Accuracy of computed radiography in osteometry: a comparison of digital imaging techniques and the effect of magnification

Abstract

Introduction Osteometric data may be obtained using digital imaging techniques, such as post-mortem computed tomography (PMCT) and digital radiography, non-invasively and without ethical objection. Osteometric data from PMCT is known to be as accurate as taking direct measurements. Digital radiography is more accessible and affordable than PMCT but is limited due to the magnification of its subjects.

Objectives To investigate the viability of implementing digital radiographic techniques for measurement of long bones, and to establish whether magnification can be accurately corrected for.

Materials and Methods Twenty hind pig (*Sus scrofa*) legs were imaged using computed radiography (CR) and PMCT, and osteometric data obtained from the digital images and 3D CT volume reconstructions. Direct measurements were taken following maceration. A calibration object was imaged using CR, to provide magnification correction factors.

Results Accuracy was determined by mean absolute error (AE), giving values of 3.3 mm ± 2.5 mm for PMCT (MPR), 2.4 mm ± 1.3 mm for PMCT (3D), 11.1 mm ± 7.4 mm for CR (PA), and 18.3 mm ± 14.5 mm for CR (LAT). PMCT data was more accurate than CR data. Through applying correction for magnification, CR data became closer to the direct measurement data, and stature estimation was substantially altered.

Conclusion Magnification in computed radiography affects osteometric data and resulting stature estimations. Applying correction factors may be a viable option for improving accuracy. For digital radiography to be used reliably in forensic anthropology, further empirical research is needed to validate a magnification correction method.

1 Introduction

Forming an accurate biological profile is vital for attaining positive identification of unknown human remains in forensic investigations. Osteometric data obtained via digital imaging has been shown to be accurate and reliable [1-5]. However, due to the difference between acquisition techniques (such as cost, mobility or magnification), there is a question as to whether digital radiography can be used as reliably as more advanced techniques. Furthermore, there is currently a lack of empirical evidence for the reliability of digital radiographic measurement data in forensic anthropology. This research focuses on the application of stature estimation anthropometry from long bones utilising post-mortem computed tomography (PMCT) and digital radiography (in this case computed radiography, CR); techniques that are non-invasive, effective at visualising skeletal elements and viable for performing virtual anthropological analysis.

Forensic anthropology is primarily concerned with the identification of individuals whose remains are skeletonised, heavily decayed or disrupted, and can assist medico-legal investigations or non-governmental organisations with identifying human remains via assessment of skeletal elements to build a biological profile [6, 7]. A post-mortem (PM) biological profile may include age-at-death, sex, stature and ancestry [7-9], which can be compared with ante-mortem (AM) data to seek identification of a deceased individual [9]. Traditional anthropological techniques require skeletal elements to be fully skeletonised (dry/without flesh) to enable direct measurement using callipers and osteometric boards [7, 9]. If remains are not already skeletonised, examination will encompass invasive maceration procedures that can be hazardous, especially so if the remains have been exposed to a CBRN type incident [10, 11]. Moreover, there are a number of ethical concerns regarding unnecessary maceration procedures, a philosophy that became prominent following a report by Lord Justice Clarke in 2001, regarding the Marchioness-Bowbelle disaster in the UK [12, 13]. Several religious and ethnic groups also oppose further disturbance to human remains (including traditional autopsy examination) [11, 14, 15]. Furthermore, disaster victim identification (DVI) management protocols often advise against further disruption where possible [16, 17]. Given the advances in digital imaging and virtual anthropology the practice can generally be avoided. It is for these reasons that forensic anthropologists should look to emerging digital technologies for assistance with ethical analysis of the deceased wherever possible.

Forensic imaging is a rapidly growing discipline within forensic science, with a marked increase in research published in this area in the last decade [18, 19]. Techniques include digital or computed radiography, fluoroscopy, computed tomography (CT) and magnetic resonance imaging, in addition to surface scanning technologies such as laser scanning, structured light scanning and photogrammetry [7, 14, 20]. PMCT is common practice in several countries for unnatural death investigation [11, 19, 21-23], for example in Leicester, UK [24], Bern, Switzerland [4] and Victoria, Australia [25]. A number of mass fatality incidents are known to have utilised PMCT [7] for examinations, for instance there are two documented cases in the UK [26-28] and one in Australia [25]. In addition to being non-invasive and non-contact, digital imaging generates a permanent virtual copy of a sample, one which can be re-evaluated, sent remotely to experts [7] and also protect the original sample from damage [29, 30]. Limitations to PMCT include: streak artefacts from metallic objects (such as dental amalgam, bullets or fragments) [7, 14]; radiation exposure with living patients, meaning that clinical CT scans are not always ideal for forensic three-dimensional (3D) reconstruction [31]; the cost associated with clinical CT software and hardware; and, the expert user-knowledge required. Nevertheless, the accessibility of PMCT is increasing [7, 30], along with the knowledge that anthropologists need training in digital analysis [12].

Digital radiography is frequently employed in forensic investigations, particularly in the field following mass disasters [11, 23, 32, 33], it can be favoured over PMCT due to the

substantial costs and limitations to mobility with PMCT [11, 34]. Digital radiography can be used for taking long bone measurements for stature estimation, particularly using femur or tibia length [7, 35-37]. However, a major limitation to digital radiography is radiographic distortion. Since digital radiography uses a point source, divergence occurs at the ends of the radiation beams whereby magnification is incurred, which can be exacerbated through difficulty in positioning a subject [2, 37-39]. To circumvent this issue, a user can use a slit beam digital x-ray, calibrate the digital x-ray machine to account for magnification [7, 40], or alternatively calculate a magnification correction factor (CF) through measuring subject tissue thickness to determine the distance of the skeletal elements from the image receptor [39, 41]. To minimise magnification, the object-to-image receptor distance (d) must be as small as possible and the tube-to-image receptor distance (D) as large as possible. A large tube-to-image receptor distance with a supine subject is impracticable and thus a distance of 100 cm is used as a standard [35] in order to keep the magnification constant. However, it should be noted that when magnification needs to be minimised (such as chest x-ray or orthognathic image of the lateral facial bones) the subject is erect, and a tube-to-image receptor distance of 180 cm is used.

Digital imaging and virtual anthropology have been investigated for their accuracy and application in forensic cases by forensic anthropologists, radiographers and pathologists alike [5, 21, 42-47]. The multi-detectors present in PMCT facilitate imaging without distortion, due to the parallel acquisition of multiple x-rays [39], thus allowing accurate osteometric data to be obtained from PMCT scans in an equivalent manner to dry bones using macroscopic methods [3, 21]. For example, a difference of 5 mm (with 95% confidence interval) was reported between PMCT models and direct measurements from the macerated bones in 2008 [1]. Stull, et al. [48] provide 2.0 mm as an accepted error range for anthropological measurements. More recently, a difference of 1-2 mm was found between PMCT and direct calliper measurements on dry bones [4]. Similarly, a comparison of PMCT models from bones scanned pre-maceration, bones scanned post-maceration and direct calliper measurements from the macerated bones gave average percentage differences of 0.6% to 3.7% [48]. Measurements from multi-planar (two-dimensional, 2D) reconstructions of PMCT data have also been shown to be as reliable as direct osteometric measurements [4, 5], but also less reliable than data from 3D models [49]. Furthermore, measurements obtained via PMCT have been found to be more accurate than those from computed or plain film radiography [2, 50-52]. Reid, et al. [2] compared data from PMCT, CR and direct measurements, and found PMCT to be both more accurate than CR and as accurate as direct callipers [2]. Additionally, magnification in CR has been identified as affecting osteometric data collection, with an average overestimation of 7.97% by Reid, et al. [2].

When dealing with human remains that are in poor condition (e.g. burnt or fragmented and reconstructed), the accuracy of stature estimation relies heavily on using reliable methods [35]. Indeed, forensic science has shifted in recent years to focus on using reliable, validated methods that have been generated through empirical research [53, 54]. Recent work conducted by the Forensic Science Regulator in the UK [55] and the Organization of Scientific Area Committees for Forensic Science (OSAC) [56] in the USA, is beginning to develop standards for forensic anthropology, however there is currently a lack of emphasis on digital data. Through creating an evidence-base for digital imaging in forensic anthropology, the methods used can be shown to be reliable and therefore better able to withstand scrutiny, such as court room cross-examination [49, 54].

