Forensic Imaging 24 (2021) 200424

Contents lists available at ScienceDirect

Forensic Imaging

journal homepage: www.sciencedirect.com/journal/forensic-imaging

Can forensic anthropologists accurately detect skeletal trauma using radiological imaging?

Amy Joy Spies [a,*], Maryna Steyn [a], Daniel Nicholas Prince [b], Desiré Brits [a]

[a] Human Variation and Identification Research Unit (HVIRU), School of Anatomical Sciences, Faculty of Health Sciences, University of the Witwatersrand, 7 York Road, Parktown, Johannesburg 2193, South Africa

[b] Department of Diagnostic Radiology, Faculty of Health Sciences, University of the Witwatersrand, 7 York Road, Parktown, Johannesburg 2193, South Africa

ARTICLE INFO

Keywords:
Forensic imaging
forensic anthropology
skeletal trauma
fracture detection ability
Computed Tomography (CT)
X-ray
Lodox®

ABSTRACT

Radiological imaging is often used to aid in forensic anthropological examinations. While forensic anthropologists may need to detect skeletal trauma virtually, they generally have very little to no radiological training. The aim of this study was to assess the ability of forensic anthropologists to accurately detect fractures using radiology. Five piglets were struck with a mallet, post-mortem, and then scanned using CT, X-ray and Lodox®. The number of fractures detected on each modality was recorded by a senior radiology resident (SRR), a novice forensic anthropologist (NFA) with brief radiological training, and more experienced forensic anthropologists (EFA) with limited to no radiological training. The piglets were macerated and the number of fractures present on the dry bone were counted. The percentage of fractures detected by each observer using each method in comparison to the dry bone was calculated. Cohen's kappa revealed that agreement between the NFA and SRR was considered substantial to almost perfect (k=0.77-1.00), while agreement between the EFA and SRR was fair to almost perfect (k=0.21-1.00). However, while the SRR and NFA had comparable sensitivities (0.0-83.7% and 0.0-89.3% respectively), the EFA detected on average 22.0% less fractures than the NFA (0.0-62.4%). This indicates that even brief radiological training increases the sensitivity of fracture detection, and that the ability of forensic anthropologists to detect fractures virtually without training is not sufficient when evaluating cases with suspected trauma. It is therefore suggested that all forensic anthropologists, both novice and more experienced, should receive radiological training in order to accurately detect fractures.

Introduction

Both forensic pathologists and forensic anthropologists are involved in the examination and identification of deceased individuals. Forensic pathologists conduct autopsies in order to determine cause and manner of death and are also commonly relied upon to determine the identification of unknown decedents [1]. Forensic anthropology is concerned with the analysis and identification of decomposed or skeletal remains, which includes the construction of a biological profile and the evaluation of bone pathologies and trauma [2, 3]. In recent times, these disciplines have started to shift toward a more virtual approach, with examinations being conducted less invasively, using various radiographic tools [3].

Forensic odontologists, pathologists and anthropologists also rely on radiological methods to assist in the identification of unknown decedents by comparing ante- and post-mortem records of dentition [4-6], paranasal sinuses [1, 7-9] and unique skeletal variations, pathology or trauma [1, 3, 10, 11]. These methods of identification are particularly useful when dealing with decomposed [12], skeletonized [3] or burnt remains [13, 14], as well as during disaster victim identification [3, 5,

Author contributions: A.J Spies, M. Steyn and D. Brits contributed to the study conception and design; all authors contributed to the acquisition of data, and the revision and final approval of the article; A.J. Spies conducted the analysis and interpretation of data, and drafted the article.

Declarations of interest: none.

Roles of funding sources: funding for this research was provided by the National Research Foundation (NRF) (grant number: 118149) and the NRF/Deutscher Akademischer Austausch Dienst (DAAD) (grant number: 117943). These funding sources however did not have any involvement in study design, collection, analysis and interpretation of data, writing of the report, or in the decision to submit the article for publication.

* **Corresponding author:** Amy Joy Spies; Human Variation and Identification Research Unit (HVIRU), School of Anatomical Sciences, Faculty of Health Sciences, University of the Witwatersrand, 7 York Road, Parktown, Johannesburg 2193, South Africa

E-mail addresses: amy.spies@students.wits.ac.za (A.J. Spies), maryna.steyn@wits.ac.za (M. Steyn), desire.brits@wits.ac.za (D. Brits).

https://doi.org/10.1016/j.fri.2020.200424
Received 14 September 2020; Received in revised form 25 November 2020; Accepted 27 November 2020
Available online 29 November 2020

8], when DNA or fingerprint analysis is not possible or feasible.

Forensic imaging is being used more frequently in the construction of biological profiles during forensic anthropological analyses. Methods for the estimation of sex, age-at-death and living stature have been conducted using magnetic resonance imaging (MRI) scans [15-17], computed tomography (CT) scans [18-20], conventional X-rays [21-23] and low-dose full-body X-ray (Lodox®) scans [24, 25]. Skull radiographs have also been used to aid in the construction of craniofacial approximations [26].

Forensic pathologists increasingly rely on the use of imaging modalities as a complementary tool during traditional autopsy procedures and in some institutions, these modalities are even used to conduct virtual autopsies, or "VIRTOPSIES®" [27, 28]. Conventional autopsies are extremely time-consuming, labour-intensive, invasive, and have many cultural, religious and emotional implications, and the use of radiological imaging to conduct virtual autopsies have therefore garnered much interest [3, 27, 28]. The assessment of both soft tissue and the skeleton can be conducted using radiological methods such as CT and MRI scans, and is particularly important for the detection of skeletal trauma which is regularly assessed during autopsies and forensic anthropological examinations in order to establish the circumstances surrounding death, and to identify victims of abuse [3]. Fractures, which are common in victims of physical abuse and violent crime [29], and the detection of which is critical in these cases, are regularly missed during traditional autopsies, with between 9% and 69% of fractures going unnoticed [30, 31]. Fractures sustained peri-mortem and which have not yet begun to heal and form calluses are frequently undetected during autopsy procedures [30]. In addition, fractures in regions particularly difficult to assess during autopsy are also regularly undetected [31, 32]. While remains may be macerated so that forensic anthropologists can conduct skeletal analyses to detect trauma, this process is time-consuming, expensive and not always feasible [33]. Furthermore, in living individuals, skeletal trauma can only be visualized radiologically. The use of imaging tools aids in the rapid and accurate detection of bone fractures [3], which could assist in identifying both victims and perpetrators of abuse and violent crime.

