

Assessment of osteopenia from spine radiographs using two different methods: the Chingford study

¹T MASUD, MA, MRCP, ²I MOOTOOSAMY, FRCP, ³E V MCCLOSKEY, MRCP, ⁴M P O'SULLIVAN, FRCP, ⁵E P WHITBY, MB, BCh, BSc, ⁵D KING, FRCP, ⁶M B MATSON, MRCP, ²D V DOYLE, MD, FRCP and ⁶T D SPECTOR, MD, MSc, MRCP

¹City Hospital NHS Trust, Nottingham NG5 1PB, UK, ²Whipps Cross Hospital, London E11 1NR, UK, ³University of Sheffield Medical School, Sheffield S10 2RX, UK, ⁴Mater Misericordiae Hospital, Dublin 7, Ireland, ⁵Royal Hallamshire Hospital, Sheffield S10 2JF, and ⁶St Thomas' Hospital, London SE1 7EH, UK

Abstract

Two methods for diagnosing radiological osteopenia in thoracic (TS) and lumbar (LS) spine radiographs were assessed: a subjective conventional method (A) and a semiquantitative method (B), by comparing them with bone mineral density (BMD) measured by dual energy X-ray absorptiometry (DEXA), in a population of "normal" women aged 45–70 years ($n=818$). For both methods there was good intraobserver and interobserver reproducibility. BMDs were significantly lower with increasing radiological osteopenia grades ($p<0.001$), and remained lower after adjustment for age and body mass index ($p<0.01$). The proportion of subjects with DEXA-defined osteoporosis rose with increasing radiological osteopenia grades for both methods. The worst osteopenia categories identified 29.7–55.3% of women with DEXA-defined osteoporosis, compared with 6.1–11.7% in the "normal" categories. Both methods, however, showed a large degree of overlap of BMDs between the various radiological osteopenia grades. The sensitivity and specificity of method A in diagnosing osteoporosis were 45.3% and 78.4%, respectively, for the TS and 19.0% and 94.3%, respectively, for the LS. For method B the sensitivities and specificities were 8.8% and 96.1%, respectively (TS), and 10.2% and 95.6%, respectively (LS). Although both methods have poor sensitivities, "definite" or "high" grade osteopenia should be an indication for bone densitometry. The high specificities suggest that a "normal" (no osteopenia) X-ray is unlikely to have a significantly low BMD.

Bone mass is a major measurable determinant of future fracture risk [1]. There are now a number of treatments which can prevent bone loss and reduce fracture risk, but their use depends on accurate diagnosis of osteoporosis [2, 3]. Dual energy X-ray absorptiometry (DEXA) is now the preferred method of estimating bone mass [4], but at present in most countries it is confined to specialist centres. Spine radiographs are commonly reported as showing radiological osteopenia, although traditionally osteoporosis is only believed to be detectable when at least 30% of bone has been lost [5–7]. Previous small studies attempting to use conventional radiographs to assess bone density have shown conflicting results [8–14].

The aim of this study was to assess the sensitivity and specificity of diagnosing radiological osteopenia by comparison with DEXA-measured bone mineral density (BMD). Two different methods of diagnosing radiological osteopenia were assessed; a subjective conventional reporting method and a semiquantitative method.

Methods

The subjects used were from the Chingford population study which has been described in detail elsewhere [15].

Received 31 August 1995 and in revised form 8 January 1996, accepted 8 February 1996.

The study involves 1035 Caucasian women (age range 45–70 years) seen initially between 1988 and 1990. This population is similar to an equivalent age group in the UK population in terms of social class, height, weight and smoking habits. Ethical approval for the study was obtained from the Waltham Forest District Ethical Committee.

DEXA (Hologic QDR/1000W) was used to measure lumbar spine BMD (L1–L4). The precision of this method given by the coefficient of variation in this centre was 0.66%, based on repeated measurements on 20 subjects.

Lateral thoracic and lumbar spine radiographs (two films) were performed by the same radiographer, using the same apparatus, at a standard target-to-film distance of 101 cm centred at T9 and L3, respectively. The thoracic and lumbar radiographs were categorized into osteopenia grades using two separate methods carried out by different experienced radiologists blinded to the bone density and clinical data.