While radiographic techniques are known to be affected by magnification errors, there is a lack of empirical data investigating these effects at a clinical or end-stage level, such as in anthropological assessments. To investigate the viability of implementing radiographic techniques to aid human identification through obtaining osteometric data, this empirical research compared the use of two imaging techniques, CR and PMCT for the measurement of long bones. The aim of this study was to address the following research questions:

1. Is PMCT more accurate than computed radiography?
2. Can the magnification in CR data be observed?
3. Does any magnification affect stature estimation?
4. Can CR data be accurately corrected for magnification using a correction factor?

2   Materials and Methods

2.1   Experimental Protocol

Fresh juvenile pig (*Sus scrofa*) hind limbs (n=20) were selected (fourteen right and six left), providing n=60 bones (n=20 each for femora, tibiae and fibulae). The distal end of the pig legs had been removed prior to purchase, with the tibiae and fibulae sectioned at the mid-shaft point. Maximum length and mid-shaft diameter measurements were taken using the pseudo ends for the tibiae and fibulae.

Each leg was imaged using computed radiography on a Toshiba Kalare system (Toshiba Medical System 1998) (Canon Medical Systems USA, Inc. 2441 Michelle Drive Tustin, CA 92780) with Konica Minolta CR plates and reader (Konica Minolta, Inc. No.1 Sakura-machi, Hino-shi Tokyo 191-8511, Japan). Posterior-anterior (PA) and lateral (LAT) projections were taken using a constant tube-to-image receptor (plate) distance of 101.6 cm, 80 kV, and 8 mAs (LAT) and 10, 12.6 or 14.2 mAs (PA). Where the legs did not fit onto one CR image, separate radiographs were taken of the proximal and distal ends. Secondly, each leg was scanned using a Toshiba Aquilion 64 slice Computed Tomography unit with slice thickness of 0.5 mm, 100 kV, modulated mA, and Bone Standard and High-Resolution reconstruction algorithms (FC30, FC81). All data was sent electronically to the Picture Archiving and Communication System (PACS) and copied onto CD for remote viewing. The CR data opened automatically with built-in software EXAM-viewer (CoActiv Medical, 900 Ethan Allen Highway, Ridgefield, CT 06877), and the PMCT data was viewed using OsiriX 32-bit, open-source version & plug-ins, version .3.9.2 (Pixmeo SARL, 266 Rue de Bernex, CH-1233 Bernex, Switzerland). PMCT data was measured from the 2D multiplanar reconstruction (MPR) mode (using axial, coronal and sagittal views) and as a virtual 3D volume reconstruction.

Following imaging, the pig legs were macerated using accepted methods [1, 12, 57], which involved simmering and the use of enzymes as described by King and Birch [58]. Firstly, the majority of flesh was removed using a boning knife, taking care not to score the bones. Secondly, the legs were labelled, placed in mesh bags, and the bags submerged in a large pot containing tap water and biological laundry detergent, initially at a temperature of 40-60˚C for 2-3 hours, and subsequently in a cold detergent solution for 12 hours. Finally, the exposed bones were left to air-dry for 24 hours in a controlled environment (19˚C with 57% humidity).

Anatomical landmarks were identified on the pig bones in order to obtain measurements similar to those correspondingly taken by anthropologists on human bones [59, 60] (Figure 1). The measurements used in this study were: bicondylar length (femora); total maximum length (tibiae and fibulae; including intercondylar eminence for tibiae); epicondylar breadth (femora); plateau breadth (tibiae) and diameter at mid-shaft (for all bones). Direct measurements were obtained using standard anthropological measurement methods using an osteometric board and manual sliding callipers [59]. These methods were adapted for the digital images, as in Verhoff, et al. [45] for the CR images (Figure 2) and 2D MPR projections (Figure 3), and as in Brough, et al. [5] for the 3D reconstructions (Figure 4). All measurement data was obtained by one observer and repeated n=3 for CR and osteometric, and n=2 for PMCT measurements.
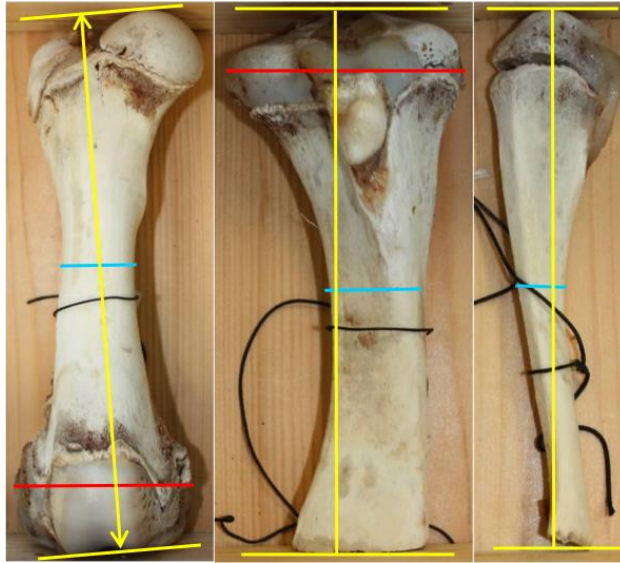
Figure 1 Dry bone femur, tibia and fibula (left to right). Yellow line illustrating length measurement, red line breadth, and blue line diameter.
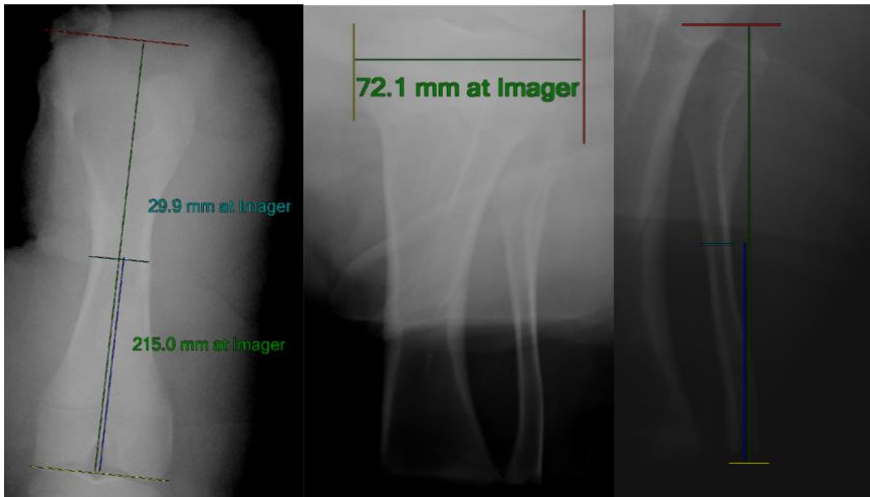

