

Experimental Assessment of the Surface Quality of 3D Printed Bones

Rachael M. Carew*^{a,b,c}, Ruth M. Morgan^{a,b}, Carolyn Rando^c

*^a Department of Security and Crime Science, University College London, London
United Kingdom*

^b Centre for the Forensic Sciences, University College London, London United Kingdom

^c Institute of Archaeology University College London, London United Kingdom

rachael.carew.16@ucl.ac.uk *corresponding author

ORCID 000-0003-3451-0394

ruth.morgan@ucl.ac.uk

c.rando@ucl.ac.uk

Word count: 4414

Experimental Assessment of the Surface Quality of 3D Printed Bones

3D printed replicas of human remains are useful tools in courtroom demonstrations. Presently, little published research has investigated the surface quality of printed replicas for use in the presentation of forensic anthropology evidence. In this study, 3D printed replicas of nine human bones were reconstructed from computed tomography (CT) scan data using selective laser sintering (SLS). A three-phased approach assessed: i) the metric accuracy of the 3D prints; ii) the viability of applying age and sex estimation methods (with multiple observers (n=8)); and, iii) the surface quality using a customised scoring method (with multiple observers (n=8)). The results confirmed that the prints in this study were accurate to within 2.0 mm of the original dry bone. Observers were able to confidently assess the gross features of the prints; however fine surface details were not always well represented compared to the dry bones. These findings confirm the applicability of 3D printed replicas for courtroom exhibition of gross features but offer caution against their use when fine detailing is important for evaluative interpretation.

Keywords: forensic science; forensic anthropology; evidence reconstruction; 3D printing

Introduction

3D printed human remains have been used in courts of law in several countries internationally¹. They are often used as visual aids or demonstrative evidence to supplement expert witness testimony by assisting jurors and judges in their understanding of complex technical topics^{1,2}, and there is some evidence that 3D printed replicas could aid in juror comprehension of expert testimony^{3,4}. By providing a physical 3D aid it is anticipated that courts may better understand the evidence presented to them^{1,5,6}, since a physical 3D printed replica provides a tool that observers can touch, feel, rotate, orientate into anatomical position and even use to mimic mechanisms of injury^{1,6,7}. Additionally, taking human remains into a courtroom or showing graphic photographs of human remains can be distressing to jurors^{3,4}, whereas

a 3D print is potentially less emotionally disturbing (compared to a photograph for example)³. A 3D print also has the potential to be less prone to causing bias in favour of the prosecution by providing a minimally graphic alternative^{2, 8-11}. However, more research is needed to fully understand this complex issue of evaluative interpretation by different actors in the forensic science process as well as the impact of different formats of presentation of evidence in court^{1, 3, 6, 12, 13}.

Any 3D printed replicas admitted into a court of law as demonstrative evidence would be subject to local evidence procedures and rules, as detailed by Carew, Morgan and Rando¹. The 3D printed replicas would be required to be accurate representations of the original material. In order to be an accurate representation, it is important to assess the accuracy of the overall shape of a print but also the finer surface details, which may be needed to interpret and assess the print and any pertinent features such as fracture lines. A preliminary investigation into the accuracy of 3D printing human remains showed that 3D printed bones produced from computed tomography (CT) scan data were accurate to the source bone¹. Indeed, 3D printed replicas have been found to be reproducible and robust anatomical models^{5, 14-16}. Previous published research has also investigated the gross quality and applicability of 3D printed anatomical specimens by gauging opinion from medical or anatomy experts, who found prints to be effective and instinctive tools^{6, 14}. However, there is currently little published research that has closely examined the quality of the surfaces of 3D printed bones, particularly within forensic anthropology. The quality of a 3D printed bone replica can be measured by whether a print is suitable for its intended purpose^{17, 18}, thus in this context as a visual courtroom aid, a high-quality print would need to be of sufficient quality whereby all of the features referred to in the testimony of an expert witness, were visible on the print.

Very fine surface details such as fracture lines, porosity and texture of bony features are important aspects in forensic anthropological analysis of human remains¹⁹,²⁰. For example, in the application of age and sex estimation methods, fine details are examined on surfaces for the evaluation of features (such as the pubic symphysis, the auricular surface and points of epiphyseal fusion). Furthermore, it may be important to demonstrate fine osteological features within a courtroom as evidence of various scenarios, such as fractures lines following a trauma⁶, gunshot wounds, toolmarks or porosity as evidence of disease²³. A forensic anthropologist may need to demonstrate such fine features as part of their expert testimony within a courtroom, and a 3D printed replica could be used as a visual aid to support their testimony.

A study by Hjalgrim, Lynnerup, Liversage and Rosenklint²⁴ in 1995, looked at the potential application of 3D printing in forensic anthropology. They found that observations seen on a skull CT reconstruction were confirmed on a stereolithic 3D printed replica, including the visualisation of muscle attachments, mastoid processes, gonial angles, teeth attrition and sutures, and also determined that the replication of finer surface details was limited²⁴. More recently, a study evaluated the surface of 3D printed blunt force cranial injuries, using two crania samples that were captured using surface scanning technologies to identify statistically significant differences between 3D printed bone lesions and their dry bone counterpart²⁵. However, all differences were within 2.0 mm and the 3D printed lesions were found to closely visually resemble the original bone lesions²⁵, it was also determined that 3D printed trauma lesions were sufficient for archiving and teaching purposes, but less reliable for fracture analysis

since the very fine details (such as conclusion of fracture lines and bone porosity) were less accurately represented on the prints using these methods ²⁵.

