



Deep learning in sex estimation from a peripheral quantitative computed tomography scan of the fourth lumbar vertebra—a proof-of-concept study

Petteri Oura^{1,2,3} · Niina Korpinen⁴ · Allison L. Machnicki⁵ · Juho-Antti Junno^{4,6,7}

Accepted: 25 January 2023 / Published online: 11 February 2023
© The Author(s) 2023

Abstract

Sex estimation is a key element in the analysis of unknown skeletal remains. The vertebrae display clear sex discrepancy and have proven accurate in conventional morphometric sex estimation. This proof-of-concept study aimed to investigate the possibility to develop a deep learning algorithm for sex estimation even from a single peripheral quantitative computed tomography (pQCT) slice of the fourth lumbar vertebra (L4). The study utilized a total of 117 vertebrae from the Terry Anatomical Collection. There were 58 male and 59 female cadavers, all of the white ethnicity, with the average age at death 49 years and a range of 24 to 77 years. A coronal pQCT scan was taken from the midway of the L4 corpus. Sex estimation was performed in a total of 19 neural network architectures implemented in the *AIDeveloper* software. Of the explored architectures, a *LeNet5*-based algorithm reached the highest accuracy of 86.4% in the test set. Sex-specific classification rates were 90.9% among males and 81.8% among females. This preliminary finding advances the field by encouraging and directing future research on artificial intelligence-based methods in sex estimation from individual skeletal traits such as the vertebrae. Combining quickly obtained imaging data with automated deep learning algorithms may establish a valuable pipeline for forensic anthropology and provide aid when combined with traditional methods.

Keywords Sex estimation; Vertebra · Lumbar spine · Peripheral quantitative computed tomography · Imaging; Artificial intelligence · Deep learning; Forensic anthropology · Terry collection

Introduction

Sex estimation, i.e., estimating the biological genotype of an individual [1], remains a key element in the analysis of unknown skeletal remains of human origin [2, 3]. Of help is the fact that skeletal features such as bone size, shape, and structure are subject to varying levels of sex discrepancy [4, 5]. A number of skeletal sites, such as the cranium, pelvis, and extremities, have been tested in sex estimation using conventional morphometric methods. Classification rates have varied widely from ~40 to ~100% [2, 6]; however, comparisons between studies are complicated by differences in the study population, skeletal site, and sex estimation method used. An acceptable threshold for an accurate sexing tool is considered to be 80 to 95% [2, 7].

In a wide range of forensic scenarios, e.g., severe trauma, burns, and other fragmentary conditions, a limited number of skeletal elements or fragments may be recovered for examination. These cases in particular may benefit from novel sex estimation methods. The vertebrae display clear sex

✉ Petteri Oura
petteri.oura@helsinki.fi

¹ Department of Forensic Medicine, Faculty of Medicine, University of Helsinki, P.O. Box 21, Helsinki 00014, Finland

² Forensic Medicine Unit, Finnish Institute for Health and Welfare, Helsinki, Finland

³ Faculty of Medicine, Research Unit of Health Sciences and Technology, University of Oulu, Oulu, Finland

⁴ Department of Archaeology, Faculty of Humanities, University of Oulu, Oulu, Finland

⁵ Center for Functional Anatomy and Evolution, Johns Hopkins School of Medicine, Baltimore, MD, USA

⁶ Cancer and Translational Medicine Research Unit, Faculty of Medicine, University of Oulu, Oulu, Finland

⁷ Archaeology, Faculty of Arts, University of Helsinki, Helsinki, Finland

discrepancy in the adult spine [8] and have proven accurate in conventional morphometric sex estimation (classification rates ~85 to ~90%, maximum 94.5%) [9–20]. However, as conventional osteology methods may be laborious and only able to focus on a limited number of parameters at a time, the application of modern image recognition techniques may increase accuracy and expedite processes.

