Osteoporos Int (1999) 9:38–44

Osteoporosis International

Original Article

Development and Evaluation of a Phantom for Morphometric X-ray Absorptiometry

S. A. Steel[1], J. A. Thorpe[1], R. Walker[2], S. Howey[2] and C. M. Langton[3]

[1]Centre for Metabolic Bone Disease and [2]Department of Medical Physics, Royal Hull Hospitals Trust; and [3]Academic Department of Medical Physics, University of Hull, Hull, UK

Abstract. Morphometric X-ray absorptiometry (MXA) provides the potential to assess vertebral deformity using a technique with much lower radiation dose to the patient than standard radiographic procedures. MXA overcomes many other limitations such as cone beam distortion observed in conventional plane radiographs. A phantom has been designed to assess the accuracy of the MXA technique, to monitor long-term precision and to assess inter- and intra-operator variability. The phantom consists of two columns of 12 cylinders representing the vertebral bodies, one of regular components and one representing vertebral deformities. Each column may be inserted, as required, into a Perspex torso-mimicking block. Initial assessment on the Lunar Expert-XL demonstrates that the phantom provides image parameters reflecting those found clinically. Measurement of vertebral height was found to be consistently underestimated by 4.9%. Operator precision ranged from 0.6% for posterior height measurement to 1.0% for middle height measurement of the regular component column. The corresponding precision range for the column representing vertebral deformation was 0.6% (posterior) and 1.1% (middle). Analysis of 10 scans of each column by two independent operators demonstrated a few significant differences in height assessment confined to the 'thoracic' region of the regular column. However, inter-operator variability was found to increase with increasing complexity of vertebral shape producing several differences, particularly in posterior height assessment of the deformed column.

Correspondence and offprint requests to: S. A. Steel, Centre for Metabolic Bone Disease, Royal Hull Hospitals Trust, Hull UK.

Keywords: Morphometric X-ray absorptiometry; Osteoporosis; Phantom; Quality assurance; Spine; Vertebral morphometry

Introduction

Osteoporosis-related fractures usually occur in the wrist, hip or spine. Those of the wrist and hip are usually readily confirmed using standard X-ray procedures, but those of the spine rarely present for clinical diagnosis. Vertebral deformity fractures are identified as a change in vertebral shape and may be classified as wedge, biconcave or crush and as grade 1, 2 or 3 [1,2] (Fig. 1). The grading may be based on percentage reduction in anterior, middle and/or posterior height [1], deviation from a defined normal range [2] or as a ratio of height to predicted height based on that of adjacent vertebrae [3]. Generally a 25% reduction in height or 3 SD from normal or predicted values is used as the cut-off point for defining a fracture.

The standard method of assessing vertebral deformity using lateral radiographs of the lumbar and thoracic spine is associated with a radiation dose to the patient of about 800 μSv [4]. Two exposures are necessary due to differences in attenuation in the lumbar and thoracic regions. The images produced are prone to magnification and distortion due to the cone beam geometry. The consequent elliptical appearance of the upper and lower borders of each vertebra and the relative distortion of vertebral height can make measurement problematic and therefore assessment of vertebral compression difficult. Patients are placed in the decubitus position for this technique making positioning difficult and repeat films

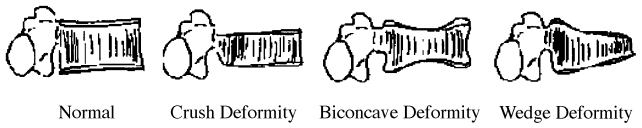

Fig. 1. Classification of vertebral deformities.

frequent, which adds to the radiation dose. Interpretation of lateral spine radiographs is generally a visual semi-quantitative assessment by a radiologist. More quantitative assessment may be performed by digitizing the radiograph and obtaining the anterior, posterior and mid-vertebral height, and several studies have investigated the correlation with readings performed by radiologists [5,6].

The recent development of morphometric X-ray absorptiometry (MXA) potentially overcomes some of the limitations of conventional plane radiographs. One of the machines providing this capability is the Lunar Expert-XL (Lunar, Madison, WI). This uses a fan beam of X-rays and a solid state detector. The X-ray tube and detector assembly are mounted on a C-arm which may be rotated to enable lateral imaging of the spine with the patient in a supine position, hence reducing distortion due to vertebral sagging in the lumbar region sometimes seen on conventional X-ray images. The C-arm scans axially from the level of L5 up to T4 producing a single image of the spine, making identification of vertebral levels easier. The superior to inferior distortion with this method should be negligible and any anteroposterior (AP) magnification should not interfere with assessment of vertebral height. We have assessed the effective dose from this technique to be 71 μSv [7], about 9% of that for conventional lateral spine radiographs. However, due to the lower resolution of the Expert-XL of approximately 1 line pair per millimeter (measured on our system), compared with the 3.5 line pairs per millimeter of conventional radiography [8], X-ray images are considered mandatory where other pathology is suspected.

