Calcif Tissue Int (1999) 64:191–199

Calcified
Tissue
International

Bone Density Reduction in Various Measurement Sites in Men and Women with Osteoporotic Fractures of Spine and Hip: The European Quantitation of Osteoporosis Study

H. Kröger,[1,2] M. Lunt,[3] J. Reeve,[2,3] J. Dequeker,[4] J. E. Adams,[5] J. C. Birkenhager,[6] M. Diaz Curiel,[7] D. Felsenberg,[8] L. Hyldstrup,[9] P. Kotzki,[10] A.-M. Laval-Jeantet,[11] P. Lips,[12] O. Louis,[13] R. Perez Cano,[14] C. Reiners,[15] C. Ribot,[16] P. Ruegsegger,[17] P. Schneider,[18] P. Braillon,[19] J. Pearson*

[1]Department of Surgery/Orthopaedics, Kuopio University Hospital, FIN-70210 Kuopio, Finland
[2]Department of Medicine, Addenbrooke's Hospital, Cambridge, UK
[3]University of Cambridge, Institute of Public Health, Cambridge, UK
[4]Arthritis and Metabolic Bone Disease Research Unit, K.U. Leuven, Belgium
[5]University of Manchester, Department of Diagnostic Radiology, Manchester, UK
[6]University Hospital Rotterdam, Department of Medicine III, Rotterdam, The Netherlands
[7]Fundation Jimenez Diaz, Servicio de Mineralomet. Un. Metabolica, Madrid, Spain
[8]Freie Universität Berlin, Klinikum Benjamin Franklin, Berlin, Germany
[9]Hvidovre Hospital Mineral and Metabolism Research Group, Department of Endocrinology, Hvidovre, Denmark
[10]University de Montpellier, Hopital Lapeyronie-Serv. de Med. Nucleaire, Montpellier, France
[11]Inserm U18, Laboratoire de Radiologie, Paris, France
[12]Academisch Ziekenhuis, Vrije University of Amsterdam, Department of Endocrinologie, Amsterdam, The Netherlands
[13]Akademisch Ziekenhuis, Vrije Universiteit Brussel, Department of Radiology, Brussels, Belgium
[14]University Hospital 'Virgin Macarena,' Departamento Medicine, Sevilla, Spain
[15]Clinic for Nuclear Medicine, University of Essen, Germany
[16]C.H.U. Purpan Service d'endocrinologie, U.F. Maladies Osseus et Metaboliques, Toulouse, France
[17]Institut für biomedizinische Technik, Universität und ETH Zurich, Switzerland
[18]Clinic for Nuclear Medicine, University of Wurtzburg, Germany
[19]Centre d'Absorptiometrie Osseuse et Department de Radiologie, Hospital E. Herriot, Lyon, France

Received: 21 May 1997 / Accepted: 24 June 1998

Abstract. We have measured bone mineral density (BMD) using dual X-ray absorptiometry (DXA) of the spine and hip, spinal quantitative computed tomography (QCTspi), and peripheral radial quantitative computed tomography (pQCTrad) in 334 spine and 51 hip fracture patients. The standardized hip and spine BMD for each patient was calculated and compared with the combined reference ranges published previously, each densitometer having been cross-calibrated with the prototype European Spine Phantom (ESPp) or the European Forearm Phantom (EFP).

Male and female fracture cases had similar BMD values after adjusting for body size, where appropriate. This suggests that the relationship between bone density (mass per unit volume) and fracture risk is similar between men and women. However, compared with age-matched controls, mean decreases in BMD ranged from 0.78 SD units (women with hip fracture, DXAspi) to 2.57 SD units (men with spine fractures, QCTspi).

The proportion of spine and hip fracture patients falling below the cutoff for osteoporosis (T-score <−2.5 SD) proposed by the World Health Organization (WHO) study group varied according to different BMD measurement procedures (range 18–94%). This finding suggests that the WHO definition requires different thresholds when used with non-DXA BMD measurement techniques.

Receiver operator characteristic (ROC) analysis was used to compare measurement techniques for their ability to discriminate between cases and controls. Among DXA sites, the proximal femur was preferred when evaluating generalized bone loss, particularly in elderly people. An additional spinal BMD measurement may add clinical value if spine fracture risk assessment has a high priority. Both axial and peripheral QCT techniques performed comparably to DXA in spinal osteoporosis, so investigators and clinicians may use any of the three technologies with similar degrees of confidence for the diagnosis of generalized or site-specific bone loss providing straightforward clinical guidelines are followed.

Key words: Dual X-ray absorptiometry — Quantitative computed tomography — Hip fracture — Spine fracture — Osteoporosis—Normal ranges.

Prospective studies have suggested that bone mineral density (BMD) is an important predictor of fracture [1, 2]. Furthermore, several studies have suggested that densitometry of a potential fracture site should provide the most accurate estimate of the risk of future fracture [3, 4]. The risk of a fracture is further increased independently after an osteoporotic fracture [5, 6]; however, the risk of a second fracture is critically dependent on BMD as well [7]. In order to offer a prognosis or to assess the effectiveness of treatments aimed at secondary fracture prevention, it is therefore important to know the degree of reduction in bone mass in patients who have already suffered a hip or a spine fracture. Whether it is necessary to measure both spine and hip for

Correspondence to: H. Kröger
* *Current address:* Pfizer Ltd, Sandwich, Kent, UK

prediction of the first fracture in an elderly population has been challenged [8]. It is not clear whether this holds for subsequent fractures in patients who have already suffered a fracture.

It seemed possible that there may be differences in BMD between male and female fracture patients as there are between normal men and women over the age of 50. Studies including both spine and hip fracture patients, enabling the comparison of BMD differences between the fracture groups, are scarce, as are previous comparisons of bone density in male as well as female patients with osteoporosis. Centers participating in the EU Concerted Action ''Quantitative Assessment of Osteoporosis'' (QAO) collected BMD data from spine and hip fracture patients with the aim of assembling a large database on cases of established osteoporosis in both sexes. The further aims were to develop guidelines relating to choice of measurement site and to assess the degree of reduction in BMD to be expected in different measurement sites in patients presenting with spine and hip fractures. The study design enabled us to compare the performance of trabecular and integral bone measurement techniques in cases presenting to specialist clinics in different European countries. Our observations are also relevant to the interpretation of the WHO study group guidelines on the diagnosis of established osteoporosis in these two fracture groups [9, 10].