Artificial intelligence has the aim of mimicking cognitive functions by means of trained algorithms [21]. Deep learning uses neural networks in tasks such as image recognition and classification [22–24]. This is particularly beneficial in complex datasets such as those used in sex estimation from vertebral imaging data. To date, however, we only found one study that applied deep learning in this context. In their study, Malatong et al. [25] analyzed the first through fifth lumbar vertebrae (L1–L5) in a sample of 220 Thai individuals. The classification rate of a deep learning-based algorithm (92.5%) exceeded that of a digital caliper-based method (82.7%) and a morphometric image analysis method (90.0%). The findings show promise but are preliminary in nature; sex estimation methods are mostly population-specific and algorithms need external validation in large datasets. Future research is thus clearly needed. To bring additional benefit over conventional osteoarcheological methods, artificial intelligence and imaging-based sex estimation processes should be portable, quick, and automated.

With these considerations, this proof-of-concept study aimed to develop a deep learning algorithm for sex

estimation from a single midcoronal peripheral quantitative computed tomography (pQCT) slice of the fourth lumbar vertebra (L4). The study utilized a total of 117 L4 vertebrae from the Terry Anatomical Collection [26].

Materials and methods

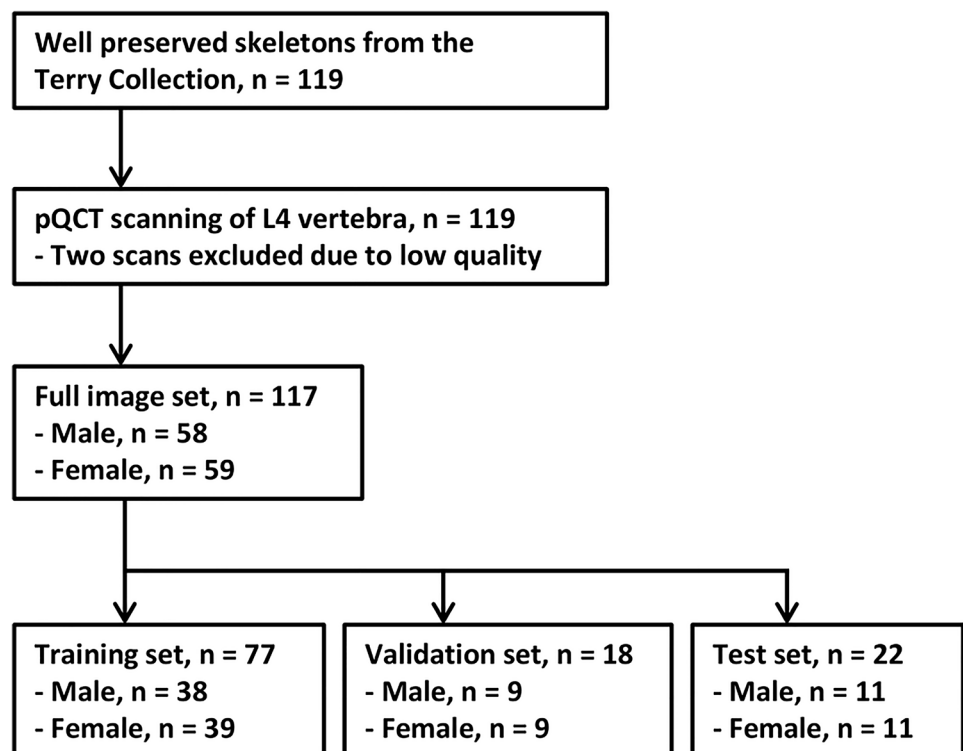
Material

The Robert J. Terry Anatomical Collection is maintained by the Department of Anthropology, National Museum of Natural History, Smithsonian Institution. The collection originates from the late nineteenth and early twentieth century and comprises the early industrial working class population of Anglo and African origins [26]. To conduct this study, L4 vertebrae from 119 well-preserved skeletons were selected to undergo pQCT scanning (Fig. 1). As two scans were later excluded due to low quality, the final sample consisted of 117 scans. There were 58 male and 59 female cadavers, all of whom were ethnically identified as white. They had died in the twentieth century at an average age of 49 years (range 24 to 77 years).