A phantom is required to assess the accuracy of the MXA technique, to monitor long-term precision and to assess inter- and intra-operator variability. Currently there is no suitable phantom available. The aluminum spine phantom supplied with the machine is in the form of a step wedge and designed to measure accuracy and precision of AP bone mineral density (BMD) only. The more anthropomorphic Hologic spine phantom (Hologic, Waltham, MA) is similarly designed for assessment of AP BMD and, although when viewed laterally has the appearance of normal vertebrae, the components are of uniform density and would not provide a clinically representative image for morphometric analysis. Also, only four vertebrae (L1 to L4) are represented in the phantom. The European Spine Phantom simulates only L1 to L3 [9] and Felsenberg et al. [8] found the image quality obtained not entirely suitable for morphometry. All three of these phantoms have insufficient vertebrae for a full MXA procedure.

A phantom for use in MXA should be constructed of materials with similar attenuating properties to bone and soft tissue at the X-ray energies used. The design of the phantom should reflect clinically observed vertebral shape and range of deformities. It should enable validation and monitoring of the dependent parameters for MXA. These include mechanical parameters such as the alignment of the X-ray tube and detector, rotational position and longitudinal movement of the C-arm and registration of the image. Misalignment of the X-ray tube, in addition to affecting image quality, may result in the X-ray beam becoming non-parallel to the endplates of the vertebrae. This would result in an overestimate of vertebral height on morphometric images. The same effect would occur with misalignment of the C-arm relative to the central axis of the bed. If alignment changes with time, it may affect longitudinal assessment of vertebral deformity. Changes in the speed of movement of the C-arm during the scan procedure or in the system's image registration could result in misdiagnosis of vertebral collapse. For example, if the synchronization ratio between image position and actual position is correct at the beginning of the scan (1:1) but changes during the scan to a ratio of 1:2 at the end, then the thoracic vertebrae will appear foreshortened by a factor of 2.

Reproducibility of bed height should have minimal effect on vertebral height assessment but may result in distortion of the vertebral shape relative to a previous scan. Finally, in cases of vertebral deformity, accuracy confirmation is required, i.e., that the degree of wedging or collapse determined from the anterior/posterior or anterior/middle height ratios correlates with the true values.

Materials and Methods

Design

The phantom consists of three components: two aluminum and Perspex insert columns and a Perspex block with a drilled core (Fig. 2). The columns are constructed of 12 cylinders of aluminum, with phosphor-bronze endplates and a Perspex core, interspersed with Perspex disks. Aluminum was chosen for its bone-equivalent attenuation properties at the X-ray energies used and phosphor-bronze for its higher density and machinability to achieve thin endplates. Sharply defined endplates are required to maximize image clarity for the assessment of precision. The thicknesses of aluminum and phosphor-bronze were determined experimentally to give a similar morphometric image appearance as seen clinically.

The first column (column 1) is designed to check alignment, positioning and mechanical movements of the bed and C-arm. The cylinders are of the same height, 28 (± 0.05) mm, including the endplate thickness of 0.05 mm (± 0.005 mm), and are separated by an 'inter-vertebral space' of 7 mm. The second column (column

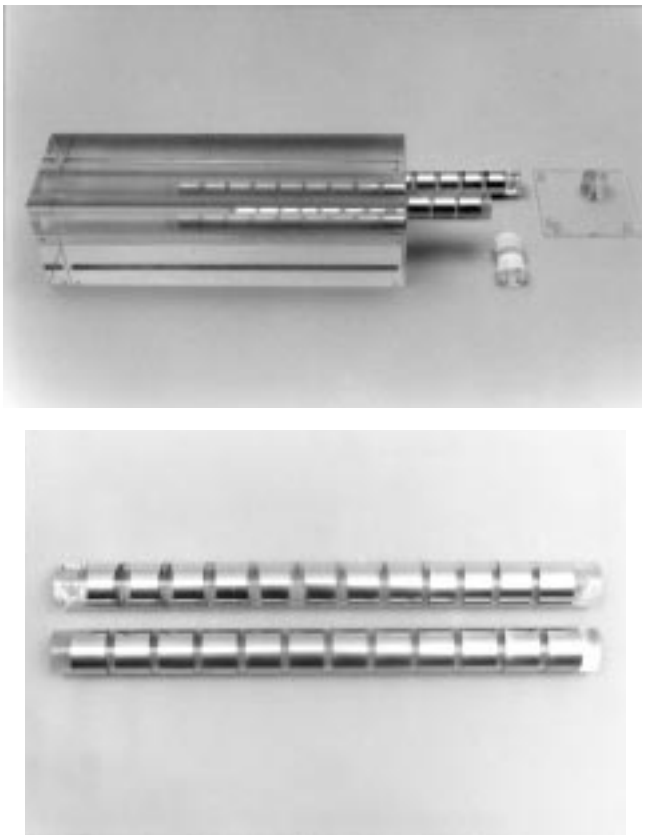

Fig. 2. a Phantom for use in morphometric X-ray absorptiometry. **b** Details of columns.

