ELSEVIER

Bone Vol. 17, No. 3
September 1995:247–254

Dual X-Ray Absorptiometry—Cross-Calibration and Normative Reference Ranges for the Spine: Results of a European Community Concerted Action

J. DEQUEKER,[1] J. PEARSON,[2] J. REEVE,[2] M. HENLEY,[2] J. BRIGHT,[2] D. FELSENBERG,[3] W. KALENDER,[4] A-M. LAVAL-JEANTET,[5] P. RUEGSEGGER,[6] J. ADAMS,[7] M. DIAZ CURIEL,[8] M. FISCHER,[9] F. GALAN,[10] P. GEUSENS,[1] L. HYLDSTRUP,[11] P. JAEGER,[12] P. KOTZKI,[13] H. KRÖGER,[14] P. LIPS,[15] A. MITCHELL,[2] O. LOUIS,[16] R. PEREZ CANO,[10] H. POLS,[17] D. M. REID,[18] C. RIBOT,[19] P. SCHNEIDER,[20] and M. LUNT[21]

[1] *Arthritic and Metabolic Bone Disease Research Unit, K.U. Leuven, Belgium*
[2] *MRC Clinical Research Centre and Northwick Park Hospital, Harrow, UK*
[3] *Freie Universität Berlin, Klinik Steglitz, Berlin, Germany*
[4] *Siemens Medical Systems and University of Erlangen, Erlangen, Germany*
[5] *Inserm 18, Laboratoire de Radiologie, Paris, France*
[6] *Universität und ETH Zürich, Inst. für Biomedizinische Technik, Zurich, Switzerland*
[7] *University of Manchester, Department of Diagnostic Radiology, Manchester, UK*
[8] *Fundacion Jimenez Diaz, Servicio de Mineralomet. Un. Metabolica, Madrid, Spain*
[9] *Kassel Städtische Kliniken, Zentrum für Radiologie, Kassel, Germany*
[10] *University Hospital "Virgin Macarena," Departamento de Medicina, Sevilla, Spain*
[11] *Hvidovre Hospital Mineral and Metabolism Research Group, Department of Endocrinology, Hvidovre, Denmark*
[12] *University of Bern, Department of Diagnostic Radiology, Bern, Switzerland*
[13] *University de Montpellier, Hôpital Lapeyronie-Serv. de Méd. Nucléaire, Montpellier, France*
[14] *Kuopio University Central Hospital, Department of Surgery, Kuopio, Finland*
[15] *Academisch Ziekenhuis, Vrije University of Amsterdam, Department of Endocrinologie, Amsterdam, The Netherlands*
[16] *Akademisch Ziekenhuis, Vrije, Universitat Brussel, Department of Radiology, Brussels, Belgium*
[17] *University Hospital Rotterdam, Department of Medicine III, Clinic of Endocrinology, Rotterdam, The Netherlands*
[18] *Grampian Health Board, Geriatric and Specialist Services Unit, City Hospital, Aberdeen, Scotland, UK*
[19] *C.H.U. Purpan Service d'endocrinologie, U.F. Maladies Osseuses et Métaboliques, Toulouse, France*
[20] *Klinik und Poliklinik für Nuklearmedi, Wurzburg, Germany*
[21] *Institute of Public Health, Cambridge, UK*

Bone density measurements by dual X-ray absorptiometry (DXA) of the spine can now be made precisely, but there is no uniformity in reporting results and in presenting reference data. A European Union Concerted Action therefore devised a uniform procedure for cross-calibrating and standardizing instruments, using the European spine phantom (ESP) prototype. This phantom differs in a number of respects from the final version of the ESP. Eighteen centers in nine countries obtained 1619 records (1035 women) from Caucasian subjects, aged 20–80 years, drawn from normal populations. The DXA machines used were made by the Hologic, Lunar, and Norland companies. Highly statistically significant differences were evident between populations, both in apparent rates of bone loss with age and in the spread of values about the age-adjusted means. There were small residual differences in the results obtained with the three machine brands which could have been due to the relatively large between-center population differences we observed. The alternative or additional explanation that they were attributable, in part, to the design differences between the ESP prototype and the definitive ESP, which became available after this study was completed, was shown to be a valid possibility. Results from postmenopausal women reported in relation to the years that have elapsed since menopause showed reduced population variance when compared with conventional reporting in relation to age. After cross-calibration, the center with the highest age-adjusted normal density value averaged 23% more than the center with the lowest. It is therefore crucially important to select appropriate reference data in clinical and epidemiological studies. These results provide a basis for designing protocols for multicenter studies using currently installed densitometers. (*Bone 17:247–254; 1995*)

Key Words: Bone densitometry; Osteoporosis; Standardization; Epidemiology; Clinical trials.

Address for correspondence and reprints: Dr. J. Reeve, University Department of Medicine, Addenbrooke's Hospital (Box 157), Hills Road, Cambridge CB2 2QQ, UK. Fax: 44 1223 330105.