Peripheral quantitative computed tomography

pQCT scans of L4 were obtained using a Stratec XCT Research SA scanner and the corresponding software

Fig. 1 Flow-chart of the study (pQCT denotes peripheral quantitative computed tomography)



version 6.20 (Stratec Medizintechnik GmbH, Pforzheim, Germany). Slice thickness was set at 0.7 mm and pixel size 0.1 mm. L4 vertebrae were scanned through the vertebral corpus in a coronal plane, at the anteroposterior midpoint of the corpus. This slice was selected as it follows the primary direction of vertebral loading in humans [27, 28]. The scans were obtained by a member of the research group (J.-A.J.) who was blinded to the sex and other background information of the cadavers. Major vertebral pathologies such as crushing, wedging, osteoarthritis, and osteophyte formation were visually ruled out prior to scanning.

Image curation

The midcoronal pQCT slices of L4 were exported as BMP files and opened in Photoshop Elements version 2022 (Adobe, San Jose, CA, USA). The vertebrae were systematically extracted with a rectangular cutter following a ratio of approximately 175×110 pixels. The images were ensured to be in grayscale format and contrast was systematically adjusted with the “Auto Contrast” tool. The curated images were saved in GIF format. Image curation was performed by a member of the research group (P.O.) who was blinded to the sex and other background information of the cadavers. Please see Supplementary Fig. S1 for samples of curated images.

Deep learning procedure

AIDeveloper software version 0.1.2 [24] was used for the deep learning procedure (i.e., training, validation, and testing of algorithms). *AIDeveloper* is an open-source deep learning software that has a number of neural network architectures available for image classification.

First, the curated images were randomly assigned to training, validation, and test sets, following an approximate 70%–15%–15% ratio in the division (Fig. 1). A random number generator tool of SPSS Statistics version 26 (IBM, Armonk, NY, USA) was used in the randomization. Then, the training and validation image sets were uploaded to *AIDeveloper*, and training and validation rounds were run. A list of the explored network architectures is given in Table 1; they were selected as they were readily implemented in the software package. A detailed configuration of the settings and parameters is given in Supplementary Table S1. The best algorithm from each architecture was taken forward to the testing round which used the test image set as material.

An a priori decision was made to evaluate the algorithms based on the following performance metrics [24, 29]:

- Accuracy (i.e., true positives + true negatives)/total)
- Precision (i.e., true positives/(true positives + false positives))
- Recall (i.e., true positives/(true positives + false negatives))
- F1 value (i.e., $2 \times (\text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$)

Table 1 List of explored network architectures and performance metrics of best models

Neural network architecture	Training and validation set			Testing set	
	Best model	Training accuracy	Validation accuracy	Testing accuracy	Correct per class (male/female, %)
LeNet5	13	0.83	0.77	0.86	90.9/81.8
MLP_64_32_16	18	0.72	0.83	0.77	63.6/90.9
MLP_16_8_16	118	1.00	0.73	0.77	90.9/63.6
MLP_8_8_8	76	0.72	0.80	0.73	54.5/90.9
LeNet5_bn_do	41	0.89	0.93	0.68	90.9/45.5
MLP_24_16_24_skipcon	39	0.83	0.90	0.68	45.5/90.9
MhNet1_bn_do_skipcon	1	0.60	0.87	0.68	100.0/36.4
LeNet5_do	28	0.73	0.83	0.68	90.9/45.5
MLP_64_80_32	42	0.93	0.80	0.68	45.5/90.9
MLP_72_64_48_48	29	0.75	0.80	0.68	45.5/90.9
MLP_256_128_64_do	148	0.76	0.70	0.68	45.5/90.9
MLP_24_16_24	24	0.79	0.80	0.63	36.4/90.9
MhNet2_bn_do_skipcon	1	0.61	0.87	0.59	90.9/27.3
MLP_4_4_4	13	0.53	0.57	0.59	100.0/18.2
MLP_72_48_24_32	74	0.98	0.83	0.55	36.4/72.7
TinyCNN	1448	0.91	0.83	0.55	90.9/18.2
TinyResNet	44	0.65	0.73	0.55	90.9/18.2
MLP_72_80_32	199	0.94	0.80	0.50	0.0/100.0
LeNet5_bn_do_skipcon	156	0.87	0.83	0.45	72.7/18.2

Table 2 True and predicted classes on test set data according to the best algorithm (*LeNet5*)

	Predicted class	
	Class 0 (male)	Class 1 (female)
True class		
Class 0 (male)	90.9% (10/11)	9.1% (1/11)
Class 1 (female)	18.2% (2/11)	81.8% (9/11)

Results

A total of 117 L4 images were used in this study. The images were divided into training (65.8%), validation (15.4%), and test sets (18.8%) (Fig. 1).

