



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/241598/>

Version: Published Version

Article:

Gibson, F.G., Ding, Z., Paggiosi, M.A. et al. (2026) Strong correlation between phantomless and inline phantom-based densitometric calibration of vertebral properties from CT scans of healthy volunteers. *Frontiers in Bioengineering and Biotechnology*, 14. 1781532.

<https://doi.org/10.3389/fbioe.2026.1781532>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



OPEN ACCESS

EDITED BY

Christopher G. Provatidis,
National Technical University of Athens,
Greece

REVIEWED BY

Henrik Lauridsen,
Aarhus University, Denmark
Jiri Chmelik,
Brno University of Technology, Czechia

*CORRESPONDENCE

Stefaan W. Verbruggen,
✉ s.verbruggen@qmul.ac.uk

RECEIVED 05 January 2026

REVISED 29 April 2026

ACCEPTED 04 May 2026

PUBLISHED 29 May 2026

CITATION

Gibson FG, Ding Z, Paggiosi MA,
Handforth C, Brown JE, Li X, Dall'Ara E and
Verbruggen SW (2026) Strong correlation
between phantomless and inline
phantom-based densitometric calibration
of vertebral properties from CT scans of
healthy volunteers.

Front. Bioeng. Biotechnol. 14:1781532.

doi: 10.3389/fbioe.2026.1781532

COPYRIGHT

© 2026 Gibson, Ding, Paggiosi, Handforth,
Brown, Li, Dall'Ara and Verbruggen. This is
an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Strong correlation between phantomless and inline phantom-based densitometric calibration of vertebral properties from CT scans of healthy volunteers

Fiona G. Gibson^{1,2}, Zichu Ding³, Margaret A. Paggiosi⁴,
Catherine Handforth^{4,5}, Janet E. Brown⁴, Xinshan Li^{1,2},
Enrico Dall'Ara^{2,4} and Stefaan W. Verbruggen^{1,2,3,6*}

¹School of Mechanical, Aerospace and Civil Engineering, The University of Sheffield, Sheffield, United Kingdom, ²Insigneo Institute, The University of Sheffield, Sheffield, United Kingdom, ³Centre for Bioengineering, School of Engineering and Materials Science, Queen Mary University of London, London, United Kingdom, ⁴Division of Clinical Medicine, School of Medicine and Population Health, The University of Sheffield, Sheffield, United Kingdom, ⁵Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom, ⁶Digital Environment Research Institute, Queen Mary University of London, London, United Kingdom

Phantom calibration is currently the gold standard for calibrating CT scans and for calculating material properties of dense tissues for computational models. However, in Oncology departments and low-resource settings, it is not routine to include a calibration phantom within the scanning protocol. Therefore, retrospective scan datasets are challenging to calibrate for biomechanical investigations, precluding detailed measurements of material and mechanical properties. In this study, we compared the results from a phantomless calibration technique, where the density within each scan was independently calibrated based on known tissue densities captured within each scan (e.g., air), with those from a traditional inline phantom calibration. To do so we used scans from a cohort of healthy volunteers from the control arm of a clinical trial dataset (ANTELOPE) in which inline calibration phantoms were included. We found that, when selecting air and the aorta as regions for calibration within individual CT scans, a strong individual-specific correlation existed between bone mineral density measured in the phantomless and phantom calibrations. This indicates that the phantomless calibration method can be a useful and reliable tool for quantifying the densitometric material properties of healthy human vertebrae, and provides the opportunity for further analysis of spinal CT scans in either retrospective datasets or in low-resource clinical settings.

KEYWORDS

BMD, bone mineral density, CT scan, phantomless calibration, vertebral biomechanics

1 Introduction

To assess the patient's risk of vertebral fracture in clinics, two imaging modalities are typically deployed, Dual Energy X-rays Absorptiometry (DXA) and Quantitative Computed Tomography (QCT). At present, DXA remains the gold standard in the clinic due to ease of use, speed, cost and the low radiation exposure (5–20 μ Sv by DXA vs. 60–90 μ Sv by QCT)

(Kabayel, 2016). QCT provides images obtained from a CT scanner, where either an inline or offline (retrospective, same scanner and scanning protocol) calibration phantom could also be scanned. Calibration phantoms are used to convert the image's Hounsfield units into equivalent bone mineral density (BMD) through a set of equations.

The role of CT technology has grown extensively due to its effectiveness in skeletal assessment for diagnosis and continuous monitoring of cancers (Duvaferrier et al., 2013). This imaging technique provides essential information for assessing spinal stability by allowing for the identification of osteopenia, lytic lesions, soft-tissue involvement and fractures (Mahnken et al., 2002). The main concern with CT use is the exposure to significantly higher doses of radiation in comparison to standard radiographs (Winterbottom and Shaw, 2009). Despite this, evidence suggests that low-dose whole body CT is effective in producing high resolution images that provide the information necessary for assessing spinal stability (Gleeson et al., 2009).

The densitometric calibration of QCT images, required to compare data across different types of scanners and protocols (Pickhardt et al., 2013; Carpenter et al., 2014), usually utilises an external phantom; however, routine QCT scans are often conducted without a densitometric calibration phantom as its usage increases the logistical burden of clinical imaging (Lee et al., 2017). To account for this, numerous methods for phantomless calibration have been developed (Michalski et al., 2020; Bartenschlager et al., 2022). One approach is to pre-calibrate the scanner using either DXA measurements or a calibration phantom and apply this general pre-calibration factor to prospective QCT scans (Budoff et al., 2013; Pickhardt et al., 2015). Even though this is an improvement from not performing any densitometric calibration, it does not consider the patient-specific differences as well as scanner and protocol changes.