This article is a Concerted Action of the European Community's COMAC-BME Programme 1989–1992. *Project management group:* J. Dequeker (project coordinator, Leuven), J. Reeve (Harrow), J. Pearson (Harrow), D. Felsenberg (Berlin), W. Kalender (Erlangen), C. Langton (Sheffield), A-M. Laval-Jeantet (Paris), P. Ruegsegger (Zürich), and G. Van der Perre (Leuven). *Other participants and centers:* J. E. Adams (Manchester), J. C. Birkenhäger (Rotterdam), J.-P. Bonjour (Geneva), P. Braillon (Lyon), M. Diaz Curiel (Madrid), M. Fischer (Kassel), F. Galan (Sevilla), C. Gennari (Siena), P. Geusens (Leuven), A. Hemmingsson (Uppsala), L. Hyldstrup (Hvidovre), P. Jaeger (Bern), R. Jonson (Göteborg), J. Kalef-Ezra (Ioannina), P. O. Kotzki (Montpellier), H. Kröger (Kuopio), P. Lips (Amsterdam), M. Osteaux (Brussels), R. Reid (Aberdeen), C. Reiners (Essen); C. Ribot (Toulouse), and P. Schneider (Würzburg).

8756-3282/95/$9.50
SSDI 8756-3282(95)00224-2

Bone Vol. 17, No. 3
September 1995:247–254

Introduction

Measurements of bone mineral density predict the risk of fracture.[3,4,8] New methods to quantitate bone mineral in vivo developed in the last decade are used widely for evaluation of risk for osteoporotic fracture and for quantitating objectively the effect of bone active drugs. Although measurements of bone density can now be made precisely, there is no uniformity in reporting bone mass measurements and there are great problems in comparing results from one center to another.[12,16]

To resolve these difficulties, the European Concerted Action "Quantitative Assessment of Osteoporosis" aimed to standardize bone densitometry to allow intercomparisons of results obtained with machines from different manufacturers.[10] There are currently four manufacturers of DXA equipment who all use copyrighted software. Each manufacturer makes different assumptions regarding the detection of bone edges and adjustment for the fat content of the marrow, and uses different calibration procedures.

In a previous study we demonstrated that the semianthropomorphic European spine phantom prototype (ESP) can be used to develop conversion equations for obtaining densities on a standard scale from the results produced by the manufacturer's own software.[19] Here we present the reference ranges obtained for *posteroanterior* (PA) measurements of the spine after converting the results to standardized densities. Comparisons between centers for differences between populations are presented and the effect of changing menopausal status was also examined. Finally, because the ESP prototype has now been replaced by a final version of the ESP, we have calculated the effects of using the prototype in this study rather than the definitive version of the ESP on the standardized densities from a posteriori measurements on both phantoms.

Methods

Subjects

In the course of the Concerted Action, 18 centers obtained 1619 technically acceptable records from Caucasian subjects defined as normal or representative of the normal population (see below); 1035 of these subjects were women. Five centers selected their normal subjects randomly from a register of voters, a register of citizens, or from family practitioner lists which covered >95% of the population. These centers were: Berlin, Kassel (Germany), Harrow, Manchester, and Aberdeen (UK). Two centers advertised in the media for normal subjects; these were Amsterdam and Rotterdam (the Netherlands). Kuopio (Finland) recruited from large employers, societies, and clubs and through advertisements. Six centers recruited from hospital staff and attenders; these were: Hvidovre (Denmark), Brussels (Belgium), Seville and Madrid (Spain), Montpellier (France) and Wurzburg (Germany). Bern (Switzerland) selected normals by writing to the staff of different companies, households, and needlework groups. Leuven (Belgium) recruited subjects aged over 50 randomly from a register of voters and under 50 by writing to family members of patients. Toulouse (France) selected women at random from a menopause clinic. Subjects with known structural and metabolic bone diseases, arthritis, or neurologic conditions were not approached. Spinal malformations were not specifically sought or excluded. Subjects who were recruited and reported previous hyperthyroidism; women with a hysterectomy or oophorectomy before the age of 50; or subjects who had been treated with androgens, fluoride therapy, Vitamin D, or corticosteroids were not included in the reference ranges or the comparison of centers. Women taking estrogens were excluded if known to be postmenopausal. Results for subjects not accompanied by information on age, sex, and diagnosis were excluded. Total exclusions for women were 16% on this basis and 4% for men.

Measurement Procedures

Each center measured their subjects using the procedures recommended by the individual DXA manufacturer for the PA view working to quality control guidelines provided by the manufacturer, supplemented by operational guidance provided by the project management group. Regions of interest were machine defined after identifying the vertebrae to be measured. Software used was the same version for subject and phantom measurements. Results for bone mineral content (BMC) and bone mineral areal mass (BMA),[9] which is the same as bone mineral density (BMD, in grams per square centimeter), used by many authors, were recorded. The average measured BMA for L-2, L-3, and L-4 was analyzed statistically and will be subsequently referred to as *density*.

To ensure the quality of the data each instrument was subjected to daily checks of machine performance, including measurement of the manufacturer's phantom, according to the manufacturer's instructions. If possible, the L1–4 region was scanned. All technicians in each center were trained to operate to exactly the same standards. Artifacts, including movement artifacts, led to rejection of a scan and, if possible, it was repeated.

Statistical Methods

All bone mineral densities were converted to standardized densities using the methods of Pearson et al.[19] "Specified" densities are those of the three vertebrae in the ESP, which were specified in the manufacturing process. The "standardized" densities are values obtained from observed subject measurements after conversion using the machine-specific formulas relating observed phantom measurements to the specified densities. For the Lunar and Hologic machines, these formulas were exponentials, and for the Norland machines they were linear. Standardized densities are referred to in what follows as BMA values. The procedures for measuring the phantom reproducibly on scanners of all three manufacturers have already been described in detail.[19]

Comparison of Centers

The aim of this comparison was to determine whether the data from different centers could be combined to produce the reference ranges or whether allowance had to be made for between-center differences. For each combination of instrument manufacturer and gender the centers were compared using a general linear model. Comparison of the scanners used in different centers, including scanners made by the same manufacturer, has already been the subject of a previous extensive analysis.[19]

In this part of the analysis, centers which measured only a few reference subjects, defined as less than eight subjects per decade of age, were first excluded. The statistical modeling strategy was to use a conventional covariance analysis.