2) consists of three sets of cylinders: four of decreasing height, four of increasing anterior wedge and four of increasing biconcavity. The types and degree of deformities represent those found in the early stages of established osteoporosis, as described by others [2,10]. More severe deformities were not represented as diagnosis in these cases is generally beyond dispute. Each column may be inserted into the Perspex torso-mimicking block as required. Locating pins enable column 2 to be positioned at two predetermined angles relative to the plane of the X-ray beam, allowing the simulation of vertebral column rotation. All components of the phantom were carefully machined to ensure no air gaps, which could interfere with image quality by reducing X-ray attenuation in the 'soft tissue' regions.

The Perspex block incorporates radio-opaque markers at each end for monitoring X-ray beam alignment, C-arm position and bed height. An aluminum strip is incorporated posterior to the central core to simulate the posterior processes of the vertebrae.

Measurement

The phantom was placed along the central axis of the bed, mattress removed, with accurate positioning assured by aligning the locating markers with the positioning lines on the bed. Scanning was performed using the standard Lunar Expert-XL MXA acquisition with a bed height of –15 cm and scan length of 49 cm. Each column was scanned 10 times with repositioning between scans. Analysis was performed using the standard semi-automated MXA technique with manual adjustment of endplate markers as required, with points being longitudinally adjusted to the middle of the brightest pixel rows of the vertebral endplates, as currently recommended in the manufacturer's documentation. The scans were analyzed independently by two medical technical officers following the protocol recommended by the manufacturers. An image magnification of 300% was used to aid manual adjustment of the markers during analysis. To prevent fatigue errors, half the scans were analyzed from inferior to superior and half from superior to inferior.

Statistical Analysis

Analysis was performed using Microsoft Excel (version 5.0). Precision is expressed as percent coefficient of variation (CV %) with incorporation of RMS values for standard deviation [11]. Inter-operator variability was determined using Student's paired *t*-test, with a p value of less than 0.05 considered significant.

Results

The mean BMD of the 'vertebrae', assessed using a standard AP spine scan, was found to be 0.97 g/cm^2. MXA images of the two columns achieved with the Lunar Expert-XL are shown in Fig. 3. Using the semi-

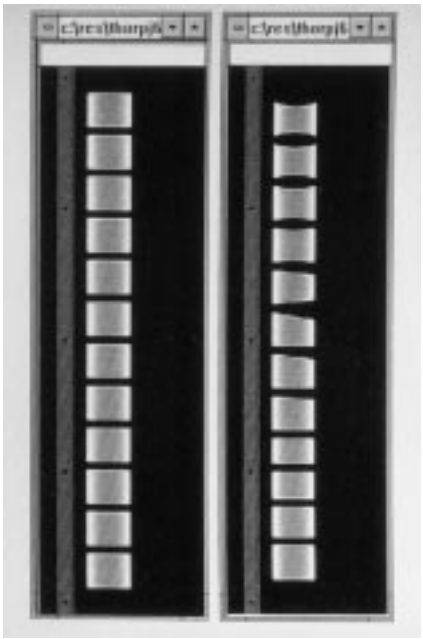

Fig. 3. Images of column 1 (left) and column 2 (right) obtained using the Lunar Expert-XL MXA.

automated analysis it was found that the software did not place the circular locating markers in the centre of the 'vertebrae'. The markers were generally placed anterior to the vertebrae and were incorrectly spaced, particularly in the upper part of the phantom corresponding to the thoracic region (Fig. 4). Manual adjustment was hence required to centralize the markers. The software then uses edge detection algorithms to locate the endplates and place identification points at the corners of each vertebra and in the centre of each endplate. Although the images are clearer than those often seen in clinical practice, the point placement was inaccurate and required manual adjustment to all 'vertebrae', again particularly in the 'thoracic' region. However, in our