The most widely used phantomless approach is to utilise internal tissues as the reference materials. The choice of tissues has been varied and depend on scan location, with some authors chose air, fat and blood (Lee et al., 2017; Schwaiger et al., 2017), while others used fat and muscle (Weaver et al., 2015; Saffarzadeh et al., 2016) and showed similar results. In particular, Bartenschlager et al. compared different combinations of two internal tissues when calibrating for vertebrae BMD, reporting the lowest error for any combination with air (<5%), particularly air and blood and the highest errors arising when using muscle in the combination (Bartenschlager et al., 2022). These authors also concluded that the use of different CT scanners did not result in significant differences in calibration outcome. However, as their study was conducted in female patients with osteoporosis, it remains to be demonstrated whether these same settings can be generalised to apply to healthy or male participants. Additionally, their study was carried out on lumbar vertebrae, and as such it remains unclear if these findings apply to other spinal sites.

Therefore, the aim of this study was to investigate whether phantomless densitometric calibration of QCT scans can assess bone mineral density in the vertebra of healthy participants similarly to current gold standard approaches based on phantom calibration. Furthermore, we aimed to assess whether the accuracy holds over multiple timepoints and multiple scanning systems. We explored this by applying these techniques to healthy male volunteers in the control cohort of a clinical trial. Furthermore, this allowed us to

compare these calibration methods across multiple time points (baseline and follow-up scans) and with two different scanners.

2 Materials and methods

2.1 Participant data

The dataset analysed in this study is a time series QCT dataset from the ANTELOPE clinical trial (Gibson et al., 2025). Full details of the ANTELOPE trial, which investigated the skeletal effects of androgen deprivation therapy on prostate cancer patients, and demographics of the participants have been described elsewhere (Handforth et al., 2024; Gibson et al., 2025). Ethical approval was obtained from the South Yorkshire Research Ethics Committee (IRAS ID 206171). As the focus of this study was to identify the applicability of phantomless calibration in healthy bone, the 25 male healthy volunteers in the trial's control group were taken as the cohort for this study. While the trial investigated the effects of androgen deprivation therapy in prostate cancer patients (i.e., the treatment group), none of these cancer patients were included in this study. Moreover, there was no presence of vertebral fracture in any of the participants' T12 vertebra assessed in this study. A total of 50 scans was analysed, with two scans for each of the 25 participants at baseline (0 months) and follow-up (12 months). Participant details are shown in Table 1.

2.2 Phantom calibration

The QCT scan protocol for this trial included a solid inline calibration phantom (Image Analysis, Inc., Columbia, KY, USA) containing rods of 0, 0.075, and 0.15 g/cm³ equivalent concentration of calcium hydroxyapatite. This was carried out as per manufacturer instructions, and following well-established standards described in depth elsewhere (Burghardt et al., 2010). The first cohort from 2017 (29 participants) was scanned at baseline using the GE LightSpeed VCT (GE Healthcare, Milwaukee, WI) in the radiology department at the Northern General Hospital, Sheffield, UK, whilst the follow-up scans in 2018 along with all second cohort scans (25 participants who completed all assessments) were scanned using the Toshiba Aquilion ONE (Toshiba Medical Systems, Tokyo, Japan) at the same hospital. Quality assurance was performed once per month using a Mindways phantom (Mindways Software, Inc., Austin, TX, USA) on both scanners. All scans were performed in the anteroposterior position, using the same noise index. The QCT protocol included a single scan from the cranial endplate of the T12 vertebra to the T12/L1 margin. For the GE scanner, the tube voltage was 120 kV and the mean tube current was set at 360 mA, with a voxel size of 0.937 × 0.937 × 0.625 mm³. For the Toshiba scanner, the tube voltage was also 120 kV, the mean tube current was set at 250 mA and a voxel size of 0.976 × 0.976 × 0.5 mm³.

The densitometric calibration was computed using a standard approach, which assumes a linear relationship between the average Hounsfield units (HU) and the known equivalent mean values of equivalent BMD of each rod. To do so, one region of interest (ROI) (black square boxes, Figure 1A) was defined manually within each insertion of the phantom (ImageJ) (Rasband, 1997; Schneider et al., 2012). ImageJ is a widely-used, open-source image processing and

TABLE 1 Participant demographics data for the 25 male healthy volunteer in the control group of the ANTELOPE clinical trial including age, height, and BMI.

Participant ID	Age (years)	Height (cm)	BMI (kg/m ²)
C01	74	192.3	31.4
C02	76	169.0	22.3
C03	77	180.1	30.1
C04	73	173.3	26.2
C05	78	170.4	26.9
C06	68	182.8	27.1
C08	53	184.4	26.9
C09	63	188.1	29.6
C10	70	161.0	29.3
C11	79	175.0	30.7
C12	78	179.0	32.1
C14	74	168.2	26.8
C15	80	175.2	26.3
C16	82	169.0	29.7
C17	78	169.0	24.9
C19	75	180.8	34.9
C20	73	178.0	24.2
C21	78	159.8	22.7
C23	77	163.2	31.8
C24	73	184.5	23.9
C25	71	182.0	24.9
C26	71	173.0	20.4
C30	75	181.0	25.5
C31	71	171.0	27.6
C32	64	171.7	23.5
Average (±SD)	73 ±6	175.9 ± 8.0	27.1 ± 3.5

analysis program originally developed at the National Institutes of Health (NIH), written in Java to serve as a highly flexible and standard tool in biomedical research (Rasband, 1997; Schneider et al., 2012). The ROIs were defined as square regions centred within each calibration rod with length equal to half the edge length of the rod (12.5 mm). For each of the three rods, mean HU values over the same 10 slices were used to perform the linear regression analysis for calibration, Equation 1.

$$\rho_{QCT}^p = a + bHU \quad (1)$$

where ρ_{QCT}^p represents the QCT equivalent BMD calculated using the calibration phantom, HU represents the Hounsfield unit values of the densitometric calibration law and a and b are constants from

the linear regression analysis performed. This equation was applied to estimate the equivalent BMD in each voxel.