The first test was whether the linear relationship between BMA and age differed between centers and a significant result was recorded under the "Age and center" column of **Table 1.** If there was no significant difference at the 5% level, the average bone densities after adjustment for age were then compared. The model assumed that the BMA values were normally distributed

Table 1. Summary of p values from a general, linear model comparing centers making measurements on each brand of DXA machine

Sex/brand of machine	Number of subjects	Transformation	Age and center	Center (after adjustment for age)	Variability
Men					
Hologic	155	ln	0.002	—	<0.001
Lunar	206	ln	0.08	0.07	0.6
Norland	179	None	0.9	0.9	0.9
Women					
Hologic	315	ln	0.4	0.003	<0.001
Lunar	466	None	0.010	—	0.007
Norland	213	ln	0.08	<0.001	0.2

and that their variances were similar for each center and did not change with age. When these assumptions were violated, natural logarithms of the data were analyzed instead.

Reference Ranges

Reference ranges, incorporating changes in BMA with age, were derived for each sex using the following method.[22]

First, polynomial regression was used to model the relationship between bone density and age. A model of the following form was used:

$$\text{Predicted BMA} = a + (b * \text{age}) + (c * \text{age}^2),$$

where a is the regression constant and b and c are the regression coefficients. The parameters b and c were only included in the model if they led to a significant improvement in fit. Normal probability plots were used to check the assumption that the densities were normally distributed. Whenever the BMA values were not normally distributed, they were transformed by taking natural logarithms (ln transformation).

Second, BMA values were plotted against age to assess whether the standard deviation of BMA changed with age. In addition, the residuals (the difference between the observed and predicted bone densities) were split into five groups by age, containing approximately equal numbers of observations. In each group, the standard deviation of the residuals was calculated and the midpoint of the age group determined. A linear regression was then used to determine whether there was a relationship between the standard deviation of the residuals and age.

The 95% reference ranges were determined by the following formulas:

$$\text{Upper limit} = \text{predicted BMA} + (1.96 * \text{SD residual})$$
$$\text{Lower limit} = \text{predicted BMA} - (1.96 * \text{SD residual})$$

where "SD residual" is the standard deviation of the residuals. If the standard deviation of the residuals changed with age, the predicted SD residual was used instead of SD residual. Whenever the data were transformed, the reference ranges were transformed back to the original scale.

The results of this analysis were presented by plotting BMA against age, with the reference ranges superimposed. The percentage of normal subjects lying above the upper limit and below the lower limit was determined for each reference range. To check that these reference ranges fitted the data for each manufacturer's standardized BMA, the reference ranges were superimposed on the data for each manufacturer and the percentages above and below the range determined.

For a large majority of the women, information was available on the menopausal status (pre or post) and on the date of menopause, from which the years since menopause (YSM) was derived.

Reference ranges were also produced in pre- and postmenopausal women separately, relating BMA to age in both pre- and postmenopausal women and to YSM in postmenopausal women, using the methods previously described.

In postmenopausal women only, regression analysis was used to determine whether BMA was more closely associated with age or YSM. This was done by testing the significance of age and YSM separately and then by testing the significance of adding age into a model containing YSM and adding YSM into a model containing age.

The Definitive ESP

After completion of the main part of the study, the final version of the ESP,[11] which differs in a number of respects from the prototype, was measured five times each on three Lunar and three Hologic machines and the results from the two phantoms compared.

Results

Comparison of Centers

All data except for men measured on Norland machines and women measured on Lunar machines required a natural logarithm transformation. A summary of the p values for the comparison of centers is shown in Table 1. There were clear differences between centers which are both statistically and clinically significant. At the age of 55, which was near the mean age for most centers, mean values for each center in men ranged from 1.08 g/cm^{-2} (Leuven) to 1.20 (Harrow) with a grand mean of 1.125 (SD 0.046). In women, at the same age of 55, mean values for each center ranged from 0.932 g/cm^{-2} (Leuven) to 1.15 (Bern) with a grand mean of 1.053 (SD 0.058).

The individual centers using Hologic machines showed differing relationships between BMA and age ($p = 0.002$) in men; Seville showed evidence of a decrease in BMA with age while no significant trend was evident in Madrid or Harrow. In women measured using Hologic machines, the relationship between BMA and age did not differ significantly between centers ($p = 0.4$), but the average BMA values differed between centers ($p = 0.003$). Bern had higher BMA values while Seville had lower values than average. In the centers using Lunar machines, the relationship between BMA and age in women differed between centers ($p = 0.01$), with BMA decreasing faster than average

Bone Vol. 17, No. 3
September 1995:247–254

with age in women measured at Leuven, Toulouse, and Manchester. There were no significant differences in men measured on Norland machines between Hvidovre, Amsterdam, and Aberdeen. However, for women measured on Norland machines, those in Aberdeen had a lower BMA than women in Hvidovre and Amsterdam ($p < 0.001$).