3D printed test cases exhibiting pathology have also been examined for their morphological surface features after being replicated using a fused deposition modelling (FDM) printer ²⁶. It was found that these features were acceptably replicated given that the diagnostic features were visible on the print ²⁶. Similarly, an evaluation of 3D printing in dental anthropology identified that surface details and distinct morphologies were well produced in 3D printed dental casts, although with some differences observed between printing techniques ²⁷. Indeed, numerous variables are known to impact the quality of 3D replicas, from model development (such as the method of data capture, the resolution of the capture method, the 3D modelling process, the software used and modelling post-processing) and the 3D printing itself (e.g. the 3D printing method used, printer parameters, print post-processing and removal of support structures) ^{1, 12, 25, 28, 29}. In forensic anthropology, the acceptable error range applied in osteometry has previously been stated as ± 2.0 mm ^{1, 30, 31} for macromorphoscopic features. Similarly, a mean absolute error of less than 1.5 mm has been said to be clinically insignificant ³², while Edwards and Rogers ²⁵ used a value of more 1.0 mm for differences to be statistically significant. Thus, it is proposed that an error of ± 2.0 mm is generally considered to be acceptable in osteometry within forensic anthropology, and ± 1.0 mm is ideal ^{1, 25}.

While there is evidence to demonstrate that a final printed replica is an accurate representation overall, it has not yet been established whether the accuracy of finer surface details and the quality of features such as fractures, porosity or texture on

replicated bones is sufficient for evaluative interpretation by a juror in a courtroom, since a juror would need to be able to see any feature that the expert witness was referring to. Therefore, this study sought to assess the surface quality of 3D printed bones, to establish whether the quality of the 3D prints was sufficient to show fine surface details such as porosity and texture. This was investigated using a three-phased approach:

- (1) carrying out a quality control check; to ensure that the prints had not incurred any major dimensional errors, and to identify any major surface similarities or differences
- (2) assessing (subjectively) fine surface details that exist on important features (such as those used in aging and sexing) to establish if they were observable on the prints
- (3) visually scoring the surface quality of the 3D printed bones in a quantitative assessment

Methods and Materials

Nine archaeological human bones were selected for this study from a teaching collection within the Institute of Archaeology at University College London. Each sample presented characteristics such as curved angles, thin cross-sections or surface features that were highly textured and/or important in aging or sexing methods (Table 1).

[Table 1 near here].

Samples were CT scanned using a Toshiba Aquilion ONE Vision Edition (Canon Medical Systems Corporation, Otawara, Japan) helical multidetector CT scanner at University College London Hospital (UCLH). Scanning parameters included

1.0 mm slice thickness at 0.7 mm intervals, 120 kVp, data collection diameter 400 mm, tube current (mA) 50 and field of view (FOV) 220.312 mm for samples 3 to 9. Samples 1 and 2 were scanned previously using 0.5 mm slice thickness at 0.25 mm intervals, 120 kVp, data collection diameter 240 mm, mAs 266 and field of view (FOV) 220.321. Hard/bone reconstruction kernels were used (FC35 for samples 3-9 and FC30 for samples 1-2). The slice thicknesses used to CT scan the bones (0.5 mm and 0.1 mm) were selected by the clinical CT-radiographer based upon their working knowledge and are reflective of that used in clinical CT and post-mortem CT scanning^{33, 34}. This study did not attempt to scan using the best possible resolution or technique, rather using a scan protocol that followed from Carew, Morgan and Rando¹, and was reflective of real-world scenarios.

The CT data was stored in Digital Imaging and Communications in Medicine (DICOM) format and reconstructed using 3D Slicer (Version 4.8.0, 3D Slicer, Brigham Women's Hospital, Boston, MA, US)^{35, 36}. The 3D models were exported as STL files that were post-processed in Blender (Version 2.78, Stichting Blender Foundation, Amsterdam, the Netherlands)³⁷ to remove any background artefacts (such as background noise or the scanner bed that was included in the thresholding) and a smoothing algorithm applied by a factor of 0.5 iterated 10 times using the Smooth Modifier tool.

The 3D prints in this study were produced following the recommendations made by Carew, Morgan and Rando¹ for producing prints from CT scan data, including the use of a hard CT reconstruction filter, the application of a smoothing algorithm to the 3D models and using a selective laser sintering (SLS) printer. Each smoothed STL

model was 3D printed using the SLS powder bed fusion method. SLS uses a laser to selectively fuse (or sinter) a powder-based material into the build object layer-by-layer, the powder is contained within a chamber, and the un-fused powder supports the build as it develops³⁸. An EOSINT P100 (EOS GmbH Electro Optical Systems, Germany) printer was used with a white powder material (PA2200, nylon 12) at a resolution of 100micron layers (0.1mm).

Phase (i)

Linear measurements (detailed in Table 2) were taken as an additional measure to assess the accuracy of the 3D printed replicas. Each measurement was taken twice from the 3D prints and bones on separate occasions by one individual. Measurements were recorded using manual spreading calipers (precise to nearest millimetre) and digital sliding calipers (to nearest 100th millimetre). This level of precision was sufficient within the accuracy threshold of ± 2.0 mm.

[Table 2 near here].