Subjects and Methods

Patients

In the course of the QAO Concerted Action, investigators in 13 centers recruited patients referred by their family practitioners for specialist clinical assessment of suspected vertebral osteoporosis. In still another center (Berlin) subjects with radiologically defined vertebral deformities were recruited from a population-based epidemiological study, the European Vertebral Osteoporosis Study (EVOS) [11]. Two hundred and ninety-four acceptable dual-X-ray absorptiometry (DXA) records were obtained from spine fracture patients and in addition, six centers collected data from 51 hip fracture patients treated in their own centers. Seven centers provided 122 acceptable spinal quantitative computed tomography (QCT) records and five centers provided 97 acceptable peripheral QCT (pQCT) records from the spine fracture patients. Finally, three centers obtained five spinal QCT results and two centers provided fourteen radial pQCT results from hip fracture patients. For the purposes of the study, spinal osteoporosis was defined by the presence of a 20% or greater deformation of at least one vertebra, with no evidence of old epiphysitis as defined by Eastell et al. [12]. The spine fracture group included 14 patients with long-term corticosteroid treatment and three patients with thyroid excess. Otherwise they were classified as having osteoporosis due to sex hormone deficiency or primary osteoporosis.

No exclusion criteria, except local bone disease (e.g., invasive cancer, Paget's disease), were used in the recruitment of the hip fracture patients, who were normal referrals to participating hospitals. All hip fracture patients were measured postoperatively (3–14 days) before their discharge from the hospital. Owing to the modest number of hip fracture patients, no attempt was made to subdivide them into femoral neck and trochanteric fracture groups. Because of old age, many hip fracture patients had several chronic diseases typical of the elderly. There was no requirement that the hip fracture cases have radiological investigation of the spine to identify the presence of spine fractures and there was no overlap between the hip and spine fracture cases.

Measurement Procedures and Reference Ranges

Each center measured its subjects and patients using the procedures recommended by the individual DXA (Hologic, Lunar, Norland), QCT (IGE, Elscint, Siemens), and pQCT (Densiscan, Stratec) manufacturers and set out in their operations manuals. Table 1 shows the densitometers used in the study. Generally, for the lumbar spine, L2–L4 were measured and the results were averaged. In cases where there was one or more fractured lumbar vertebrae, results for the remaining unfractured vertebrae were analyzed. In hip fracture cases the contralateral proximal femur was measured. In previous work we have described how we calculated standardized densities from measured densities using the results from DXA equipment made by three different manufacturers [13]. We then presented the results from normal subjects of various ages for the hip [14] and spine [15].

Table 1. Densitometers used in the study and the number of centers

DXA densitometers	
Hologic QDR-1000	3
Hologic QDR-1000W	3
Hologic QDR-2000	
Lunar DPX	5
Lunar DPX-L	2
Norland XR26	2
QCT densitometers	
IGE 9000-series	3
Elscint	1
Siemens	4
pQCT densitometers	
Densiscan	1
Stratec	4

In the same way, the standardized hip and spine densities for each patient were calculated and compared with the combined reference ranges published previously.

Analogous procedures, including standardization with the ESP prototype phantom or the European Forearm Phantom (EFP), were adopted for calculating standardized densities from measured densities with the radial pQCT [16, 17] and spinal QCT data [13]. The normal data for spinal QCT are presented in this paper. QCT techniques used in this study have been described in previous publications [13, 16, 17].

Normal Ranges for Spinal QCT Measurements

Normal men and women were recruited to obtain normal reference ranges for subjects between 20 and 80 years of age. After excluding 63 women who reported diseases or treatments that may affect bone metabolism, as was done for our DXA normal ranges [15], 128 normal women aged 22–80 were included. They were measured on IGE, Siemens and Elscint machines which were standardized with the ESP prototype as previously described. Similarly, 68 men aged 26–81 years were available (after 9 exclusions) for generating normal range in men. The statistical procedures for establishing the normal range included tests for between-center differences and normality of distribution about regressions against age. For women, separate analyses were carried out for the whole data and for the postmenopausal women, with years since menopause as the independent variate. There was no significant skewness in the distribution of the residuals in the regressions and no significant tendency for variability to increase with age ($P > 0.4$), so the results are presented as simple linear regressions. The data used to generate Z-scores for the women patients were obtained from the regression equation relating bone density to age for all women.

Statistical Analysis

The BMD results are presented as both means with 95% confi-

Table 2. Normal results for QCT spine expressed in standarized BMD (mg/cm^3)

	Intercept at age 50 or at zero time since menopause ($mg.cm^{-3}$)	Regression coefficient ($mg.cm^{-3}.yr^{-1}$)	SD residual $mg.cm^{-3}$)
All women ($N = 128$)	126	2.06	28
Postmenopausal women ($N = 70$)	124	−1.90	29
Men ($N = 68$)	140	−1.47	28

dence intervals and as T- and Z-scores relative to the combined European reference ranges for each sex presented in previous publications [14, 15, 17]. The Z-score shows the patient's result as a difference from the mean of the age-matched controls divided by the standard deviation among these controls. The T-score refers to the peak bone mass of the young normal adults (20–40 years) and is calculated similarly to the Z-scores. Since the BMD values in the fracture groups were not always normally distributed, we chose the Mann-Whitney U-test to analyze the differences between the fracture groups and sexes. Analysis of covariance (weight and height as covariates) was used to test sex differences in BMD values. The between-center differences in BMD values were expressed as coefficients of variation (%CV). The proportions of fracture patients demonstrating T-scores equal to or lower than −2.5, which is the definition of severe osteoporosis in women according to the WHO study group [9], were calculated.