The variability at any age was similar in centers measuring men on Lunar and Norland machines. There were differences between centers for Hologic machines in men; in general, subjects measured in Madrid were much less variable than those in Seville or Harrow. In women measured on Hologic machines, both Berlin and Harrow showed more variability than other centers. For women measured on the Lunar machines, subjects measured in Toulouse were less variable than those at other centers. There was no evidence of a difference in variability between centers for women measured with the Norland machines.

Reference Ranges

Table 2 gives a summary of the p values for the polynomial regression of BMA on age and for the regression of standard deviations of the residuals when grouped by age on the age midpoint of the group. **Table 3** shows the reference ranges for men and women based on a combination of data from all manufacturers. For both men and women, after standardization to achieve cross-calibration between machines, the data were log transformed. In men, the log of bone density decreased linearly with age, showing an average decrease of 1.7% (95% confidence interval [CI] 0.9–2.6%) every decade. In women, the relationship between the log of BMA and age was curvilinear. In both men and women, the variability of the measurements increased with age and this was reflected in the resulting reference ranges. **Figures 1 and 2** show these reference ranges superimposed on all the data and on the data for each manufacturer. These figures show that subjects which lie outside the reference ranges are distributed evenly over the entire age range. The reference ranges fit the data well. The percentage of subjects outside the reference ranges is shown in **Table 4.** There is evidence that, for the Hologic machines, slightly more subjects than expected lie above the upper reference limit for both men and women.

The age-related reference ranges for pre- and postmenopausal women are also shown in Table 3, as are the YSM-related reference ranges for postmenopausal women. In this table, the data are shown fitted to a quadric regression after a log transformation. Where a linear fit to the transformed data is adequate, the quadratic coefficient is reported "NS" and the fit is linear. The residual SD in some cases was age or YSM dependent; however, where this is not so, β is reported as "NS." The individual data are shown in **Figures 3 and 4.** For each of the reference ranges the relationship between the log of BMA and age, or YSM, was not significantly different from linear. In premenopausal women, the apparent loss of bone density with age over a decade averaged 1.1% (95% CI −0.2% to +2.3%) and just failed to reach statistical significance, whereas in postmenopausal women the loss of BMA with age was much greater than in premenopausal women, the women losing on average 7.1% (95% CI 5.4–8.9%) with every decade of age. Similar results were obtained for each 10 years after the menopause, but the residual SD and therefore the range of normal values about the mean density at the mean age at 1 year after menopause (51 years) was considerably narrower when the data were expressed relative to YSM (0.132 vs. 0.185 g/cm^2).

Table 2. Summary of p values for polynomial regression of BMA on age (or YSM) and for regression of SDs of residuals on age (or YSM) midpoint

Group (age or YSM)	Number of subjects	Polynomial regression		Regression SD on age (linear)
		Linear	Quadratic	
Men, all (age)	584	<0.001	0.1	0.004
Women, all (age)	1035	<0.001	0.002	0.02
Premenopausal (age)	428	0.09	0.9	0.3
Postmenopausal (age)	499	<0.001	0.06	0.06
Postmenopausal (YSM)	499	<0.001	0.08	0.04

YSM = years since menopause.

Comparison of the ESP Prototype with the ESP to Investigate the Apparent Excess of Subjects Lying Above the Reference Ranges When Measured on Hologic Machines

When the final version of the ESP was measured on Hologic machines, measured BMA for the low density "vertebra" was 5.7% higher, for the middle density vertebra it was 2.7% higher, and for the highest density vertebra it was 9.8% higher than the equivalent readings obtained with the ESP prototype. The equivalent changes for the Lunar machines were +9.2%, −1.3%, and +2.3%. The effects on standardized densities measured on human subjects were with both brands of densitometer to lower them, particularly with the Hologic at values of greater than 1.0 g/cm^{-2}. Values of 0.500 g/cm^{-2} calibrated with the prototype, when recalibrated with the final version of the ESP, came down with the Hologic to 0.495 and with the Lunar to 0.484 g/cm^{-2}. Values of 1.000 g/cm^{-2} calibrated with the prototype came down similarly with the Hologic to 0.946 and with the Lunar to 0.978 g/cm^{-2}. Finally, values of 1.500 g/cm^{-2} calibrated with the prototype came down with the Hologic to 1.349 and with the Lunar to 1.473 g/cm^{-2}.

Discussion

With this European Concerted Action we have assembled the largest DXA data base so far published on normal populations of women and men, over the full adult age range using the PA spine view. Previously, some large single country and comparative studies of male and female normal ranges have been published.[7,13–15,18] The primary motivation behind this project was to standardize the results obtained with different commercial DXA densitometers so that clinicians and investigators could more reliably interpret them. We have succeeded in showing that Hologic, Lunar, and Norland machines gave very similar results after their results were standardized. There are some residual differences between the three manufacturer's machines, which could have been due either to true differences between the individual populations or residual differences between the three brands of machine uncorrected by calibration, or both. The second (or third) possibility is supported by our comparison of the two versions of the ESP and also by the results of Genant et al.[5] Our comparison of the two phantoms suggests that about half the male values outside the upper 95% CI (i.e., most or all of the excess) would have been brought inside it if the final version of the ESP had been available, while the Lunar results would have been only slightly affected.