The reliability of the measurement data was assessed using descriptive statistics including within-subject standard deviation (wSD; square root of the average standard deviation). The accuracy of the 3D printed replicas was evaluated by calculating the difference in mean observed data between the print and bone measurements, with ± 2.0 mm being the threshold for accuracy^{1, 30, 31}. A two-tailed paired t-test was performed to identify if there was a statistical difference between the mean print data and the mean bone data. All analysis was performed in Microsoft Excel version 16.9 for Mac (Microsoft, Redmond, WA, US). In addition, a simple visual comparison of the bones, the 3D virtual models and the 3D printed replicas was also carried out to identify any observable similarities or differences between the surfaces.

Phase (ii)

Multiple observers (n=8) performed traditional qualitative age and sex estimation methods on four of the 3D printed replicas (the cranium, the mandible, the os coxa and the clavicle). Traditional methods (described in Table 3) were used to establish if observers were able to sufficiently visualise various features on the 3D prints. Ten methods were tested: four categorised as using gross features, five using fine surface details and one being a combination. Methods were split into their separate parts to identify which features were or were not observable, rather than just comparing the overall score (see Table 3 for scoring details). To compare whether the observers were able to visualise the necessary features on the prints, the results were marked as complete, failed to complete (no data), attempted to complete (incomplete data), or feature unobservable (as appropriate).

[Table 3 near here].

Each observer also recorded how confident they were in the method applied using a Likert five-point categorical scale of 1-5 (1 being very low; 2 low; 3 medium; 4 high; and, 5 very high)^{39, 40}, as in ^{41, 42}. A high confidence score would indicate that the observers were confident that they could see the features on the prints. At this stage, seven of the observers had not previously seen the original dry bones (to reduce extraneous information that could impact their scoring). One observer had seen the original bones which may have influenced the evaluation of the prints in favour of finding high-quality surfaces.

Phase (iii)

The same multiple observers (n=8) scored the nine 3D printed bones using the customised ranked surface quality scoring method from Edwards and Rogers²⁵ (Table 4). This method was developed as an objective visual assessment for scoring the 3D printed bones. Scores were divided by general morphology, detailed morphology, texture and fracture pattern, and ranged from 1-5 for each category¹⁵. The scoring method was adapted for this study to exclude the ‘fracture pattern’ element as this was not relevant with these samples. Thus, in this study observers were scoring out of 15 (not 20). A score of less than 4 in each category (or combined score of less than 12) would indicate that the print was not an accurate representation of the surface of the bone (examples given in Figure 1).

[Table 4 near here].

Results

Phase (i)

Intra-observer differences within print data ranged from -2.6 mm to +2.3 mm (wSD 0.6 mm, absolute mean 0.6 mm), and within bone data from -1.4 mm to 0.9 mm (wSD 0.5 mm, absolute mean 0.4 mm) (Table 5). Differences between the mean print data and mean bone data per bone measurement ranged from -1.0 mm to +1.2 mm, with an overall absolute mean difference of 0.5 mm (Table 5). A paired two-tailed t-test did not find a statistically significant difference between the mean print data and the mean bone data (p-value 0.75).

[Table 5 near here].

The visual comparison of the bones, the 3D virtual models and the 3D printed replicas (Figure 2) led to a clear difference being observed on the trabecular bone from within

the tibia fragment (sample number 6). While this complex micro-structure was present and identifiable, it was not accurately portrayed in the 3D model or subsequent print.

Other features that were noted included that:

- The sutures on the cranium were present and well represented on the 3D print
- The fracture line on the neonate cranial fragment was well represented on the 3D print
- The mandible had the correct morphology and dentition present, there was some loss of detail at the cemento-enamel junction
- The porosity present on the phalanx and rib head was absent on the 3D prints
- The hamate 3D print looked good overall, but perhaps not so well defined at the edges
- The macromorphoscopic features of the os coxa were very well represented on the 3D print

Phase (ii)

All methods that used gross features were successfully carried by all observers and thus were scored as 'complete', with one anomaly (Table 6). More varied results were seen with methods that used fine surface features, with two of these methods (b and e, in Table 6) being scored as being completed by all observers, and method (e) (auricular surface) was a mixture of observers attempting to perform the method (for example, marking some but not all features, or scoring features as observable but following these with a question mark) and failing to record any observable features. [Table 6 near here].

Confidence data from the age and sex methods assessment (Table 7) consistently showed very low confidence with method (e) (auricular surface morphology), method (j) (clavicle epiphyseal fusion) varied with high and low confidence grades given, while other methods were generally confident. Mean confidence per method was predominantly around scores 3 to 4 (Table 7), with variance ranging from 0.5 to 2.0. Mean confidence per observer ranged from 2.9 to 4.3 (except for observer 6) (Table 8),

with variance ranging from 0.4 to 2.4. Low confidence was frequently recorded by observer 6 (experience 0-2 years). Observer 1 recorded higher confidence on average.

[Table 7 near here]

[Table 8 near here]

Phase (iii)

The range of the customised rank scores are shown in Figure 3; observers gave higher scores for general morphology, followed by detailed morphology, and texture received the lowest scores. 3D prints 1, 2, 4, 6 and 8 received all scores >3 by at least one observer. Prints 7 and 9 (the phalanx and hamate) failed to receive a total score >11 by any observer (Table 9).

Observers 3 and 7 gave the lowest average scores of 8 (both with 2-4 years of experience), and observers 5 and 8 marked the highest average scores of 12 (0-2 years and 2-4 years of experience respectively) (Table 9). On average, 3D prints 1 and 2 (cranium and mandible) were accurate with average total scores >11 .

[Table 9 near here].