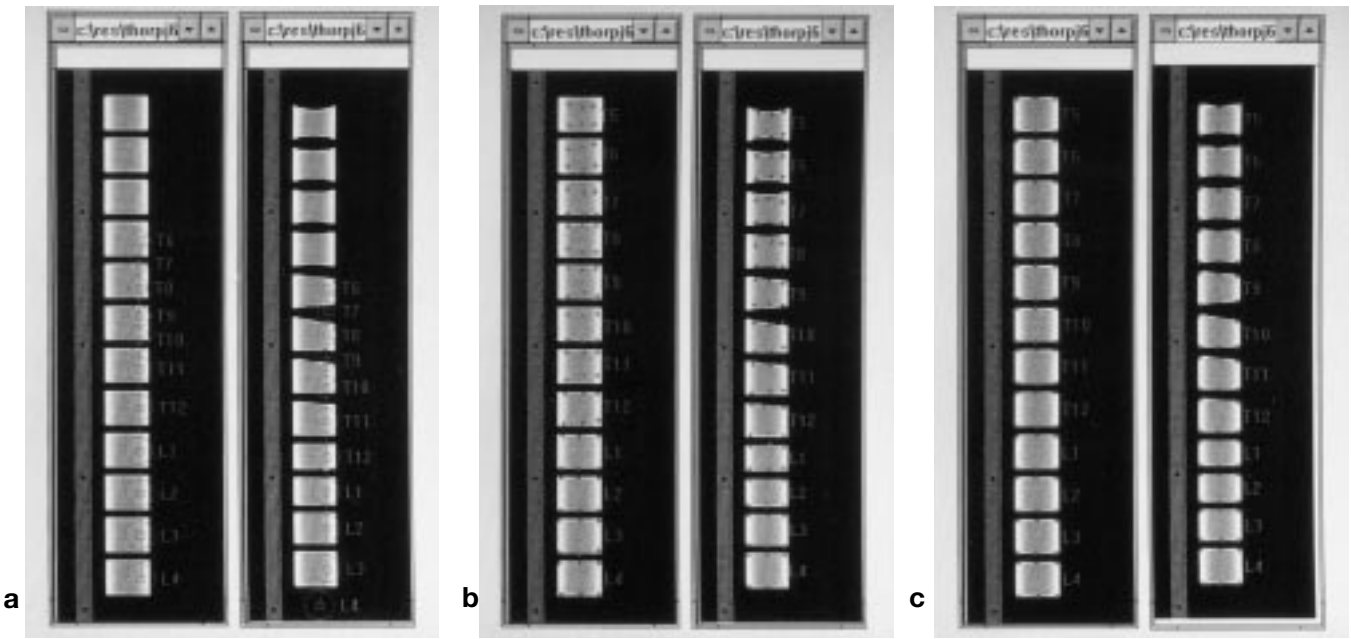

Fig. 4. Stages during MXA standard analysis showing: **a** automatic placement of circular locating markers, **b** subsequent automatic placement of endplate points and **c** Manually adjusted endplate points.