In longitudinal studies it remains unsatisfactory to compare measurements made on one machine with those made at a later date on a different brand or even the same brand of machine. However, the differences in mean values obtained by different

Table 3. Parameter estimates and their standard errors, defining the reference ranges for both men and women: ln (BMA) = $a + (b \times \text{age}) + (c \times \text{age}^2) + \xi_i$; SD residual = $\alpha + (\beta \times \text{age or YSM})$

Group (age or YSM)	a	SE(a)	b	SE(b)	c	SE(c)	α	SE(α)	β	SE(β)	SD at age 51
Men, all (age)	0.2157	0.0251	−0.0018	0.0004	NS	—	0.0897	0.0101	0.0014	0.0002	0.161
Women, all (age)	0.2728	0.0520	−0.00042	0.00210	−0.000063	0.000020	0.0504	0.0261	0.0021	0.0005	0.158
Premenopausal (age)	0.1803	0.006	NS	—	NS	—	0.1253	—	NS	—	0.125
Postmenopausal (age)	0.4474	0.0595	−0.00741	0.00095	NS	—	0.1848	—	NS	—	0.185
Postmenopausal (YSM)	0.0813	0.0134	−0.00729	0.00082	NS	—	0.1281	0.0167	0.0038	0.0010	—

NS = not significant at 5% level; YSM = years since menopause. ξ_i: term to represent statistical uncertainty.
95% reference ranges = $\exp\{a + (b \times \text{age}) + (c \times \text{age}^2) \pm [1.96 \times \text{SD residual}]\}$.

brands which, for Hologic and Lunar on the unstandardized data are on average 12% for men and 14% for women, are substantially reduced to an average of 6% and 4% by standardization with the ESP.

It did not prove possible to recruit all the subjects from random samples of family practitioner age–gender registers or similar samples of voter registers. This is the most desirable way of achieving a reliable cross section of the population under study. Consequently, we expected some statistically significant differences between populations in apparent (cross-sectional) rates of bone loss with age and in the spread of values about the means adjusted for age. Even so, the differences between centers in

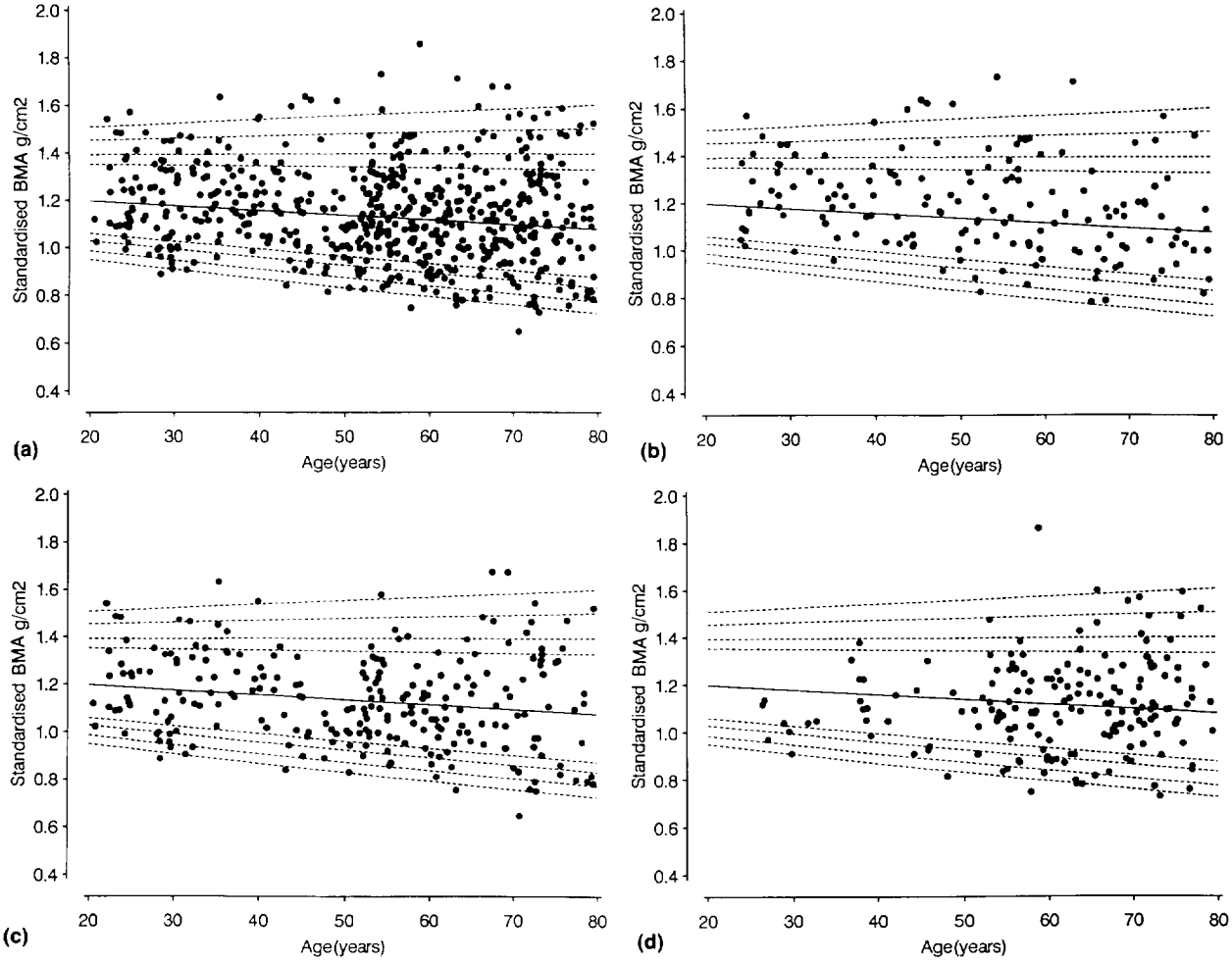

Figure 1. Normal measurements of DXA spine for men plotted against age. The lines are the regression line of BMA on age, the 95%, 90%, 80%, and 70% reference ranges. (a) All men; (b)–(d) men measured on Hologic, Norland, and Lunar machines, respectively.