Method A

The radiographs were categorized into "normal" (no osteopenia) (NO), "possible osteopenia" (PO) and "definite osteopenia" (DO). The radiologist used traditional subjective reporting methods assessing diffuse demineralization (grey levels), without atlas or specific criteria, designed to represent "current practice".

Method B

A previously used semiquantitative method [16] which assesses both demineralization and vertebral trabecular pattern changes was used as follows:

Grade 0 — normal bone density.

Grade 1 — minimal loss of bone density; end plates begin to stand out giving a stencilled effect.

Grade 2 — vertical striation is more obvious; end plates are thinner.

Grade 3 — more severe loss of bone density than Grade 2, end plates becoming less visible.

Grade 4 — ghost-like vertebral bodies. Density is no greater than soft tissue. No trabecular pattern is visible (when the vertebral outline is covered, the body is invisible).

The reproducibility of both methods was determined by two radiologists reading 30 radiographs, of varying osteopenia grades, blind to clinical details and the results of each other. One radiologist re-read the films at least 2 months later, blinded to previous readings. Results were analysed using a kappa (κ) statistical package adapted for a personal computer (PC) [17].

The number of subjects with at least one vertebral deformity was determined by a researcher working independently from the above radiologists. The anterior, posterior and central heights of T4-L5 vertebrae were measured and a vertebral deformity was defined as at least a 25% reduction in vertebral height calculated from the dimensions of adjacent vertebrae. Subjects with vertebral deformities were excluded from further analysis as the radiologists may have been biased into classifying a radiograph into a higher osteopenia grade in the presence of vertebral deformities.

Results were analysed using the statistical package SPSS. The Spearman's rank correlation coefficients (r) were calculated between osteopenic grades and BMD measurements. Mean BMD, ages and body mass index (BMI) measurements in each radiological osteopenia grade were calculated for both methods. Analysis of covariance was used to adjust the means for confounding variables such as age and BMI. For the purposes of this study osteoporosis was defined as DEXA-measured bone mass less than 2 standard deviations below the mean peak bone mass, which in the population studied was 0.827 g cm^{-2} for the lumbar spine.

This arbitrary BMD level was used to calculate the percentage of subjects who had osteoporosis in each radiological osteopenia grade and the sensitivities and specificities of both X-ray methods. To calculate sensitivities and specificities of both methods in diagnosing osteoporosis, osteoporosis was defined as "definite osteopenia" only for method A and as "Grade 3 or more" for method B. The sensitivities and specificities were recalculated using more relaxed criteria for defining osteoporosis: "osteoporosis = possible or definite osteopenia"

for method A, and "osteoporosis = Grades 2 or more" for method B.

Results

The intraobserver and interobserver reproducibilities of both methods for the thoracic and lumbar spines are shown in Table I. For method A the intraobserver reproducibilities for the thoracic and lumbar spine and interobserver reproducibility for the lumbar spine were good whereas interobserver reproducibility for the thoracic spine was moderate. Method B showed moderate intraobserver and interobserver reproducibilities for thoracic and lumbar spines.

Complete interpretable data on DEXA, X-ray osteopenia readings and vertebral deformity assessments were present in 841 subjects (mean age 54.3 ± 6.0 years). Of these, 23 subjects (2.8%) had at least one vertebral deformity and were therefore excluded from further analysis.

The frequencies, mean ages, BMI, BMDs and percentage of subjects with DEXA-defined osteoporosis in each osteopenia grade for both methods are shown in Table II. In method B only one subject in the studied population was classified into Grade 4 for the thoracic and lumbar spines (same subject, $\text{BMD} = 0.736 \text{ g cm}^{-2}$) and therefore Grades 3 and 4 were combined (Grade 3-plus). Osteopenia grades were significantly correlated with age for both method A ($r=0.14$ (thoracic), $r=0.15$ (lumbar) and method B ($r=0.21$ (thoracic), $r=0.20$ (lumbar) ($p<0.001$). Thoracic spine osteopenia grades were significantly correlated with BMI in both methods A and B ($r=-0.11$, $p<0.001$ and $r=-0.10$, $p<0.01$, respectively). However, there was no significant correlation between BMI and lumbar spine grades for either method.