Ali et al. [1] suggest that forensic pathologists with limited radiological experience can accurately compare ante- and post-mortem records of skull features in order to identify unknown decedents. However, the detection of trauma in the rest of the skeleton may be more challenging and may require the expertise of a radiologist, which is not always practical during forensic analyses where forensic anthropologists who typically have no to limited radiological training and experience, may be required to interpret radiological scans themselves [3]. Although many authors agree that the use of virtual methods to detect skeletal trauma is a critical part of modern forensic anthropology [3, 10, 11, 34], the ability of forensic anthropologists to accurately identify this trauma radiologically is not currently known. The aim of this study was therefore to determine the ability of both novice and experienced forensic anthropologists, in comparison to a radiologist, to accurately detect the number of skeletal fractures on CT, X-ray and Lodox® scans. For this purpose, a pig model was used.

Materials and methods

The Animal Research Ethics Committee at the University of the Witwatersrand (Clearance Certificate No 2019/04/27/O) and the South African Department of Agriculture, Forestry and Fisheries granted ethical permission for the completion of this study.

This paper forms part of a larger study by Spies et al. [35, 36] which investigated the sensitivities of various imaging modalities in detecting skeletal trauma. In order to assess whether forensic anthropologists can accurately detect skeletal trauma, the trunk (including the ribs and vertebrae), the anterior limbs and the posterior limbs of five piglets (5-15 kg each) were each hit between 1 and 10 times with a mallet post-mortem. Piglets were then radiographed. Helical computed tomography (CT) scans were taken at the Wits Donald Gordon Medical Centre in Johannesburg using a 64 Slice Philips Brilliance CT scanner, with parameters of 120 kV, 150 mA, 3.0 mm slice thickness with a 1.5 mm slice increment, 0.391 pitch, and a reconstruction kernel C: 500 W: 1500 on a D filter (bone kernel). Conventional X-rays (58kV, 2.2-2.5mA) were taken at the Central Animal Services at the University of the Witwatersrand, and included ventrodorsal and left and right lateral views of each body region, as well as oblique views of both the left and right ribs. Lodox® scans (80 kV, 160 mA) were taken at the Johannesburg Forensic Pathology Services and included ventrodorsal and right and left lateral views. Lodox® scanners are frequently used both in South Africa and worldwide, producing full-body X-ray images with very low doses of radiation and at very high speeds, and are therefore invaluable for performing initial trauma screening procedures [37]. For a more detailed description of the Lodox® scanner, the reader is referred to work by du Plessis et al. [37].

Four observers, including a fairly novice forensic anthropologist, more experienced forensic anthropologists, and a senior radiology resident read each scan. For the more experienced forensic anthropologists, one read the X-rays and another the CT and Lodox® scans. Two experienced forensic anthropologist observers, both with very limited radiological experience, were used due to the availability of each observer. While all forensic anthropologist observers had limited radiological experience, the novice received brief training in the basics of interpreting radiological images prior to the start of this study. This training amounted to approximately 10 hours, and included two workshops on detecting skeletal fractures and anomalies, as well as training by a radiologist on how to differentiate between anatomical and non-anatomical disruptions in bone cortex, and how to interpret CT scans using multiplanar reconstruction (MPR), maximum intensity projection (MIP) and volume rendering (3-dimensional reconstructions). The novice also had prior experience with sub-adult pig anatomy and osteology, which both the radiologist and experienced forensic anthropologists lacked.

Each observer recorded the exact number and location of fractures present in each body region as detectable on CT, X-ray and Lodox® scans. Fractures were detected based on non-anatomical linear radiolucencies and disruptions of the bone cortex. The skulls were not assessed as some were highly fragmented, and exact fracture numbers could not be determined as with the other body regions. All observers read the scans of each imaging modality using RadiAnt DICOM Viewer, and for CT scans, both 2-dimensional and 3-dimensional images were assessed.

The inter-observer repeatability between each of the observers for detecting the number and location of fractures was assessed using Cohen's kappa, where the level of agreement was considered slight with a kappa score of 0.00 - 0.20, fair with a score of 0.21 - 0.40, moderate with a score of 0.41 - 0.60, substantial with a score of 0.61 - 0.80, and almost perfect with a score >0.81 [38]. Intra-observer repeatability for the primary investigator, who was the novice forensic anthropologist, was also assessed using Cohen's kappa. As the aim of this study was specifically to assess inter-observer repeatability, the intra-observer repeatability for each other observer was not assessed.

The cleaned skeletons, which were considered as the standard of comparison in this study, were also assessed following maceration and skeletal reconstruction of each piglet, and the number of fractures present in the trunk, anterior and posterior limbs was determined. The sensitivity and specificity of each observer for detecting these fractures using each type of scanning method was calculated. Sensitivity is the correct radiological detection of a fracture which is also present on the dry bone, while specificity refers to detecting no fracture radiologically which is also not present osteologically [39]. Sensitivity is represented by the number of true-positive fractures divided by the total number of true-positives and false-negatives, whereas specificity is represented by dividing the number of true-negatives by the combined total of the number of true-negative and false-positive fractures [39]. IBM® SPSS®

version 25.0 was used to perform all statistical analyses.

Results

A total of 279 post-cranial fractures were present on the dry bones, which included 178 rib fractures, 29 vertebral fractures, 31 anterior limb fractures and 41 posterior limb fractures. The overall sensitivities for the senior radiology resident (SRR) were 73.1% for CTs, 51.3% for X-rays, and 45.9% for Lodox® scans. The novice forensic anthropologist (NFA) had sensitivities of 77.4%, 54.5% and 47.7% for CTs, X-rays and Lodox® scans respectively, while the sensitivities for the more experienced forensic anthropologists (EFA) were 53.4%, 28.3% and 31.5% respectively. Overall specificities ranged between 95.1% and 99.2% for all observers. Examples of frequently undetected fractures are presented in Figures 1-3.

The overall detection of fractures was highly repeatable between the SRR and NFA, with kappa values indicating almost perfect agreement between the two observers (CT 0.91; X-ray 0.86; Lodox® 0.89). Agreement between the SRR and the EFA (CT 0.64; X-ray 0.45; Lodox® 0.62) was similar to that between the NFA and EFA (CT 0.62; X-ray 0.48; Lodox® 0.62), and both were considered moderate to substantial agreement. Intra-observer repeatability for the NFA was considered substantial to almost perfect (CT 0.92; X-ray 0.78; Lodox® 0.86).