Discussion

The level of surface detail presented on the nine 3D printed bones in this research was investigated using a three-phased assessment. Through a comparison of linear osteometric measurements taken from the original dry bones and the 3D printed replicas, the phase (i) analysis confirmed that no observable scaling error had occurred during the modelling and printing processes and that the gross morphology of the 3D replicas was accurate. Differences between the print and bone data were all within ± 2.0 mm, with only one data point outside of ± 1.0 mm (cranium NLH with a difference of $+1.2$ mm). The phase (i) results were consistent with previous work using similar

modelling and printing methods ^{1, 16} and maintained that 3D prints were accurate representations of the bones, as with previous studies ^{15, 16, 27}.

Phase (ii) of this study sought to evaluate the extent to which the features necessary in applying age and sex estimation methods were observable and how confident the observers were in their assessments (not how accurately 3D prints can be used to estimate age and sex of individuals). By asking the observers to assign how confident they were in their score for aging and sexing, it was possible to identify which features could be seen on the 3D prints. All gross features were successfully analysed, except in one instance with method (a) where observer 6 failed to complete one feature, which is thought to be due to lack of knowledge of the feature rather than the feature not being present. Additionally, good levels of confidence were generally observed for eight of the ten methods, with low variance in scoring. The phase (ii) assessment used traditional methods that observers were familiar with, although some did have low levels of experience that could have affected their abilities. The higher confidence recorded by observer 1 could be expected given that they had previously seen the dry bones. It is acknowledged that the use of primarily postgraduate student participants in this work is a limitation given the relative lack of experience in handling bones and making assessments using forensic anthropology methods, however the initial insights gained from this study do offer valuable findings that can be developed in future research with professional forensic anthropologists with greater experience.

Three of the observers commented that the clavicle was difficult to assess, due to the nature of the feature or that it was difficult to differentiate osteological surface billowing from step artefacts from the CT scan. Indeed, the quality of the surface of the

clavicle was poor with scan stepping visible on the surface, which was likely due to poor segmentation of the CT data or a result of positioning during the scanning. As noted, there are multiple sources of error that may occur at any point through the modelling process when producing 3D printed bones. To minimise the potential of these effects and for consistency, this study used the segmentation and smoothing protocol recommended by Carew, Morgan and Rando ¹.

Method (d) (pubic symphysis degeneration) was well completed, despite it being a method looking at fine detail. This may be explained by the fact that the pubic symphysis selected was from a young individual and thus the features were prominent and so well manifested in the 3D print. Observers were the least confident and least able to assess the auricular surface of the os coxa for method (e), which may be because this method requires observation of bone porosity (including micro-porosity and granulation) on the os coxa that was found not to be adequately manifested on the print. Indeed, five observers (including observer 1) commented that the auricular surface was especially poorly represented on the 3D print, and in particular that it was not possible to distinguish the texture, porosity or margins of the surface.

Using the customised ranked surface qualitative scoring method from Edwards and Rogers ²⁵ in phase (iii), the cranium and mandible were found to exhibit the highest quality surfaces on average overall and especially so for the general morphology. The results concur with previous studies that have found gross features to be accurately printed ^{24, 26}, but that finer surface details are not always observable ^{24, 25}. It was unsurprising that the phalanx and hamate failed to receive accurate total scores, they were very small bones with a high degree of porosity that was not expected to be fully captured during scanning or printing given the parameters used. The tibia bone with

exposed trabecular bone was scored as accurate (>11) by two of the nine observers. In general, the lack of detail seen on the printed surface features may be due to the CT scan and provides evidence for the limitations of the quality of 3D printed bones from CT scan parameters used in clinical settings, which could be important if applied in post-mortem CT scenarios. CT scans are more appropriate for scanning internal features, the use of micro-CT or perhaps a surface scanning method might produce better results. In this study, the scan resolution determined the resolution of the printed replica, since this was larger than the printer resolution ¹.

It is important to acknowledge that only one sample of each bone was used that represented one stage of growth or degeneration of one individual. Different samples may be expressed as a 3D print to differing degrees given how the morphology of the landscape of these surface features changes with the age of an individual. While the results from this study are reflective of 3D printing, it must be noted that these are specific to both the acquisition method (CT) and the printing technique (SLS) used in this instance. The scoring method ²⁵ worked well as an accessible method that all observers could follow with the exception of one aspect, it was noted that the ‘texture’ portion of the method was less clear, for example if a print exhibited good similarity in porosity but not rugosity, it was difficult to know which score to mark on the scale (Table 4).

The SLS printer and material used were selected for having previously exhibited good aesthetic appearance ¹, although a disadvantage of this powder-based material is that it does have a powdery or grainy feel to the touch. Nevertheless, the replicas produced were sturdy and robust, something that was not seen by Barrera, Silvestro,

Calle-Toro, Scribano, Wood, Henry and Andronikou ⁶ when using an ABS plastic material. The printer material will have an impact on both the aesthetics and feel of a 3D print and is an important factor to consider that may differ depending on the desired use of the replicas ²⁵. The experimental research presented here, provides new data for the assessment of the capabilities of SLS 3D printers to accurately replicate human bones for potential use as visual aids within a courtroom. These results have shown that the surface quality of the prints was not always sufficiently accurate or representative of the original bones. Care must be taken when preparing 3D printed replica bones for courtroom use, each step from acquisition through to printing must be carefully considered and justified, the final replicas must be checked for their accuracy to represent the original material and any features needed (such as fracture lines), and prints also ought to be presented alongside photographs or original scan renderings, as previously recommended ^{1,2}. This study addresses an important aspect of 3D modelling and 3D printing processes in terms of considering the small scale surface features in addition to the macro morphological features and contributes to an important issue in the emerging field of 3D printing in forensic anthropology evidence assessments and crime reconstruction ¹.

