Bone xxx (2008) xxx-xxx

Contents lists available at ScienceDirect

Bone

journal homepage: www.elsevier.com/locate/bone

Geographical variation in DXA bone mineral density in young European men and women. Results from the Network in Europe on male osteoporosis (NEMO) study

Stephen Kaptoge [a,*], Jose A. da Silva [b], Kim Brixen [c], David M. Reid [d], Heikki Kröger [e], Torben L. Nielsen [c], Marianne Andersen [c], Claus Hagen [c], Roman Lorenc [f], Steven Boonen [g], Marie-Christine de Vernejoul [h], Jan J. Stepan [i], Judith Adams [j], Jean-Marc Kaufman [k], Jonathan Reeve [a]

[a] Department of Medicine and Strangeways Research Laboratory, University of Cambridge, Wort's Causeway, Cambridge CB1 8RN, United Kingdom
[b] Department of Rheumatology, University of Coimbra, Coimbra, Portugal
[c] Department of Endocrinology, Odense University Hospital, DK-5000 Odense C, Denmark
[d] Deparment of Medicine and Therapeutics, University of Aberdeen, Aberdeen, United Kingdom
[e] Department of Surgery, Bone and Cartilage Research Unit, Kuopio University Hospital, Kuopio, Finland
[f] Department of Biochemistry and Experimental Medicine, The Children's Memorial Health Institute, Warsaw, Poland
[g] Department of Medicine, University Hospital, Leuven, Belgium
[h] INSERM U606 and Department of Rheumatology, Hôpital Lariboisière, 75010 Paris, France
[i] Department of Internal Medicine, Charles University, Prague, Czech-Republic
[j] Clinical Radiology, Imaging Science and Biomedical Engineering, University of Manchester, Manchester, UK
[k] Deparment of Endocrinology, University Hospital, Ghent, Belgium

ARTICLE INFO

Article history:
Received 11 September 2007
Revised 28 March 2008
Accepted 1 April 2008
Available online xxxx

Edited by: Harry Genant

Keywords:
Bone mineral density (BMD)
Dual energy X-ray absorptiometry (DXA)
Europe
T-score
Osteoporosis
Young normal

ABSTRACT

We collected population-based young normal hip and spine BMD data from 17 centres across Europe to assess between centre differences and to compare reference values with the US NHANES-III data. There was strong evidence of between country heterogeneity, but not between centres within countries. Hip BMD mean values were lower in European women, but SD's differed little from the NHANES-III USA results in both sexes. It may be necessary to adjust NHANES-III based T-scores by adding/subtracting a country-specific adjustment factor.

Introduction: It remains unclear whether young normal BMD reference values specific to an American population can be validly used for T-score calculation in Europeans.

Methods: We collected population based BMD data from 1163 men and 329 women aged 19-29 years from 17 centres across Europe to compare mean and SD values with the NHANES-III study USA results. BMD(g/cm^2) was measured at the hip and spine using DXA densitometers cross-calibrated with the European Spine Phantom (ESP). The only exclusions were for technically inadequate scans. A linear regression model was used to derive reference values. To allow for direct comparison with published NHANES III study data, the cross-calibrated BMD values were converted using the ESP equations to Hologic QDR 1000 units.

Results: In men, the overall mean(SD) BMD values expressed in Hologic-QDR1000 units of measurement, were: femoral neck 0.912(0.132); trochanter 0.793(0.124); and L2-L4 spine 1.027(0.123). The respective estimates in women were: 0.826(0.115); 0.670(0.093); and 0.983(0.107). However the I^2 statistic for heterogeneity indicated moderate to strong evidence of between-centre heterogeneity. There was, however, no significant heterogeneity observed between centres within countries, suggesting that this variation arose from national differences. Compared to the NHANES III population-based US data, the mean values in women were significantly lower at both sites due to some lower national European means. However, at all sites and in both sexes the SD's were very similar between the US and Europe. There was some evidence that recruiting volunteers resulted in biased values in women.

Conclusion: Our T-score normal values for the lumbar spine (L2-L4) should be more reliable for spine-specific risk assessment than some non-representative normal ranges, and should be evaluated for that purpose in Europe. If T-scores are to be used to compare individual data with ranges seen in normal young subjects of the same nationality, it may be necessary to adjust femoral NHANES III-based T-scores by adding (or subtracting) a country-specific adjustment factor. In risk assessment it is probably sufficient to use NHANES III–based hip T-scores, as supplied for the hip by densitometer manufacturers, interpreting them in light of recent international meta-analysis data on the relationship between BMD and fracture risk.