The kappa values indicating agreement levels between observers for each body region using each imaging modality are presented in Table 1. Repeatability between the SRR and NFA was consistently the highest, with substantial to almost perfect agreement for all regions and all radiological methods (kappa values between 0.77 and 1.00). Agreement between the SRR and EFA, and between the NFA and EFA, were comparable across all body regions and imaging modalities, with kappa values ranging between 0.21 for rib fractures detected using X-rays and 1.00 for vertebral fractures detected using both X-ray and Lodox® scans. In general, CT scans had the highest agreement between observers for all body regions, except the vertebrae, where X-rays and Lodox® had higher repeatabilities, and the anterior limbs for which repeatability was highest using X-rays. Repeatability using Lodox® scans was higher than X-rays for the ribs and posterior limbs, but repeatability using X-rays was higher than Lodox® scans for the anterior limbs. In general, regardless of imaging modality used, the highest agreement between all observers was for fractures of the vertebrae (k=0.70-1.00), while the lowest agreement between all observers was for the ribs (k=0.21-0.87).

The sensitivities for each observer in detecting fractures in each body region using each radiological method is presented in Table 2. The NFA had the highest sensitivities for all body regions using all diagnostic tools, with a range of 0.0% to 89.3%, and were comparable to the sensitivities of the SRR (0.0% to 83.7%). The EFA had the lowest sensitivities for detecting fractures across all regions and using all modalities, ranging between 0.0% and 62.4%. Sensitivities were highest for CT scans for all observers and all body regions (31.0-89.3%), and in general Lodox® scans had the lowest sensitivities (0.0-56.2%). In general, the ribs were the region with the highest sensitivity for all observers using all modalities (25.8-89.3%), while the spine had the lowest sensitivities (0.0-37.9%). For the SRR, NFA and EFA respectively, specificities ranged between 92.2-100.0%, 94.5-100.0% and 92.7-97.9% for CT scans; 83.0-100.0%, 93.8-100.0% and 89.8-98.3% for X-rays; and 97.9-100.0%, 94.0-100.0% and 91.5-100.0% for Lodox® scans.

Discussion

Radiological imaging is often used during forensic anthropological analyses, which includes the assessment of skeletal trauma. Two papers from the same larger study compared the sensitivities of CT, X-ray and Lodox® scans in detecting fractures of the skull, trunk, forelimbs and hindlimbs of piglets subjected to blunt force trauma [35, 36]. The results of these papers showed that many fractures were missed using all imaging modalities, but that CTs consistently had the highest sensitivities

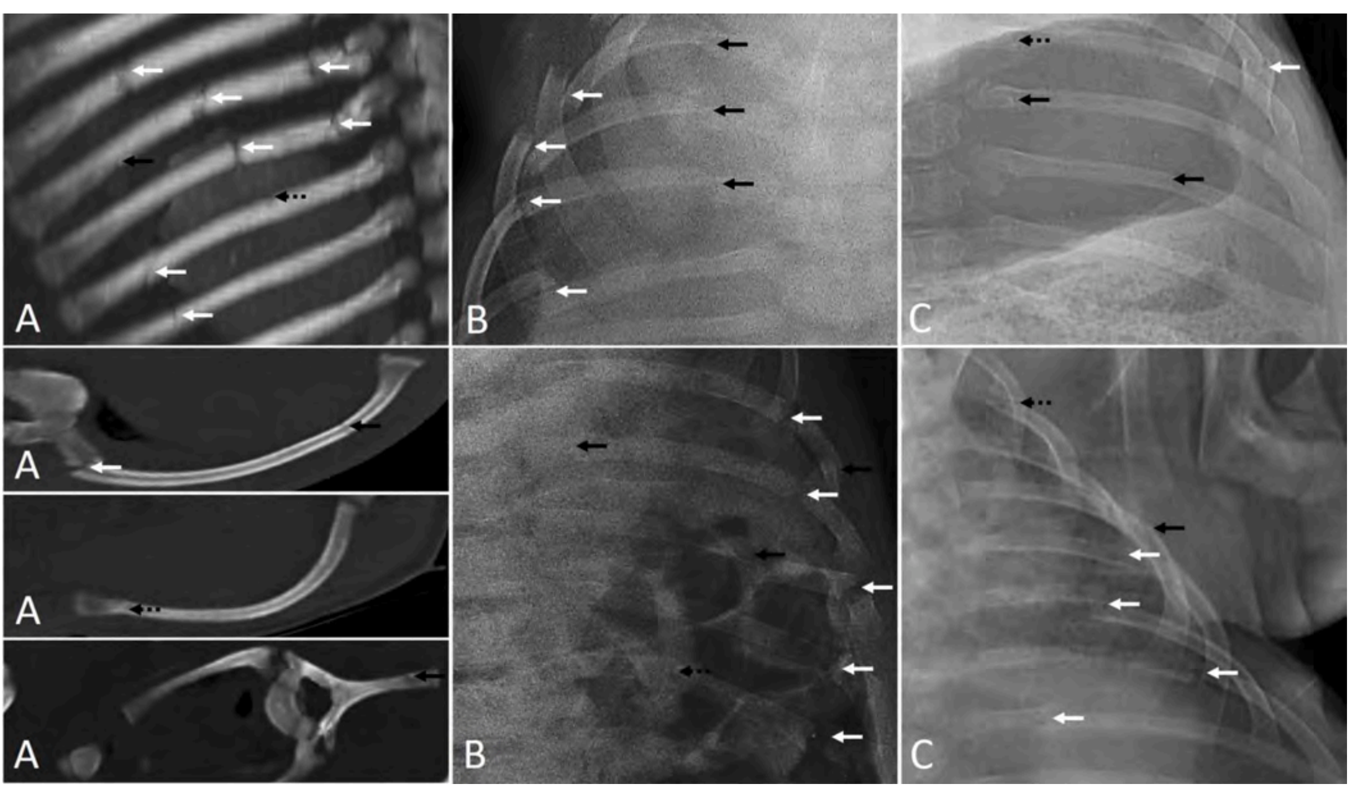

Fig. 1. Frequently unidentified fractures of the ribs and vertebrae as shown with CT (A), X-ray (B) and Lodox® (C) scans. Solid white arrows indicate fractures identified by all observers; solid black arrows indicate fractures identified only by the SRR and NFA; dashed black arrows indicate fractures identified only by the NFA.