To investigate retrospectively the ability of the various measures of bone density to identify cases of fracture, nonparametric receiver operating characteristic curves (ROC) were produced [18]. The area under a ROC curve represents the probability that, given a randomly selected case and a randomly selected control, the case will have lower bone density than the control. Thus, an area of 1 represents perfect discrimination and an area of 0.5 represents no discrimination. Since the sample sizes were quite small, making the verification of distributional assumptions difficult, the nonparametric methods of Delong et al. [19] were used to calculate confidence intervals around the estimate of the area under the ROC curve, and to perform matched comparisons of the areas under two curves. Matched comparisons (i.e., comparisons in which only those subjects who had been measured by both methods were included) were used since they have far greater power to detect differences in area than unmatched comparisons, in the same way that a paired *t*-test is more powerful than an unpaired *t*-test when the data being tested is paired. Owing to the study design it was not possible to apply all measurement techniques to all patients. Direct comparisons between groups could only be done through the ROC analyses.

Results

Normal ranges for spinal QCT data were described by the regression equations shown in Table 2.

Fracture Group and Sex Differences in Standardized BMDs (Tables 3 and 4)

As expected, hip fracture patients were significantly older than spine fracture patients. For both hip and spine fractures the male cases were younger than the female.

In women, spinal BMDdxa was significantly lower in the patients with spine fractures than in those with hip fractures. In men, the equivalent difference was in the same direction but did not reach statistical significance. On the other hand, femoral neck BMDdxa was lower in female hip fracture patients than in female spine fracture patients. No significant differences in trochanteric BMDdxa between the fracture groups were found (Tables 3 and 4).

In the spine fracture group, the measured BMDdxa values were lower in women than in men (*P*-values for DXAspi, DXAneck, and DXAtroc were 0.002, 0.03, and 0.001, respectively). However, when these values were adjusted for weight and height, the differences no longer reached statistical significance (*P*-values from 0.136 to 0.634). Similarly, no differences in spinal BMDqct values between the sexes were found ($P = 0.629$) (Tables 3 and 4).

Z-scores

Mean Z-scores at all sites in both sexes and both fracture categories were significantly reduced compared with normals (Figs. 1 and 2, Tables 3 and 4) with the exception of DXAspi Z-score in female hip fracture cases. DXAspi Z-score was significantly lower in patients with spine fractures than in patients with hip fractures. There were no differences in proximal femoral Z-scores between the fracture groups (Tables 3 and 4). Men with spine fractures showed significantly lower QCTspi and pQCTrad Z-scores than women with spine fractures (Figs. 1 and 2). Since men and women with spinal fractures had similar spinal BMDqct values, this difference in Z scores primarily reflected the sex difference in QCT values resulting from normal postmenopausal bone loss in female controls.

Odds of a Case Having a Lower Bone Density Than His or Her Control According to Measurement Site

Statistical comparisons were based on the ROC analyses in groups of subjects who had paired measurements, that it was desirable to contrast (Table 5). DXAspi was a slightly better discriminator of spine fracture than DXAneck ($P = 0.07$) or DXAtroc ($P = 0.03$) in women (Fig. 3). In men, the areas under the ROC curves did not differ significantly (Fig. 4).

For male hip fracture patients the areas under the ROC curves were significantly lower for DXAspi than DXAneck ($P = 0.03$) or DXAtroc ($P = 0.04$). In female hip fracture cases the differences between the ROC curves did not reach statistical significance (Table 5).

When the discriminatory abilities of DXAspi and QCTspi for detecting spine fractures were compared, no significant differences in the areas under the ROC curves could be found in men ($P = 0.96$) or in women ($P = 0.08$).

Separate ROC analyses for QCTspi and pQCTrad measurements were performed in male and female spine fracture patients. As to QCTspi measurements, the areas under the ROC curve were 0.97 and 0.85 for male and female spine fracture patients, respectively. In pQCTrad the corresponding areas were 0.88 and 0.84 for male and female, respectively.

Definition of Severe Osteoporosis

In the following, ''severe osteoporosis'' is defined as a T-score of <2.5 SD units associated with at least one spine or hip fracture. Of the female spine fracture cases, 71.2% had a spinal BMDdxa below the −2.5 SD limit for young normal women, whereas only 33.8% and 24.9%, respectively, of the cases were classified as having severe osteo-

Table 3. Comparison of demographic (mean with SD) and bone densitometry data (mean with 95% CI) between spine and hip fracture groups in men

	Spine Fx (n = 50)	Hip Fx (n = 12)	*P*-value
Age (years)	57.2 (10.9)	65.9 (9.7)	0.017
Weight (kg)	69.2 (12.3)	73.8 (11.8)	0.254
Height (m)	1.71 (0.10)	1.72 (0.06)	0.826
Spinal BMDdxa (g/cm^2)	0.849 (0.799 to 0.899)[a]	0.965 (0.827 to 1.104)	0.124
Femoral neck BMDdxa (g/cm^2)	0.687 (0.648 to 0.726)[b]	0.662 (0.537 to 0.788)[c]	0.249
Trochanteric BMDdxa (g/cm^2)	0.623 (0.587 to 0.659)[b]	0.616 (0.486 to 0.746)[c]	0.646
Spinal BMDqct (mg/cm^3)	55.6 (46.0 to 65.2)[d]	—[f]	—
Radial BMDqct (mg/cm^3)	92.4 (75.0 to 109.8)[e]	—[f]	—
DXAspi Z-score	−1.75 (−2.10 to −1.39)	−0.87 (−1.61 to −0.13)	0.037
DXAneck Z-score	−1.70 (−2.04 to −1.37)	−1.81 (−2.94 to −0.68)	0.501
DXAtroch Z-score	−1.54 (−1.80 to −1.28)	−1.51 (−2.49 to −0.53)	0.638
QCTspi Z-score	−2.57 (−2.92 to −2.23)	—	—
pQCTrad Z-score	−2.19 (−3.02 to −1.35)	—	—

The reference equations from the studies on normal subjects of Pearson et al. [10, 11, 13] were used to calculate these Z-score results without rounding the regression constants, regression coefficients, or SDs. Use of these coefficients after rounding would have led to changes in the group Z-scores of 0.07 units
[a] n = 44; [b]n = 41; [c]n = 11; [d]n = 18; [e]n = 17; [f]not analyzed because too few measurements made