Figure 2 CR images showing femur bicondylar length (left), tibia plateau breadth (middle), fibula maximum length and mid-shaft diameter (right).
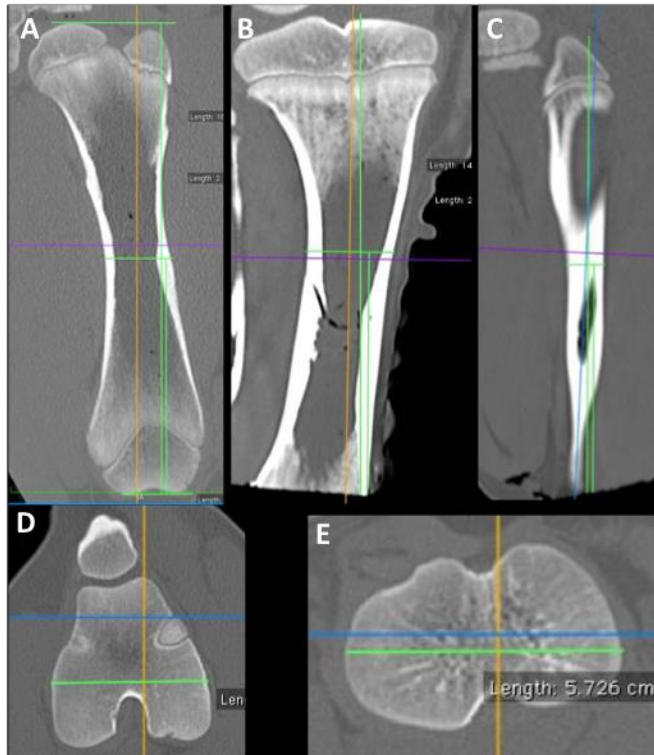
Figure 3 PMCT (MPR) images showing linear measurements (green line). A: femur bicondylar length; B: tibia total maximum length; C: fibula maximum length; D femur epicondylar breadth; E: tibia plateau breadth.
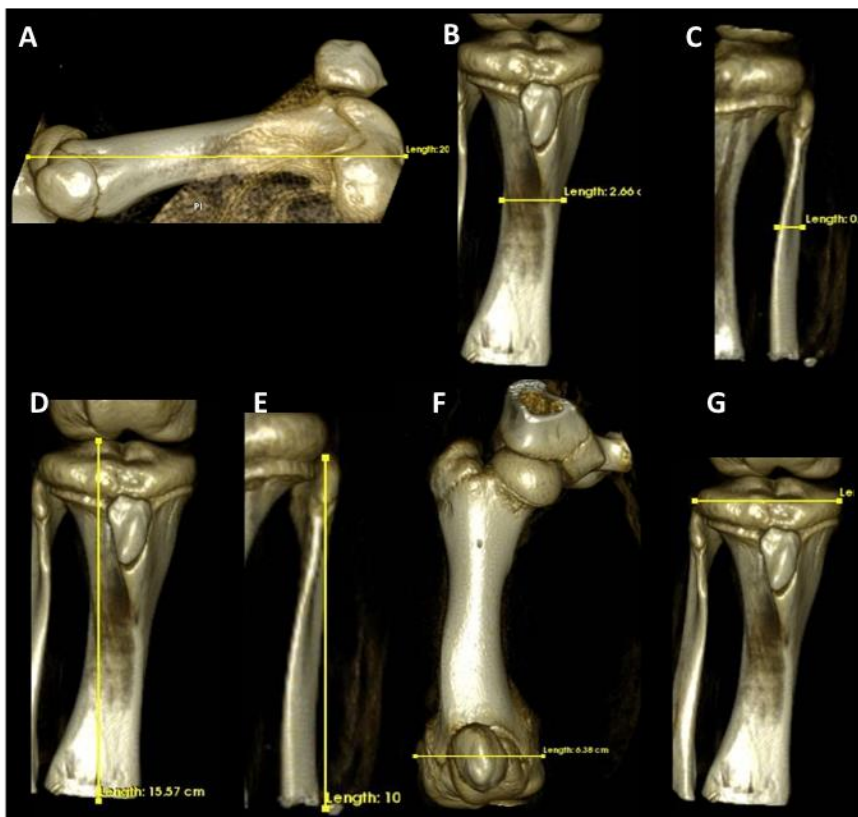

Figure 4 PMCT (3D) images showing linear measurements (yellow line). A: femur bicondylar length; B: tibia mid-shaft diameter; C: fibula mid-shaft diameter; D: tibia total maximum length; E: fibula maximum length; F: femur epicondylar breadth; G: tibia plateau breadth.

2.2 Magnification correction

The mean CR measurement data (measured image dimension, I) were converted into 'corrected measurement' data (true object dimensions, O) using the conversion calculation $O = CF \times I$, whereby $CF = \frac{(D-d)}{D}$ [39] (see Figure 5). To calculate average object-to-image receptor distances (d), tissue thickness measurements were taken from four PMCT images in MPR mode, from both axial and sagittal views and added to the image receptor-to-table distance (9 cm) (Figure 6).

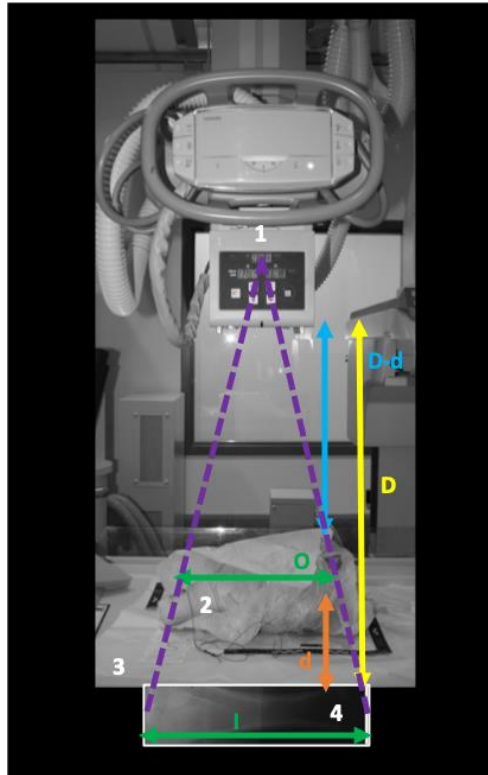Figure 5 Divergence of x-ray beams: 1 tube (focus), 2 object, 3 table, 4 image receptor (film/plate), D tube-to-image receptor distance, D-d tube-to-object distance, d object-image receptor-distance, O true object dimension, I measured image dimension [39].
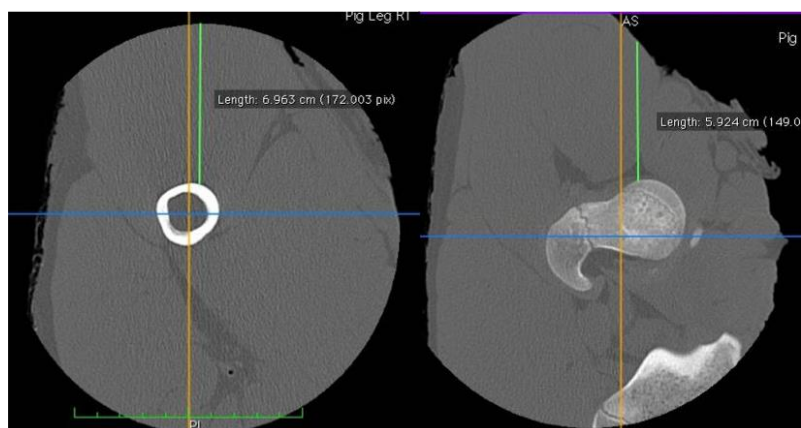

Figure 6 PMCT (MPR) screenshots of pig leg tissue thickness measurement (green line) in the PA position. Mid-shaft measurement (left) and proximal measurement (right).

2.3   Magnification experimental calibration

The following procedure was carried out to experimentally quantify the effect of the CR magnification on the equipment used in this study: 1) a calibration piece of metal 10.0 mm in length was imaged using CR placed directly on top of the CR plate; 2) a block of foam was cut at 1 cm intervals vertically from its base, from 1 cm to 24 cm; 3) the same piece of metal was CR imaged whilst positioned in each cut in the foam, thus at increasing intervals from the plate; 4) the length of the metal piece was measured from the CR images at each interval. The measured image length (I) of the calibration piece of metal was divided by the true object length (O) and all multiplied by 100, to calculate the percentage change (%) and generate magnification factors (MF) that could be used at varying object-to-image receptor distances (d).

Subsequently the same level of magnification used to correct the CR pig images was used to artificially magnify a human femoral length (as stated in Hasegawa, et al. [37]) using the experimentally derived magnification factor (MF) with the equation $measured\ image\ dimension\ (I) = \left(\frac{O}{100}\right) \times (100 + MF)$. Finally, stature ranges for this human femur based on its original and magnified lengths were obtained using standard methods [61] (for white, male, femur).