* Corresponding author. Strangeways Research Laboratory, Wort's Causeway, Cambridge CB1 8RN, UK. Fax: +44 1223 741618.
E-mail address: stephen@srl.cam.ac.uk (S. Kaptoge).

8756-3282/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.bone.2008.04.001

Introduction

Currently there is no population-based reference database for spine BMD that has won approval for use across all brands of DXA densitometers. Whether young normal femoral BMD reference values specific to the US non-black, non-hispanic population can be validly used for calculating T-scores in European populations also remains unclear, even though all brands of densitometers offer the NHANES III data as referent. Moreover, there is increasing evidence that the approaches developed to diagnose and treat osteoporosis in women may be equally useful in men [1]. Therefore with the aim of strengthening the approach of using young normal data as DXA reference for both hip and spine, we assembled young normal BMD data from 1163 male and 329 female participants aged 19-29 years from two multicentre population based cohort studies across Europe and from a further 4 single-centre population based studies following an extended call for data in the FP5 Thematic Network in Europe on Male Osteoporosis (NEMO).

The two multicentre studies were: a concerted action of the European Community's COMAC-BME FP2 programme (5 centres) [2,3] and the Polish EPOLOS osteoporosis study (6 centres). Principal investigators in Coimbra (Portugal), Ghent (Belgium), Kuopio (Finland), Prague (Czech Republic), and Odense (Denmark) contributed further population-based young normal BMD data. Young normal non-population based data was also received from Ghent (Belgium), Paris (France) and additional centres in the COMAC-BME study.

The first aim was to compare the data from different investigational centres to determine whether differences in young normal BMD values at the commonly measured sites of proximal femur and lumbar spine were related to country of origin or centre. A further aim was to compare BMD means and standard deviations (SD's) with the NHANES III study [4] to determine whether current use of US-derived T-scores is generally appropriate to European populations.

A subsidiary objective of the study was to determine the size of any biasing effect of recruiting volunteer control subjects, since historically this method has been frequently used to measure the size of the effect of various bone diseases on BMD.

Subjects and Methods

Subjects

The principal group of population-based subjects (Table 1) were all selected from age-sex registers such as complete population listings, voter lists, or primary medical care lists of patients in countries with comprehensive health services and were approached by letter or telephone. Each subject was asked to attend once for a densitometry measurement of the hip and/or spine and height and weight measurement. Each centre had first received ethics permission to proceed with their study from their local Research Ethics

Table 1
Population-based young normal BMD data contributions by cohort study and centre to the NEMO collaboration