Table 4. Comparison of demographic (mean with SD) and bone densitometry data (mean with 95% CI) between spine and hip fracture groups in women

	Spine Fx (n = 284)	Hip Fx (n = 39)	*P*-value
Age (years)	64.2 (8.9)	76.4 (9.9)	0.001
Weight (kg)	60.5 (10.6)	58.7 (10.4)	0.306
Height (m)	1.58 (0.07)	1.57 (0.07)	0.638
Spinal BMDdxa (g/cm^2)	0.764 (0.743 to 0.785)[a]	0.849 (0.797 to 0.902	0.002
Femoral neck BMDdxa (g/cm^2)	0.637 (0.618 to 0.656)[b]	0.602 (0.548 to 0.661)[c]	0.044
Trochanteric BMDdxa (g/cm^2)	0.536 (0.519 to 0.552)[b]	0.515 (0.471 to 0.558)[c]	0.155
Spinal BMDqct (mg/cm^3)	60.3 (53.9 to 66.7)[d]	—[f]	—
Radial BMDqct (mg/cm^3)	64.2 (56.2 to 72.1)[e]	—[f]	—
DXAspi Z-score	−1.49 (−1.64 to −1.33)	−0.78 (−1.90 to −0.35)	0.001
DXAneck Z-score	−1.38 (−1.59 to −1.16)	−1.15 (−1.81 to −0.49)	0.641
DXAtroc Z-score	−1.34 (−1.53 to −1.16)	−1.01 (−1.48 to −0.54)	0.208
QCTspi Z-score	−1.38 (−1.58 to −1.17)	—	—
pQCTrad Z-score	−0.99 (−1.15 to −0.82)	—	—

[a] n = 251; [b]n = 193; [c]n = 36; [d]n = 104; [e]n = 79; [f]not analyzed because too few measurements made

porosis when femoral neck BMDdxa or trochanteric BMDdxa were assessed. In female hip fracture patients the DXAspi T-score was below the −2.5 limit in 48.7% of the cases, whereas DXAneck and DXAtroc T-scores were below the limit in 52.8% and 36.1% of the cases. Up to 94.2% of spine fracture cases had their QCTspi below the cutoff, but the specificity of the cutoff was only moderate (Table 6). The same cutoff applied to males gave generally lower sensitivity but higher specificity than in females.

We also examined the alternative DXAspi −2.0 SD cutoff for women. It appeared that although the sensitivity increased to 85.2% for spine fractures and 64.1% for hip fractures, the specificity decreased to 77.0%.

Between-Center Differences

Patients with Hip Fractures. Although the CV of mean bone density values between the centers varied from 5.5 to 15.5% in women with hip fractures and from 23.2 to 31.1% in men with hip fractures, none of the between-center differences in mean bone density were significant.

Patients with Spine Fractures. Mean values for spinal BMDdxa differed significantly between centers in men ($P < 0.05$, CV = 11.6%) and in women ($P < 0.001$, CV = 9.6%). The differences disappeared after adjustment for age, weight, and height in men, but not in women ($P = 0.001$). Similarly, femoral neck ($P < 0.001$) and trochanteric BMDdxa ($P < 0.001$) differed significantly between centers in female spine fracture cases (CV = 7.8–8.4%) but not in men. These differences remained even after age, weight, and height adjustment ($P = 0.003$ and $P = 0.046$, respectively). The differences in the spine fracture group slightly diminished (*P*-values for spinal, femoral neck, and trochanteric BMDdxa were 0.016, 0.006, and 0.049, respectively) after the exclusion of the Berlin center, which recruited spine fracture cases from a population-based epidemiologi-

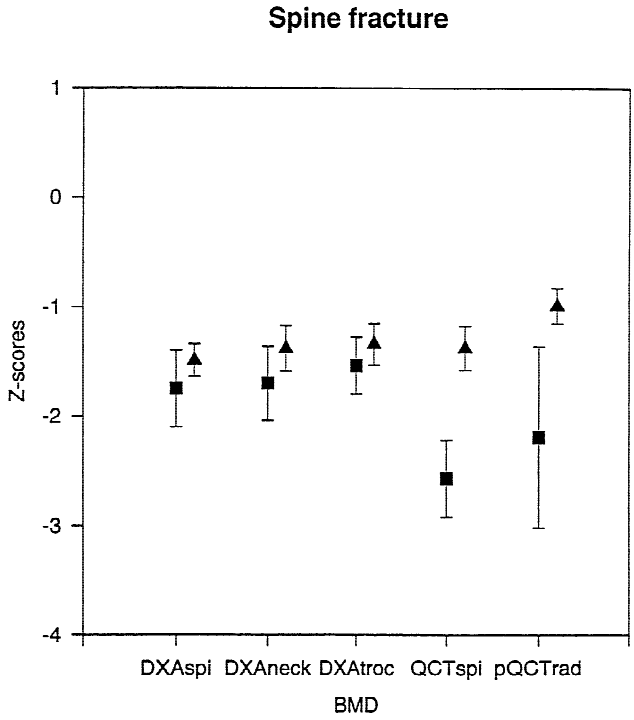

Fig. 1. Spinal (DXA and QCTspi), femoral neck (DXA), trochanteric (DXA), and radial (PQCTrad) Z-scores (mean with 95% CI) for men (squares) and women (triangles) with spine fractures. All Z-scores differed significantly from normals. The difference between sexes was statistically significant at QCTspi BMD (P = 0.001) and pQCTrad BMD (P = 0.001).