Mean BMDs significantly decreased with increasing radiological osteopenia grades for both methods (Method A: $r=-0.22$ (thoracic), $r=-0.35$ (lumbar); Method B: $r=-0.21$ (thoracic), $r=-0.28$ (lumbar) ($p<0.001$), even when adjusted for age and BMI using analysis of covariance (Method A: $p<0.001$ (thoracic and lumbar); Method B: $p=0.003$ (thoracic), $p<0.001$ (lumbar)). For method A the mean BMD was 9.4%

Table I. Interobserver and intraobserver agreement (Ag) and reproducibility (kappa) of methods A and B

		Thoracic Spine		Lumbar Spine	
		Ag	Kappa (95%CI)	Ag	Kappa (95%CI)
Method A	Intra	80%	0.73 (0.69-0.77)	83%	0.77 (0.74-0.80)
	Inter	70%	0.59 (0.55-0.63)	73%	0.63 (0.59-0.67)
Method B	Intra	63%	0.60 (0.56-0.64)	53%	0.43 (0.39-0.47)
	Inter	57%	0.55 (0.52-0.58)	60%	0.42 (0.38-0.47)

Table II. Mean (m) age (years), BMI ($\text{kg}^2 \text{cm}^{-1}$), BMD (g cm^{-2}) and percentage of subjects with DEXA-defined osteoporosis (% DEXA-OP) by osteopenia (OP) grades: methods A and B

Method A	OP Grade	n	Age (SD)	BMI (SD)	m BMD (SD)	m BMD ^a (SD)	% DEXA-OP
Thoracic spine	No OP	240	53.4 (6.1)	26.0 (4.4)	1.01 (0.16)	1.00 (0.16)	11.7 (28)
	Possible OP	369	53.8 (6.1)	25.5 (4.5)	0.98 (0.15)	0.97 (0.15)	12.7 (47)
	Definite OP	209	55.7 (5.4)	24.7 (3.7)	0.91 (0.16)	0.93 (0.16)	29.7 (62)
Lumbar spine	No OP	496	53.4 (5.9)	25.7 (4.5)	1.01 (0.15)	1.00 (0.15)	8.9 (44)
	Possible OP	275	55.3 (6.0)	25.1 (4.0)	0.92 (0.14)	0.93 (0.14)	24.4 (7)
	Definite OP	47	55.6 (5.6)	25.3 (4.0)	0.83 (0.15)	0.85 (0.15)	55.3 (26)

Method B	OP Grade	n	Age (SD)	BMI (SD)	m BMD (SD)	m BMD ^a (SD)	%DEXA-OP
Thoracic spine	0	28	50.9 (4.9)	26.6 (5.8)	1.08 (0.17)	1.04 (0.17)	10.7 (3)
	1	373	53.2 (5.8)	25.8 (4.7)	0.99 (0.15)	0.99 (0.15)	11.0 (41)
	2	385	55.1 (6.0)	25.1 (3.7)	0.95 (0.16)	0.96 (0.16)	21.0 (81)
	3-plus ^b	32	57.2 (5.7)	24.9 (3.2)	0.88 (0.16)	0.91 (0.16)	37.5 (12)
Lumbar spine	0	82	52.1 (6.1)	25.9 (4.8)	1.05 (0.15)	1.03 (0.15)	6.1 (5)
	1	494	53.7 (5.8)	25.6 (4.4)	0.99 (0.14)	0.98 (0.14)	11.1 (55)
	2	206	55.8 (6.0)	25.0 (3.8)	0.91 (0.16)	0.93 (0.16)	30.6 (63)
	3-plus ^b	36	56.5 (5.0)	25.8 (3.4)	0.89 (0.18)	0.91 (0.18)	38.9 (14)

^a Adjusted for age and BMI.

^b Only one subject classified as Grade 4 (0.736 g cm^{-2}).

(95%CI=9.1–9.7), lower in the definite osteopenia group compared with the no osteopenia group for the thoracic spine, and 17.7% (95%CI=17.4–18.0) lower for the lumbar spine. For method B Grade 3-plus had a mean BMD 18.8% (95%CI=18.3–19.3) lower than Grade 0 for the thoracic spine, and 15.2% (95%CI=14.8–15.6) lower for the lumbar spine.

The proportion of subjects with DEXA-defined osteoporosis rose with increasing radiological osteopenia grade ranging from 11.7%–29.7% (thoracic) and 8.9%–55.3% (lumbar) in method A and 10.7%–37.5% (thoracic) and 6.1%–38.9% (lumbar) for method B (Table II).