Table 1 lists the performance metrics of the explored algorithms. Classification rates in the training and validation sets ranged from 0.53 to 1.00 and from 0.57 to 0.93, respectively. Testing accuracies were generally lower than training and validation accuracies, ranging from 0.45 to 0.86.

A *LeNet5*-based algorithm reached the highest overall testing accuracy, classifying 90.9% of males and 81.8% of females correctly (Tables 2 and 3). The neural network structure of the algorithm is given in Supplementary Table S2, and a visual overview of its performance in the test image set is shown in Supplementary Fig. S1.

Discussion

This preliminary study proved the concept of using a single pQCT slice of the L4 vertebra in artificial intelligence-based sex estimation. Altogether 19 neural network architectures were explored and their performance varied significantly. The best algorithm estimated sex correctly in 86.4% of the test set cases; sex-specific classification rates were 90.9% among males and 81.8% among females.

Most previous literature has described conventional morphometric methods to estimate sex from the vertebrae. Classification rates have mainly fluctuated between 85 and 90% [9–20], and the accuracy of the current best algorithm also falls within this range. In comparison to a previous Thai study [25] that developed a deep learning algorithm for

sex estimation from L1 to L5, the accuracy was somewhat lower (86.4% vs. 92.5%). However, the present algorithm was developed on the basis of one vertebra only (L4), and both external shape and internal morphology were captured in the midcoronal scans. In contrast, the Thai study was mostly based on surface plate images and external morphology. Moreover, comparisons between studies should be made with caution, bearing in mind differences in source population, sample size and heterogeneity, vertebral parameters, and data analysis methods. Of note is the fact that the current best algorithm clearly exceeded the 80% threshold for an acceptable sexing tool, encouraging further research on the topic.

This preliminary study explored 19 neural network architectures that were readily implemented in the deep learning software. There were little prior data and thus no a priori hypotheses regarding their performance relative to the study question. Notably, the performance of individual architectures varied significantly from the equivalent of a coin toss to the level of an acceptable sexing tool. While the authors believe that the differences are primarily explained by technical differences between architectures, future studies are required to investigate further. In the current dataset, some *LeNet5*- and multilayer perceptron-based architectures appeared to perform best.

As conventional osteology methods may be laborious and able to focus on a limited number of parameters at a time, novel sex estimation methods should be portable, quick, and automated. Despite being a three-dimensional imaging modality with an accurate depiction of bone geometry and microstructure, pQCT is designed for compactness and portability [30]. It is also swift to operate. In forensic scenarios, such as mass disasters in remote areas, the portability of the equipment and diagnostic devices would be a key advantage. Combining quickly obtained imaging data with automated image recognition algorithms may establish a valuable pipeline for forensic anthropology. Automated image classification also reduces subjectivity and measurement errors [25]. Naturally, large datasets and careful external validation will be necessary.

Several aspects require clarification in the future. First, studies should aim to explore whether the axial, sagittal, or coronal plane carries the most information value in sex estimation and compare these planes to a true three-dimensional profile of the vertebra. In a Finnish study using conventional metric methods, vertebral width and depth had clearly higher predictive values than vertebral height in all studied age groups [12]. It would also be interesting to explore whether geometry or microstructure possesses higher information value in sex estimation. Second, the drivers of misclassification should be identified. Males have generally larger vertebrae than females [8]; however, several lifestyle-related factors may influence the vertebrae and increase inter-individual

Table 3 Performance metrics of the best algorithm on test set data (*LeNet5*)

Metric	Class 0 (male)	Class 1 (female)
Testing accuracy	0.86	0.86
F1	0.87	0.86
Precision	0.83	0.90
Recall	0.91	0.82

variation in vertebral geometry and microstructure (e.g., age [12, 31], body size [32, 33], physical activity [34, 35], nutrition [36, 37]). Although major vertebral pathologies were ruled out, degenerative changes may also have affected the sex estimation accuracy. Third, robust algorithms should be based on large, representative, and contemporary samples. Careful external validation will be required prior to routine forensic use. On the basis of the current results, a similar pipeline utilizing post-mortem CT scans could be achievable; however, larger spinal segments or other skeletal elements should be explored as material, as they may well outperform L4. Naturally, careful optimization and validation would be necessary prior to routine implementation.