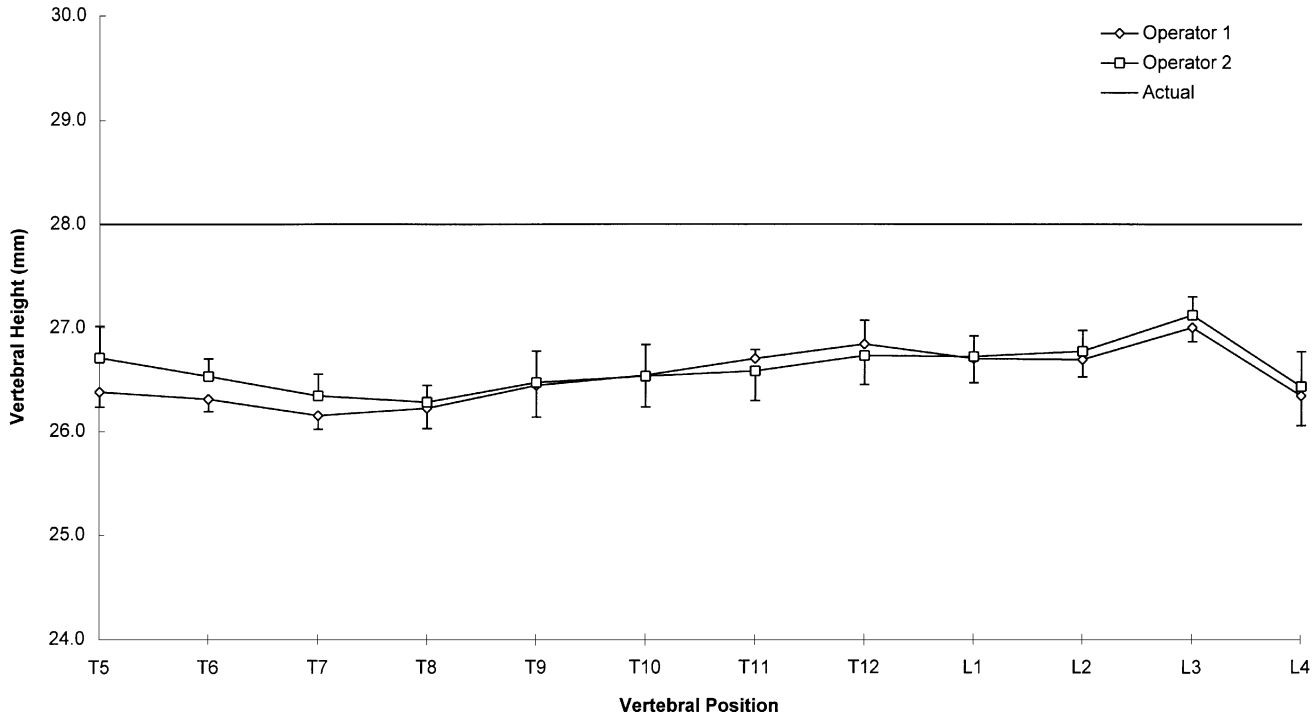

Fig. 5. Measured heights (mean ± SD) for each component of column 1 compared with actual heights. Results shown are the mean of anterior, middle and posterior heights from 10 scans performed using the Lunar Expert-XL MXA.

experience, the Lunar semi-automated analysis technique makes similar errors with clinical images, frequently misidentifying vertebrae and misplacing identification points.

The mean heights (anterior, middle and posterior) of column 1 were on average 1.4 mm (4.9 %) ± 0.8 mm less than the actual heights (Fig. 5). The height and precision of anterior, mid and posterior height for each vertebra are shown in Table 1. There were few significant differences in height assessment between the operators for column 1, with those that were observed tending to be associated with the upper thoracic vertebrae.

Table 1. Mean and precision (CV%) of measured heights determined by each operator for column 1

	n	Operator 1	Operator 2	Mean difference
Anterior height				
T5	10	26.3 (1.08)	26.9 (1.15)	–0.56**
T6	10	26.5 (0.51)	26.4 (0.62)	0.09
T7	10	26.5 (0.70)	26.5 (0.65)	–0.06
T8	10	25.9 (0.95)	25.9 (0.75)	0.07
T9	10	26.8 (0.99)	26.6 (0.85)	0.19
T10	10	26.5 (0.31)	26.3 (0.92)	0.16
T11	10	26.8 (1.18)	26.6 (1.40)	0.27
T12	10	26.9 (1.29)	26.5 (0.91)	0.37*
L1	10	27.1 (0.82)	27.1 (0.47)	–0.08
L2	10	26.4 (0.55)	26.5 (0.95)	–0.06
L3	10	27.1 (0.42)	27.2 (0.63)	–0.05
L4	10	26.5 (0.33)	26.5 (0.90)	0.03
Overall	120	26.6 (0.83)	26.6 (0.88)	0.03
Middle height				
T5	10	26.4 (1.16)	27.1 (1.22)	–0.73**
T6	10	26.4 (0.57)	26.5 (0.70)	–0.05
T7	10	26.5 (0.75)	26.7 (0.99)	–0.15
T8	10	26.0 (1.07)	26.1 (0.83)	–0.04
T9	10	26.9 (0.99)	26.8 (1.38)	0.07
T10	10	26.6 (0.33)	26.6 (1.40)	–0.02
T11	10	26.9 (1.30)	26.8 (0.87)	0.12
T12	10	27.0 (1.19)	26.7 (0.82)	0.30
L1	10	27.2 (0.69)	27.3 (0.85)	–0.07
L2	10	26.5 (0.55)	26.6 (0.87)	–0.06
L3	10	27.2 (0.38)	27.2 (0.74)	–0.01
L4	10	26.6 (0.80)	26.6 (1.45)	–0.06
Overall	120	26.7 (0.87)	26.7 (1.04)	–0.06
Posterior height				
T5	10	26.4 (1.01)	26.7 (1.39)	–0.33*
T6	10	26.3 (0.46)	26.5 (0.71)	–0.22*
T7	10	26.2 (0.37)	26.3 (0.60)	–0.19**
T8	10	26.2 (0.59)	26.3 (0.74)	–0.06
T9	10	26.4 (0.54)	26.5 (0.84)	–0.03
T10	10	26.5 (0.36)	26.5 (0.82)	0.01
T11	10	26.7 (0.75)	26.6 (0.53)	0.12
T12	10	26.8 (0.83)	26.7 (0.31)	0.11
L1	10	26.7 (0.64)	26.7 (0.39)	–0.02
L2	10	26.7 (0.37)	26.8 (0.50)	–0.08
L3	10	27.0 (0.49)	27.1 (0.49)	–0.12
L4	10	26.3 (0.37)	26.4 (1.03)	–0.09
Overall	120	26.5 (0.60)	26.6 (0.75)	–0.08**

The actual anterior, middle and posterior heights in all cases is 28 mm. The phantom was scanned 10 times with repositioning on a Lunar Expert-XL and analyzed using semi-automated MXA analysis
*$p < 0.05$; **$p < 0.01$ using paired *t*-test.