Bone Vol. 17, No. 3
September 1995:247–254

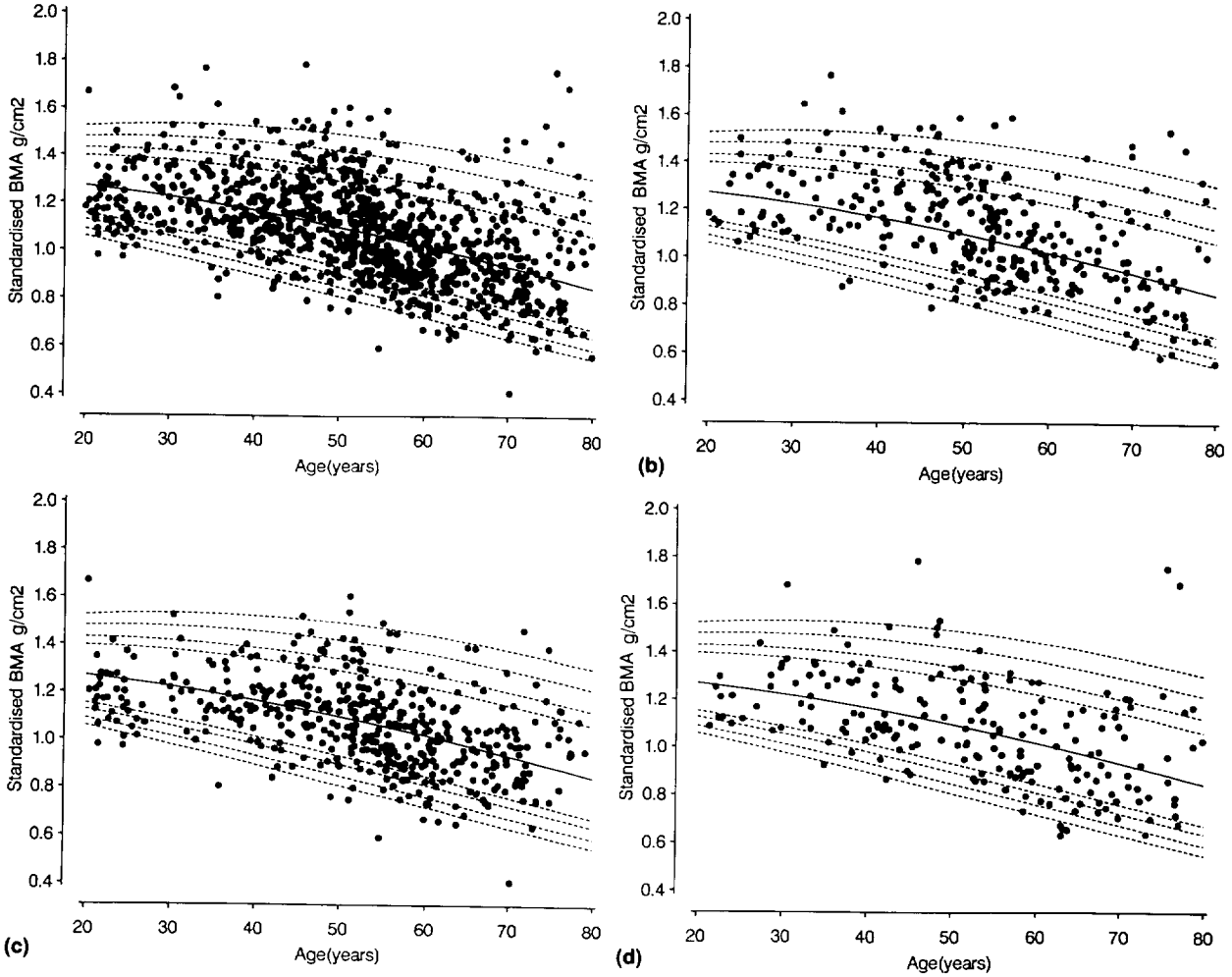

Figure 2. Normal measurements of DXA spine for women plotted against age. The lines are the regression line of BMA on age, the 95%, 90%, 80%, and 70% reference ranges. (a) All women; (b)–(d) women measured on Hologic, Lunar, and Norland machines, respectively.

mean values remain impressive, the center with the highest BMA values in women at age 55 having a mean that was 23% higher than the center with the lowest. Many of these apparent differences between centers were difficult to attribute to differences in recruitment method. For example, the two Spanish populations sampled in Madrid and Seville were both measured on Hologic machines and sampled according to the same technique, yet each showed significant differences in the spread of values about their respective age-adjusted means and in the apparent rate of loss of bone with age in both sexes. The reason for this is conjectural, but it is of interest that environmental factors reflected in measures of social class and occupation are associated with differences in the bone density distributions seen in the population in the Spanish province of Catalonia, as shown by the study of del Rio Barquero et al.[21]