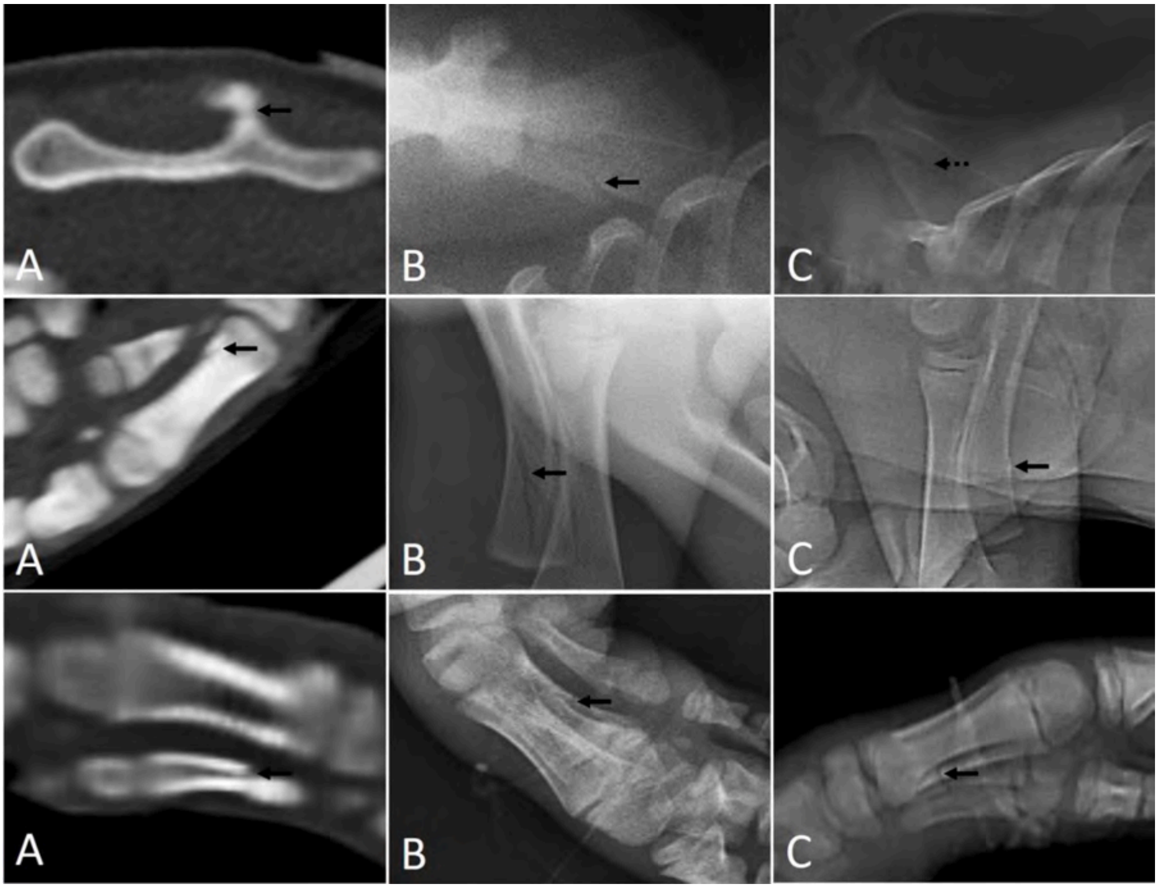

Fig. 2. Frequently unidentified fractures of the anterior limb as shown by CT (A) (top: scapula; middle: metacarpal; bottom: metacarpal), X-ray (B) (top: scapula; middle: ulna; bottom: metacarpal) and Lodox® (C) (top: scapula; middle: ulna; bottom: metacarpal). Solid black arrows indicate fractures identified only by the SRR and NFA; dashed black arrows indicate fractures identified only by the NFA. A lack of white arrows indicates that none of the fractures were identified by all thee observers.

while plain radiographs were insufficient in detecting fractures in these cases. The aim of the present study was therefore not to compare the sensitivities of the imaging modalities themselves or to assess how efficient they are, but to compare the ability of different observers, including forensic anthropologists, to detect fractures using these methods. While the use of a pig model is challenging in terms of the anatomy and osteology, and may have reduced the overall detection of fractures, the senior radiology resident used in the present study was as unfamiliar with pig anatomy as most forensic anthropologists, and therefore the results can still be considered accurate.

The number of fractures detected by the NFA was highly comparable to those detected by the SRR for all body regions using all radiological methods (k=0.77-1.00). The level of agreement between these two observers was substantial to almost perfect [38], indicating that the NFA was able to identify fractures visible radiologically at the same level as the SRR. The number of fractures detected by the more experienced forensic anthropologists (EFA) was less comparable with those identified by both the SRR (k=0.21-1.00) and the NFA (k=0.28-1.00). However, agreement between the EFA and SRR was still considered fair to almost perfect. The much higher repeatability of the NFA may be because this observer received brief radiological training prior to the start of this study, and therefore had slightly more experience with reading scans in comparison to the EFA. The NFA also had the highest sensitivities for detecting fractures in all body regions using all imaging modalities (0.0-89.3%). Sensitivities of the NFA were on average 3.0% higher than that of the SRR (0.0-83.7%) and 22.0% higher than that of the EFA (0.0-62.4%). The higher sensitivities of the NFA are, when compared with the EFA, likely due to the fact that the NFA had a slightly higher level of radiological experience. In addition, the NFA had a much more detailed knowledge of piglet osteology and anatomy, which explains why the NFA had higher sensitivities than both the SRR and EFA who are less familiar with pig anatomy. It is therefore likely that the sensitivities for detecting fractures by both the SRR and the EFA may increase when reading scans of human infants and children, with which all observers are much more familiar.

These results, which indicate that observer experience does indeed affect the number of fractures detected using radiological scans, are similar to those of Hong et al. [40] who reported an increase in sensitivity of pediatric rib fracture detection from 29% to 46% for X-rays and 52% to 85% for CT scans when a more senior radiologist read the scans in comparison to a less experienced observer. However, Shelmerdine et al. [41] found that observer experience was not a significant predictor of sensitivity when detecting rib fractures in children.