For column 2, expected ratios were calculated for each of the deformities. The degree of crush deformity was determined as a ratio of the posterior height to the mean measured height of L4 (non-compressed), for wedge deformities as anterior/posterior height and for biconcave deformities as middle/posterior height. The crush deformities (L4 to L1) were generally overestimated by 2% and the wedge deformities underestimated by 2%. On average, biconcavity was underestimated by 8% (Fig. 6). Mean and precision of measured heights for column 2 are shown in Table 2. There was found to be an increase in inter-operator variability, as would be expected, with the increased complexity of shape. Several significant

Table 2. Mean and precision (CV%) of measured heights determined by each operator for column 2

	n	Actual	Operator 1	Operator 2	Mean difference
Anterior height					
T5	10	28.0	24.5 (0.84)	24.5 (0.98)	0.00
T6	10	28.0	24.5 (0.77)	24.8 (0.94)	–0.31*
T7	10	28.0	25.2 (0.84)	25.3 (0.95)	–0.08
T8	10	28.0	24.9 (0.91)	25.2 (0.75)	–0.22*
T9	10	21.0	19.2 (1.04)	19.6 (1.09)	–0.39**
T10	10	22.4	20.0 (0.76)	20.5 (1.03)	–0.46**
T11	10	23.8	21.8 (0.76)	22.2 (1.98)	–0.41*
T12	10	25.2	23.4 (1.03)	23.8 (0.68)	–0.35**
L1	10	21.0	19.5 (0.64)	19.5 (0.65)	–0.07
L2	10	22.4	21.5 (0.53)	21.5 (0.88)	–0.06
L3	10	25.2	23.7 (0.53)	23.9 (0.68)	–0.20**
L4	10	28.0	27.0 (0.44)	27.1 (0.43)	–0.15*
Overall	120	25.1	22.9 (0.78)	23.2 (0.99)	–0.23**
Middle height					
T5	10	21.0	20.2 (0.71)	20.3 (1.42)	–0.02
T6	10	22.4	21.6 (0.90)	21.5 (1.05)	0.04
T7	10	23.8	23.0 (1.12)	23.2 (1.23)	–0.18
T8	10	25.2	23.7 (0.36)	23.7 (1.04)	0.01
T9	10	24.5	21.7 (1.11)	22.0 (1.09)	–0.30*
T10	10	25.2	22.1 (0.67)	22.2 (1.08)	–0.08
T11	10	25.9	23.6 (0.63)	23.8 (0.80)	–0.26**
T12	10	26.6	24.4 (0.64)	24.6 (1.50)	–0.21
L1	10	21.0	19.4 (0.50)	19.7 (1.30)	–0.22*
L2	10	22.4	21.6 (0.45)	21.6 (0.70)	–0.00
L3	10	25.2	23.7 (0.40)	23.9 (0.86)	–0.17*
L4	10	28.0	27.2 (0.59)	27.2 (0.36)	0.07
Overall	120	24.3	22.7 (0.72)	22.8 (1.08)	–0.11**
Posterior height					
T5	10	28.0	24.4 (0.59)	24.7 (1.48)	–0.27
T6	10	28.0	24.6 (0.38)	24.9 (1.07)	–0.34**
T7	10	28.0	24.8 (0.47)	25.1 (0.75)	–0.26**
T8	10	28.0	24.8 (0.50)	25.2 (0.82)	–0.37**
T9	10	28.0	24.8 (0.43)	25.2 (0.81)	–0.46**
T10	10	28.0	24.9 (0.46)	25.4 (0.72)	–0.50**
T11	10	28.0	25.2 (0.56)	25.6 (0.55)	–0.43**
T12	10	28.0	25.6 (0.87)	25.8 (0.81)	–0.22
L1	10	21.0	19.3 (0.60)	19.6 (0.92)	–0.29**
L2	10	22.4	21.3 (0.84)	21.5 (0.54)	–0.25**
L3	10	25.2	23.6 (0.45)	23.8 (0.62)	–0.25**
L4	10	28.0	26.9 (0.64)	27.0 (0.40)	–0.13*
Overall	120	26.7	24.2 (0.59)	24.5 (0.84)	–0.31**

Actual heights are given in the table. The column was scanned and analyzed as per column 1.
*$p < 0.05$; **$p < 0.01$ using paired *t*-test.