This equation can then be applied to estimate the BMD of the bone (Equations 2, 3). From this equation, the apparent density and the elastic modulus of the tissue are calculated using the following relationships. The density determined by CT (ρ_{QCT}) and ashing density (ρ_{ash}) is highly correlated (Schileo et al., 2008), therefore it is assumed that $\rho_{QCT} = \rho_{ash}$ (Schileo et al., 2008).

$$\rho_{ash} = \rho_{app} \times 0.6 [g/cm^3] \quad (2)$$

where ρ_{QCT} is the QCT equivalent BMD and ρ_{app} is the apparent density (Morgan et al., 2003).

$$E = 4730\rho_{app}^{1.56} [MPa] \quad (3)$$

where E defines the elastic modulus as a function of the apparent density, ρ_{app} . This allows mechanical properties, reliant upon the densitometric conversion, to be calculated as previously (Gibson et al., 2025).

For the assessment of BMD, from each QCT an ellipse shaped ROI was identified (ImageJ) in the anterior most region of the vertebral body. The ellipse was identified by creating a circular region of interest in the vertebral body, ensuring the cortical portion was included. HU values within the ROI for all slices were converted into BMD using the densitometric calibration identified as described above. To capture the integral BMD (g/cm^3), the FE software Ansys Workbench (2021R1) was used to select a ROI for all the elements in each vertebral body after segmentation and reconstruction as described elsewhere (Gibson et al., 2025), excluding the posterior elements and processes, incorporating both the cortical and trabecular regions. The integral BMD was then calculated as the sum of the individual element's bone mineral content (element BMD multiplied by element volume) divided by the total volume of the vertebral body ROI.

2.3 Phantomless calibration

To conduct the phantomless calibration, a combination of internal materials (IM) was used. In depth explanations of phantomless calibrations are available elsewhere (Lee et al., 2017). Briefly, from each scan, tissue ROIs for subcutaneous adipose tissue (SAT), air, aortic blood, and the erector spinae skeletal muscle (ES) were manually sampled from the scan field-of-view, as depicted in Figure 1B. To reduce influence of variations in tissue HUs across the scan field-of-view, the ROIs were placed adjacent to the bones of interest (T12 vertebra) for each tissue, and the mean HUs were determined from the tissue sample aggregated histograms of ten 2D slices, providing a compromise between the size of the ROI and size of the anatomical feature of interest. Using mass absorption coefficients (Table 2) obtained from the National Institute of Standards and Technology (www.nist.gov National Institute of Standards and Technology, NISTIR 4999), the scan effective energy was estimated by iteratively correlating the ROI-specified HUs and corresponding mass absorption coefficient at each energy level and maximizing the coefficient of determination (Millner et al., 1978), as shown in Figure 1C. This calibration for scan effective energy is necessary, as different local geometry and material properties affect the local attenuation of the X-ray beam, and this

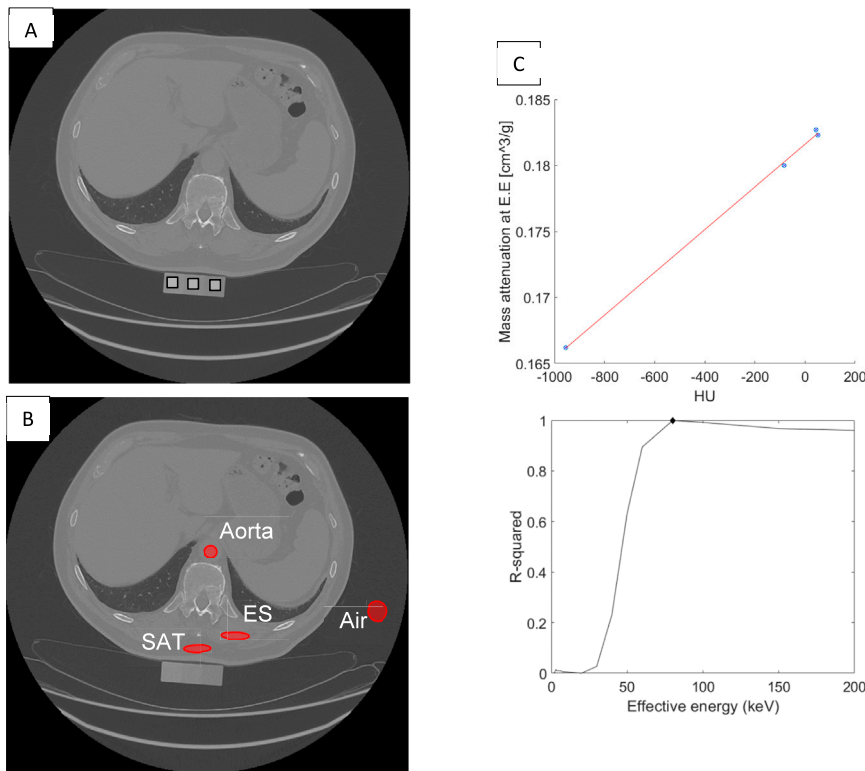


FIGURE 1 Density calibration methods for quantitative CT analysis. **(A)** Phantom-based calibration uses the phantom. Each calibration rod is sampled from the image to determine linear conversion between HUs and equivalent density (black square boxes). **(B)** In scan tissues of reference (adipose (SAT), air, blood (aorta), and skeletal muscle (ES)) are sampled adjacent to the bone of interest for internal calibration. **(C)** HUs and mass attenuation coefficients for each tissue (circle data points) are correlated by iterating at each effective energy (EE). Scan effective energy is determined by maximizing the coefficient of determination across all effective energies (black diamond).

procedure allows more accurate correlation of the local HUs to the material properties (Millner et al., 1978). Additionally, the coefficient of determination can vary greatly at different effective energies, as demonstrated in Figure 1D, and therefore emphasises the importance of maximising this coefficient.