Conclusion

These data provide insights into the degree to which it is possible to assess the surface qualities of 3D printed replicas. The findings from this study established that:

- (1) The overall morphology of the 3D printed replicas were metrically accurate.

Macromorphoscopic features, sutures, and to some extent porosity and fracture lines could be observed, but the latter two are particularly dependent on the print resolution.

- (2) It was possible to confidently assess the gross surface features on the 3D printed bones, with less success for fine features.
- (3) The surface quality of fine details was poorly represented on the majority of 3D prints, especially for texture.

Therefore, these findings contribute new insights to inform the use of 3D prints in forensic anthropology evidence reconstruction and continue to support the use of 3D prints for courtroom display of the general morphology of human remains. However, the limitations of the scanning and printing techniques demonstrate that caution is warranted if 3D prints are to be used for demonstrating fine features. Further research is essential to determine if there are approaches that can enhance 3D printed replicas of bones adequately in order to exhibit intricate features that are common in crime reconstruction scenarios, such as those deriving from trauma and taphonomy.

References

1. Carew RM, Morgan RM, Rando C. A preliminary investigation into the accuracy of 3d modeling and 3d printing in forensic anthropology evidence reconstruction. *J Forensic Sci* 2019;64(2): 342-352.
2. Baier W, Warnett JM, Payne M, Williams MA. Introducing 3d printed models as demonstrative evidence at criminal trials. *J Forensic Sci* 2018;63(4): 1298-1302.
3. Blau S, Phillips E, O'Donnell C, Markowsky G. Evaluating the impact of different formats in the presentation of trauma evidence in court: A pilot study. *Journal*.
4. Errickson D, Fawcett H, Thompson TJU, Campbell A. The effect of different imaging techniques for the visualisation of evidence in court on jury comprehension. *Int J Legal Med* 2019.
5. Ebert LC, Thali MJ, Ross S. Getting in touch--3d printing in forensic imaging. *Forensic Sci Int* 2011;211(1-3): e1-6.
6. Barrera CA, Silvestro E, Calle-Toro JS, Scribano PV, Wood JN, Henry MK, Andronikou S. Three-dimensional printed models of the rib cage in children with non-accidental injury as an effective visual-aid tool. *Pediatr Radiol* 2019.
7. Bizzotto N, Tami I, Tami A, Spiegel A, Romani D, Corain M, Adani R, Magnan B. 3d printed models of distal radius fractures. *Injury* 2016;47(4): 976-978.
8. Errickson D, Thompson TJU, Rankin BWJ. The application of 3d visualization of osteological trauma for the courtroom: A critical review. *J. Forensic Radiol. Imaging* 2014;2(3): 132-137.
9. Bright DA, Goodman-Delahunty J. Gruesome evidence and emotion: Anger, blame, and jury decision-making. *Law Hum Behav* 2006;30(2): 183-202.

10. Chaudhary R, Doggalli N, Chandrakant HV, Patil K. Current and evolving applications of three-dimensional printing in forensic odontology: A review. *International Journal of Forensic Odontology* 2018;3(2): 59.
11. Urbanova P, Vojtisek T, Frishons J, Sandor O, Jurda M, Krajsa J. Applying 3d prints to reconstructing postmortem craniofacial features damaged by devastating head injuries. *Leg Med (Tokyo)* 2018;33: 48-52.
12. Carew RM, Errickson D. Imaging in forensic science: Five years on. *Journal of Forensic Radiology and Imaging* 2019;16: 24-33.
13. Weiss D, McLeod-Henning D, Waltke H. Using advanced imaging technologies to enhance autopsy practices. In. *Using advanced imaging technologies to enhance autopsy practices. NIJ Journal*; 3rd July 2018; 2018.
14. Lozano MTU, D'Amato R, Ruggiero A, Manzoor S, Haro FB, Méndez JAJ. A study evaluating the level of satisfaction of the students of health sciences about the use of 3d printed bone models. 2018: 368-372.
15. McMenamin PG, Quayle MR, McHenry CR, Adams JW. The production of anatomical teaching resources using three-dimensional (3d) printing technology. *Anat Sci Educ* 2014;7(6): 479-486.
16. Smith EJ, Anstey JA, Venne G, Ellis RE. Using additive manufacturing in accuracy evaluation of reconstructions from computed tomography. *Proc Inst Mech Eng H* 2013;227(5): 551-559.
17. Christensen A, Rybicki FJ. Maintaining safety and efficacy for 3d printing in medicine. *3D Print Med* 2017;3(1): 1.
18. FDA. Technical considerations for additive manufactured devices - draft. Guidance for industry and food and drug administration staff. In. *Technical considerations for additive manufactured devices - draft. Guidance for industry and food and drug administration staff*; 11th February 2020; 2016.
19. Buckberry JL, Chamberlain AT. Age estimation from the auricular surface of the ilium: A revised method. *Am J Phys Anthropol* 2002;119(3): 231-239.
20. Baez-Molgado S, Bartelink EJ, Jellema LM, Spurlock L, Sholts SB. Classification of pelvic ring fractures in skeletonized human remains. *J Forensic Sci* 2015;60 Suppl 1: S171-176.
21. Villa C, Buckberry J, Cattaneo C, Frohlich B, Lynnerup N. Quantitative analysis of the morphological changes of the pubic symphyseal face and the auricular surface and implications for age at death estimation. *J Forensic Sci* 2015;60(3): 556-565.
22. Marquez-Grant N. An overview of age estimation in forensic anthropology: Perspectives and practical considerations. *Ann Hum Biol* 2015;42(4): 308-322.
23. Ruddy GN, Brough A, Biggs MJ, Robinson C, Lawes SD, Hainsworth SV. The role of micro-computed tomography in forensic investigations. *Forensic Sci Int* 2013;225(1-3): 60-66.
24. Hjalgrim H, Lynnerup N, Liversage M, Rosenklint A. Stereolithography, potential applications in anthropological studies. *Am J Phys Anthropol.* 1995;97: 329-333.
25. Edwards J, Rogers T. The accuracy and applicability of 3d modeling and printing blunt force cranial injuries. *J Forensic Sci* 2018;63(3): 683-691.
26. Cole G, Kingham E, Waldron T. Printing pathology: A case study in presenting pathological human skeletal remains for education and display. *Journal of the Institute of Conservation* 2018;42(1): 18-33.