The value of calculating reference ranges separately for women who are pre- and postmenopausal in densitometry of spinal trabecular bone by quantitative computed tomography (QCT) has been appreciated for some time.[1] There appears to be a relatively greater effect on menopause on trabecular than cortical bone loss in the early years after menstruation ceases.[6] Perhaps, in consequence, in assessing DXA bone density clinically in postmenopausal women, it is not yet current practice to report the results relative to those seen in women of the same

Table 4. Percentages of reference subjects that lie above the upper (U) and below the lower (L) reference limits

Group (age or YSM)	Number of subjects	Width of reference range (percentage that should lie outside each limit)			
		90% (5%)		70% (15%)	
		U	L	U	L
Men, all data	584	5.0	4.8	14.6	16.1
Hologic	160	6.9	1.9	22.5	9.4
Lunar	240	5.0	5.0	13.3	17.9
Norland	184	3.3	7.1	9.2	19.6
Women, all data	1035	4.9	4.7	15.0	14.7
Hologic	341	7.6	4.1	20.8	11.7
Lunar	479	3.1	5.4	11.3	15.5
Norland	213	4.7	4.2	14.1	17.8
Premenopausal (women)	428	5.1	4.9	14.0	12.6
Postmenopausal (age)	499	4.4	4.2	14.0	13.2
Postmenopausal (YSM)	499	6.0	3.8	15.0	15.4

YSM = years since menopause.

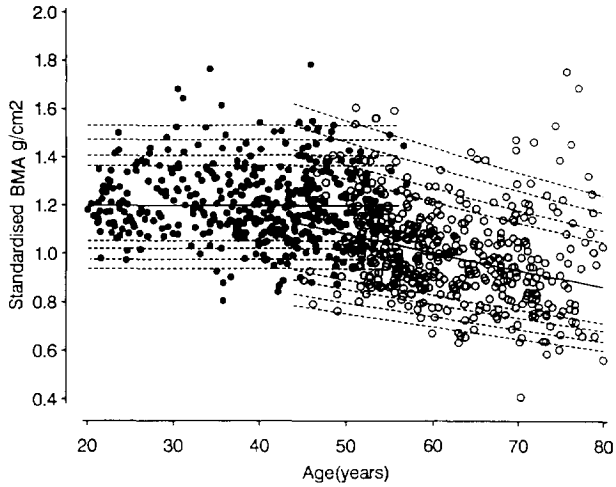

Figure 3. Normal measurements of DXA spine for pre- (●) and postmenopausal (○) women plotted against age. The lines are the regression line, 95%, 90%, 80%, and 70% reference ranges obtained separately for pre- and postmenopausal women.

menopausal, as distinct from chronological, age. However, an important result of the present study has been to show that reference to years since menopause in the early postmenopause results in a substantial narrowing (by up to 30%) in the reference range. The longitudinal study of Pouilles et al.[20] suggested that a rapid phase of bone loss occurs immediately after menopause. An early separation of pre- from postmenopausal women of the same age with respect to BMD of the spine is therefore to be expected. We recommend that, in the future, in women within the first 15 years of menopause, results are related to both the premenopausal reference range and the range appropriate to years since menopause, rather than the present practice of relating results to the young reference range and the chronologically matched reference range.

In men and postmenopausal women, we have shown a trend to increasing dispersion of BMA values with age. Burger et al.[2] found, in a recent study in Rotterdam, that very high values were associated with spinal osteoarthritis in subjects randomly selected from the general population. That this has not always been commented on in previous studies leads to suspicion that some of these studies were affected by selection bias to exclude subjects with asymptomatic osteoarthritis and/or osteoporosis.

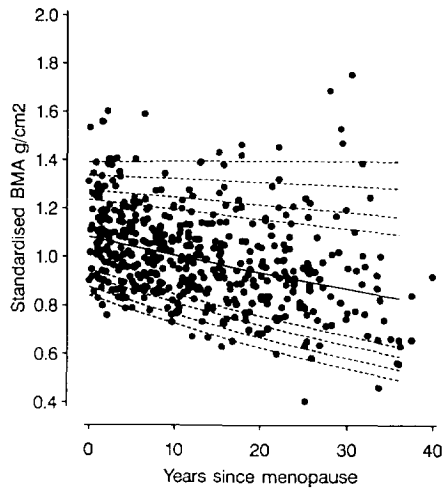

Figure 4. Normal measurements of DXA spine for postmenopausal women plotted against years since menopause (YSM). The lines are the regression line of BMA on YSM, 95%, 90%, 80%, and 70% reference ranges.

To use the standardized reference range it was concluded that a DXA machine should be standardized with at least five measurements of the ESP.[19] This allows simple equations to be calculated for converting units of measurement into standardized units. Remeasurement of the ESP is necessary if there is a source, detector, software, or other significant change during maintenance that is likely to affect the absolute values of the results generated, providing that stability can be assured by using the manufacturer's instructions for quality assurance with regular measurements of the manufacturer's quality control phantom. The ESP has advantages over some other phantoms in monitoring stability as it allows measurement of three different densities on each occasion. In a companion study, we have shown that low densities are more imprecisely measured and might, in principle, be more susceptible than high densities to drift with one brand of machine.[19] Ideally, therefore, weekly measurements of the ESP should be performed.