Study/Centre (ref)	Country	DXA brand	Male							Female						
			Age Mean (Range)	Femoral neck BMD		Trochanter BMD		L2-L4 Spine BMD		Age Mean (Range)	Femoral neck BMD		Trochanter BMD		L2-L4 Spine BMD	
				n	Mean (SD)	n	Mean (SD)	n	Mean (SD)		n	Mean (SD)	n	Mean (SD)	n	Mean (SD)
COMAC [2,14]																
Aberdeen	UK	Norland XR26	29 (26 - 30)	3	0.991 (0.171)	3	0.835 (0.112)	3	1.011 (0.103)	-	-	-	-	-	-	-
Leuven	Belgium	Lunar DPX	24 (22 - 28)	8	0.981 (0.134)	8	0.840 (0.124)	8	1.167 (0.131)	23 (20 - 28)	15	1.024 (0.161)	15	0.789 (0.144)	15	1.250 (0.143)
Manchester	UK	Lunar DPX	28 (27 - 29)	2	1.102 (0.065)	2	0.875 (0.046)	2	1.290 (0.053)	-	-	-	-	-	-	-
Harrow	UK	Hologic QDR1000	26 (24 - 29)	6	1.069 (0.103)	6	0.880 (0.105)	6	1.305 (0.173)	25 (21 - 30)	8	0.921 (0.172)	8	0.773 (0.074)	8	1.228 (0.156)
Kuopio	Finland	Lunar DPX	25 (22 - 30)	12	1.092 (0.219)	12	0.966 (0.203)	12	1.243 (0.165)	23 (21 - 27)	10	0.934 (0.177)	10	0.774 (0.155)	10	1.163 (0.090)
Total COMAC			26 (22 - 30)	31	1.050 (0.168)	31	0.898 (0.157)	31	1.216 (0.164)	24 (20 - 30)	33	0.972 (0.170)	33	0.781 (0.131)	33	1.218 (0.134)
EPOLOS																
wwa1	Poland	Lunar DPX	25 (20 - 29)	24	1.009 (0.205)	24	0.813 (0.145)	24	1.104 (0.151)	25 (20 - 30)	65	0.894 (0.117)	65	0.686 (0.102)	65	1.070 (0.103)
wwa2	Poland	Lunar DPX	25 (22 - 29)	6	0.968 (0.178)	6	0.740 (0.179)	6	1.018 (0.145)	27 (23 - 30)	6	0.833 (0.135)	6	0.620 (0.047)	6	1.006 (0.067)
pozn	Poland	Lunar DPX	25 (20 - 30)	7	1.007 (0.217)	7	0.740 (0.146)	7	0.993 (0.104)	26 (22 - 29)	10	0.977 (0.173)	10	0.727 (0.060)	10	1.027 (0.088)
lodz	Poland	Lunar DPX	24 (20 - 30)	11	0.973 (0.143)	11	0.728 (0.097)	11	1.043 (0.162)	24 (20 - 30)	28	0.942 (0.136)	28	0.683 (0.093)	28	1.058 (0.164)
krak	Poland	Lunar DPX	25 (20 - 29)	11	1.008 (0.135)	11	0.798 (0.113)	11	1.044 (0.070)	24 (20 - 30)	11	0.909 (0.073)	11	0.691 (0.060)	11	1.038 (0.117)
wroc	Poland	Lunar DPX	27 (22 - 30)	9	0.980 (0.138)	9	0.796 (0.124)	8	1.072 (0.139)	26 (21 - 30)	10	0.891 (0.117)	10	0.698 (0.113)	10	1.092 (0.106)
Total EPOLOS			25 (20 - 30)	68	0.995 (0.171)	68	0.781 (0.134)	67	1.061 (0.137)	25 (20 - 30)	130	0.909 (0.126)	130	0.687 (0.094)	130	1.060 (0.118)
Single centres																
Coimbra-JD	Portugal	Hologic QDR4500A	24 (19 - 30)	111	0.962 (0.147)	111	0.802 (0.122)	111	1.039 (0.136)	24 (19 - 30)	166	0.825 (0.114)	166	0.684 (0.099)	166	1.024 (0.113)
Ghent-JMK	Belgium	Hologic QDR4500	28 (24 - 30)	126	0.949 (0.143)	126	0.836 (0.135)	126	1.124 (0.145)	-	-	-	-	-	-	-
Kuopio-HK [15]	Finland	Lunar DPX	26 (21 - 29)	50	1.047 (0.154)	50	0.902 (0.150)	50	1.208 (0.125)	-	-	-	-	-	-	-
Prague-JS	Czech-Rep	Hologic QDR4500A	27 (27 - 27)	1	1.014 (-)	1	0.712 (-)	-	-	-	-	-	-	-	-	-
Odense-KB [16]	Denmark	Hologic QDR4500A	25 (20 – 30)	776	0.971 (0.160)	776	0.838 (0.144)	777	1.115 (0.149)	-	-	-	-	-	-	-
Total all centres			26 (19 - 30)	1163	0.974 (0.159)	1163	0.835 (0.143)	1162	1.112 (0.150)	25 (19 - 30)	329	0.873 (0.135)	329	0.695 (0.105)	329	1.058 (0.130)

Please cite this article as: Kaptoge S, et al, Geographical variation in DXA bone mineral density in young European men and women. Results from the Network in Europe on male osteoporosis (NEMO) study, Bone (2008), doi:10.1016/j.bone.2008.04.001

Committee according to the Declaration of Helsinki. The non-population-based subjects were recruited as described in Table 2, again according to the Helsinki Declaration.