The main strength of the study was a novel approach, using a single pQCT slice of L4 in artificial intelligence-based sex estimation. Importantly, the best algorithm reached an accuracy of 86.4%. The proof-of-concept aim of the study was fulfilled. The main limitations were a relatively small sample size, wide age range, and lack of external validation. Although inter-observer bias cannot be fully ruled out, manual work was only required in few strictly standardized steps during the process (i.e., positioning of the vertebra for pQCT and image curation). However, in future applications, the pipeline would ideally be fully automated.

The lack of transparency remains a great limitation of artificial intelligence-based methods, especially in the legal context. As a standalone concept, artificial intelligence lacks reliability, since it may be subject to varying kinds of bias, but cannot be examined similarly to a human witness in the courtroom [38]. The authors, therefore, suggest that algorithms should primarily serve as additional tools for forensic anthropologists. The final interpretation of a case would always remain with the expert, preferably combining artificial intelligence-based tools with conventional evidence.

Conclusion

In this study of 117 midcoronal pQCT slices of the L4 vertebra, a *LeNet5*-based deep learning algorithm reached a sex estimation accuracy of 86.4%. This preliminary finding advances the field by encouraging and directing future research on artificial intelligence-based methods in sex estimation from individual skeletal elements such as vertebrae. Combining quickly obtained imaging data with automated deep learning algorithms may establish a valuable pipeline for forensic anthropology and provide aid when combined with traditional methods.

Key points

1. The vertebrae display clear sex discrepancy and have proven accurate in conventional morphometric sex estimation.
2. This study tested the application of several deep learning algorithms for sex estimation from a single computed tomography slice of the fourth lumbar vertebra (L4).
3. A *LeNet5*-based algorithm reached the highest accuracy of 86.4% in the test set.
4. This finding encourages future studies on artificial intelligence methods in sex estimation from individual skeletal traits such as the vertebrae.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12024-023-00586-6>.

Author contribution Conceptualization and methodology: PO, ALM, JAJ. Formal analysis and investigation: PO. Writing—original draft preparation: PO, NK, JAJ. Writing—review and editing: PO, NK, ALM, JAJ. All authors read and approved the final manuscript.

Funding Open Access funding provided by University of Helsinki including Helsinki University Central Hospital.

Data availability The datasets generated and analyzed during the study are available from the corresponding author upon reasonable request.

Code availability Not applicable.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interests The authors declare no competing interests.

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

References

1. Goodfellow PN, Darling SM. Genetics of sex determination in man and mouse. *Development*. 1988;102(2):251–8. <https://doi.org/10.1242/dev.102.2.251>.