Both methods showed a wide variation of BMDs around the means and a large degree of overlap between the various radiological osteopenia grades (Figure 1).

Using the “definite osteopenia” criteria for diagnosing osteoporosis radiologically, the sensitivity and specificity of method A in diagnosing osteoporosis were 45.3% and 78.4%, respectively, for the thoracic spine and 19.0% and 94.3%, respectively, for the lumbar spine. False positives and false negatives were 70.3% and 12.3%, respectively (TS) and 44.7% and 14.4%, respectively (LS). For method B, using “Grade 3-plus” to diagnose osteoporosis radiologically, the sensitivity and specificity were 8.8% and 96.1%, respectively (TS), and 10.2% and 95.6%, respectively (LS). The false positives and false negatives were 62.5% and 15.9%, respectively (TS), and 61.1% and 15.7%, respectively (LS). Positive predictive values for method A in diagnosing osteoporosis were 29.7% (TS) and 55.3% (LS); and for method B were 37.5% (TS) and 38.9% (LS).

Using the more relaxed “possible or definite osteopenia” criteria for diagnosing osteoporosis radiologically, for method A, the sensitivity and specificity were 79.6% and 29.3%, respectively (TS), and 67.9% and 60.6%, respectively (LS). False positives and false negatives were 81.1% and 11.7%, respectively (TS), and 71.1% and 8.9%, respectively (LS). Similarly the relaxed “criteria” for method B (“Grade 2 or more”) changes the sensitivity and specificity to 67.9% and 49.0%, respectively (TS), and 56.2% and 70.4% (LS), and the false positives and false negatives to 77.7% and 11.0%, respectively (TS), and 68.2% and 10.4%, respectively (LS). With these relaxed criteria the positive predictive values are reduced to 18.9% (TS) and 28.9% (LS) for method A and to 22.3% (TS) and 31.9% (LS) for method B.

Discussion

Conventional spinal radiography is a common investigation for patients with backpain. Some X-rays are reported as showing apparent osteopenia with or without evidence of vertebral deformity. Previous smaller studies assessing the usefulness of conventional radiography in diagnosing osteopenia have shown conflicting results.

In 132 women, Krolner et al found that their radiographic grade of osteopenia significantly correlated with bone mineral content of L2–L4 using dual photon absorptiometry (DPA) [8]. Using quantitative computerized tomography to measure bone mass in 41 subjects, Michel et al found significant correlations between bone mineral content (BMC) and lateral lumbar X-ray criteria for osteopenia including “vertebral density *versus* soft tissue”, “amount of trabeculations” and “overall