27. Fiorenza L, Yong R, Ranjitkar S, Hughes T, Quayle M, McMenamin PG, Kaidonis J, Townsend GC, Adams JW. Technical note: The use of 3d printing in dental anthropology collections. *Am J Phys Anthropol* 2018;167(2): 400-406.
28. Guyomarc'h P, Santos F, Dutailly B, Desbarats P, Bou C, Coqueugniot H. Three-dimensional computer-assisted craniometrics: A comparison of the uncertainty in measurement induced by surface reconstruction performed by two computer programs. *Forensic Sci Int* 2012;219(1-3): 221-227.
29. Hodgdon T, Danrad R, Patel MJ, Smith SE, Richardson ML, Ballard DH, Ali S, Trace AP, DeBenedectis CM, Zygmunt ME et al. Logistics of three-dimensional printing: Primer for radiologists. *Acad Radiol* 2018;25(1): 40-51.
30. Morgan B, Ford ALJ, Smith MJ. Standard methods for creating digital skeletal models using structure-from-motion photogrammetry. *Am J Phys Anthropol* 2019;169(1): 152-160.
31. Stull KE, Tise ML, Ali Z, Fowler DR. Accuracy and reliability of measurements obtained from computed tomography 3d volume rendered images. *Forensic Sci Int* 2014;238: 133-140.
32. Fourie Z, Damstra J, Gerrits PO, Ren Y. Evaluation of anthropometric accuracy and reliability using different three-dimensional scanning systems. *Forensic Sci Int* 2011;207(1-3): 127-134.
33. Ruddy GN, Biggs MJP, Brough A, Morgan B, Webster P, Heathcote A, Dolan J, Robinson C. Remote post-mortem radiology reporting in disaster victim identification: Experience gained in the 2017 grenfell tower disaster. *Int J Legal Med* 2019.
34. O'Donnell C, Iino M, Mansharan K, Leditscke J, Woodford N. Contribution of postmortem multidetector ct scanning to identification of the deceased in a mass disaster: Experience gained from the 2009 victorian bushfires. *Forensic Sci Int* 2011;205(1-3): 15-28.
35. 3d slicer (Version 4.8.0) [Computer software]. Brigham Women's Hospital, Boston, MA, US: Slicer Community.
36. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M et al. 3d slicer as an image computing platform for the quantitative imaging network. *Magn Reson Imaging* 2012;30(9): 1323-1341.
37. Blender (Version 2.78) [Computer software]. Amsterdam, Netherlands: Blender Institute.
38. Silva DN, Gerhardt de Oliveira M, Meurer E, Meurer MI, Lopes da Silva JV, Santa-Barbara A. Dimensional error in selective laser sintering and 3d-printing of models for craniomaxillary anatomy reconstruction. *J Craniomaxillofac Surg* 2008;36(8): 443-449.
39. Streiner DL, Norman GR, Cairney J. *Health measurement scales: A practical guide to their development and use*. Oxford University Press.
40. Jamieson S. Likert scales: How to (ab)use them. *Med Educ* 2004;38(12): 1217-1218.
41. Ampanozi G, Zimmermann D, Hatch GM, Ruder TD, Ross S, Flach PM, Thali MJ, Ebert LC. Format preferences of district attorneys for post-mortem medical imaging reports: Understandability, cost effectiveness, and suitability for the courtroom: A questionnaire based study. *Leg Med (Tokyo)* 2012;14(3): 116-120.