2. 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. <https://doi.org/10.1016/j.forsciint.2016.02.007>.
3. Franklin D. Forensic age estimation in human skeletal remains: current concepts and future directions. *Leg Med (Tokyo).* 2010;12(1):1–7. <https://doi.org/10.1016/j.legalmed.2009.09.001>.
4. Krogman W. *The human skeleton in forensic medicine.* Springfield: Charles C Thomas; 1962.
5. Seeman E. Clinical review 137: Sexual dimorphism in skeletal size, density, and strength. *J Clin Endocrinol Metab.* 2001;86(10):4576–84. <https://doi.org/10.1210/jcem.86.10.7960>.
6. Spradley MK, Jantz RL. Sex estimation in forensic anthropology: skull versus postcranial elements. *J Forensic Sci.* 2011;56(2):289–96. <https://doi.org/10.1111/j.1556-4029.2010.01635.x>.
7. Rogers TL. A visual method of determining the sex of skeletal remains using the distal humerus. *J Forensic Sci.* 1999;44(1):57–60.
8. MacLaughlin SM, Oldale KNM. Vertebral body diameters and sex prediction. *Ann Hum Biol.* 1992;19(3):285–92. <https://doi.org/10.1080/03014469200002152>.
9. Hou WB, Cheng KL, Tian SY, Lu YQ, Han YY, Lai Y, et al. Metric method for sex determination based on the 12th thoracic vertebra in contemporary north-easterners in China. *J Forensic Leg Med.* 2012;19(3):137–43. <https://doi.org/10.1016/j.jflm.2011.12.012>.
10. Yu SB, Lee UY, Kwak DS, Ahn YW, Jin CZ, Zhao J, et al. Determination of sex for the 12th thoracic vertebra by morphometry of three-dimensional reconstructed vertebral models. *J Forensic Sci.* 2008;53(3):620–5. <https://doi.org/10.1111/j.1556-4029.2008.00701.x>.
11. Ostrofsky KR, Churchill SE. Sex determination by discriminant function analysis of lumbar vertebrae. *J Forensic Sci.* 2015;60(1):21–8. <https://doi.org/10.1111/1556-4029.12543>.
12. Oura P, Karppinen J, Niinimäki J, Junno JA. Sex estimation from dimensions of the fourth lumbar vertebra in Northern Finns of 20, 30, and 46 years of age. *Forensic Sci Int.* 2018;290:350.e1–350.e6. <https://doi.org/10.1016/j.forsciint.2018.07.011>.
13. Amores A, Botella MC, Alemán I. Sexual dimorphism in the 7th cervical and 12th thoracic vertebrae from a Mediterranean population. *J Forensic Sci.* 2014;59(2):301–5. <https://doi.org/10.1111/1556-4029.12320>.
14. Bozdog M, Karaman G. Virtual morphometry of the first lumbar vertebrae for estimation of sex using computed tomography data in the Turkish population. *Cureus.* 2021;13(7):e16597. <https://doi.org/10.7759/cureus.16597>.
15. Ekizoglu O, Hocaoglu E, Inci E, Karaman G, Garcia-Donas J, Kranioti E, et al. Virtual morphometric method using seven cervical vertebrae for sex estimation on the Turkish population. *Int J Legal Med.* 2021;135(5):1953–64. <https://doi.org/10.1007/s00414-021-02510-5>.
16. Garoufi N, Bertsatos A, Chovalopoulou ME, Villa C. Forensic sex estimation using the vertebrae: an evaluation on two European populations. *Int J Legal Med.* 2020;134(6):2307–18. <https://doi.org/10.1007/s00414-020-02430-w>.
17. Azofra-Monge A, Alemán Aguilera I. Morphometric research and sex estimation of lumbar vertebrae in a contemporary Spanish population. *Forensic Sci Med Pathol.* 2020;16(2):216–25. <https://doi.org/10.1007/s12024-020-00231-6>.
18. Decker SJ, Foley R, Hazelton JM, Ford JM. 3D analysis of computed tomography (CT)-derived lumbar spine models for the estimation of sex. *Int J Legal Med.* 2019;133(5):1497–506. <https://doi.org/10.1007/s00414-019-02001-8>.
19. Rozendaal AS, Scott S, Peckmann TR, Meek S. Estimating sex from the seven cervical vertebrae: an analysis of two European skeletal populations. *Forensic Sci Int.* 2020;306:110072. <https://doi.org/10.1016/j.forsciint.2019.110072>.
20. Tsubaki S, Morishita J, Usumoto Y, Sakaguchi K, Matsunobu Y, Kawazoe Y, et al. Sex determination based on a thoracic vertebra and ribs evaluation using clinical chest radiography. *Leg Med (Tokyo).* 2017;27:19–24. <https://doi.org/10.1016/j.legalmed.2017.06.003>.