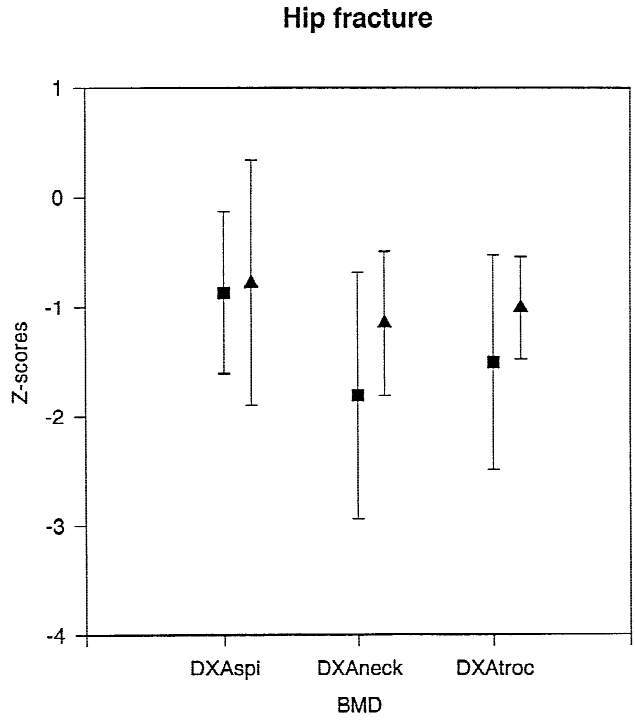

Fig. 2. Spinal, femoral neck, and trochanteric Z-scores (mean with 95% CI) for men (squares) and women (triangles) with hip fractures. The Z-scores, except spinal Z-score in women, differed significantly from normals. There were no statistical differences between sexes.

cal study cohort rather than from a specialist clinic. As to BMDqct data, there were similar statistical between-center differences for both spinal BMDqct (CV = 24.2%) and radial BMDqct (CV = 10.6%) in female spine fracture cases ($P < 0.001$) and for spinal BMDqct in males (P = 0.04, CV = 22.9%).

Discussion

Utility and Application of T-Scores

T-scores refer to the peak bone mass of young normal adults and a negative T-score represents assumed bone loss relative to the young healthy state. According to the WHO study group classification, 71.2% of the female spine fracture patients were classified as having severe osteoporosis when spinal BMDdxa was assessed. Similarly, 52.8% of the hip fracture patients were deemed to have severe osteoporosis in terms of femoral neck BMDdxa. This means that at a minimum, nearly half of the hip fracture cases and 30% of the spine fracture cases had fractures, though they did not have a sufficiently low bone density to be classified as having osteoporosis prior to fracture. In the case of spine fractures, one possibility was that the classification method chosen for diagnosing spine fracture was imperfect. A tighter definition for a vertebral fracture might have resulted in lower BMDs in the fracture group; however, the exclusion of moderate deformity fractures in this way is controversial. Another factor contributing to the moderate sensitivity of the WHO threshold may have been the presence of confounding degenerative changes in the spinal DXA images [20–22]. We did not exclude any patients for the presence of osteoarthritis (OA) in the present study. As to hip fracture cases, another bias might have arisen from our inability to subdivide hip fracture cases into femoral neck and trochanteric groups. Some studies have suggested that patients with trochanteric fractures are older and have lower femoral neck BMDdxa, and particularly lower spinal BMDdxa than patients with femoral neck fracture [23, 24]. Furthermore, hip fracture patients who consent to be studied may have higher BMD than those too ill or confused to participate.

QCTspi showed better sensitivity than DXA, but the specificity for the −2.5 SD cutoff was much lower, suggesting a relatively low positive predictive value which might be improved by use of a different cutoff. On the other hand, the WHO cutoff proved to be insufficiently sensitive for pQCTrad. Similarly, a −2.5 SD cutoff applied to males showed low sensitivity. It seems that the new WHO study group definition for osteoporosis may require different thresholds when used with different BMD measurement techniques. The optimal definitions for men are also uncertain.

The choice of thresholds for case-finding strategies and decision making is a complex process, but in essence it depends on available resources, the prevalence of disease, and the relationship of sensitivity to specificity (as illustrated graphically in a ROC curve). If there is an overall high risk of fracture in a population it may be acceptable to designate a lower specificity and hence make more positive treatment decisions in people without disease than if there is a low risk. The choice of cutoff also depends on costs and

Table 5. ROC analyses (areas under the ROC curves) for different densitometry methods for discrimination of spine and hip fracture cases from controls in men and women

Method 1	Method 2	Cases	Controls	Area 1	Area 2	*P*-value for difference
			Male spine fracture patients			
DXAspi	DXAneck	41	540	0.89	0.88	0.6
DXAspi	DXAtroc	41	540	0.89	0.87	0.8
DXAneck	DXAtroc	41	540	0.87	0.88	0.7
DXAspi	QCTspi	16	70	0.97	0.97	0.9
DXAspi	pQCTrad	12	133	0.82	0.91	0.2
DXAneck	QCTspi	16	68	0.91	0.97	0.1
DXAneck	pQCTrad	11	129	0.94	0.93	0.9
			Female spine fracture patients			
DXAspi	DXAneck	192	1016	0.89	0.86	0.07
DXAspi	DXAtroc	192	1016	0.89	0.86	0.03
DXAneck	DXAtroc	192	1016	0.86	0.86	0.8
DXAspi	QCTspi	88	142	0.85	0.89	0.08
DXAspi	pQCTrad	51	286	0.97	0.84	0.3
DXAneck	QCTspi	52	107	0.90	0.89	0.6
DXAneck	pQCTrad	52	272	0.87	0.84	0.3
QCTspi	pQCTrad	32	84	0.73	0.81	0.2
			Male hip fracture patients			
DXAspi	DXAneck	11	540	0.72	0.85	0.03
DXAspi	DXAtroc	11	540	0.72	0.81	0.04
DXAneck	DXAtroc	11	540	0.85	0.81	0.4
			Female hip fracture patients			
DXAspi	DXAneck	36	1016	0.82	0.86	0.08
DXAspi	DXAtroc	36	1016	0.82	0.87	0.1
DXAneck	DXAtroc	36	1016	0.86	0.87	0.9