42. Chen S, Pan Z, Wu Y, Gu Z, Li M, Liang Z, Zhu H, Yao Y, Shui W, Shen Z et al. The role of three-dimensional printed models of skull in anatomy education: A randomized controlled trail. *Sci Rep* 2017;7(1): 575.
43. Buikstra JE, Ubelaker DH. Standards for data collection from human skeletal remains. *Arkansas Archaeological Survey Research Series No. 44*. 1994.
44. Christensen AM, Passalacqua NV, Bartelink EJ. Age estimation. In: Christensen AM, Passalacqua NV and Bartelink EJ editors. *Age estimation. Forensic anthropology: Current methods and practice*; Oxford: Elsevier Science; 2014.
45. Mahon TJ, Friedling LJ, Gordon GM. Spheno-occipital synchondrosis: Examining the degree of fusion in a south african black skeletal sample. *Forensic Sci Int* 2017;278: 408 e401-408 e405.
46. Cunningham C, Scheuer L, Black S. The pelvic girdle. In. *The pelvic girdle. Developmental juvenile osteology*; Elsevier Science; 2016.
47. Cunningham C, Scheuer L, Black S. The thorax. In. *The thorax. Developmental juvenile osteology*; Elsevier Science; 2016.
48. Christensen AM, Passalacqua NV, Bartelink EJ. Sex estimation. In: Christensen AM, Passalacqua NV and Bartelink EJ editors. *Sex estimation. Forensic anthropology: Current methods and practice*; Oxford: Elsevier Science; 2014.
49. Rowbotham. SK. Anthropological estimation of sex. In: Blau S and Ubelaker DH editors. *Anthropological estimation of sex. Handbook of forensic anthropology and archaeology*; Routledge 2016.
50. Brooks S, Suchey JM. Skeletal age determination based on the os pubis: A comparison of the acsádi-nemeskéri and suchey-brooks methods. *Human Evolution* 1990;5(3): 227-238.
51. Lovejoy CO, Meindl RS, Pryzbeck TR, Mensforth RP. Chronological metamorphosis of the auricular surface of the ilium: A new method for the determination of adult skeletal age at death. *American Journal of Physical Anthropology* 1985;68(1): 15-28.
52. Phenice TW. A newly developed visual method of sexing the os pubis. *American Journal of Physical Anthropology* 1969;30(2): 297-301.

Table 1. Osteological samples and their characteristics.

Bone	Feature
1	Cranium (male) Curved surfaces, important anatomical landmarks, sutures
2	Mandible (male) Unusual shape, important anatomical landmarks, dentition
3	Os coxa (left) Unusual shape, curvature, auricular surface and pubic symphysis
4	Clavicle (left) Unfused epiphysis
5	Neonate frontal (right) Thin bone and curved surface
6	Tibia fragment (proximal; right) Sharp broken edges and exposed trabecular bone
7	Distal phalanx (hand) Small size, porous surface
8	Rib Curvature, thin cross-section and porosity
9	Hamate Unusual shape, small size and porous surfaces

Table 2. Osteometric measurement points on each bone ⁴³.

Bone	Measurement
1 Cranium	a maximum cranial length (g-op): glabella (g) - opisthocranium (op)
	b nasal height (n-ns): nasion (n) to nasospinale (ns)
	c foramen magnum length (ba-o): (ba) to opisthion (o)
2 Mandible	d bicondylar breadth (cdl-cdl): left outer condyle to right outer condyle
	e maximum right ramus height: condyle to gonion (go)
	f chin height (in-gn): infradentale (id) to gnathion (gn)
3 Os Coxa	g Iliac breadth: anterior-superior iliac spine to posterior-superior iliac spine
	h pubic symphyseal face width
	i pubic symphyseal face length
4 Clavicle	j maximum length
	k maximum superior-inferior diameter
5 Neonate frontal	l anterior - posterior length
	m medial - lateral length
6 Tibia fragment	n maximum length (including intercondylar eminence)
	o maximum distal epiphyseal width
	p lateral-medial width at distal end
7 Phalanx	q maximum length
	r proximal width
8 Rib	s maximum length
	t maximum superior-inferior diameter
9 Hamate	u maximum proximal - distal width
	v maximum medial - lateral width

Table 3. Traditional qualitative age and sex estimation methods categorised as using gross or fine features. Aging methods from: several sources ⁴⁴⁻⁴⁷. Sexing methods from: Christensen, Passalacqua and Bartelink ⁴⁸. U = unobservable, M = male, F = female.

Bone	Technique	Method	Scoring	Code	Gross/fine
Cranium and mandible	Sexing	Shape: Buikstra and Ubelaker (1990) ⁴³	1-5 or U	a	Gross
	Aging	Sphenoidal synchondrosis fusion	1-3 or U	b	Fine
Mandible Os coxa	Sexing	Shape ⁴⁹	F-M, or U	c	Gross
	Aging	Pubic symphysis degeneration: Suchey-Brooks (1990) ⁵⁰	Select the features that are present and/or observable	d	Fine
	Aging	Auricular surface morphology: Lovejoy et al (1985) ⁵¹	Select the features that are present and/or observable	e	Fine
	Aging	Iliac crest ossification	Open, partial, complete or U	f	Fine
	Sexing	Shape (excluding sacral features) ²⁹	M or F, or U	g	Gross
	Sexing	Features: Phenice (1969) ⁵²	Present, absent or U	h	Gross & fine

	Sexing	Sciatic notch: Buikstra and Ubelaker (1990) ⁴³	1-5 or U	i	Gross
Clavicle	Aging	Epiphyseal fusion	Tick which phased observed	j	Fine

Table 4. Customised ranked scores used for visual assessment of 3D prints. Adapted from Edwards and Rogers²⁵.

Scores	Description
General Morphology (Basic Shape of Subject)	
1	No similarity in basic morphological shape
2	Little to no visible resemblance in one or more areas
3	Some visible resemblance in one or more areas
4	Some visible resemblance in all areas
5	Clear and definite resemblance in overall shape
Detailed Morphology (Landmarks, Individuated Features)	
1	No similarity in detailed morphological features
2	Little to no visible resemblance in detailed morphological features
3	Some visible resemblance in one or more areas
4	Some visible resemblance in all areas
5	Clear and definite resemblance in detailed morphology
Texture (Porosity, Rugosity, or Smoothness)	
1	No textural resemblance
2	General similarity in rugosity
3	General similarity in porosity
4	General similarity in both porosity and amount of rugosity or smoothness
5	Clear and definite resemblance in all textural similarities

Table 5. Linear measurement data from the 3D printed replicas and the original bones to one decimal place. Mean, standard deviation (SD) and difference given within category for intra-observer error. The difference between the print and bone mean data given as ‘print minus bone difference’.