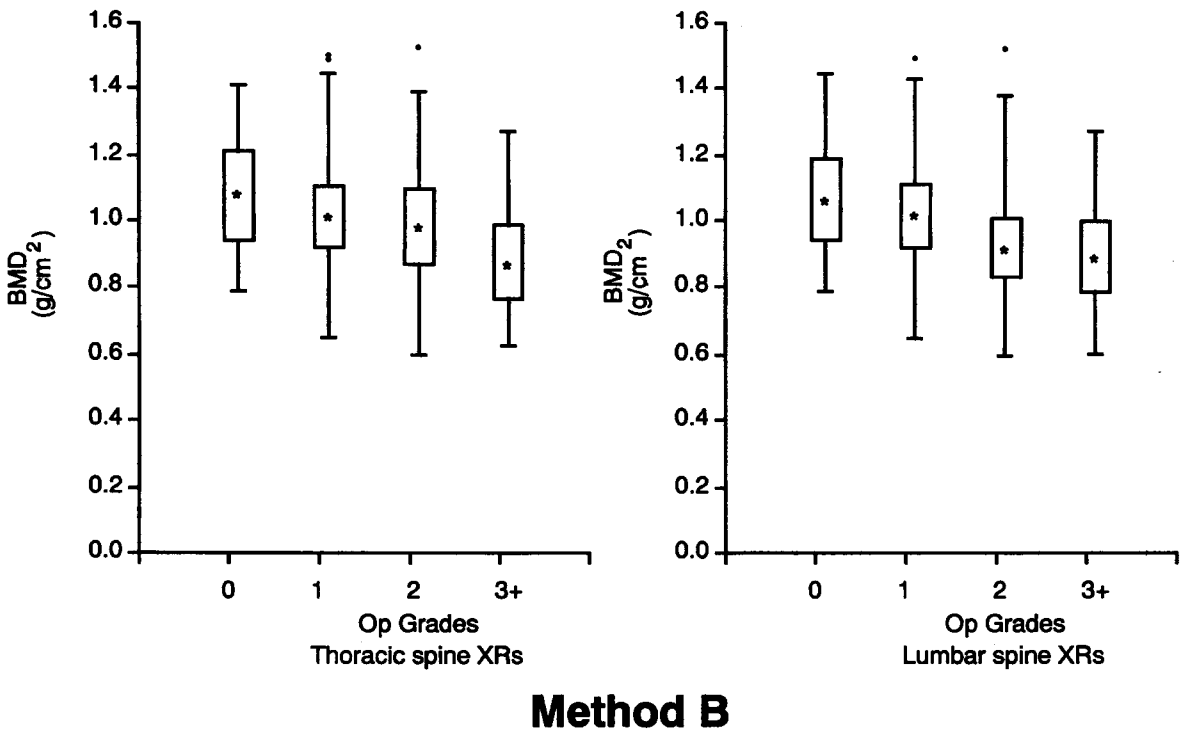
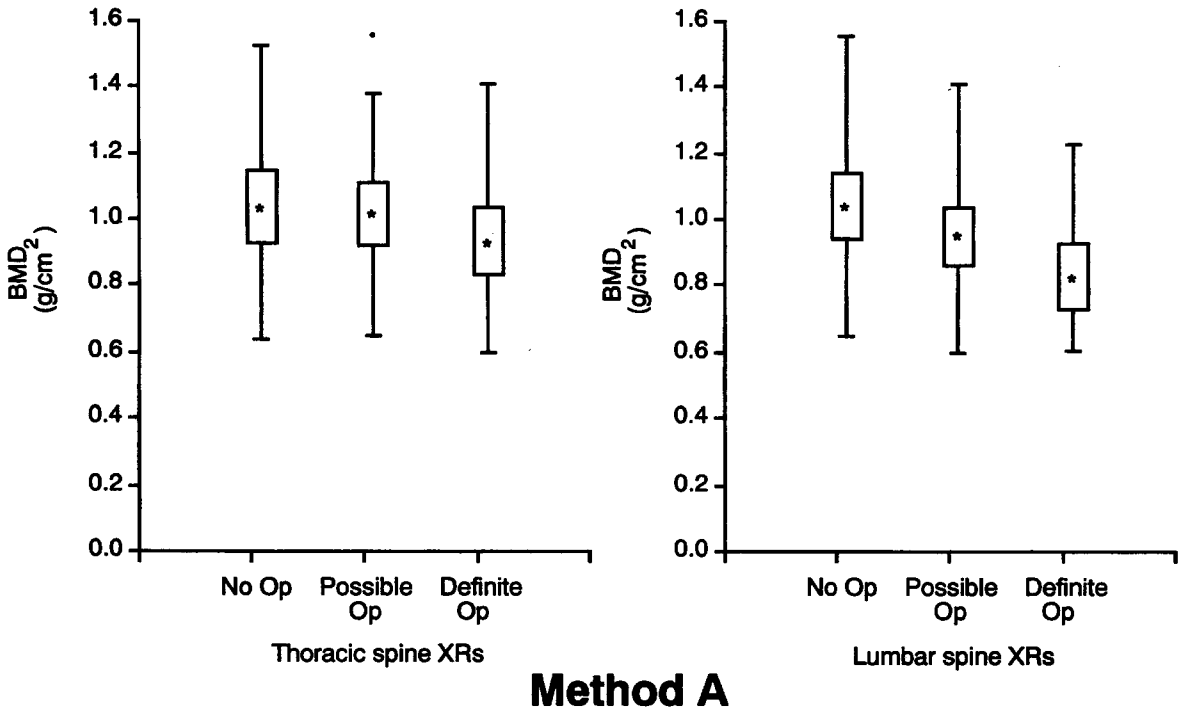


Figure 1. BMD by osteopenia grades for methods A and B (interquartile range (box) with range; * = median).

estimate of osteopenia", although changes in mineral content over a 5 year period could not be detected [9]. Garton et al [10] found that DEXA measured BMD was significantly related to visually estimated osteopenia in 200 men and women (mean age 64.9 years, range 52–90 years). Those without apparent radiological osteopenia had a low risk (9–15%) compared with those with definite osteopenia (64–86%) of being within the lowest quartile of BMD [10]. A vertebral trabecular pattern index as a semiquantitative assessment of bone loss was found to have a significant correlation with DEXA measured BMD in 106 subjects (mean age 53.7 years, range 24–73 years) [11].