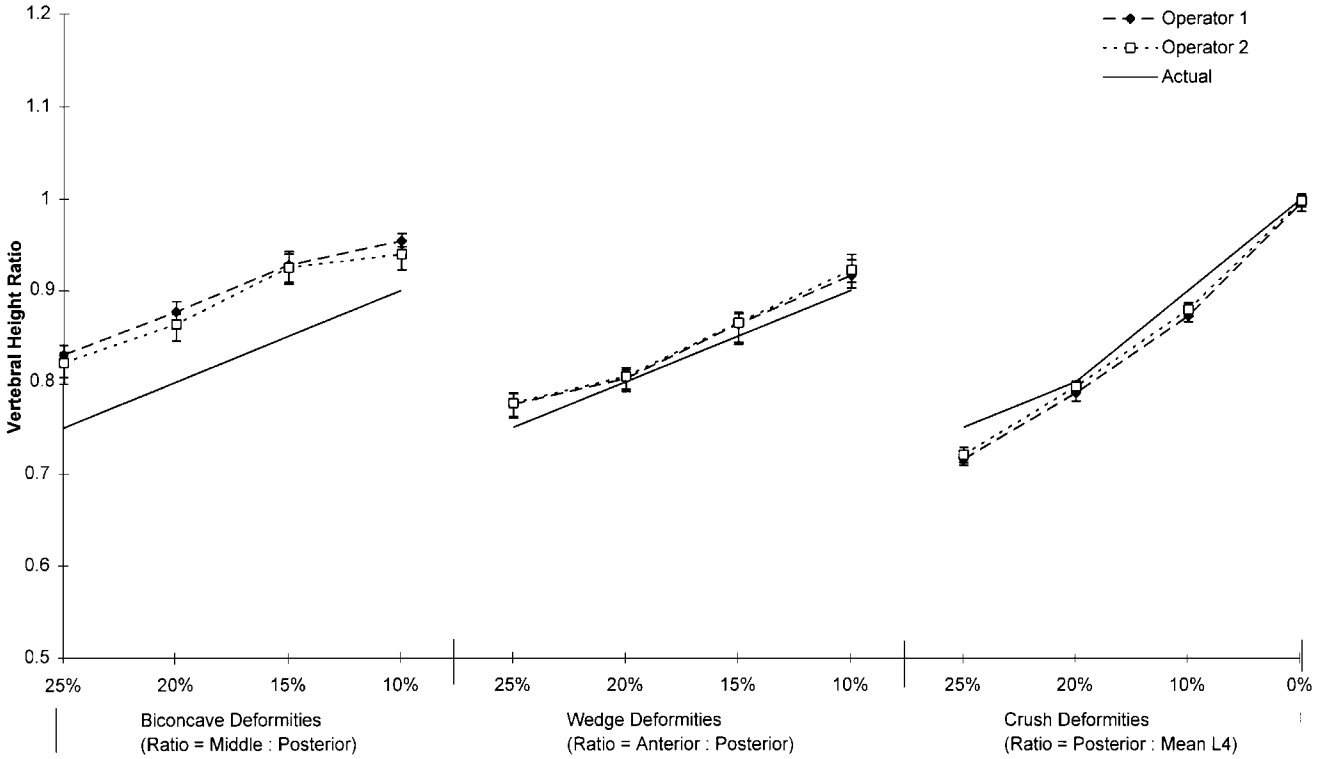

Fig. 6. Measured height ratios (mean + SD) form MXA of column 2 compared with actual: middle/posterior (biconcavities), anterior/posterior (wedges), posterior/posterior height L4 (crush).

differences were found in vertebral height measurement, particularly in posterior height assessment.

Discussion

A phantom has been designed that is suitable for determining the accuracy and precision of MXA vertebral height assessment and for monitoring long-term equipment stability. The phantom was designed for use on the Lunar Expert-XL but should be suitable for similar MXA machines and could provide a comparison with radiographic morphometric procedures. This is not a truly anthropomorphic phantom and has not been designed to validate the aspects of the proprietary software concerned with comparison to reference data. However, the phantom could be used to investigate inter-machine variability, which is a critical factor where reference data are to be accumulated or multicenter studies conducted.

The columns currently constructed have BMDs in the osteopenic range. MXA images obtained in osteoporotic subjects are generally of poorer quality, making analysis more difficult and reducing precision. However, the modular design of the phantom enables the construction and insertion of columns representing a range of densities and deformities, as found in clinical practice.