Bone	Measurement	3D Print			Bone			Print minus bone
		Mean	SD	Difference	Mean	SD	Difference	Difference
1	a	178.0	0.0	0.0	178.0	0.0	0.0	0.0
	b	49.9	0.3	-0.4	48.7	0.6	0.9	1.2
	c	35.6	0.5	0.8	36.0	0.0	0.0	-0.4
2	d	115.9	0.3	0.5	115.2	0.2	-0.3	0.7
	e	62.1	0.1	-0.1	62.3	0.5	-0.8	-0.3
	f	29.8	0.4	-0.6	29.8	0.2	0.3	0.0
3	g	155.0	0.0	0.1	156.0	0.0	0.1	-1.0
	h	12.7	0.5	-0.7	12.6	0.1	0.2	0.2
	i	36.9	1.8	-2.6	37.3	1.0	-1.4	-0.4
4	j	133.8	0.2	-0.3	133.6	0.2	-0.3	0.2
	k	8.1	0.1	-0.1	8.9	0.1	-0.1	-0.8
5	l	62.3	0.2	-0.3	62.1	0.0	-0.1	0.2
	m	49.3	0.6	0.9	49.2	0.0	0.0	0.0
6	n	98.1	0.0	0.0	97.4	0.3	-0.5	0.8
	o	70.0	0.0	0.0	70.4	0.0	0.0	-0.4
	p	22.3	0.1	0.1	22.2	0.1	0.1	0.1
7	q	19.3	0.3	0.4	18.6	0.5	0.7	0.7
	r	10.6	0.7	1.0	11.2	0.5	-0.7	-0.6
8	s	128.4	1.6	2.3	129.4	0.8	-1.1	-0.9
	t	12.4	0.4	0.6	12.9	0.1	0.1	-0.5
9	u	17.9	0.8	-1.2	17.3	0.6	0.9	0.6
	v	10.3	0.1	-0.1	10.5	0.2	-0.3	-0.2

Table 6. Ability of observers (Obs) to complete each method. Tick = completed, cross = failed, A = attempted (some but not all), U = marked some features as unobservable, * = anomaly.

Method	Gross/fine	Obs1	Obs2	Obs3	Obs4	Obs5	Obs6	Obs7	Obs8
a	Gross	✓	✓	✓	✓	✓	✓/U*	✓	✓
b	Fine	✓	✓	✓	✓	✓	✓	✓	✓
c	Gross	✓	✓	✓	✓	✓	✓	✓	✓
d	Fine	✓	✓	✓	✓	✓	✓	✓	✓
e	Fine	×	×	A	A	A	×	A	A
f	Fine	✓	✓	✓	✓	✓	✓/U	✓	✓/U
g	Gross	✓	✓	✓	✓	✓	✓	✓	✓
h	Gross/fine	✓	✓/U	✓	✓/U	✓	✓/U	✓	✓
i	Gross	✓	✓	✓	✓	✓	✓	✓	✓
j	Fine	✓	×	✓	✓	✓	✓	✓	×

Table 7. Descriptive statistics from confidence data for all observers per method (SD, standard deviation; Var, variance)

Method	Mean	Min	Max	SD	Var
a	3.9	3.0	5.0	0.6	0.4
b	3.6	2.0	5.0	1.1	1.1
c	3.5	2.0	5.0	0.9	0.9
d	3.5	2.0	5.0	0.9	0.9
e	1.0	1.0	1.0	0.0	0.0
f	3.1	1.0	4.0	1.2	1.5
g	3.0	2.0	5.0	1.0	1.0
h	3.4	2.0	5.0	1.1	1.1
i	4.3	3.0	5.0	0.7	0.5
j	2.9	1.0	4.0	1.6	2.4

Table 8. Descriptive statistics from confidence data for all methods per observer (SD, standard deviation; Var, variance); with observer experience in years.

	Mean	Min	Max	SD	Var	Experience (years)
Observer_1	4.3	1.0	5.0	1.3	1.6	4+
Observer_2	3.3	1.0	5.0	1.4	2.0	0-2
Observer_3	3.3	1.0	4.0	0.9	0.9	2-4
Observer_4	2.9	1.0	4.0	1.3	1.7	4+
Observer_5	3.9	3.0	5.0	0.8	0.6	0-2
Observer_6	1.9	1.0	3.0	0.7	0.5	0-2
Observer_7	3.4	1.0	4.0	1.0	0.9	2-4
Observer_8	3.0	1.0	4.0	1.1	1.1	2-4

Table 9. Sum of total scores given by each observer (1-8), from customised ranked scoring of prints (1-9)

Print Number	Observer Number								Average
	1	2	3	4	5	6	7	8	
1	14	12	10	11	15	12	10	11	12
2	13	12	10	11	13	13	9	12	12
3	10	10	9	9	12	9	8	11	10
4	11	10	9	9	13	8	9	13	10
5	12	11	8	10	12	8	6	12	10
6	12	11	6	9	13	9	8	11	10
7	9	8	5	7	10	7	6	11	8
8	12	12	6	13	11	10	9	12	11
9	8	11	6	9	10	7	8	11	9
Average	11	11	8	10	12	9	8	12	

Figure 1



Figure 2