In conclusion, our results represent an important advance in the practice of DXA-based bone densitometry. They allow, for the first time, epidemiological and clinical studies requiring substantial numbers of subjects to employ different brands of densitometer providing each individual subject remains associated with only one machine of guaranteed stability. Our results will permit clinicians to confidently assess results obtained by their colleagues using DXA equipment different from that used in their own departments. The principle behind this approach has received approval from all four DXA manufacturers.[17] In achieving the objective of more rigorous cross-calibration and standardization, this study has enhanced the future of DXA measurements of the spine in research and clinical practice. However, these cross-sectional results, like those of all cross-sectional studies, cannot be used to infer rates of bone loss because of cohort effects and differences in recruitment rates at different ages.

References

1. Block, J. E., Smith, R., Glueer, C.-C., et al. Models of spinal trabecular bone loss as determined by quantitative computed tomography. J Bone Miner Res 4:249–257; 1989.
2. Burger, H., van Daek, P. L. A., Algra, D., van den Ouvreland, F. A., Grobbee, D. E., Hofman, A., van Kuijk, C., Schutte, H. E., Birkenhager, J. C., and Pols, H. A. P. The association between age and bone mineral density in men and women aged 55 years and over: The Rotterdam study. Bone Miner 25:1–13; 1994.
3. Cummings, S. R., Black, D. M., Nevitt, M. C., et al. Bone density at various sites for prediction of hip fractures. Lancet 341:72–75; 1993.
4. Gardsell, P., Johnell, O., and Nilsson, B. E. Predicting fractures in women by using forearm bone densitometry. Calcif Tissue Int 46:235–242; 1989.
5. Genant, H. K., Grampp, S., Glüer, C. C., Faulkner, K. G., Jergan, M., Engelke, K., et al. Universal standardisation for dual x-ray absorptiometry: Patient and phantom cross-calibration results. J Bone Miner Res 9:1503–1514; 1994.
6. Geusens, P., Dequeker, J., Verstraeten, A., and Nijs, J. Age-, sex-, and menopause-related changes of vertebral and peripheral bone: Population study using dual and single photon absorptiometry and radiogrammetry. J Nucl Med 27:1540–1549; 1986.
7. Hogiwara, S., Miki, T., Nishizawa, Y., Ochi, H., Onoyama, Y., and Morii, H. Quantification of bone mineral content using dual-photon absorptiometry in a normal Japanese population. J Bone Miner Res 4:217–222; 1989.
8. Hui, S. L., Slemenda, C. W., and Johnston, S. S. Baseline measurement of bone mass predicts fracture in white women. Ann Intern Med 111:355–361; 1989.

9. Jonson, R. Mass attenuation coefficients, quantities and units for use in bone mineral determinations. Osteopor Int 3:103–106; 1993.
10. Kalender, W. A. and Fischer, M. Quality control and standardisation of absorptiometric and computer tomographic measurements of bone mineral density. Radiat Protect Dosimetry 49:229–233; 1993.
11. Kalender, W. A., Felsenberg, D., Genant, H. K., Fischer, M., Dequeker, J., and Reeve, J. The European Spine Phantom—A tool for standardisation and quality control in spinal bone mineral measurements by DXA and QCT. Eur J Radiol (in press).
12. Kelly, T. L., Slovik, D. M., and Neer, R. M. Calibration and standardization of bone mineral densitometers. J Bone Miner Res 4:663–669; 1989.
13. Kroger, H., Heikkinen, J., Laitinen, K., and Kotaniemi, A. Dual-energy x-ray absorptiometry in normal women: a cross-sectional study of 717 Finnish volunteers. Osteopor Int 2:135–140; 1992.
14. Kroger, H. and Laitinen, K. Bone mineral density measured by dual-energy x-ray absorptiometry in normal men. Eur J Clin Invest 22:454–460; 1992.
15. Mazess, R. B., Barden, H. S., Ettinger, M., Johnston, C., Dawson-Hughes, B., Baran, D., et al. Spine and femur density using dual-photon absorptiometry in US white women. Bone Miner 2:211–219; 1987.
16. Morita, R., Orimo, H., Yamamoto, I., et al. Some problems of dual energy X-ray absorptiometry in clinical use. Osteopor Int 3(suppl):S87–S90; 1993.
17. Nord, R. H., Stein, J., Mazess, R. B., and Pommet, R. Letter to the editor. Osteopor Int 1:94; 1991.
18. Pocock, N. A., Eisman, J. A., Mazess, R. B., Sambrook, P. N., Yeates, M. G., and Freund, J. Bone mineral density in Australia compared with the United States. J Bone Miner Res 3:601–604; 1988.
19. Pearson, J., Dequeker, J., Henley, M., Bright, J., Reeve, J., Kalender, W., et al. European semi-anthropomorphic spine phantom for the calibration of bone densitometers. Assessment of precision, stability and accuracy. Osteopor Int 5:174–184; 1995.
20. Pouilles, J. M., Tremollieres, F., and Ribot, C. The effects of menopause on longitudinal bone loss from the spine. Calcif Tiss Int 52:340–343; 1993.
21. del Rio Barquero, L., Baures, M. R., Segura, J. P., et al. Bone mineral density in two different socio-economic population groups. Bone Miner 18: 159–168; 1992.
22. Royston, J. P. Constructing time-specific reference ranges. Stat Med 10:675–690; 1991.

Date Received: June 21, 1994
Date Revised: April 25, 1995
Date Accepted: May 11, 1995