Both inter-observer repeatability as well as fracture detection sensitivity was generally highest for CT scans, for all observers and all body regions. This indicates that CTs are not only more sensitive than both X-ray and Lodox® scans in detecting fractures, but that CT scans are also much easier to read and interpret regardless of observer experience or level of training, as was seen by Spies et al. [35, 36]. While the superimposition of bones in X-ray and Lodox® scans may cause the detection of fractures to be much more challenging, the three-dimensional renderings of each bone available with CT scans contributes to fractures being detected much more easily by observers of all levels of experience. Hong et al. [40] suggest that the use of CT scans in addition to X-rays may only be of benefit for observers with less radiological experience, as more senior observers may have a high

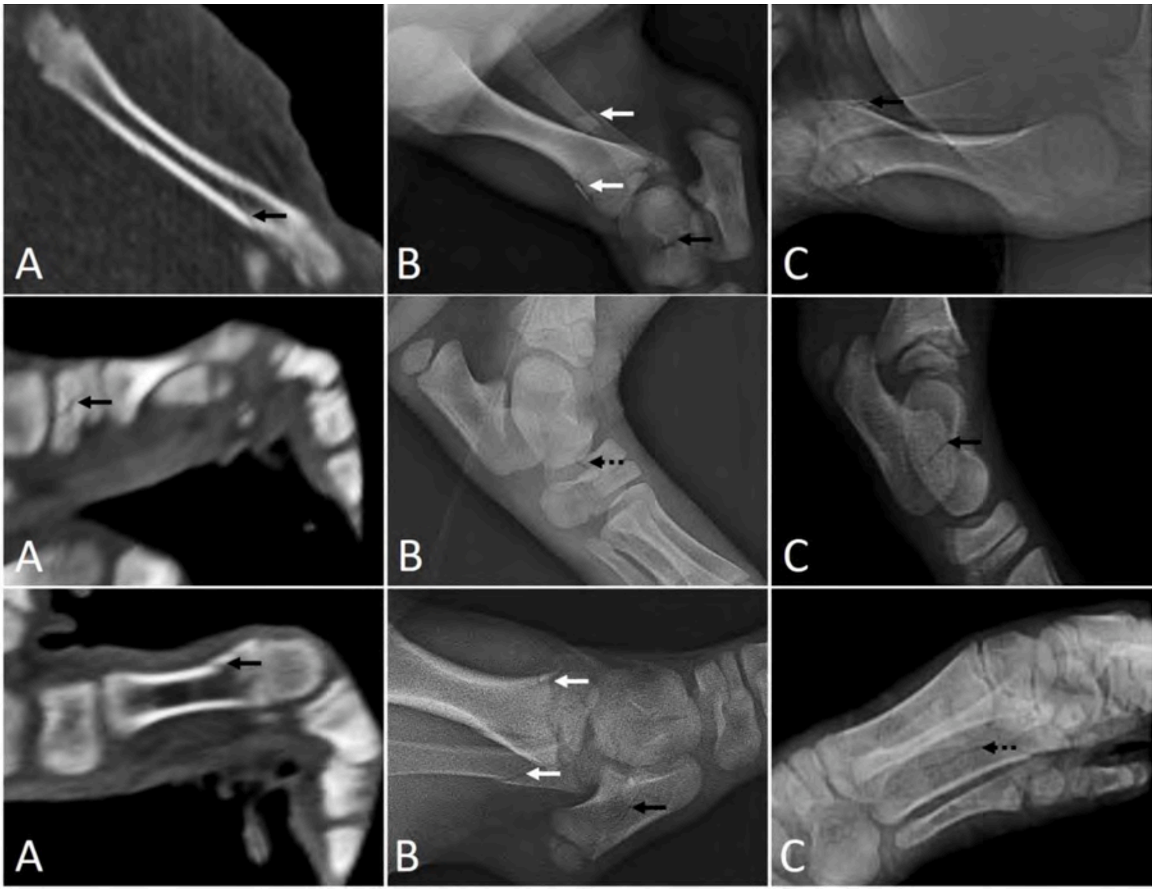

Fig. 3. Frequently unidentified fractures of the posterior limb as shown by CT (A) (top: fibula; middle: tarsal; bottom: metatarsal), X-ray (B) (top: tibia, fibula and talus; middle: tarsal; bottom: tibia, fibula and calcaneus) and Lodox® (C) (top: fibula; middle: talus; bottom: metatarsal). Solid white arrows indicate fractures identified by all observers; solid black arrows indicate fractures identified only by the SRR and NFA; dashed black arrows indicate fractures identified only by the NFA.

Table 1
Kappa values indicating level of agreement between the number of fractures detected by the senior radiology resident (SRR), novice forensic anthropologist (NFA) and experienced forensic anthropologist (EFA) in each body region using each imaging modality.

	Ribs			**Vertebrae**			**Anterior limbs**			**Posterior limbs**		
	CT	X-ray	Lodox	CT	X-ray	Lodox	CT	X-ray	Lodox	CT	X-ray	Lodox
SRR-NFA	0.87	0.77	0.83	0.97	1.00	1.00	0.86	1.00	0.95	0.97	0.90	0.88
SRR-EFA	0.54	0.21	0.53	0.70	1.00	1.00	0.59	0.69	0.53	0.67	0.56	0.66
NFA-EFA	0.52	0.28	0.54	0.72	1.00	1.00	0.60	0.69	0.49	0.64	0.57	0.63

Table 2
Sensitivities (%) of the senior radiology resident (SRR), novice forensic anthropologist (NFA) and experienced forensic anthropologist (EFA) in detecting fractures in each body region using each imaging modality.

	Ribs			**Vertebrae**			**Anterior limbs**			**Posterior limbs**		
	CT	X-ray	Lodox	CT	X-ray	Lodox	CT	X-ray	Lodox	CT	X-ray	Lodox
SRR	83.7	60.7	55.6	37.9	0.0	0.0	64.5	48.4	45.2	58.5	48.8	36.6
NFA	89.3	64.6	56.2	37.9	0.0	0.0	67.8	48.4	48.4	61.0	53.7	43.9
EFA	62.4	25.8	36.5	31.0	0.0	0.0	41.9	38.7	29.0	39.0	26.8	34.1

enough diagnostic accuracy using X-rays alone. However, the results of this study, as well as the study by Shelmerdine et al. [41], show that all observers have a higher diagnostic accuracy using CT scans than when using plain radiographs. In addition to higher sensitivities with CTs, scanning times are also much quicker than conventional X-rays, and only one scan needs to be taken in comparison to three scans for Lodox®, and between 10 and 15 images for X-rays, depending on the size of the image cassettes. Furthermore, there is the added benefit of using 3-dimensional reconstructions, and even 3-dimensional printed models of the CT scans, which may be used as evidence in court during criminal trials [42]. Therefore, CTs are preferred over X-ray or Lodox® scans regardless of observer experience.