For compounds such as hydroxyapatite (Ca₁₀(PO₄)₆(OH)₂) that are not tabulated in NIST, mass absorption coefficients can be calculated if the atomic mass fractions and the mass densities are known. When a compound like hydroxyapatite (HA), is not listed in standard databases like NIST, its mass absorption coefficient must be calculated using the mixture rule. First, the weight fraction of each element (Ca, P, O, H) was calculated by dividing its total atomic mass in the formula by the total molecular weight of HA. Once the scan effective energy was determined for the scan, the mass absorption coefficients, equivalent density and measured HU values for each material were used in a two-component mass fraction model (Genant and Boyd, 1977) to calculate the associated calibration equation, Equation 4.

$$\rho_{QCT}^{pl} = \frac{\left(\frac{\mu}{\rho}\right)_1 \rho_1 \frac{HU - CT_2}{CT_1 - CT_2} + \left(\frac{\mu}{\rho}\right)_2 \rho_2 \frac{HU - CT_1}{CT_2 - CT_1} - \left(\frac{\mu}{\rho}\right)_w \rho_w}{\left(\frac{\mu}{\rho}\right)_{HA} - \left(\frac{\mu}{\rho}\right)_w \frac{\rho_w}{\rho_{PHA}}} \quad (4)$$

where ρ_{QCT}^{pl} represents the BMD calculated using the phantomless calibration method, HU represents the Hounsfield unit values of the

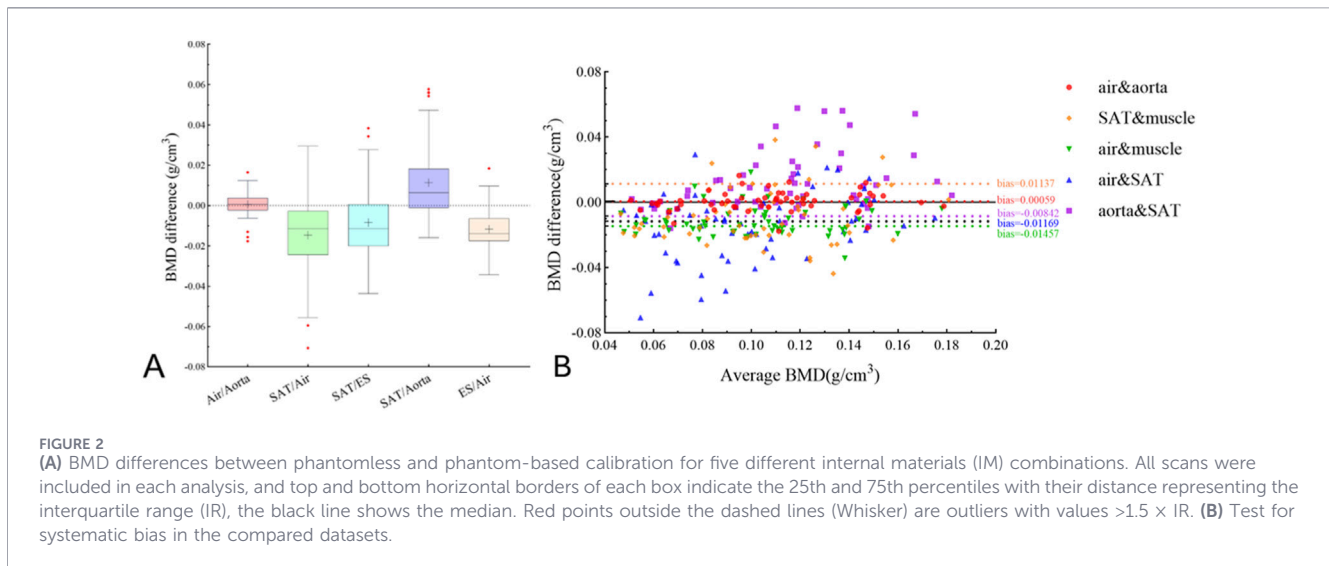
TABLE 2 Mass densities of internal calibration materials obtained from the National Institute of Standards and Technology (NIST) database.

Material	Density (g/cm ³)
Blood	1.06
Air	1.205 × 10 ⁻³
ES	1.05
SAT	0.95

ES, erector spinae muscle, SAT, subcutaneous adipose tissue.

densitometric calibration law, CT_1 and CT_2 represent the averaged grey value of each internal material, ρ_1 and ρ_2 represent the density of each internal material, ρ_w and ρ_{HA} represent the density of water and hydroxyapatite respectively and $\left(\frac{\mu}{\rho}\right)_1$ and $\left(\frac{\mu}{\rho}\right)_2$ represent the mass absorption coefficient for the internal material, $\left(\frac{\mu}{\rho}\right)_w$ and $\left(\frac{\mu}{\rho}\right)_{HA}$ represent the mass absorption coefficients for water and hydroxyapatite respectively.

The total $\left(\frac{\mu}{\rho}\right)_{HA}$ is the sum of the individual coefficients of its constituent elements, each multiplied by its respective mass fraction. Because $\left(\frac{\mu}{\rho}\right)$ varies with X-ray energy, the calculation must be performed at the scan effective energy (the monoenergetic



equivalent of the polychromatic X-ray beam). The [Genant and Boyd \(1977\)](#) model assumes that any given volume (voxel) in the bone can be represented as a mixture of two primary components: bone mineral (hydroxyapatite) and soft tissue (represented as water). The terms involving, for example, $HU - CT_1$, represent a linear interpolation between two known internal reference materials that have been measured in HU values (e.g., fat), and by combining with the density values, the terms essentially convert the abstract HU values back into physical attenuation values based on the specific scanner's performance during that scan. The term $-(\frac{\mu}{\rho})_w \rho_w$ subtracts the expected attenuation of water. This isolates the "excess" attenuation caused by the presence of bone mineral. Thus the final calculated value ρ_{QCT}^{pl} provides a quantitative measure of BMD that is corrected for the specific energy of the CT scan and the patient's own internal tissue characteristics, allowing for "phantomless" screening using routine clinical scans.

The two calculated BMD values were compared directly in order to estimate the accuracy errors for each of the 50 scans, with measured trabecular CT values of the T12 vertebra converted to BMD using both the phantom and phantomless calibration equations. Then for each scan, the difference ($\Delta BMDs$) between the phantom based and phantomless approaches was determined.