2.4   Analysis

Analyses were performed using Microsoft Excel version 16.9 for Mac (Microsoft, Redmond, WA, US) and SPSS Statistics version 25 for Mac (IBM, Armonk, NY, US). The spread of the data was assessed using descriptive statistics and reliability investigated using within-subject standard deviation (wSD) with 95% repeatability [1, 5], followed by intraclass correlation coefficients (ICC) [62]. The ICC analysis used a two-way mixed model and absolute agreement [63, 64] with Landis and Koch's strength criteria as a scale to assign agreement: <0 = poor, 0.0-2.0 = slight, 0.21-0.40 = fair, 0.41-0.60 = moderate, 0.61-0.80 = substantial, and 0.81-1.00 = almost perfect agreement [65]. The accuracy of the techniques was investigated using means of the absolute error (AE) and absolute percentage errors (APE) [62], with the mean of the direct measurements providing the reference (true) values. AE was calculated by subtracting the virtually recorded measurement from the direct measurement, and APE calculated by dividing the absolute error by the reference values, all multiplied by 100 [62]. Scatterplots were produced to assess the accuracy of the mean recorded virtual measurement data and 'corrected' CR data, compared against the mean direct measurement data for each bone and measurement type.

3 Results

3.1 Osteometric data

Osteometric measurement data was obtained from the CR images (in PA and LAT positions), PMCT MPR images, PMCT 3D volume reconstructions and directly from the dry bones. CR (LAT) data could not be obtained for the fibula diameter. Reliability assessment found that standard error and standard deviations were generally higher for length measurements, and for those taken with CR (Table 1). Within-subject standard deviation (wSD) and 95% repeatability for each dataset are given in Table 2. wSD values ranged from 0.6 mm (with 95% repeatability of 1.7 mm) to 32.3 mm (with 95% repeatability of 89.6 mm). wSD values were higher for CR data compared with PMCT data, and for measurement type, length data were higher and diameter the lowest.

		Direct			CR (PA)			CR (LAT)			PMCT (MPR)			PMCT (3D)		
		Mean		SD	Mean		SD	Mean		SD	Mean		SD	Mean		SD
		Statistic	SE	Statistic	Statistic	SE	Statistic	Statistic	SE	Statistic	Statistic	SE	Statistic	Statistic	SE	Statistic
Femur	Length	193.9	1.5	6.5	219.4	2.5	11.4	219.3	1.7	7.7	185.5	1.0	4.5	194.8	1.4	6.1
	Breadth	60.7	0.5	2.3	70.2	0.5	2.2	106.8	0.7	3.2	58.6	0.4	1.6	60.1	0.4	1.7
	Diameter	26.2	0.4	1.6	30.9	0.5	2.4	30.6	0.5	2.1	25.7	0.3	1.5	27.3	0.4	1.6
Tibia	Length	137.9	3.8	17.1	130.4	5.4	24.2	148.6	4.2	18.7	134.2	3.7	16.5	137.2	3.7	16.8
	Breadth	61.9	0.5	2.2	69.4	0.7	3.0	83.5	1.5	6.6	58.8	0.5	2.1	60.5	0.4	1.9
	Diameter	26.2	0.5	2.1	32.4	1.0	4.5	31.9	1.2	5.6	26.2	0.4	1.8	27.6	0.5	2.3
Fibula	Length	113.3	4.2	18.7	100.4	5.3	23.8	125.6	3.7	16.8	117.1	3.8	16.9	117.5	3.8	16.9
	Diameter	13.1	0.4	2.0	8.8	0.3	1.3	-	-	-	12.2	0.6	2.7	10.2	0.3	1.2

Table 1 Descriptive statistics for osteometric data from each source (mm). Mean, standard error (SE) of the mean and standard deviation (SD).

		CR (PA)	CR (LAT)	PMCT (MPR)	PMCT (3D)	Direct
Length	wSD	24.4	32.3	4.6	15.6	6.4
	95% Repeatability	67.5	89.6	12.9	43.2	17.9
Breadth	wSD	4.1	10.1	1.6	0.8	0.6
	95% Repeatability	11.3	28.1	4.5	2.1	1.7
Diameter	wSD	0.9	2.3	2.8	0.7	2.1
	95% Repeatability	2.5	6.4	7.8	2.0	5.9

Table 2 Within-subject standard deviation (wSD) (mm) and 95% repeatability (%), per measurement-type for each technique

Intraclass correlation coefficient (ICC) values indicated substantial to almost perfect agreement for each dataset except for CR (PA) breadth, which showed moderate agreement (Table 3). Between-measurement-type, the ICC values showed that length data had the highest agreement (ICC 0.98-1.00), followed by diameter (ICC 0.88-0.99) and breadth data had the lowest agreement (ICC 0.58-0.95). Between-techniques, the direct data showed the highest agreement (ICC 0.89-1.00), followed by CR (LAT) (ICC 0.88-0.98), PMCT (3D) (ICC 0.78-0.99), PMCT (MPR) (ICC 0.61-1.00) and CR (PA) the least (ICC 0.58-0.99).

	CR (PA)		CR (LAT)		PMCT (MPR)		PMCT (3D)		Direct	
	ICC	95% CI	ICC	95% CI	ICC	95% CI	ICC	95% CI	ICC	95% CI
Length	0.99	0.99-1.00	0.98	0.97-0.99	1.00	0.99-1.00	0.99	0.98-0.99	1.00	0.99-1.00
Breadth	0.58	0.40-0.73	0.95	0.92-0.97	0.61	0.36-0.77	0.78	0.63-0.88	0.89	0.82-0.94
Diameter	0.99	0.99-1.00	0.88	0.81-0.93	0.94	0.90-0.97	0.99	0.98-0.99	0.95	0.92-0.97

Table 3 Intraclass correlation coefficients (ICC) with 95% confidence intervals (CI) per measurement-type for each technique

PMCT (3D) data showed the highest accuracy overall with average AE of 2.4 ±1.3 mm, closely followed by PMCT (MPR) data (average AE 3.3 ±2.5 mm) (Table 4). CR datasets showed very low accuracy for both CR (PA) (average AE 11.1 ±7.4 mm) and CR (LAT)

(average AE 18.3 ±14.5 mm). Mean APE values (Table 5) were the least accurate for the CR techniques; for CR (PA) average APE was 19.3 ±13.3%, for CR (LAT) 18.5 ±12.3%, for PMCT (MPR) 5.5 ±4.4%, and for PMCT (3D) average APE was 7.0 ±9.8%.

Pig leg	CR (PA)		CR (LAT)		PMCT (MPR)		PMCT (3D)	
	Mean AE	SD	Mean AE	SD	Mean AE	SD	Mean AE	SD
A	26.0	25.9	22.8	18.6	2.3	2.4	1.6	1.3
B	11.0	8.4	13.6	13.8	3.5	3.9	3.1	3.1
C	6.2	3.2	17.0	16.3	3.5	2.9	1.6	1.0
D	20.5	20.3	19.1	16.8	3.4	3.1	2.4	1.4
E	10.2	8.1	18.2	13.2	0.9	0.7	1.6	0.6
F	16.0	15.6	17.2	15.3	2.4	2.0	1.1	1.0
G	7.5	8.8	19.6	15.2	3.9	3.2	2.6	3.1
H	8.0	9.7	18.4	14.7	3.8	3.1	2.1	1.3
I	11.8	8.3	19.7	13.5	3.8	3.7	2.4	2.5
J	8.7	11.2	17.2	14.5	2.3	2.2	2.0	1.5
K	9.1	8.5	18.8	14.7	3.6	3.0	3.6	3.7
L	9.3	11.1	17.2	14.8	3.2	3.1	1.9	0.8
M	10.1	8.6	14.6	13.5	3.6	2.7	2.5	1.4
N	6.5	5.4	18.6	17.2	4.0	5.0	3.8	5.1
O	10.4	9.7	19.0	14.5	3.7	3.1	4.5	6.3
P	8.1	3.8	18.7	14.2	3.5	3.7	3.0	2.9
Q	9.1	8.5	19.3	12.1	3.5	3.5	3.3	3.4
R	13.0	11.3	18.7	16.2	2.6	2.5	1.5	1.3
S	12.8	12.0	19.8	16.3	5.2	5.7	2.6	4.1
T	8.2	5.6	18.4	15.9	3.8	5.0	1.9	1.4
Average	**11.1**	**7.4**	**18.3**	**14.5**	**3.3**	**2.5**	**2.4**	**1.3**