Repeatability between all observers was lowest for the ribs using all imaging modalities. This is likely due to difficulties in counting the ribs, which range between 14 and 16 in pigs, as well as superimposition of the ribs in X-ray and Lodox® scans. In addition, abdominal gas often obscured the lower ribs, making fracture detection challenging. Sensitivities, however, were lowest for the spine, likely due to superimposition of the vertebrae in X-ray and Lodox® scans which makes it challenging to identify fractures of the transverse and spinous processes. Sensitivities were also relatively low for the limbs, which may be due to difficulty in detecting long bone metaphyseal fractures, as well as superimposition of some bones. While long bones such as the humerus and femur were easily visualised, smaller bones like the metacarpals and metatarsals were highly superimposed, making fractures difficult to detect. This is evident as 19.0-56.3% and 26.1-63.6% of all limb fractures undetected by observers using X-rays and Lodox® respectively, were those of the metacarpals/metatarsals. Again, increased radiological experience and training, as well as an increase in knowledge of piglet osteology may result in both improved inter-observer repeatability and sensitivity for these body regions. Furthermore, in human infants and children, where the bones of the hands and feet are not superimposed, sensitivities for detecting fractures in these regions are expected to be higher for all observers.

Specificities were extremely high for all observers (95.1-99.2%), indicating very few false-positive fracture identifications. This is significant since it suggests that all observers, irrespective of level of training, are not likely to incorrectly identify victims of abuse. For all observers, specificities were comparable between CT, X-ray and Lodox® scans, indicating that the level of observer experience does not affect the likelihood of making false-positive identifications using any imaging method. Hong et al. [40], however, suggest that limited radiological experience may result in lower specificities for CTs in comparison to X-rays, and that with increased experience and training, the risk of false-positive identifications can be reduced. The results of the present study, however, contradict this.

The results of both inter-observer repeatability and fracture detection sensitivity in the present study show that an increase in observer experience results in an increase in the number of fractures detected using all imaging methods, as expected. The NFA received brief training in the basics of radiological imaging interpretation prior to the start of this study, while the EFA had little to no radiological experience. This training resulted in the NFA having a sensitivity for fracture detection comparable to that of the SRR, and which was on average 22.0% higher than the EFA. This indicates that even very brief radiological training and experience can result in a drastic increase in sensitivity, and that the ability of forensic anthropologists without this training to detect fractures is limited and not sufficient when evaluating cases with skeletal trauma such as in cases of physical child abuse.

A potential limitation of this study is the fact that a CT slice thickness of 3.0 mm with a slice increment of 1.5 mm is relatively thick and may have made the detection of very fine fractures more difficult, thus resulting in lower sensitivities. In addition, the fact that the NFA had a much more detailed knowledge of piglet anatomy and osteology may have resulted in higher fracture detection sensitivities and skewed the results in favour of the NFA. Therefore, including more observers familiar with piglet anatomy may improve results. Furthermore, including more observers of all levels of experience, including specialist pediatric radiologists, as well as increasing the sample size, may increase the validity of the results.

Conclusion

The field of forensic anthropology is becoming much more virtual in its approach and is making use of radiological tools to conduct forensic anthropological analyses, including the interpretation of trauma [3]. However, it was previously unknown whether forensic anthropologists, who generally have very limited to no radiological experience and training, are able to accurately identify skeletal trauma using various imaging modalities. Results of this study show that observer experience does indeed affect the number of fractures detected using CT, X-ray and Lodox® scans, with sensitivity increasing as level of experience increases. This indicates that even very brief radiological training, as was received by the NFA in this study, can drastically improve the ability of the observer to detect skeletal trauma. While forensic anthropologists do have fair to almost perfect agreement with the number of fractures detected by the SRR, fracture detection sensitivities of forensic anthropologists without radiological training are low and not sufficient for assessing cases of suspected trauma, such as physical child abuse or violent crime. The authors therefore suggest that all forensic anthropologists, both novice and experienced, should receive radiological training, either included in undergraduate courses or offered as professional courses for more advanced anthropologists before they comment on fractures observed on radiographic material.

Acknowledgements

The authors would like to extend thanks to GHB Farms in South Africa for donating the piglets, and to Mr Meyer for inflicting the trauma on the piglets. For assistance with CT scanning, thanks to Dr Haagensen and Mrs Bussy from the Wits Donald Gordon Medical Centre; for assistance with X-ray scanning, thanks to Ms Khoza and Ms Rammekwa from the Central Animal Services at the University of the Witwatersrand; and for assistance with Lodox® scanning, thanks to Ms Nkosi from Lodox® Systems (Pty) Ltd, and Ms Reindorp from the Johannesburg Forensic Pathology Services. For assistance with maceration of the piglets, thanks are extended to Mr Matjila, Ms Joubert, Ms Swiegers and Ms Mamabolo. This work is based on the research supported in part by the National Research Foundation (NRF) of South Africa (Grant number: 118149) and the NRF/Deutscher Akademischer Austausch Dienst (DAAD) (Grant number: 117943).

References

[1] Z. Ali, N. Mourtzinos, B.B. Ali, D.R. Fowler, A pilot study comparing postmortem and antemortem CT for the identification of unknowns: could a forensic pathologist do it? J. Forensic Sci 65 (2) (2020) 492–499. https://doi.org/10.1111/1556-4029.14199.

[2] T.D. Stewart, Essentials of forensic anthropology: especially as developed in the United States, Charles C Thomas Publisher Ltd, Illinois, 1979.

[3] D. Franklin, L. Swift, A. Flavel, Virtual anthropology' and radiographic imaging in the Forensic Medical Sciences, Egypt, J. Forensic Sci. 6 (2) (2016) 31–43. https://doi.org/10.1016/j.ejfs.2016.05.011.