2.4 Statistical analysis

The linear regression analysis was used to compare the values of BMD calculated for each vertebra at both time points using the phantom and the phantomless calibration methods. For statistically significant regressions ($p < 0.05$) the Spearman's correlation coefficient (r) was reported. An *a priori* power analysis was conducted to determine the required sample size for a Spearman's rank-order correlation. The calculation was based on a two-tailed test, an alpha (α) level of 0.5, and 80% power ($1 - \beta$) of 0.80. To achieve these parameters to detect a strong effect size ($\rho \approx 0.62$), an $n = 20$ is required. This sample size is sufficient to identify meaningful monotonic relationships while allowing for non-parametric nature of the

data. Steiger's z-test was used to compare the statistical significance of differences (z) between correlations established using regression analyses.

3 Results

The systematic bias and random variability for BMD difference between phantom based and phantomless calibrations are shown in [Figure 2](#). Here, each datapoint represents the average BMD for an individual scan as calculated if calibrated with the indicated combination of internal materials. In [Figure 2A](#), each boxplot contains 50 values for a given internal-material combination, while in [Figure 2B](#), each point represents one scan, where the y-axis is the difference between phantomless and phantom-based mean T12 BMD, and the x-axis is the average of those two scan-level mean BMD values. The largest biases were found for the combination of air and SAT (-0.0146 g/cm^3 , underestimation) and aorta and SAT (0.0114 g/cm^3 , overestimation). The air and aorta combination demonstrated the highest accuracy, with a negligible bias of 0.0006 g/cm^3 and the lowest variability ($SD = 0.006 \text{ g/cm}^3$). These correspond to relative errors of 0.6% (air/aorta), 10.6% (aorta/SAT), -13.6% (air/SAT), -10.9% (air/muscle), and -7.9% (SAT/muscle).

[Figure 3](#) reports the linear regressions between the phantom based and phantomless BMD values for all internal materials combinations. It can be seen that all regression analyses were statistically significant ($p < 0.01$ in all cases) and that the maximum variation between baseline and follow-up datasets was 0.09 (for SAT/ES). The least variation in regression between baseline and follow-up data was for Air/Aorta, with an $r = 0.98$ for both datasets.

Here we found that calibration using a combination of the air in the scan, and the blood inside the aorta, provided the strongest correlation with standard phantom calibration ($r = 0.98$, $p < 0.01$). Additionally, using Steiger's z-test to compare the statistical significance of differences between correlations ([Table 3](#)), we

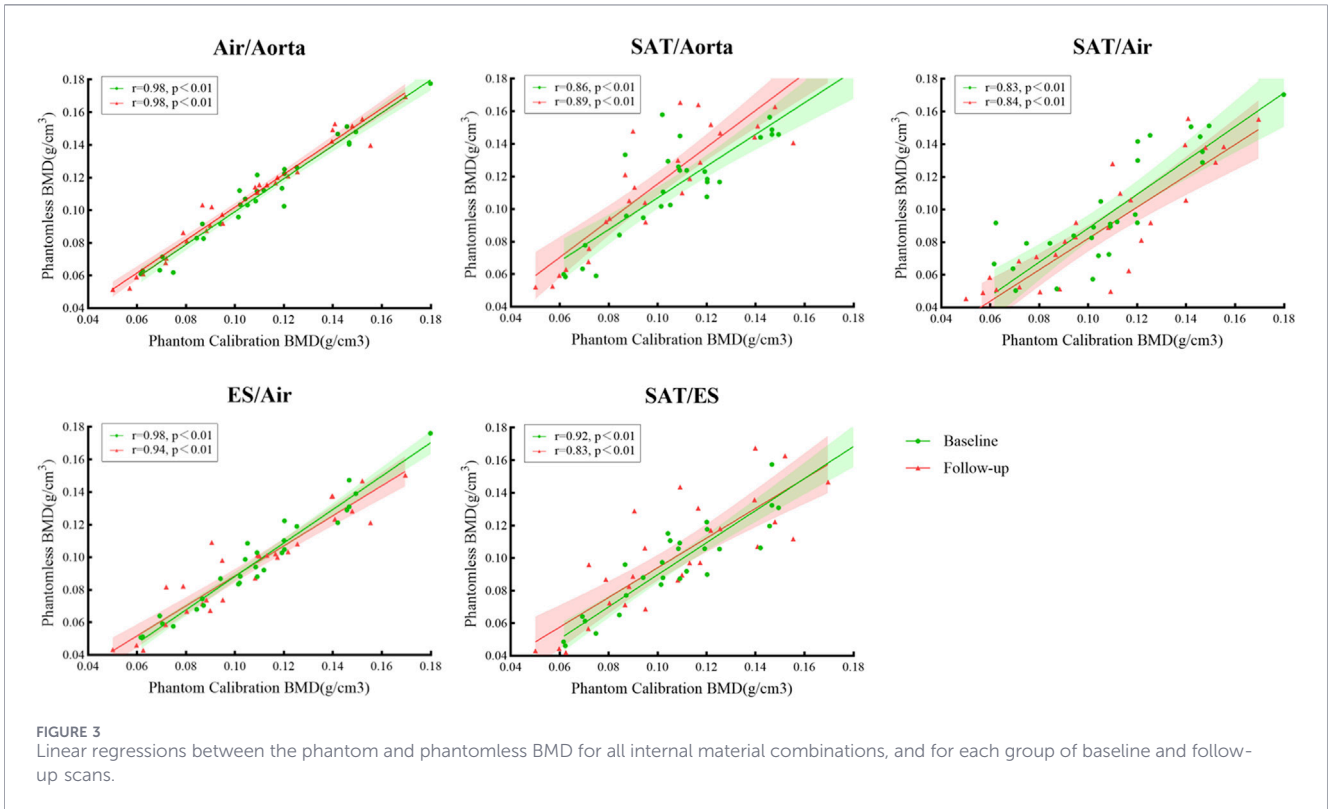
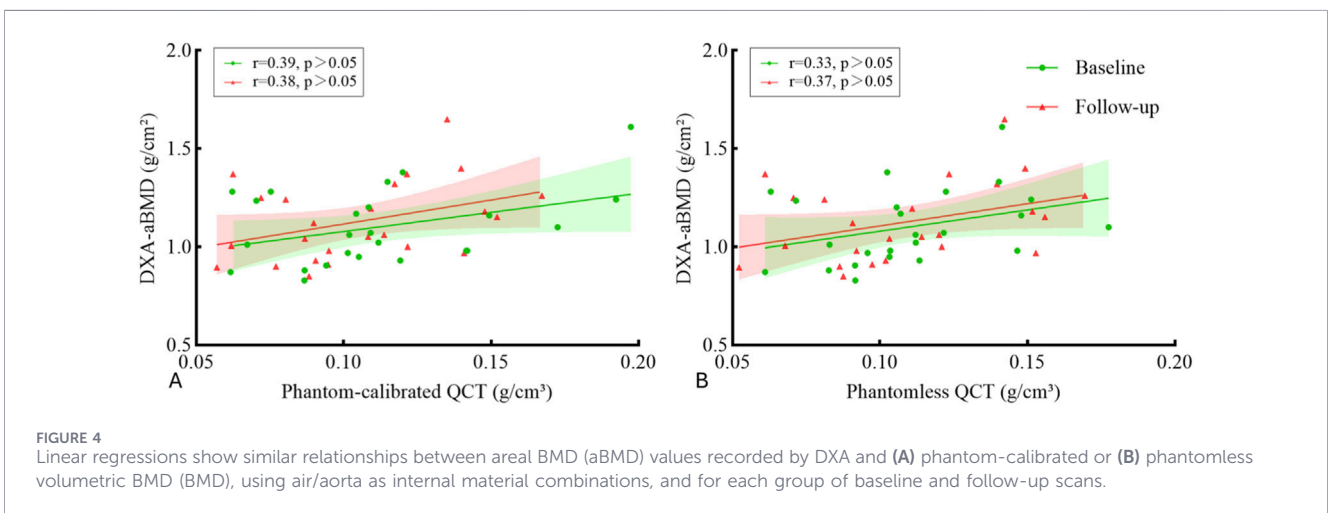