Table 4 Mean absolute errors (AE) (mm) for each technique

Pig leg	CR (PA)		CR (LAT)		PMCT (MPR)		PMCT (3D)	
	Mean APE	SD	Mean APE	SD	Mean APE	SD	Mean APE	SD
A	35.2	34.0	19.6	15.4	6.4	11.6	3.7	5.7
B	25.0	27.9	16.0	13.3	8.9	17.4	12.7	28.0
C	14.0	12.6	17.5	14.3	6.9	8.5	5.7	9.4
D	29.1	24.2	18.1	14.4	4.6	3.5	5.6	5.1
E	18.0	15.6	19.4	12.3	2.9	5.1	4.8	6.2
F	20.3	14.6	17.3	15.3	6.6	10.8	2.6	3.5
G	13.5	12.3	17.3	13.4	8.5	12.7	7.9	13.7
H	15.2	17.7	16.9	13.9	4.6	2.7	5.3	6.1
I	21.4	22.5	21.2	12.0	5.1	4.1	13.9	29.9
J	13.3	12.7	17.7	13.9	4.3	4.8	7.5	12.2
K	23.5	35.4	20.5	12.8	6.9	6.0	12.2	22.7
L	17.8	21.2	17.5	13.0	4.0	2.1	6.5	9.7
M	25.1	36.2	16.6	12.7	6.1	4.4	9.7	17.4
N	14.8	15.8	18.8	13.8	5.1	4.8	8.2	10.3
O	14.0	8.6	17.7	12.4	4.4	2.0	4.7	4.3
P	17.2	14.1	20.1	13.3	5.3	3.6	8.3	11.3
Q	17.4	17.4	21.8	10.7	6.0	6.0	7.4	6.9
R	18.8	12.8	18.8	13.5	5.2	6.9	2.8	3.0
S	18.3	14.2	19.1	14.3	4.4	3.2	5.4	5.7
T	13.6	8.2	19.0	13.8	4.3	2.5	4.8	5.2
Average	**19.3**	**13.3**	**18.5**	**12.3**	**5.5**	**4.4**	**7.0**	**9.8**

Table 5 Mean absolute percentage errors (APE) (%) and standard deviations for each technique

3.2 Magnification correction

The average tissue thickness for the pig femora in the PA position was 62.6 mm (or 6.26 cm), and in the lateral position 53.2 mm (or 5.32 cm) (Table 6). The image receptor-to-table distance (9 cm) plus tissue thickness, provided object-to-image receptor distances (d) of 15 cm for CR PA and 14 cm for CR LAT (to two significant figures). The mean CR measurement data was corrected for magnification using the conversion calculation to estimate the true object dimensions (O).

Bone ref	Location	PA		Lateral	
		Tissue thickness	Mean	Tissue thickness	Mean
A1	Proximal	62.2		49.2	

1

	Mid-shaft	83.5		60.2	
	Distal	55.4		39.9	
			67.0		49.8
C7	Proximal	59.1		48.9	
	Mid-shaft	69.6		46.4	
	Distal	36.4		34.8	
			55.1		43.4
E13	Proximal	73.6		66.0	
	Mid-shaft	82.3		77.2	
	Distal	60.2		53.4	
			72.0		65.5
G19	Proximal	59.2		49.6	
	Mid-shaft	67.2		65.8	
	Distal	42.2		46.5	
			56.2		54.0
Average			**62.6**		**53.2**

Table 6 Average tissue thickness measurements from four femora taken at three locations (mm). Data measured from whole pig legs in PA and lateral aspects, using PMCT in MPR view

The length scatterplots (Figure 7) indicated that PMCT (3D and MPR) length data was accurate, and CR LAT length values were greater than the direct measurements. CR PA length data was greater than the direct data for femur, but smaller for tibia and fibula. The fibula CR PA length data became less accurate with increasing length of fibula. The breadth data scatterplots (Figure 8) indicated that the PMCT data was accurate for breadth data, while CR LAT data was not accurate (breadth values too high). CR PA breadth data was accurate to the direct measurements after correction for magnification. The diameter data scatterplots (Figure 9) indicated that the PMCT and corrected CR datasets were accurate to the direct data for diameter for the femur and tibia. The fibula diameter datasets were generally inaccurate to the direct data (too small), for fibula CT MPR was the most accurate.

CR PA data was often less accurate after correction (overcorrected) for length and breadth with tibia and fibula, and for diameter of fibula. CR LAT was well corrected for (more accurate after correction) for femur breadth but overcorrected for tibia.
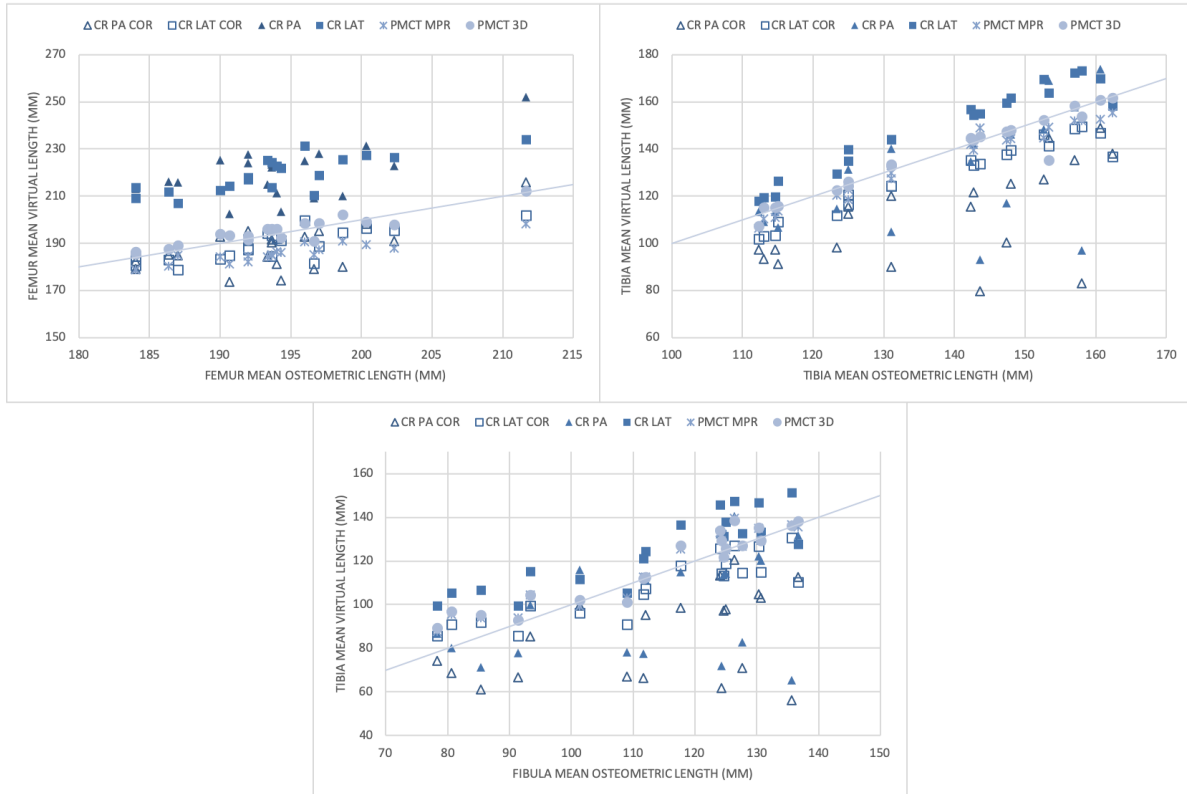
Figure 7 Length virtual data vs. osteometric (direct) data for femur, tibia and fibula. The line represents the ideal true measurement (COR = 'corrected' data).
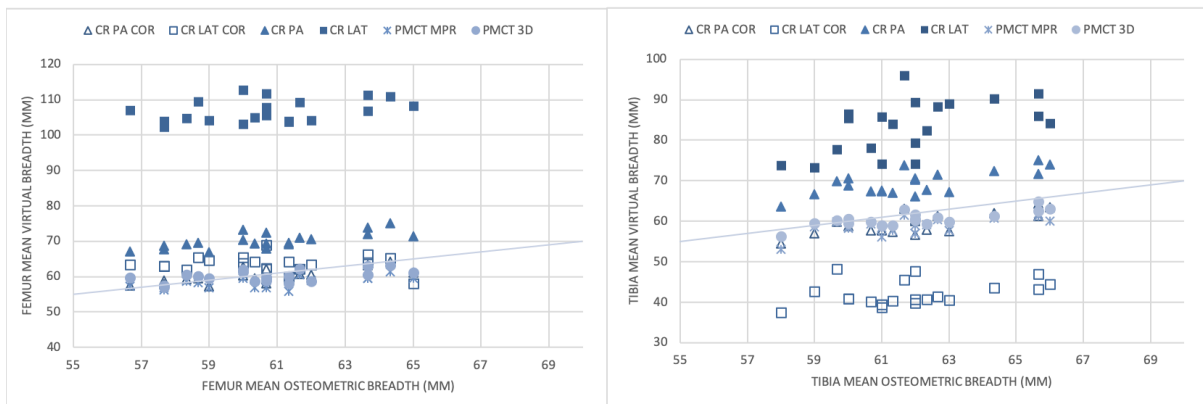