DXA Procedures

Femoral neck, trochanter, and/or L2-L4 spine BMD were measured according to the manufacturer's instructions with either pencil beam or fan beam densitometers made by three manufacturers (Hologic, Lunar, and Norland). Each centre followed manufacturers' procedures to check the stability and accuracy of their equipment using regular (daily) measurements of the manufacturer's phantom. When measurements were inadvertently made during times of equipment malfunction, the data was excluded from this analysis.

Because the three brands do not give the same results while expressing their BMD results in the units of g/cm^2 it was necessary to cross-calibrate them using the European Spine Phantom (ESP). This was done as described by Pearson et al. [3], Kalender et al. [5], and Lunt et al. [6]. The minimum acceptable data for inclusion in the study was BMD at one measurement site (from femoral neck, trochanter, or L2-L4 spine), age, weight, and height at/or near the scan date.

Supplementary analyses were conducted using universally standardised BMD values (sBMD) [7,8] to assess the extent to which sBMD accounted for between-centre differences in mean BMD values attributable to different DXA manufacturer-brands used in the centres (Hologic, Lunar, Norland). Since the use of phantom data for universal standardisation has been found to be imprecise compared to using human data [7–9], in principle universal standardisation requires that a random sample of participants have their BMD measured using all the densitometers in the study [7,8], which was not possible in a retrospective meta-analysis. Thus instead, we first estimated the mean ESP cross-calibration parameters for each DXA brand using published mean BMD values from measurements of the first 30 ESP's on each DXA brand [5]. We then used these brand-specific mean ESP parameters to convert each participant's ESP-standardised BMD value into equivalent BMD estimates as expected from measurements using other DXA brands. Then we applied the optimal universal standardisation procedure of Hui L et al [8] to generate sBMD values. In keeping with previous convention for reporting, the sBMD results are reported in mg/cm^2 units.

Statistical analysis

Between-centre differences in weight and height adjusted mean BMD levels and trends with age were assessed using analysis of variance (ANOVA)/linear regression models. The same models were used to derive age and sex specific reference ranges for the ESP-standardised BMD values. Due to small numbers in some centres, the primary analyses were done with the centres combined by country. To enable a comparison with NHANES 3 US data, all the ESP standardised BMD values were transformed to Hologic QDR 1000W equivalents using ESP cross-calibration parameters [10]. Comparison was then made with the young normal values published by the NHANES III study [4].

Prior to the main analysis a check was done to investigate possible effects on BMD in Europe of using volunteer recruits, as has been common in the published literature. A nested ANOVA model was used to test whether on average there was a statistically significant difference in the mean BMD value for population-based vs. non-population based centres relative to variation between centres under the same classification (i.e. centres nested within recruitment method) adjusted for age, weight, and height. Differences in variance of BMD in the population-based vs. non-population based centres was assessed using Levene's test for equality of variance. Each sex was analysed separately.

Overall mean BMD across the centres was determined by pooling the adjusted centre-specific mean BMD values using random effects meta-analysis, which weights the contribution of each centre by the inverse of the variance of its estimated mean BMD plus estimated between centre variance (τ^2, tau^2) in BMD. Therefore, the pooled estimate is less likely to be biased towards results from large centres in the presence of significant heterogeneity ($\tau^2 > 0$) between centres. The extent of heterogeneity in mean BMD levels across centres was tested using Cochran's Q statistic, calculated as the sum of squared differences between each centre's adjusted mean BMD and the overall mean BMD weighted by the information provided by each centre [11]. Under the null hypothesis of no heterogeneity, the Q statistic follows a chi-square distribution with degrees of freedom equal to the total number of centres less 1 and provides a formal statistical test of significance of heterogeneity. The Q statistic was supplemented by the I^2 statistic for quantifying the impact of heterogeneity[11]. I^2 values range between 0% to 100% and represent the proportion of variation in the overall pooled estimate that is attributable to true between centre heterogeneity rather than chance. The centres were further sub-grouped by country to assess the extent to which the heterogeneity could be explained by country. Finally, centres with less than 10 subjects recruited were excluded and the data re-analysed to check that the random-effects modelling strategy had effectively adjusted for centre size in calculating means, SDs, and inter-country heterogeneity.