TABLE 3 Comparison of regression equations between the best performer air/aorta and each of the other internal calibration materials using Steiger’s z-test.

Test	Air/Aorta vs. SAT/Aorta	Air/Aorta vs. SAT/Air	Air/Aorta vs. ES/Air	Air/Aorta vs. SAT/ES
z	7.5085	7.5512	2.8956	6.9338
p-value	<0.0001	<0.0001	0.0038	<0.0001



found that calibration using air and aortic blood also outperformed the second-best method (air and muscle) (Steiger’s $z = 2.90, p < 0.05$). This is also inherently sensible, as the densities of air and blood are well characterised and, compared to the density of a

mineralised tissue like bone, do not vary significantly from participant to participant.

Lastly, to investigate the effectiveness of phantomless calibration in the clinic, we directly compared both calibration methods to DXA

measurements of areal BMD from the same patients at the relevant time point within the ANTELOPE study (Figure 4). As we have found previously (Gibson et al., 2025), areal BMD values generated by DXA are not highly predictive of volumetric BMD, and therefore fracture risk. The lack of volumetric detail in DXA scans is one of the disadvantages to what is an otherwise widely available and useful technique for patient stratification. However, it is noteworthy that in a regression analysis DXA was equally predictive of QCT volumetric BMD values regardless of whether phantom- or phantomless-calibration was used. This further highlights that phantomless calibration provides equal clinical utility to the current gold-standard phantom calibration method.

4 Discussion and conclusion

This study compared between phantomless and phantom based calibration, and evaluated the use of different combinations of internal materials for the phantomless approach. The combination of aorta and air for calibrating the CT scans yielded the most similar volumetric BMD with a 0.41% difference when compared with the gold standard phantom method (high correlation of 0.98), and equally as predictive as phantom-calibration when compared against DXA measurements. These findings confirmed previous work showing that the phantomless calibration method is a useful and reliable tool for quantifying the material and mechanical properties of human vertebrae (Bartenschlager et al., 2022). As expected, this study confirmed that phantomless calibration is equally accurate in male healthy volunteers in clinical trial conditions, indicating that phantomless techniques may be applied more broadly, and providing the opportunity for further analysis of spinal CT scans in either retrospective datasets or in low-resource clinical settings.

An important limitation to note is that, while this study analyses healthy volunteers from the control arm of a clinical trial, the participants are older and not representative of the general population. It would indeed be useful to expand this method to study a wider range of age and BMI. Nonetheless, as previous studies have investigated scans of osteoporotic women, it is a finding of note that our study gives strong confidence that these methods are robust in an equivalent elderly male population.

A number of previous studies have demonstrated the potential of phantomless calibration for densitometric analysis of bones using QCT scans. Early work on vertebrae found that it was possible to apply phantomless calibration to the lumbar spine scans of patients injured in traffic accidents, showing excellent agreement with phantom calibration in these patients (Weaver et al., 2015; Saffarzadeh et al., 2016). Similarly, a study comparing phantomless predictions in prostate cancer patients found strong agreement in vertebral and femoral fracture risk prediction in these patients (Schwaiger et al., 2017). Further analysis by this group showed very low inter-operator variability with phantomless calibration using femur QCT scans of patients, indicating it is a reliable form of calibration (Lee et al., 2017). Most recently, Bartenschlager et al. used a retrospective QCT database of post-menopausal

women to show that phantomless calibration was highly correlated with phantom based method, and that the best internal materials to use were values measured from the aorta and air (Bartenschlager et al., 2022). Our study expands upon these previous findings, showing that phantomless calibration can be applied to healthy volunteers in clinical trial conditions. We also confirmed that the most accurate calibration values were obtained when using the air and aorta as internal materials. Additionally, our study investigated changes across two different timepoints, with scans acquired using two different CT machines, demonstrating robust and broad applicability of this calibration method.