Figure 8 Breadth virtual data vs. osteometric (direct) data for femur and tibia. The line represents the ideal true measurement (COR = 'corrected' data).
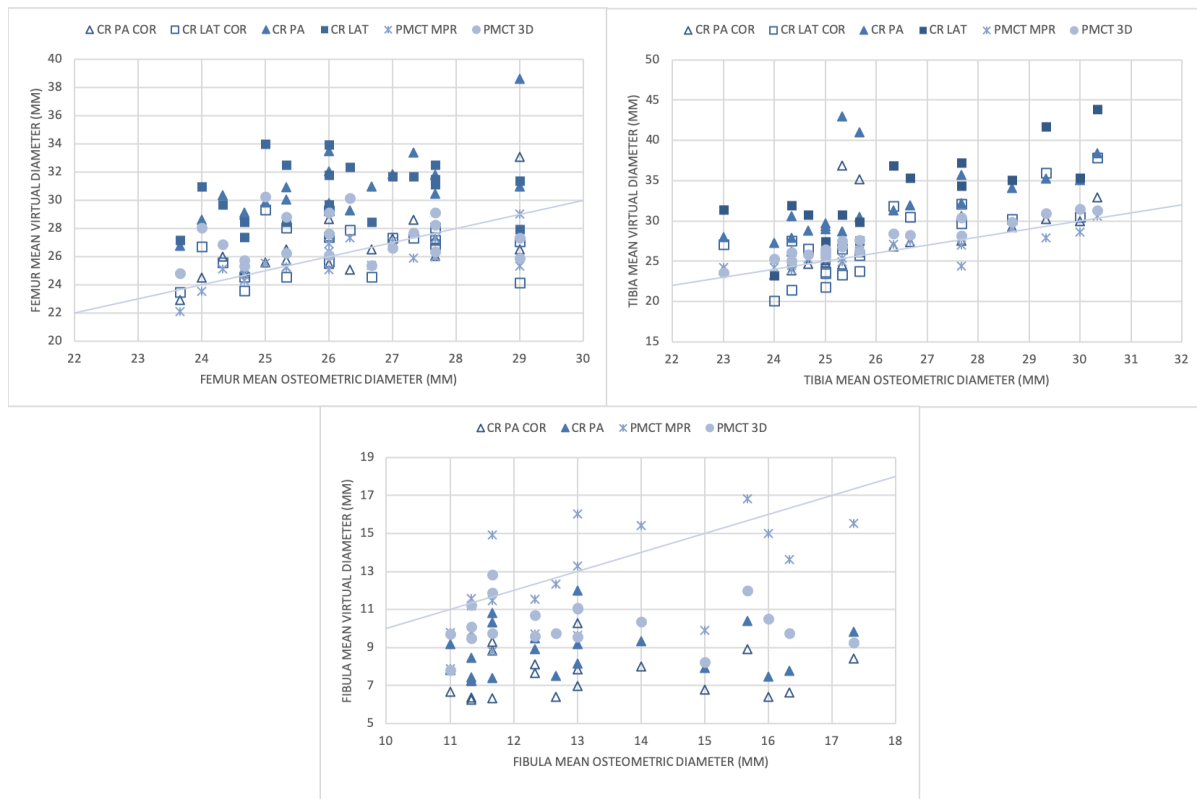
Figure 9 Diameter virtual data vs. osteometric (direct) data for femur, tibia and fibula. The line represents the ideal true measurement (COR = 'corrected' data).

3.3 Magnification experimental calibration

The measurements from the calibration object generated magnification factors (MF) for use at varying object-to-image receptor distances (d) (Table 7). The human femur length was magnified using the CR PA object-to-image receptor distance (15 cm), therefore applying the experimentally derived magnification factor at a level of 16.8% magnification. The calculated stature range for the human femur was 172.9-179.4 cm using the true value, and 192.1-198.6 cm with the magnified value. The application of magnification altered the human femur stature estimation by 19.2 cm.

Object-to-image receptor distance (cm)	Recorded length (mm)	Magnification factor (MF) (%)
0	10.1	1.0
1	10.2	1.7
2	10.3	2.6
3	10.3	3.1
4	10.4	3.9
5	10.5	5.3
6	10.7	6.5
7	10.7	7.4
8	10.9	8.7
9	10.9	9.3
10	11.1	10.6
11	11.2	12.3
12	11.4	14.1
13	11.5	15.0
14	11.6	15.9
15	11.7	16.8
16	11.8	18.4
17	11.8	17.5
18	11.8	17.9
19	12.0	19.7
20	12.2	21.7
21	12.4	23.7
21	12.7	26.6

23	12.8	28.0
24	12.9	28.5

Table 7 Object-to-image receptor distance (O) (cm) and CR measured image dimension (I) (mm) of calibration object, with resulting magnification factor (MF) (percentage change, %).

4   Discussion

4.1   Osteometric data

Advances in digital imaging have fuelled the popularity of PMCT techniques, but the advantages of CR (accessibility, mobility, cost, etc.) keep it firmly in use both clinically and in disaster scenarios for DVI. In order for osteometric measurement data from CR to be used reliably in forensic cases (and especially so if presented in a court of law), the recording of data from CR images must be empirically validated through exploring the reliability and accuracy of the methods used. A comparison of data from the various imaging techniques in this study indicated that, osteometric data obtained using PMCT images were the most reliable and most accurate to those taken directly from the dry bones, agreeing with previous studies [1-5]. The data from the CR imaging was not accurate as it was affected by magnification.

The data acquired using PMCT was slightly greater than the accepted level of error of 2.0 mm [48] and indicated that more accurate data can be obtained from 3D volume renders than those taken using the MPR (2D) mode, which was concurrent with previous findings (49). This is likely due to the fact that when measuring from an MPR image, the user scrolls through the slices to find 'the best fit' i.e. the best slice that contains all the features. Thus, when using a single slice of the bone it is more difficult to visualise all of the necessary anatomical features at once. This potentially indicates that 3D models should be used over 2D slices for osteometric data collection. The CR measurement data was considerably less accurate than data from PMCT (both 3D and MPR), with mean AE values of 11.1 ± 7.4 mm for CR (PA) and 18.3 ± 14.5 mm for CR (LAT), concurrent with the literature [2, 50-52]. These CR values are far greater than the accepted level of error and can be attributed to the magnification incurred during imaging.

There was variation in the accuracy and reliability of measurements taken from different skeletal elements and between different measurement types. The PMCT (3D) data was taken using a 'virtual osteometric board' as in Brough, et al. [5], however it is thought that more accurate linear measurements could be taken by rotating the bone model to ensure that the virtual ruler is in the correct position. The PMCT imaging process was quicker than the CR imaging, as positioning of the leg was not an issue with the PMCT. However, more time was spent manoeuvring the PMCT volume renders into the correct position for measurement, and for this reason the PMCT measurements were only taken twice. While there are several 'standards' available for virtual anthropological analysis or data gathering [5, 18], a complete set of standards for data collection from PMCT images needs to be developed and validated [1, 66]. Furthermore, virtual anthropological analysis can be seen as an unavoidable additional skillset, one that should be taught in higher education and anthropologists should be competent at performing [12, 21].