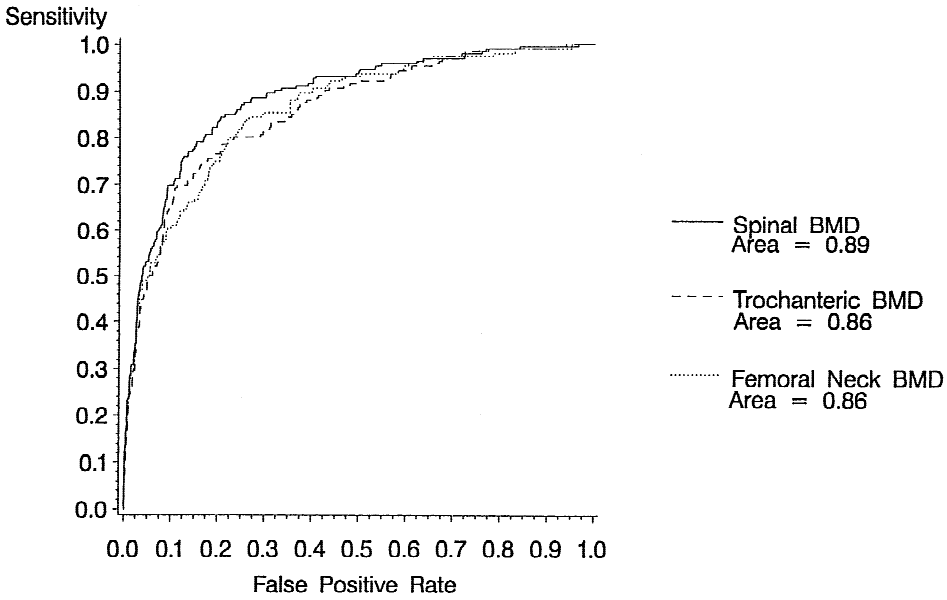

Fig. 3. ROC curves for DXA for discrimination of spine fractures from controls in women.

side effects. What our data show is that if a single treatment decision point (e.g., at a T-score of −2.5) were to be adopted across all modalities in current use for bone densitometry, the consequences for patient management would be highly dependent on choice of technology and measurement site. It seems clear from our analysis that in different communities world wide, with their varied resource constraints, available technologies, and prevalences of osteoporotic fracture, in-

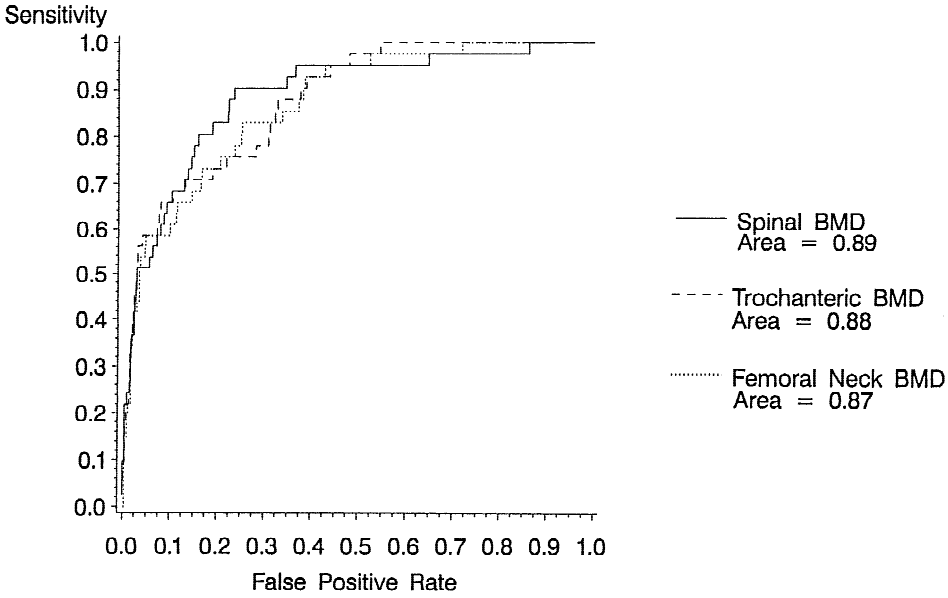

Fig. 4. ROC curves for DXA for discrimination of spine fractures from controls in men.

Table 6. Proportions of spine and hip fracture patients falling below the WHO defined female severe osteoporosis cutoff (T score < −2.5 SD) using different BMD measurements. The specificity (normals above the same cutoff) is also shown

Females	Spine Fx cases below −2.5 SD cutoff (sensitivity) (%)	Hip Fx cases below −2.5 SD cutoff (sensitivity) (%)	Normals above −2.5 SD cutoff (specificity) (%)
DXAspi	71.2	48.7	88.6
DXAneck	33.8	52.8	97.2
DXAtroc	24.9	36.1	98.2
QCTspi	94.2	—	58.6
pQCTrad	17.7	—	98.3
Males	**Spine Fx cases below −2.5 SD cutoff (sensitivity) (%)**	**Hip Fx cases below −2.5 SD cutoff (sensitivity) (%)**	**Normals above −2.5 SD cutoff (specificity) (%)**
DXAspi	50.0	25.0	97.3
DXAneck	22.0	45.5	98.3
DXAtroc	14.6	36.4	99.4
QCTspi	100.0	—	50.0
pQCTrad	23.5	—	100.0

formed local choices have to be made of the clinically appropriate and technology-specific decision points to be incorporated into disease-management guidelines.

Sex Differences and Similarities in Fracture Cases

Our female spine fracture patients had significantly lower standardized BMDdxa values than males with the corresponding fractures. However, these differences in areal bone density may be due to differences in skeletal size, because no differences in BMD values between the sexes were found after adjustment for weight and height. In agreement with this conclusion, there were also no differences between the sexes in spinal BMDqct values, which is a true volumetric estimate of density, although the male fracture patients were significantly younger than the females. Similarly, with hip fractures, no significant BMDdxa differences between the sexes could be found. These results suggest that the relationship between BMD and fracture risk is similar between men and women and that fracture risk is related to volumetric density, or so-called areal density after adjustment for body size. However, other factors, e.g., skeletal size and susceptibility to falls, are certainly important co-determinants of fracture.