[4] A. Middleton, A. Alminyah, M.A. Apostol, L.W.T. Boel, A. Brough, W. Develter, S. Heinze, Y. Makino, L. Mulcahy, C. O'Donnell, G. Gorincour, P.A.M. Hofman, M. Iino, L. Oesterhelweg, D. Ranson, C. Robinson, T. Ruder, G.N. Rutty, M.K. C. Singh, C. Villa, M.D. Viner, K. Woźniak, M. Yoshida, Forensic odontology radiography and imaging in disaster victim identification, J. Forensic Radiol 6 (2016) 28–30. https://doi.org/10.1016/j.jofri.2016.08.003.

[5] T.D. Ruder, Y.A. Thali, S.N.A. Rashid, M.T. Mund, M.J. Thali, G.M. Hatch, A. M. Christensen, S. Somaini, G. Ampanozi, Validation of post mortem dental CT for disaster victim identification, J. Forensic Radiol 5 (2016) 25–30. https://doi.org/10.1016/j.jofri.2016.01.006.

[6] E. Nguyen, E. Doyle, Dental post-mortem computed tomography for disaster victim identification: a literature review, J. Forensic Radiol 13 (2018) 5–11. https://doi.org/10.1016/j.jofri.2018.03.002.

[7] L. Kullman, B. Eklund, R. Grundin, Value of the frontal sinus in identification of unknown persons, J. Forensic Odontostomatol 8 (1) (1990) 3–10.

[8] T.D. Ruder, M. Kraehenbuehl, W.F. Gotsmy, S. Mathier, L.C. Ebert, M.J. Thali, G. M. Hatch, Radiologic identification of disaster victims: a simple and reliable method using CT of the paranasal sinuses, Eur. J. Radiol. 81 (2012) e132–e138. https://doi.org/10.1016/j.ejrad.2011.01.060.

[9] L. Deloire, I. Diallo, R. Cadieu, M. Auffret, Z. Alavi, J. Ognard, D.B. Salem, Post-mortem X-ray computed tomography (PMCT) identification using ante-mortem CT-scan of the sphenoid sinus, J. Neuroradiol 46 (2019) 248–255. https://doi.org/10.1016/j.neurad.2018.08.003.

[10] F. Dedouit, F. Savall, F-Z. Mokrane, H. Rousseau, E. Crubézy, D. Rougé, N. Telmon, Virtual anthropology and forensic identification using multidetector CT, Brit. J. Radiol. 87 (1036) (2014), 20130468. https://doi.org/10.1259/bjr.20130468.

[11] H.M. Garvin, M.K. Stock, The utility of advanced imaging in forensic anthropology, Acad. Forensic Pathol. 6 (3) (2016) 499–516. https://doi.org/10.23907/2016.050.

[12] R.F. Silva, F.F. Picoli, T. de Lucena Botelho, R.G. Resende, A. Franco, [Forensic identification of decomposed human body through comparison between ante-mortem and post-mortem CT images of frontal sinuses: case report], Acta Stomatol. Croat. 51 (3) (2017) 227–231. https://doi.org/10.15644/asc51/3/6.

[13] C.P. Campobasso, A.S. Dell'Erba, M. Belviso, G. Di Vella, Craniofacial identification by comparison of antemortem and postmortem radiographs. Two case reports

dealing with burnt bodies, Am. J. Forensic Med. Pathol. 28 (2) (2007) 182–186. https://doi.org/10.1097/PAF.0b013e31806195cb.

[14] J-B. Coty, C. Nedelcu, S. Yahya, V. Dupont, C. Rougé-Maillart, M. Verschoore, C. Ridereau Zins, C. Aubé, Burned bodies, Post-mortem computed tomography, an essential tool for modern forensic medicine, Insights into Imaging 9 (2018) 731–743. https://doi.org/10.1007/s13244-018-0633-2.

[15] D. Brits, P.R. Manger, M.A. Bidmos, The accuracy of the anatomical method for stature estimation in Black South African females, Forensic Sci. Int. 278 (2017) 409.e1–409.e10. https://doi.org/10.1016/j.forsciint.2017.06.004.

[16] P. Oura, J. Karppinen, J. Niinimäki, J-A. Junno, Sex estimation from dimensions of the fourth lumbar vertebra in Northern Finns of 20, 30, and 46 years of age, Forensic Sci. Int. 290 (2018) 350.e1–350.e6. https://doi.org/10.1016/j.forsciint.2018.07.011.

[17] V. Vieth, R. Schulz, W. Heindel, H. Pfeiffer, B. Buerke, A. Schmeling, C. Ottow, Forensic age assessment by 3.0T MRI of the knee: proposal of a new MRI classification of ossification stages, Eur, Radiol 28 (8) (2018) 3255–3262. https://doi.org/10.1007/s00330-017-5281-2.

[18] K.L. Boyd, C. Villa, N. Lynnerup, The use of CT scans in estimating age at death by examining the extent of ectocranial suture closure, J. Forensic Sci 60 (2) (2015) 363–369. https://doi.org/10.1111/1556-4029.12683.

[19] S. Torimitsu, Y. Makino, H. Saitoh, A. Sakuma, N. Ishii, D. Yajima, G. Inokuchi, A Motomura, F. Chiba, R. Yamaguchi, M. Hashimoto, Stature estimation in a contemporary Japanese population based on clavicular measurements using multidetector computed tomography, Forensic Sci. Int. 275 (2017) 316.e1–316.e6. https://doi.org/10.1016/j.forsciint.2017.02.037.

[20] D. Toneva, S. Nikolova, S. Harizanov, I. Georgiev, D. Zlatareva, V. Hadjidekov, A. Dandov, N. Lazarov, Sex estimation by size and shape of foramen magnum based on CT imaging, Legal Med 35 (2018) 50–60. https://doi.org/10.1016/j.legalmed.2018.09.009.

[21] P.J. Macaluso, J.S. Lucena, Stature estimation from radiographic sternum length in a contemporary Spanish population, Int. J. Legal Med. 128 (5) (2014) 845–851. https://doi.org/10.1007/s00414-014-0975-3.

[22] V. Bartolini, V. Pinchi, B. Gualco, S. Vanin, G. Chiaracane, G. D'Elia, G.A. Norelli, M. Focardi, The iliac crest in forensic age estimation: evaluation of three methods in pelvis X-rays, Int. J. Legal Med. 132 (1) (2018) 279–288. https://doi.org/10.1007/s00414-017-1629-z.