Measurement uncertainty in this research could have originated from multiple sources [20, 67], such as shrinkage of the bones post-maceration or the PMCT scan resolution used. The resolution used in this study was the highest available at the time and the reconstruction algorithms (filters) used to reconstruct the 3D volume renders were those appropriate for bony material. The quantification of dimensional error as a gauge for accuracy (e.g. AE and APE) as used in this research is consistent with similar studies [5, 48, 49, 62, 67]. Only one observer obtained the data in this study and further observers and/or more experienced observers could aid the reliability of the results. Nevertheless, the overall trends observed

were clear, a comparison of measurements from various techniques in this study and from the literature highlighted a clear order of accuracy, where 3D PMCT is the most accurate technique for taking osteometric data, followed by 2D MPR, followed by CR.

4.2   Magnification correction

A preliminary investigation into the feasibly of correcting for magnification indicated that the CR corrected data were generally more accurate than the original CR recorded data. The correction appeared to overcorrect length measurement data for tibia and fibula but did work well for the femur measurements. The tissue thickness data used was derived from the average of four femora, while tissue thickness will have a natural variation per bone and per individual, this data provided a useful insight into a potential solution to correct for magnification. Further research to explore human leg tissue depth could facilitate the development of a universal correction factor for use in human stature calculations when using CR images, which would be useful for forensic examinations and could also have orthopaedic applications. The findings indicated that when using digital radiography to obtain measurements from fleshed remains, any results should be given with the warning that they would be slightly higher than the actual measurements, concurring with Smith [38]. Additionally, the current validity of anthropological measurement data obtained through digital radiography has been brought into question, and a consistent, empirically validated method is needed to reliably correct for magnification. Presently, CR derived measurements should be used with caution in forensic cases [38].

4.3   Magnification experimental calibration

The effect of CR magnification on measured image dimensions was determined to address the third research aim and assess the impact of magnification on stature estimations. The experimental calibration confirmed and quantified the magnification that was occurring in the CR imaging process and validated the magnification correction procedure. The results demonstrated that the error from CR magnification was substantial enough to affect the stature estimation. The stature range given for the human femur was shifted by almost 20.0 cm, such a change could have implications in forensic human identification investigations, such as leading to a missed identification when comparing AM and PM biological profile data. It is thought that this initial exploration of altered stature estimations effectively illustrates the substantial influence of magnification and demonstrates the necessity for correcting for magnification when generating stature estimations. Although it is noted that other issues inherent within look-up tables are present. Nevertheless, if magnification could be precisely corrected for, then measurements taken using digital radiography could have improved accuracy and be used with confidence.

4.4   Conclusion

Overall, the PMCT measurement data were the closest to the direct bone measurements and more accurate than those from CR. As expected, CR measurements were affected by magnification, with CR resulting in mean APE of almost 20%. The results suggest that the influence of magnification/radiographic distortion on measurement data from digital radiography can be significant enough to alter any resulting stature estimates, but also that it is possible to correct for this distortion and potentially develop a universal correction factor. This study adds weight to the notion that it is not necessary to conduct invasive procedures or autopsies in order to perform anthropological examinations, since accurate information can be obtained from digital imaging techniques. It is advised that practitioners seek PMCT imaging or access to available PMCT images, wherever possible to avoid unnecessary and potentially unethical disruption to remains.

References

1.   Robinson C, Eisma R, Morgan B, Jeffery A, Graham EA, Black S, et al. Anthropological measurement of lower limb and foot bones using multi-detector computed tomography. J Forensic Sci. 2008;53(6):1289-95.
2.   Reid A, Schneider-Kolsky ME, O'Donnell CJ. Comparison of computed radiography and multi-detector computed tomography in the detection of post mortem metacarpal index. Forensic Sci Int. 2008;177(2-3):192-8.
3.   Brough AL, Morgan B, Black S, Adams C, Rutty GN. Postmortem computed tomography age assessment of juvenile dentition: comparison against traditional OPT assessment. Int J Legal Med. 2014;128(4):653-8.
4.   Zech WD, Hatch G, Siegenthaler L, Thali MJ, Losch S. Sex determination from os sacrum by postmortem CT. Forensic Sci Int. 2012;221(1-3):39-43.
5.   Brough AL, Bennett J, Morgan B, Black S, Rutty GN. Anthropological measurement of the juvenile clavicle using multi-detector computed tomography--affirming reliability. J Forensic Sci. 2013;58(4):946-51.
6.   Komar DA, E. BJ. Forensic Anthropology: Contemporary Theory and Practice: Oxford University Pres; 2008.
7.   Viner MD. Overview of Advances in Forensic Radiological Methods of Human Identification. In: Latham KE, Bartelink EJ, Finnegan M, editors. New Perspectives in Forensic Human Skeletal Identification. London: Academic Press; 2017. p. 217-26.
8.   Groen WJM, Márquez-Grant N, Janaway R. Forensic Archaeology: A Global Perspective: John Wiley & Son; 2015.
9.   Tersigni-Tarrant MTA, Shirley NR. Forensic Anthropology: An Introduction: CRC Press; 2012.
10.   Brough AL, Morgan B, Rutty GN. The basics of disaster victim identification. J Forensic Radiol Imaging. 2015;3(1):29-37.
11.   Brough AL, Rutty GN, Black S, Morgan B. Post-mortem computed tomography and 3D imaging: anthropological applications for juvenile remains. Forensic Sci Med Pathol. 2012;8(3):270-9.
12.   Marquez-Grant N. An overview of age estimation in forensic anthropology: perspectives and practical considerations. Ann Hum Biol. 2015;42(4):308-22.
13.   Clarke CLJ. Thames Safety Inquiry: Final Report by Lord Justice Clarke. The Stationery Office: Department of the Environment Transport and the Regions; 2001.
14.   Thali MJ, Dirnhofer R, Vock P. The Virtopsy Approach: 3D Optical and Radiological Scanning and Reconstruction in Forensic Medicine: CRC Pres; 2009.
15.   Franco do Rosário Junior A, Paulo Henrique Couto Souza WC, Patrick Thevissen, Guy Willems, Reinhilde Jacobs. Virtual autopsy in forensic sciences and its applications in the forensic odontology. Rev Odonto Cienc. 2012;27(1):5-9.
16.   INTERPOL. Disaster Victim Identification Guide: INTERPOL; 2014.
17.   Association of Chief Police Officers, Association of Chief Police Officers Scotland. Guidance on Disaster Victim Identification (Decommissioned). National Policing Improvement Agency; 2011.
18.   Lottering N, MacGregor DM, Barry MD, Reynolds MS, Gregory LS. Introducing standardized protocols for anthropological measurement of virtual subadult crania using computed tomography. J Forensic Radiol Imaging. 2014;2(1):34-8.
19.   Baglivo M, Winklhofer S, Hatch GM, Ampanozi G, Thali MJ, Ruder TD. The rise of forensic and post-mortem radiology—Analysis of the literature between the year 2000 and 2011. J Forensic Radiol Imaging. 2013;1(1):3-9.

20. Carew RM, Errickson D. Imaging in forensic science: Five years on. Journal of Forensic Radiology and Imaging. 2019;16:24-33.

21. Franklin D, Swift L, Flavel A. 'Virtual anthropology' and radiographic imaging in the Forensic Medical Sciences. Egypt J Forensic Sci. 2016;6(2):31-43.

22. Weiss D, McLeod-Henning D, Waltke H. Using Advanced Imaging Technologies to Enhance Autopsy Practices. NIJ Journal. 2018;279.

23. Viner MD, Alminyah A, Apostol M, Brough A, Develter W, O'Donnell C, et al. Use of radiography and fluoroscopy in Disaster Victim Identification: Positional statement of the members of the Disaster Victim Identification working group of the International Society of Forensic Radiology and Imaging. J Forensic Radiol Imaging. 2015;3(2):141-5.