Other researchers have found conventional X-rays less useful in estimating bone mass. Williamson et al [12] found that experienced radiologists could not accurately diagnose the degree of bone density on 45 lateral chest films when compared with the results of DPA. They concluded that to comment upon the presence or absence of osteoporosis on the basis of chest films could not be justified [12]. Using DPA to measure BMD in 33 women whose spine X-rays were reported as showing apparent osteopenia without evidence of vertebral deformation, Scane et al showed that only 66.7% of these women had a bone density below the normal range for young women (2 SD below mean BMD) [13]. Doyle et al assessed six separate radiological criteria for diagnosing spinal osteoporosis, including a subjective "reduced bone density" method, "trabecular loss" and "morphometric changes" in 34 subjects. They concluded that in the diagnosis and follow-up of patients with spinal osteoporosis, for practical purposes, the only reliable information concerns the morphology of vertebral bodies and not the BMC [14]. Their study, however, lacked any reliable methods for measuring bone density. The present study, considerably larger than previous studies, shows that although mean BMDs decreased with increasing radiological osteopenia grades, there was wide overlap of BMDs between the grades. This finding suggests that conventional X-rays should not be used as a routine way of diagnosing osteopenia (without vertebral deformation) in individual subjects. This conclusion applies to both a subjective assessment of osteopenia ("grey levels") as in method A and a semiquantitative assessment of demineralization and trabecular changes as in method B. Both methods had poor sensitivities, although reasonable specificities, in diagnosing osteoporosis if the strict criteria were used. If the criteria for diagnosing osteoporosis using X-rays were relaxed the sensitivities improved (though still less than 80%) but at the expense of specificities and therefore higher false positives. The data suggest that there is no advantage in using a semiquantitative method using trabecular changes to assess BMD, since method A was more reproducible and was simpler to perform. In comparison to thoracic spine X-rays, lumbar spine X-rays were more reproducible and had higher positive predictive values for diagnosing osteopenia with method A. For method B the lumbar spine X-rays had a higher positive predictive value than thoracic spine X-rays, but were less reproducible.

We used BMD > 2SD below the young adult mean to define osteoporosis. Although arbitrary, this definition was commonly used when the study commenced. We have retained this definition for the purposes of our study. More recent World Health Organization (WHO) criteria define osteoporosis as BMD > 2.5 SD below the young adult mean. Using this new definition made no difference to the conclusions of our study. We used DEXA measurements of BMD to define osteoporosis; however, there is growing awareness of the confounding effects of degenerative arthritis and aortic calcification on spine BMD measurements and of errors of analysis and interpretation, especially in the elderly [18]. However, because of the age group studied, these effects are unlikely to have a major impact on our results and conclusions.

In clinical practice, spinal radiography and bone densitometry should be regarded as complementary rather than alternative diagnostic procedures as the principles of the two are essentially different. The former assesses bone structure including trabeculations and vertebral deformities and may provide information on "quality" of bone. It may also provide information on conditions which may alter bone density measurements such as spondylosis [18], and other conditions which may also present with back pain such as myeloma and neoplastic disease. Bone densitometry provides information on bone mineral content which is independently associated with fracture risk.

However, in many areas bone densitometry is not yet readily available and in this situation conventional spine X-rays are often assessed for osteopenia and reported as such. Our data suggest that in the absence of vertebral deformity the finding of "possible" osteopenia or mild grades of osteopenia (Grade 2 or less in method B) are not clinically useful and should not be commented on. "Definite" osteopenia or a high grade osteopenia (Grades 3 and 4) should, however, be an indication for bone densitometry. In the absence of vertebral deformities treatment for osteoporosis should not be started solely on the basis of an X-ray report of osteopenia as many subjects (44.7%–70.3% in method A and 61.1%–62.5% in method B) would be treated unnecessarily. In such cases bone densitometry should be performed prior to treatment.