In the assessment of the degree of bone loss in our cases by comparison with our controls, we used Z-scores partly because of the significant differences in age distribution between the sexes. Z-scores were also necessary to make measurements with different techniques comparable. A limitation of our study was that for the calculation of the reference values and Z-scores for spinal QCT the numbers of normal subjects were not ideal, particularly in men, but our normal ranges for women are in good agreement with previously published data [25, 26]. However, Kalendar et al. [26] reported male values that on average were 11.9% lower

than ours. The difference may be due to the fact that their reference material was recruited from patients scheduled for a clinical CT examination unrelated to skeletal disease rather than being recruited from the normal population. There is a case for recruiting a further series of normal male subjects to resolve this uncertainty.

In both fracture groups, men showed lower Z-scores than women; however, the difference reached statistical significance only for QCTspi and pQCTrad. Therefore, when compared with their peers, the relative osteoporosis in men with fractures may be more severe than in women with the corresponding fractures. It should be noted that lower Z-scores in male fracture patients are due to the fact that age-specific bone loss is less in men than in women after the age of 50, not to bone density being lower in male than female cases in units of measurement. We suggest that this sex difference in Z-scores is an effect of the menopause which lowers normal age-matched bone density values in women. However, it is reasonable to suspect that our results might have been affected to some extent by a lower threshold for suspicion of osteoporosis in women than in men.

Effects of Selection and Choice of Equipment

There were no significant between-center differences in bone densities at various sites for hip fracture patients. However, significant between-center differences in bone densities for spine fracture patients were found. This may be due to the fact that the patient material may have varied between the centers. Although a similar radiological criterion for a spine fracture was used in all centers, this does not exclude the possibility that some centers chose more severe cases of spinal osteoporosis than others, due, for example, to differences in local referral patterns.

The differences in mean BMD values obtained by different densitometers (e.g., Lunar and Hologic) can be on average 12–14%. However, standardization with the ESP_p reduces this to a maximum of 4–6% at the high end of the spinal range for standardized BMD and less (1–3%) for the hip and lower spinal values [15]. We suggest that imperfect cross-calibration only contributed in a minor way to the between center variations seen in this study since we screened osteoporotic patients whose BMD values were low.

Choice of Measurement Site and Technique for Diagnostic Discrimination

A valid use of densitometry is in identifying patients at high risk of a future fracture. Patients with at least one previous fracture are at higher risk than those without a previous fracture, and risk in patients with and without previous fractures increases logistically as a function of the reduction in their bone density. In assessing risk, it is desirable to minimize numbers of measurements in the individual so we compared measurement sites for their power to discriminate between cases and controls. The patients with hip fracture showed significantly decreased DXAneck and DXAtroc Z-scores. However, the female hip fracture patients did not differ from age-matched normals in terms of DXAspi. This agrees with previous studies which have shown that patients with hip fractures have significantly decreased femoral Z-scores, but generally the spinal Z-scores did not differ from controls [27, 28]. Also, the QCT density of the spine in hip fracture patients has been shown to be similar to that of age-matched controls [29]. These findings may suggest preferential femoral osteopenia and conservation of spinal bone mass in some female hip fracture cases. There is, however, a second possible explanation for this finding in the spine. Postmenopausal bone loss in women results in a rapid spinal bone loss at menopause and subsequently the spinal BMDdxa reference curve declines more sharply than that for the femoral neck. Thus, more normal women than men move close to the −2.5 SD threshold with the result that actual differences in BMD between normal women and women with fractures are quite small. Also, osteoarthritis of the spine could have elevated spinal BMDdxa in our hip fracture patients although we did not do routine X-ray studies on their spines to check for this.

We interpret our results as showing that the measurement of BMD at the proximal femur was good at detecting general osteopenia, since the proximal femoral Z-scores were almost equally affected in both spine and hip fracture patients in both sexes. Our results concur with previous studies showing that proximal femur BMD is decreased in patients with vertebral fractures [4, 30, 31].

Whereas spinal BMDdxa was poor at detecting osteopenia in hip fracture patients, spinal Z-scores were significantly decreased in spine fracture cases as compared with both normal and hip fracture cases. This suggests that when low spinal BMD occurs, it may usefully be a predictor of spine fracture. Peripheral QCT measurements also appeared to detect spinal osteopenia, particularly in men. In a previous study, pQCT was shown to have a good sensitivity and specificity in discriminating postmenopausal spinal crush fracture patients from controls [21].

Our findings suggest that there may be valid choices when cost effectiveness is a criterion in setting up a densitometric service. It seems that the proximal femur may be the measurement site of choice when evaluating generalized bone loss, particularly in elderly people, who may suffer from spinal osteophytes and deformations. If spinal osteoporosis is suspected and a risk of spinal fractures needs evaluation, an additional spinal BMD measurement might add clinical value to bone densitometry, but we had insufficient statistical power to test this hypothesis. Also, according to the present results, pQCT seems to discriminate spine fracture cases from controls as well as DXA and, in agreement with Grampp et al. [32], our ROC analyses disclosed no significant difference in discriminatory ability among axial QCT, pQCT, and DXA methodologies. It should also be noted that spine fracture patients have Z-scores for proximal femur similar to that of hip fracture patients. This finding suggests a mechanism whereby patients with spinal fractures are candidates for future hip fractures, in agreement with previous epidemiological studies [33, 34] demonstrating increased risk of hip fracture in spine fracture patients.

In conclusion, the careful assessment of bone density in cases of osteoporotic fractures of the spine and hip is likely to be of considerable clinical value when treatment choices have to be made and a prognosis given for the risk of future fracture. It is not clear that any one technique (DXA, QCT, pQCT) is superior for this purpose. However, care should be taken when applying T scores, particularly in choosing appropriate thresholds using axial QCT below which patients are to be advised that they suffer from osteopenia or osteoporosis.

Acknowledgments. This study has been supported by grants from

Sigrid Juselius Foundation, The Wellcome Trust (044003/IGS/MH), The Finnish Academy and The European Community's COMAC-BME programme for 1990–1993. Dr. A. Rapado (Madrid), Mr. J. Bright, and Mr. M. Henley are acknowledged for their important contributions to study planning and data management.