24. Rutty GN, Robinson C, Morgan B, Black S, Adams C, Webster P. Fimag: the United Kingdom disaster victim/forensic identification imaging system. J Forensic Sci. 2009;54(6):1438-42.

25. 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.

26. Rutty GN, Robinson CE, BouHaidar R, Jeffery AJ, Morgan B. The role of mobile computed tomography in mass fatality incidents. J Forensic Sci. 2007;52(6):1343-9.

27. University of Leicester Press Office. Scanning technology helps bring answers to Shoreham tragedy 2015 [Available from: https://www2.le.ac.uk/offices/press/press-releases/2015/september/scanning-technology-helps-bring-answers-to-shoreham-tragedy.

28. Vallis J. The role of radiography in disaster victim identification. In: Thompson T, Errickson D, editors. Human Remains: Another Dimension: The Application of Imaging to the Study of Human Remains: Academic Press; 2017. p. 57-69.

29. Villa C, Buckberry J, Cattaneo C, Lynnerup N. Technical note: Reliability of Suchey-Brooks and Buckberry-Chamberlain methods on 3D visualizations from CT and laser scans. Am J Phys Anthropol. 2013;151(1):158-63.

30. Weber GW. Virtual anthropology. Am J Phys Anthropol. 2015;156 Suppl 59:22-42.

31. Wozniak K, Rzepecka-Wozniak E, Moskala A, Pohl J, Latacz K, Dybala B. Weapon identification using antemortem computed tomography with virtual 3D and rapid prototype modeling--a report in a case of blunt force head injury. Forensic Sci Int. 2012;222(1-3):e29-32.

32. Wagner GN. Operation Noble Eagle: Forensic and psychosocial aspects of the Armed Forces Institute of Pathology's response to the September 11 Pentagon attack. Military Medicine 2002;167(9):81-2.

33. Mohd Noor MS, Khoo LS, Zamaliana Alias WZ, Hasmi AH, Ibrahim MA, Mahmood MS. The clandestine multiple graves in Malaysia: The first mass identification operation of human skeletal remains. Forensic Sci Int. 2017;278:410 e1- e9.

34. Brogdon BG, Lichtenstein JE. Forensic Radiology in Historical Perspective. In: Thali MJ, Viner MD, Brogdon BG, editors. Brogdon's Forensic Radiology. Second ed: CRC Press; 2010. p. 9-24.

35. Petrovecki V, Mayer D, Slaus M, Strinovic D, Skavic J. Prediction of stature based on radiographic measurements of cadaver long bones: a study of the Croatian population. J Forensic Sci. 2007;52(3):547-52.

36. Dedouit F, Telmon N, Rosseau H, Crubézy E, Joddre F, Rougé D. Modern Cross-Sectional Imaging in Anthropology. In: Thali MJ, Viner MD, Brogdon BG, editors. Brogdon's Forensic Radiology. Second ed: CRC Press; 2010. p. 107-26.

37. Hasegawa I, Uenishi K, Fukunaga T, Kimura R, Osawa M. Stature estimation formulae from radiographically determined limb bone length in a modern Japanese population*. Leg Med (Tokyo). 2009;11(6):260-6.

38. Smith SL. Stature estimation of 3-10-year-old children from long bone lengths. J Forensic Sci. 2007;52(3):538-46.

39. Brogdon BG. Radiological Identification: Anthropological Parameters. In: Thali MJ, Viner MD, Brogdon BG, editors. Brogdon's Forensic Radiology. Second ed: CRC Press; 2010. p. 85-106.

40.   Viner MD. The Use of Radiology in Mass Fatality Events. In: Adams B, Byrd J, editors. Commingled Human Remains: Methods in Recovery, Analysis, and Identification: Academic Press; 2014 p. 87-122.

41.   Viner MD. Forensic radiography. In: Marquez-Grant N, Roberts J, editors. Forensic Ecology Handbook From Crime Scene to Court: Wiley-Blackwell; 2012. p. 69-84.

42.   Sidler M, Jackowski C, Dirnhofer R, Vock P, Thali M. Use of multislice computed tomography in disaster victim identification--advantages and limitations. Forensic Sci Int. 2007;169(2-3):118-28.

43.   Grassberger M, Gehl A, Puschel K, Turk EE. 3D reconstruction of emergency cranial computed tomography scans as a tool in clinical forensic radiology after survived blunt head trauma--report of two cases. Forensic Sci Int. 2011;207(1-3):e19-23.

44.   Decker SJ, Davy-Jow SL, Ford JM, Hilbelink DR. Virtual determination of sex: metric and nonmetric traits of the adult pelvis from 3D computed tomography models. J Forensic Sci. 2011;56(5):1107-14.

45.   Verhoff MA, Ramsthaler F, Krahahn J, Deml U, Gille RJ, Grabherr S, et al. Digital forensic osteology--possibilities in cooperation with the Virtopsy project. Forensic Sci Int. 2008;174(2-3):152-6.

46.   Dedouit F, Savall F, Mokrane FZ, Rousseau H, Crubezy E, Rouge D, et al. Virtual anthropology and forensic identification using multidetector CT. Br J Radiol. 2014;87(1036):20130468.

47.   Krishan K, Chatterjee PM, Kanchan T, Kaur S, Baryah N, Singh RK. A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. Forensic Sci Int. 2016;261:165 e1-8.

48.   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;238:133-40.

49.   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):342-52.

50.   Akdeniz BG, Oksan T, Kovanlikaya I, I. G. Evaluation of bone height and bone density by computed tomography and panoramic radiography for implant recipient sites. J Oral Implantol. 2000;26(2):114-9.

51.   Cole RJ, Bindra RR, Evanoff BA, Gilula LA, Yamaguchi K, Gelberman RH. Radiographic evaluation of osseous displacement following intra-articular fractures of the distal radius: reliability of plain radiography versus computed tomography. J Hand Surg Am. 1997;22(5):792-800.

52.   Davis GJ, Kapatkin AS, Craig LE, Heins GS, Wortman JA. Comparison of radiography, computed tomography, and magnetic resonance imaging for evaluation of appendicular osteosarcoma in dogs. J Am Vet Med Assoc. 2002;220(8):1171-6.

53.   Koehler JJ, Meixner Jr. JB. An empirical research agenda for the forensic sciences. Crim L & Criminology. 2016;106(1).

54.   Morgan RM. Conceptualising forensic science and forensic reconstruction. Part I: A conceptual model. Sci Justice. 2017;57(6):455-9.

55.   Forensic Science Regulator (UK). Annual Report 2015 [updated 03/10/2017. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/482248/2015_FSR_Annual_Report_v1_0_final.pdf.

56.   Organization of Scientific Area Committees for Forensic Science: Anthropology Subcommittee (US). Research and Development Needs 2017 [updated November 16, 2017. Available from: https://www.nist.gov/topics/forensic-science/organization-scientific-area-committees-osac/osac-research-development-needs#Anthropology.

57.   Simonsen KP, Rasmussen AR, Mathisen P, Petersen H, Borup F. A fast preparation of skeletal materials using enzyme maceration. J Forensic Sci. 2011;56(2):480-4.

58.   King C, Birch W. Assessment of maceration techniques used to remove soft tissue from bone in cut mark analysis. J Forensic Sci. 2015;60(1):124-35.

59.   Buikstra JE, Ubelaker DH. Standards for Data Collection from Human Skeletal Remains. Arkansas Archaeological Survey Research Series No 44. 1994.

60.   Burns KR. Forensic Anthropology Training Manual. 2nd ed. New Jersey: Pearson Education; 2007.

61.   Trotter M, Gleser GC. Estimation of stature fron long bones of american whites and negroes. Am J Phys Anthropol. 1952;10:463–514.

62.   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):127-34.

63.   Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-63.

64.   Landers RN. Computing Intraclass Correlations (ICC) as Estimates of Interrater Reliability in SPSS. The Winnower. 2015;2.

65.   Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977;33(1):159-74

66.   Brough AL, Morgan B, Robinson C, Black S, Cunningham C, Adams C, et al. A minimum data set approach to post-mortem computed tomography reporting for anthropological biological profiling. Forensic Sci Med Pathol. 2014;10(4):504-12.

67.   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):221-7.