is CEO of Dr. Pahr Ingenieurs e.U., which develops and distributes Medtool. All other authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-09063-6>.

**Correspondence** and requests for materials should be addressed to S.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



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

© The Author(s) 2022