21. Russell S, Norvig P. *Artificial intelligence: a modern approach.* 4th ed. Hoboken: Pearson; 2021.
22. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer.* 2018;18:500–10. <https://doi.org/10.1038/s41568-018-0016-5>.
23. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521(7553):436–44. <https://doi.org/10.1038/nature14539>.
24. Kräter M, Abuhattum S, Soteriou D, Jacobi A, Krüger T, Guck J, et al. AIDeveloper: Deep learning image classification in life science and beyond. *Adv Sci (Weinh).* 2021;8(11):2003743. <https://doi.org/10.1002/advs.202003743>.
25. Malatong Y, Intasuwan P, Palee P, Sinthubua A, Mahakkanukrauh P. Deep learning and morphometric approach for Sex determination of the lumbar vertebrae in a Thai population. *Med Sci Law.* 2023;63(1):14–21. <https://doi.org/10.1177/00258024221089073>.
26. Hunt DR, Albanese J. History and demographic composition of the Robert J. Terry anatomical collection. *Am J Phys Anthropol.* 2005;127(4):406–17. <https://doi.org/10.1002/ajpa.20135>.
27. Adams MA, Dolan P. Spine biomechanics. *J Biomech.* 2005;38(10):1972–83. <https://doi.org/10.1016/j.jbiomech.2005.03.028>.
28. Ferguson SJ, Steffen T. Biomechanics of the aging spine. *Eur Spine J.* 2003;12(Suppl 2):S97–103. <https://doi.org/10.1007/s00586-003-0621-0>.
29. Machin D, Campbell M, Walters S. *Medical statistics, fourth edition - A textbook for the health sciences.* Hoboken: John Wiley & Sons; 2007.
30. Wong AK. A comparison of peripheral imaging technologies for bone and muscle quantification: a technical review of image acquisition. *J Musculoskelet Neuronal Interact.* 2016;16(4):265–82.
31. Autio E, Oura P, Karppinen J, Paananen M, Niinimäki J, Junno JA. Changes in vertebral dimensions in early adulthood - A 10-year follow-up MRI-study. *Bone.* 2019;121:196–203. <https://doi.org/10.1016/j.bone.2018.08.008>.
32. Oura P, Nurkkala M, Auvinen J, Niinimäki J, Karppinen J, Junno JA. The association of body size, shape and composition with vertebral size in midlife - The Northern Finland Birth Cohort 1966 study. *Sci Rep.* 2019;9(1):3944. <https://doi.org/10.1038/s41598-019-40880-4>.
33. Oura P, Junno JA, Autio E, Karppinen J, Niinimäki J. Baseline anthropometric indices predict change in vertebral size in early adulthood - A 10-year follow-up MRI study. *Bone.* 2020;138:115506. <https://doi.org/10.1016/j.bone.2020.115506>.
34. Modarress-Sadeghi M, Oura P, Junno JA, Niemelä M, Niinimäki J, Jämsä T, et al. Objectively measured physical activity is associated with vertebral size in midlife. *Med Sci Sports Exerc.* 2019;51(8):1606–12. <https://doi.org/10.1249/MSS.0000000000001962>.
35. Oura P, Paananen M, Niinimäki J, Tammelin T, Auvinen J, Korpelainen R, et al. High-impact exercise in adulthood and vertebral dimensions in midlife - the Northern Finland Birth Cohort 1966 study. *BMC Musculoskelet Disord.* 2017;18(1):433. <https://doi.org/10.1186/s12891-017-1794-8>.
36. Oura P, Niinimäki J, Karppinen J, Nurkkala M. Eating behavior traits, weight loss attempts, and vertebral dimensions among the general Northern Finnish population. *Spine (Phila Pa 1976).* 2019;44(21):E1264–71. <https://doi.org/10.1097/BRS.0000000000003123>.

37. Oura P, Auvinen J, Paananen M, Junno JA, Niinimäki J, Karppinen J, et al. Dairy- and supplement-based calcium intake in adulthood and vertebral dimensions in midlife—the Northern Finland Birth Cohort 1966 Study. *Osteoporos Int*. 2019;30(5):985–94. <https://doi.org/10.1007/s00198-019-04843-9>.
38. Gless S. AI in the Courtroom: A comparative analysis of machine evidence in criminal trials. *Georget J Int Law*. 2020;51:195–253.

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