An advantage of QCT imaging of bone is that the resulting scans can be used to create 3D biomechanical models, using finite element (FE) analysis of the vertebra to estimate the bone strength. Patient-specific FE models have long been used to investigate the biomechanical response of bones to loading (Engelke et al., 2016; Schileo and Taddei, 2021). FE has been used increasingly in bones affected by diseases such as osteoporosis (Matsumoto et al., 2009) and different types of cancers including breast, colorectal and renal cell carcinoma (Costa et al., 2019). These models have also been used to study the effect of treatments and have been proven to predict vertebral strength more accurately than DXA in individuals without skeletal diseases (Crawford et al., 2003; Dall'Ara et al., 2012) and with osteoporosis (Imai, 2015). We have recently applied such FE modelling using the ANTELOPE trial dataset reported here, using phantom based calibration to investigate the effects of androgen deprivation therapy on vertebral biomechanics (Gibson et al., 2025). In that study we found that simple BMD changes due to treatment did not capture the full extent of the degradation in vertebral strength evidenced in the FE models. Therefore it remains to be seen whether the similar correlation between phantom and phantomless calibration shown here, and elsewhere (Bartenschlager et al., 2022), will propagate through into resulting predictions of fracture risk and thus further study in this area is required.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by South Yorkshire Research Ethics Committee (IRAS ID206171). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

FG: Writing – review and editing, Methodology, Writing – original draft, Investigation, Visualization, Data

curation, Validation. ZD: Formal Analysis, Data curation, Writing – original draft, Visualization, Writing – review and editing. MP: Supervision, Writing – review and editing, Investigation, Resources, Validation, Data curation, Project administration, Methodology, Formal Analysis. CH: Visualization, Data curation, Validation, Methodology, Investigation, Writing – review and editing, Formal Analysis. JB: Resources, Investigation, Project administration, Data curation, Validation, Writing – review and editing, Writing – original draft, Funding acquisition, Conceptualization, Supervision, Visualization, Methodology. XL: Writing – review and editing, Methodology, Supervision, Visualization, Resources, Validation, Project administration. ED'A: Project administration, Resources, Visualization, Validation, Methodology, Conceptualization, Supervision, Writing – review and editing. SV: Writing – original draft, Visualization, Resources, Conceptualization, Writing – review and editing, Project administration, Supervision, Funding acquisition, Investigation, Methodology, Validation.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was partially supported by the Weston Park Cancer Charity, Sheffield (Grant number: CA133), by the Engineering and Physical Sciences Research Council (EPSRC) Frontier Multisim Grant (EP/K03877X/1 and EP/S032940/1) and by EPSRC grant (EP/Y001842/1) (SWV). ZD was supported via a China Scholarship Council PhD studentship in collaboration with Queen Mary University of London. This work forms part of the research portfolio of the National Institute for Health Research Barts Biomedical Research Centre (#NIHR203330).

References

- Bartenschlager, S., Dankerl, P., Chaudry, O., Uder, M., and Engelke, K. (2022). BMD accuracy errors specific to phantomless calibration of CT scans of the lumbar spine. *Bone* 157, 116304. doi:10.1016/j.bone.2021.116304
- Budoff, M. J., Malpeso, J. M., Zeb, L., Gao, Y. L., Li, D., Choi, T.-Y., et al. (2013). Measurement of phantomless thoracic bone mineral density on coronary artery calcium CT scans acquired with various CT scanner models. *Radiology* 267, 830–836. doi:10.1148/radiol.13111987
- Burghardt, A. J., Buie, H. R., Laib, A., Majumdar, S., and Boyd, S. K. (2010). Reproducibility of direct quantitative measures of cortical bone microarchitecture of the distal radius and tibia by HR-pQCT. *Bone* 47, 519–528. doi:10.1016/j.bone.2010.05.034
- Carpenter, R. D., Saeed, I., Bonaretti, S., Schreck, C., Keyak, J. H., Streeper, T., et al. (2014). Inter-scanner differences in *in vivo* QCT measurements of the density and strength of the proximal femur remain after correction with anthropomorphic standardization phantoms. *Med. Eng. Phys.* 36, 1225–1232. doi:10.1016/j.medengphy.2014.06.010
- Costa, M. C., Éltes, P., Lazary, A., Varga, P. P., Viceconti, M., and Dall'Ara, E. (2019). Biomechanical assessment of vertebrae with lytic metastases with subject-specific finite element models. *J. Mech. Behav. Biomed. Mat.* 98, 268–290. doi:10.1016/j.jmbm.2019.06.027
- Crawford, R. P., Cann, C. E., and Keaveny, T. M. (2003). Finite element models predict *in vitro* vertebral body compressive strength better than quantitative computed tomography. *Bone* 33, 744–750. doi:10.1016/S8756-3282(03)00210-2
- Dall'Ara, E., Pahr, D., Varga, P., Kainberger, F., and Zysset, P. (2012). QCT-based finite element models predict human vertebral strength *in vitro* significantly better than simulated DEXA. *Osteoporos. Int.* 23, 563–572. doi:10.1007/s00198-011-1568-3
- Duvaferrier, R., Valence, M., Patrat-Delon, S., Brillet, E., Niederberger, E., Marchand, A., et al. (2013). Current role of CT and whole body MRI in multiple myeloma. *Diagn. Interv. Imaging* 94, 169–183. doi:10.1016/j.diii.2012.12.001
- Engelke, K., van Rietbergen, B., and Zysset, P. (2016). FEA to measure bone strength: a review. *Clin. Rev. Bone Min. Metab.* 14, 26–37. doi:10.1007/s12018-015-9201-1
- Genant, H. K., and Boyd, D. (1977). Quantitative bone mineral analysis using dual energy computed tomography. *Invest. Radiol.* 12, 545–551. doi:10.1097/00004424-197711000-00015
- Gibson, F., Brown, J., Li, X., Dall'ara, E., and Verbruggen, S. (2025). Altered vertebral biomechanical properties in prostate cancer patients following androgen deprivation therapy. *Bone*, 195. doi:10.1016/j.bone.2025.117465
- Gleeson, T. G., Moriarty, J., Shortt, C. P., Gleeson, J. P., Fitzpatrick, P., Byrne, B., et al. (2009). Accuracy of whole-body low-dose multidetector CT (WBLDCT) versus skeletal survey in the detection of myelomatous lesions, and correlation of disease distribution with whole-body MRI (WBMRI). *Skelet. Radiol.* 38, 225–236. doi:10.1007/s00256-008-0607-4
- Handforth, C., Paggiosi, M. A., Jacques, R., Gossiel, F., Eastell, R., Walsh, J. S., et al. (2024). The impact of androgen deprivation therapy on bone microarchitecture in men with prostate cancer: a longitudinal observational study (the ANTELOPE study). *J. Bone Oncol.* 47, 100611. doi:10.1016/j.jbo.2024.100611
- Imai, K. (2015). Computed tomography-based finite element analysis to assess fracture risk and osteoporosis treatment. *World J. Exp. Med.* 5, 182–187. doi:10.5493/wjem.v5.i3.182
- Kabayel, D. D. (2016). The role of quantitative computed tomography and magnetic resonance imaging in diagnosis and follow-up of osteoporosis: a review. *Turk. J. Phys. Med. Rehabil. Fiz. Tip. Ve Rehabil. Derg.* 62, 288–295. doi:10.5606/tftrd.2016.33407
- Lee, D. C., Hoffmann, P. F., Kopperdahl, D. L., and Keaveny, T. M. (2017). Phantomless calibration of CT scans for measurement of BMD and bone strength—Inter-operator reanalysis precision. *Bone* 103, 325–333. doi:10.1016/j.bone.2017.07.029
- Mahnken, A. H., Wildberger, J. E., Gehbauer, G., Schmitz-Rode, T., Blaum, M., Fabry, U., et al. (2002). Multidetector CT of the spine in multiple myeloma: comparison with MR imaging and radiography. *Am. J. Roentgenol.* 178, 1429–1436. doi:10.2214/ajr.178.6.1781429