References

1. Kröger H, Huopio J, Honkanen R, et al (1995) Prediction of fracture risk using axial bone mineral density in a perimenopausal population: a prospective study. J Bone Miner Res 10:302–306
2. Melton L III, Atkinson E, O'Fallon W, Wahner H, Riggs B (1993) Long-term fracture prediction by bone mineral assessed at different skeletal sites. J Bone Miner Res 8:1223–1227
3. Cummings SR, Black DM, Nevitt MC, et al (1993) Bone density at various sites for prediction of hip fractures. Lancet 341:72–75
4. Mazess R, Barden H, Ettinger M, Schultz E (1988) Bone density of the radius, spine and proximal femur in osteoporosis. J Bone Miner Res 3:13–18
5. Silman A (1995) The patient with fracture: the risk of subsequent fractures. Am J Med 98:12S–16S
6. Wasnich R, Davis J, Ross P (1994) Spine fracture risk is predicted by non-spine fractures. Osteoporosis Int 4:1–5
7. Ross PD, Davis JN, Epstein RS, Wasnich RD (1991) Preexisting fractures and bone mass predict vertebral fracture incidence in women. Ann Intern Med 114:919–923
8. Black D, Bauer D, Lu Y, Tabor H, Genant H, Cummings S (1995) Should BMD be measured at multiple sites to predict fracture risk in elderly women? J Bone Miner Res 10:S140
9. Kanis J, Melton L III, Christiansen C, Johnston C, Khaltaev N (1994) Perspective: the diagnosis of osteoporosis. J Bone Miner Res 9:1137–1141
10. WHO Study Group (1994) Assessment of fracture risk and its application to screening for postmenopausal osteoporosis WHO technical report, vol. 843. World Health Organisation, Geneva
11. O'Neill TW, Felsenberg D, Varlow J, et al (1996) The prevalence of vertebral deformity in European men and women: the European vertebral osteoporosis study. J Bone Miner Res 11: 1010–1017
12. Eastell R, Cedel SL, Wahner HW, Riggs BL, Melton LJ III (1991) Classification of vertebral fractures. J Bone Miner Res 6:207–215
13. Pearson J, Dequeker J, Henley M, et al (1995) European semi-anthropomorphic spine phantom for the calibration of bone densitometers: assessment of precision, stability and accuracy. Osteoporosis Int 5:174–184
14. Pearson J, Dequeker J, Reeve J, et al (1995) Dual X-Ray absorptiometry of the proximal femur: normal European values standardised with the European Spine Phantom. J Bone Miner Res 10:315–324
15. Dequeker J, Pearson J, Reeve J, et al (1995) Dual X-ray absorptiometry: cross-calibration and normative reference ranges for the spine. Results of a European Community Concerted Action. Bone 17:247–254
16. Pearson J, Ruegsegger P, Dequeker J, et al (1994) European semi-anthropomorphic phantom for the cross-calibration of peripheral bone densitometers: assessment of precision accuracy and stability. Bone Miner 27:109–120
17. Reeve J, Kroger H, Nijs J, et al (1996) Radial cortical and trabecular bone densities of men and women standardised with the European Forearm Phantom. Calcif Tissue Int 58: 135–143
18. Hanley J, McNeill B (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29–36
19. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the area under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845
20. Masud T, Langley S, Wiltshire P, Doyle DV, Spector TD (1993) Effect of spinal osteophytosis on bone mineral density measurements in vertebral osteoporosis. Br Med J 307:172–173
21. Schneider P, Börner W (1990) Ability of peripheral QCT at the radius in comparison to DXA of the spine to diagnose vertebral fractures in postmenopausal women. In: Christiansen C, Overgaard K (eds) Osteoporosis. Osteopress ApS, Copenhagen, pp 921–924
22. Abdel-Hamid Osman A, Bassiouni H, Koutri R, Nijs J, Geusens P, Dequeker J (1994) Aging of the thoracic spine: distinction between wedging in osteoarthritis and fracture in osteoporosis—a cross-sectional and longitudinal study. Bone 15:437–442
23. Nakamura N, Kyou T, Takaoka K, Ohzono K, Ono K (1992) Bone mineral density in the proximal femur and hip fracture type in the elderly. J Bone Miner Res 7:755–759
24. Vega E, Mautalen C, Gomez H, Garrido A, Melo L, Sahores O (1991) Bone mineral density in patients with cervical and trochanteric fractures of the proximal femur. Osteoporosis Int 1:81–86
25. Block JRS, Gluer C, Steiger P, Ettinger B, Genant H (1989) Models of spinal trabecular bone loss as determined by quantitative computed tomography. J Bone Miner Res 4:249–257
26. Kalender W, Felsenberg D, Louis O, et al (1989) Reference values for trabecular and cortical vertebral bone density in single and dual-energy quantitative computed tomography. Eur J Radiol 9:75–80
27. Aloia J, McGowan D, Erens E, Miele G (1992) Hip fractures patients have generalized osteopenia with a preferential deficit in the femur. Osteoporosis Int 2:88–93
28. Chevalley T, Rizzoli R, Nydegger V, et al (1991) Preferential low bone mineral density of the femoral neck in patients with a recent fracture of the proximal femur. Osteoporosis Int 1: 147–154
29. Harma M, Karjalainen P, Hoikka V, Alhava E (1985) Bone density in women with spinal and hip fractures. Acta Orthop Scand 56:380–385
30. Nuti R, Martini G, Gennari C (1993) Total body, spine and femur dual X-ray absorptiometry in spinal osteoporosis. Calcif Tissue Int 53:388–393
31. Turner C, Peacock M, Timmerman L, Neal J, Johnston C (1995) Calcaneal ultrasonic measurements discriminate hip fracture independently of bone mass. Osteoporosis Int 5:130–135
32. Grampp S, Genant H, Mathur A, et al (1997) Comparisons of non-invasive bone mineral measurements in assessing age-related loss, fracture discrimination and diagnostic classification. J Bone Miner Res 12:697–711
33. Gardsell P, Johnell O, Nilsson B (1989) The predictive value of fracture, disease and falling tendency for fragility fractures in women. Calcif Tissue Int 45:327–330
34. Kotovics M, Melton L III, Cooper C, Atkinson E, O'Fallon W, Riggs B (1994) Risk of hip fracture in women with vertebral fracture. J Bone Miner Res 9:599–605