[23] M. Ramezani, V. Shokri, A. Ghanbari, Z. Salehi, K.A. Niknami, Stature estimation in Iranian population from x-ray measurements of femur and tibia bones, J. Forensic Radiol 19 (2019), 100343. https://doi.org/10.1016/j.jofri.2019.100343.

[24] K.E. Stull, E.N. L'Abbé, S.D. Ousley, Using multivariate adaptive regression splines to estimate subadult age diaphyseal dimensions, Am. J. Phys. Anthropol 154 (3) (2014) 376–386. https://doi.org/10.1002/ajpa.22522.

[25] K.E. Stull, E.N. L'Abbé, S.D. Ousley, Subadult sex estimation from diaphyseal dimensions, Am. J. Phys. Anthropol 163 (1) (2017) 64–74. https://doi.org/10.1002/ajpa.23185.

[26] K. Imaizumi, K. Taniguchi, Y. Ogawa, K. Matsuzaki, H. Maekawa, T. Nagata, T. Moriyama, I. Okuda, H. Hayakawa, S. Shiotani, Development of three-dimensional facial approximation system using head CT scans of Japanese living individuals, J. Forensic Radiol. Imaging 17 (2019) 36–45. https://doi.org/10.1016/j.jofri.2019.05.004.

[27] M.J. Thali, K. Yen, W. Schweitzer, P. Vock, C. Boesch, C. Ozdoba, G. Schroth, M. Ith, M. Sonnenschein, T. Doernhoefer, E. Scheurer, T. Plattner, R. Dirnhofer, Virtopsy, a new imaging horizon in forensic pathology: virtual autopsy by postmortem multislice computed tomography (MSCT) and magnetic resonance imaging (MRI) – a feasibility study, J. Forensic Sci 48 (2) (2003) 386–403. https://doi.org/10.1520/JFS2002166.

[28] R. Dirnhofer, C. Jackowski, P. Vock, K. Potter, M.J. Thali, VIRTOPSY: minimally invasive, image-guided virtual autopsy, Radiographics 26 (2006) 1305–1333. https://doi.org/10.1148/rg.265065001.

[29] A.M. Kemp, F. Dunstan, S. Harrison, S. Morris, M. Mann, K. Rolfe, S. Datta, D. P. Thomas, J.R. Sibert, S. Maguire, Patterns of skeletal fractures in child abuse: systematic review, BMJ 337 (2008) 859–862. https://doi.org/10.1136/bmj.a1518.

[30] C. Cattaneo, E. Marinelli, A. Di Giancamillo, M. Di Giancamillo, O. Travetti, L. Vigano, P. Poppa, D. Porta, A. Gentilomo, M. Grandi, Sensitivity of autopsy and radiological examination in detecting bone fractures in an animal model: implications for the assessment of fatal child physical abuse, Forensic Sci. Int. 164 (2-3) (2006) 131–137. https://doi.org/10.1016/j.forsciint.2005.12.016.

[31] C. Chevallier, F. Doenz, P. Vaucher, C. Palmiere, A. Dominguez, S. Binaghi, P. Mangin, S. Grabherr, Postmortem computed tomography angiography vs conventional autopsy: advantages and inconveniences of each method, Int. J. Legal Med. 127 (5) (2013) 981–989. https://doi.org/10.1007/s00414-012-0814-3.

[32] C. Jacobsen, B.H. Bech, N. Lynnerup, A comparative study of cranial, blunt trauma fractures as seen at medicolegal autopsy and by Computed Tomography, BMC Med. Imaging 9 (1) (2009) 18–26. https://doi.org/10.1186/1471-2342-9-18.

[33] C. King, W. Birch, Assessment of maceration techniques used to remove soft tissue from bone in cut mark analysis, J. Forensic Sci 60 (1) (2015) 124–135. https://doi.org/10.1111/1556-4029.12582.

[34] D.H. Ubelaker, A. Shamlou, A. Kunkle, Contributions of forensic anthropology to positive scientific identification: a critical review, Forensic Sci. Res. 4 (1) (2018) 45–50. https://doi.org/10.1080/20961790.2018.1523704.

[35] A.J. Spies, M. Steyn, E. Busy, D. Brits, Forensic imaging: the sensitivities of various imaging modalities in detecting skeletal trauma in simulated cases of child abuse using a pig model, J. Forensic Leg. Med. 76 (2020), 102034, https://doi.org/10.1016/j.jflm.2020.102034.

[36] A.J. Spies, M. Steyn, D. Brits, A comparison of computed tomography, X-ray and Lodox® scans in assessing pediatric skull fractures using piglets, J. Forensic Sci. (2020), https://doi.org/10.1111/1556-4029.14614. Early View.

[37] M. du Plessis, M. Date-Chong, L. Liebenberg, Lodox®: the invaluable radiographic solution in the forensic setting, Int. J. Legal Med. 134 (2020) 655–662. https://doi.org/10.1007/s00414-019-02116-y.

[38] J.R. Landis, G.G Koch, The measurement of observer agreement for categorical data, Biometrics 33 (1) (1977) 159–174. https://doi.org/10.2307/2529310.

[39] R. Parikh, A. Mathai, S. Parikh, G.C. Sekhar, R. Thomas, Understanding and using sensitivity, specificity and predictive values, Indian J. Opthalmol. 56 (1) (2008) 45–50. https://doi.org/10.4103/0301-4738.37595.

[40] T.S. Hong, J.A. Reyes, R. Moineddin, D.A. Chiasson, W.E. Berdon, P.S. Babyn, Value of postmortem thoracic CT over radiography in imaging of pediatric rib fractures, Pediatr. Radiol. 41 (2011) 736–748. https://doi.org/10.1007/s00247-010-1953-7.

[41] S.C. Shelmerdine, D. Langan, J.C. Hutchinson, M. Hickson, K. Pawley, J. Suich, L. Palm, N.J. Sebire, A. Wade, O.J. Arthurs, Chest radiographs versus CT for the detection of rib fractures in children (DRIFT): a diagnostic accuracy observational study, Lancet Child Adolesc. Health 2 (2018) 802–811. https://doi.org/10.1016/S2352-4642(18)30274-8.

[42] W. Baier, J.M. Warnett, M. Payne, M.A. Williams, Introducing 3D printed models as demonstrative evidence at criminal trials, J. Forensic Sci 63 (4) (2018) 1298–1302. https://doi.org/10.1111/1556-4029.13700.