Acknowledgements

We would like to thank all the ANTELOPE trial participants, and the Weston Park Cancer charity for funding the ANTELOPE trial.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Matsumoto, T., Ohnishi, I., Bessho, M., Imai, K., Ohashi, S., and Nakamura, K. (2009). Prediction of vertebral strength under loading conditions occurring in activities of daily living using a computed tomography-based nonlinear finite element method. *Spine (Phila. pa. 1976)* 34, 1464–1469. doi:10.1097/BRS.0b013e3181a55636
- Michalski, A. S., Besler, B. A., Michalak, G. J., and Boyd, S. K. (2020). CT-based internal density calibration for opportunistic skeletal assessment using abdominal CT scans. *Med. Eng. Phys.* 78, 55–63. doi:10.1016/j.medengphy.2020.01.009
- Millner, M. R., Payne, W. H., Waggener, R. G., McDavid, W. D., Dennis, M. J., and Sank, V. J. (1978). Determination of effective energies in CT calibration. *Med. Phys.* 5, 543–545. doi:10.1118/1.594488
- Morgan, E. F., Bayraktar, H. H., and Keaveny, T. M. (2003). Trabecular bone modulus–density relationships depend on anatomic site. *J. Biomech.* 36, 897–904. doi:10.1016/S0021-9290(03)00071-X
- Pickhardt, P. J., Pooler, B. D., Lauder, T., del Rio, A. M., Bruce, R. J., and Binkley, N. (2013). Opportunistic screening for osteoporosis using abdominal computed tomography scans obtained for other indications. *Ann. Intern. Med.* 158, 588–595. doi:10.7326/0003-4819-158-8-201304160-00003
- Pickhardt, P. J., Bodeen, G., Brett, A., Brown, J. K., and Binkley, N. (2015). Comparison of femoral neck BMD evaluation obtained using Lunar DXA and QCT with asynchronous calibration from CT colonography. *J. Clin. Densitom.* 18, 5–12. doi:10.1016/j.jocd.2014.03.002
- Rasband, W. (1997). *ImageJ*. U. S. Natl. Institutes Heal.
- Saffarzadeh, M., Hightower, R. C., Talton, J. W., Miller, A. N., Stitzel, J. D., and Weaver, A. A. (2016). Multicenter analysis of CIREN occupant lumbar bone mineral density and correlation with age and fracture incidence. *Traffic Inj. Prev.* 17, 34–41. doi:10.1080/15389588.2016.1203429
- Schileo, E., and Taddei, F. (2021). Finite element assessment of bone fragility from clinical images. *Curr. Osteoporos. Rep.* 19, 688–698. doi:10.1007/s11914-021-00714-7
- Schileo, E., Dall'Ara, E., Taddei, F., Malandrino, A., Schotkamp, T., Baleani, M., et al. (2008). An accurate estimation of bone density improves the accuracy of subject-specific finite element models. *J. Biomech.* 41, 2483–2491. doi:10.1016/j.jbiomech.2008.05.017
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi:10.1038/nmeth.2089
- Schwaiger, B. J., Kopperdahl, D. L., Nardo, L., Facchetti, L., Gersing, A. S., Neumann, J., et al. (2017). Vertebral and femoral bone mineral density and bone strength in prostate cancer patients assessed in phantomless PET/CT examinations. *Bone* 101, 62–69. doi:10.1016/j.bone.2017.04.008
- Weaver, A. A., Beavers, K. M., Hightower, R. C., Lynch, S. K., Miller, A. N., and Stitzel, J. D. (2015). Lumbar bone mineral density phantomless computed tomography measurements and correlation with age and fracture incidence. *Traffic Inj. Prev.* 16, S153–S160. doi:10.1080/15389588.2015.1054029
- Winterbottom, A. P., and Shaw, A. S. (2009). Imaging patients with myeloma. *Clin. Radiol.* 64, 1–11. doi:10.1016/j.crad.2008.07.006