MXA of the phantom on the Lunar Expert-XL demonstrates incorrect placement of the circular locating markers by the semi-automated analysis. This suggests that the software is making assumptions about the shape of the vertebral column and relative dimensions of the vertebrae. Once the markers are manually adjusted to the correct location, the edge detection process places the endplate markers. The well-defined vertebral endplates on the image enabled accurate manual adjustment of the points where required.

Measurement of anterior, posterior and middle vertebral heights was consistently underestimated. This differs from the findings of Felsenberg et al. [8], who reported a mean overestimation of about 1.7% in vertebral height assessment on the Lunar Expert from MXA images of the European Spine Phantom. This may have been due to different scattering properties of the materials, but is more likely due to a difference in the protocols for the positioning of vertebral marker points. The observed underestimation of vertebral height may not be clinically significant, as the degree of vertebral deformity is generally determined relative to the heights of L2 to L4. However, this may have implications for the establishment of reference data. Normative data established by radiographic technique would have to be appropriately scaled. If reference data were to be accumulated on MXA machines, it would be important to ensure the comparability of results both within and between systems.

Mean precision of vertebral height measurement in vitro was found to be 0.84%. Using the method described by the British Standards Institution [12], a change in vertebral height of 2.38% would be required in

order to be reliably detected. This represents a mean change in vertebral height of 0.7 mm, which corresponds with the MXA image pixel size (0.7 mm × 0.5 mm) on the Lunar Expert-XL.

Increasing complexity of vertebral shape caused a decrease in accuracy and an increase in inter-operator variability. In particular, placement of points at the center of the endplate of a biconcavity was found to be least accurate, producing an underestimate of deformity. Also, location of the center of a wedged component is subjective and any AP displacement will directly influence height assessment.

The initial findings using this phantom suggest that the Lunar Expert-XL MXA technique provides a precise measure of vertebral height. Defined protocols for morphometric analysis may reduce operator variability. Although the absolute heights are underestimated, this is not considered detrimental as it is the ratio of heights that defines the morphometric status. If reference data are to be accumulated or multicenter studies conducted on these machines, such a phantom may be required to assess inter-system performance and provide cross-calibration.

Acknowledgements. The authors wish to acknowledge the excellent technical assistance of Mrs A Goodby, Mrs A Smith and Mrs D Langton, and to thank the local charity OSPREY for their financial support.

References

1. Genant HK, Wu CY, Van Kuijk C, Nevitt M. Vertebral fracture assessment using a semiquantitative technique. J Bone Miner Res 1993;8:1137–48.
2. Eastell R, Cedel SL,Wahner HW, Riggs BL, Melton LJ. Classification of vertebral fractures. J Bone Miner Res 1991; 6:207–15.
3. McCloskey EV, Spector TD, Eyres KS, et al. The assessment of vertebral deformity: a method for use in population studies and clinical trials. Osteoporos Int 1993;3:138–47.
4. Shrimpton PC, Wall BF, Jones DG, et al. A national survey of doses to patients undergoing a selection of routine X-ray examinations in English hospitals. National Radiological Protection Board NRPB-R200. London: HMSO, 1986.
5. Genant HK, Jergas M, Palermo L, et al. Comparison of semiquantitative visual and quantitative morphometric assessment of prevalent and incident vertebral fractures in osteoporosis. J Bone Miner Res 1996;11:984–96.
6. Leidig-Bruckner G, Genant HK, Minne HW, et al. Comparison of a semiquantitative and a quantitative method for assessing vertebral fractures in osteoporosis. Osteoporos Int 1994;4:154–61.
7. Steel SA, Baker AJ, Saunderson JR. An assessment of the radiation dose to patients and staff from a Lunar Expert-XL fan beam densitometer. Physiol Meas 1998;19:17–26.
8. Felsenberg D, Gowin W, Diessel E, Armbrust S, Mews J. Recent developments in DXA: quality of new DXA/MXA devices for densitometry and morphometry. Eur J Radiol 1995;20:179–84.
9. Kalender WA, Felsenberg D, Genant HK, Fischer M, Dequeker J, Reeve J. The European Spine Phantom: A tool for standardization and quality control in spinal bone mineral measurements by DXA and QCT. Eur J Radiol 1995;20:83–92.
10. Melton LJ III, Lane AW, Cooper C, Eastell R, O'Fallon WM, Riggs BL. Prevalence and incidence of vertebral deformities. Osteoporos Int 1993;3:113–9.
11. Gluer C, Blake G, Lu Y, Blunt BA, Jergas M, Genant HK. Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. Osteoporos Int 1995;5:262–70.
12. British Standards Institution. Guide for the determination of repeatability and reproducibility for a standard test method. In: BS5497 (1987), Precision of test methods, part 1. London: Bracknell, Technical Indexes Ltd, 1987:5–9.

Received for publication 11 June 1997
Accepted in revised form 16 April 1998