Acknowledgments

We would like to thank the City Branch of The National Osteoporosis Society for funding, Trish Wiltshire for performing the DEXA scans and the women of Chingford who took part in the study.

References

1. JOHNSTON, C C JR, SLEMENDA, C W and MELTON, L J III, Clinical use of bone densitometry, *N. Engl. J. Med.*, 324, 1105–1109 (1991).
2. RIGGS, B L and MELTON, L J III, The prevention and treatment of osteoporosis, *N. Engl. J. Med.*, 327, 620–627 (1992).
3. LINDSAY, R, Prevention and treatment of osteoporosis, *Lancet*, 341, 801–805 (1993).

4. KELLY, T L, SLOVIK, D M, SCHOENFELD, D A and NEER, R M, Quantitative digital radiography versus dual photon absorptiometry of the lumbar spine, *J. Clin. Endocrinol. Metab.*, 67, 839-844 (1988).
5. LACHMAN, E, Osteoporosis: the potentialities and limitations of its roentgenologic diagnosis, *AJR*, 74, 712-715 (1955).
6. JOHNSTON, C C and EPSTEIN, S, Clinical, biochemical, radiographic, epidemiologic, and economic features of osteoporosis, *Orthop. Clin. North. Am.*, 12, 559-569 (1981).
7. WANG, D C, KOTTAMASU, S R and KARVELIS, K, Scintigraphy in metabolic bone disease. In *Primer on the Metabolic Bone Diseases and Disorders of Mineral Metabolism*, ed. by M J Flavus (William Byrd Press, Richmond, Virginia, USA) pp. 87-93 (1990).
8. KROLNER, B, BERTHELSON, B, PORS NIELSON, S, Assessment of vertebral osteopenia: comparison of spinal radiography and dual photon absorptiometry, *Acta Radiol. Diagn.*, 23, 517-521 (1982).
9. MICHEL, B A, BJORKENGREN, A G, LAMBERT, E ET AL, Estimating lumbar bone mineral density from routine radiographs of the lumbar spine, *Clin. Rheumatol.*, 12, 49-52 (1993).
10. GARTON, M J, ROBERTSON, E M, GILBERT, F J ET AL, Can radiologists detect osteopenia on plain radiographs? *Clin. Radiol.*, 49, 118-122 (1994).
11. SCHNITZLER, C M, PITCHFORD, D G K, WILLIS, E M and GEAR, K A, Comparison of the vertebral trabecular pattern with the vertebral fracture prevalence and spinal bone density, *Osteoporosis Int.*, 3, 293-299 (1993).
12. WILLIAMSON, M R, BOYD, C M and WILLIAMSON, S L, Osteoporosis: diagnosis by plain chest films versus dual photon bone densitometry, *Skeletal Radiol.*, 19, 27-30 (1990).
13. SCANE, A C, MASUD, T, JOHNSON, F J and FRANCIS, R M, The reliability of diagnosing osteoporosis from spinal radiographs, *Age Ageing*, 23, 283-286 (1994).
14. DOYLE, F H, GUTTERIDGE, D H, JOPLIN, G F and FRASER, R, An assessment of radiological criteria used in the study of spinal osteoporosis, *Br. J. Radiol.*, 40, 241-250 (1967).
15. SPECTOR, T D, MCCLOSKEY, E V, DOYLE, D V and KANIS, J A, Prevalence of vertebral fracture in women and the relationship with bone density and symptoms: the Chingford study, *J. Bone Miner. Res.*, 8, 817-822 (1993).
16. SAVILLE, P D, A quantitative approach to simple radiographic diagnosis of osteoporosis: its application to the osteoporosis of rheumatoid arthritis, *Arthritis Rheum.*, 10, 416-422 (1967).
17. BRENNAN, P and SILMAN, A, Statistical methods for assessing observer variability in clinical measures, *Br. Med. J.*, 304, 1491-1494 (1992).
18. MASUD, T, LANGLEY, S, WILTSHIRE, P ET AL, The effect of spinal osteophytes on bone mineral density measurements in patients with vertebral osteoporosis, *Br. Med. J.*, 307, 172-173 (1993).