Results

For the men there was no significant effect of recruitment method, with the BMD difference between population-based vs. non-population based classifications being: femoral neck (-0.004 g/cm^2, P=0.32, n=1320); trochanter (-0.018 g/cm^2, P=0.89, n=1319) and L2-L4 spine (-0.011 g/cm^2, P=0.99, n=1222). However for the women, while the difference between groups was not significant at the femoral neck or trochanter (mean differences 0.094 and 0.020 g/cm^2, p=0.49 and 0.11 respectively, n=367), significant differences emerged at the L2-L4 spine measurement site (0.111 g/cm^2 p=0.01, n=390). Therefore, the

Table 2
Non population-based young normal BMD data contributions by cohort study and centre to the NEMO collaboration

Study/Centre (ref)	Country	DXA brand	Male							Female						
			Age Mean (Range)	Femoral neck BMD		Trochanter BMD		L2-L4 Spine BMD		Age Mean (Range)	Femoral neck BMD		Trochanter BMD		L2-L4 Spine BMD	
				n	Mean (SD)	n	Mean (SD)	n	Mean (SD)		n	Mean (SD)	n	Mean (SD)	n	Mean (SD)
COMAC [2,14]																
Madrid	Spain	Hologic QDR1000	27 (24 - 30)	-	-	-	-	11	1.257 (0.118)	27 (23 - 29)	-	-	-	-	13	1.249 (0.118)
Hvidore	Denmark	Norland XR26	-	-	-	-	-	-	-	25 (21 - 29)	2	0.978 (0.187)	2	0.812 (0.113)	11	1.194 (0.120)
Amsterdam	Netherlands	Norland XR26	28 (27 - 29)	3	1.000 (0.067)	3	0.917 (0.037)	3	1.048 (0.083)	25 (22 - 30)	6	0.994 (0.192)	6	0.812 (0.187)	6	1.258 (0.082)
Bern	Switzerland	Hologic QDR1000	-	-	-	-	-	-	-	26 (21 - 29)	11	1.000 (0.068)	11	0.831 (0.049)	11	1.328 (0.118)
Rotterdam-C	Netherlands	Lunar DPX	26 (21 - 30)	4	0.912 (0.139)	4	0.782 (0.095)	4	1.140 (0.162)	25 (23 - 29)	4	0.889 (0.089)	4	0.714 (0.084)	5	1.244 (0.121)
Wurzburg	Germany	Lunar DPX	27 (21 - 30)	11	0.949 (0.175)	11	0.796 (0.147)	11	1.055 (0.159)	25 (21 - 30)	15	0.928 (0.113)	15	0.710 (0.097)	15	1.143 (0.132)
Total COMAC			27 (21 - 30)	18	0.949 (0.151)	18	0.813 (0.130)	29	1.143 (0.163)	26 (21 - 30)	38	0.958 (0.119)	38	0.767 (0.114)	61	1.228 (0.131)
Single centres																
Ghent-confam	Belgium	Hologic QDR2000	24 (13 - 30)	20	0.939 (0.173)	20	0.808 (0.158)	-	-	-	-		-		-	
Ghent-conpro	Belgium	Hologic QDR2000	24 (19 - 29)	41	0.948 (0.129)	41	0.817 (0.131)	-	-	-	-		-		-	
Ghent-sons	Belgium	Hologic QDR1000+	25 (22 - 29)	52	1.056 (0.163)	52	0.910 (0.158)	-	-	-	-		-		-	
Paris-MC	France	Lunar DPXL	25 (19 - 30)	27	0.927 (0.133)	26	0.791 (0.109)	31	1.076 (0.163)	-	-		-		-	
Total all centres			25 (13 - 30)	158	0.979 (0.158)	157	0.842 (0.147)	60	1.108 (0.165)	26 (21 - 30)	38	0.958 (0.119)	38	0.767 (0.114)	61	1.228 (0.131)

Please cite this article as: Kaptoge S, et al, Geographical variation in DXA bone mineral density in young European men and women. Results from the Network in Europe on male osteoporosis (NEMO) study, Bone (2008), doi:10.1016/j.bone.2008.04.001

primary report is on the population-based data. There was no significant difference in the residual variance for BMD measurements at the three sites in the population-based centres vs. non-population based centres in either sex ($p > 0.19$).

Figs. 1 and 2 show the effects of country and centres within country on mean femoral neck and L2-L4 spine BMD respectively, with age, weight, and height adjusted to their overall mean values in the combined data. There was significant between centre heterogeneity in

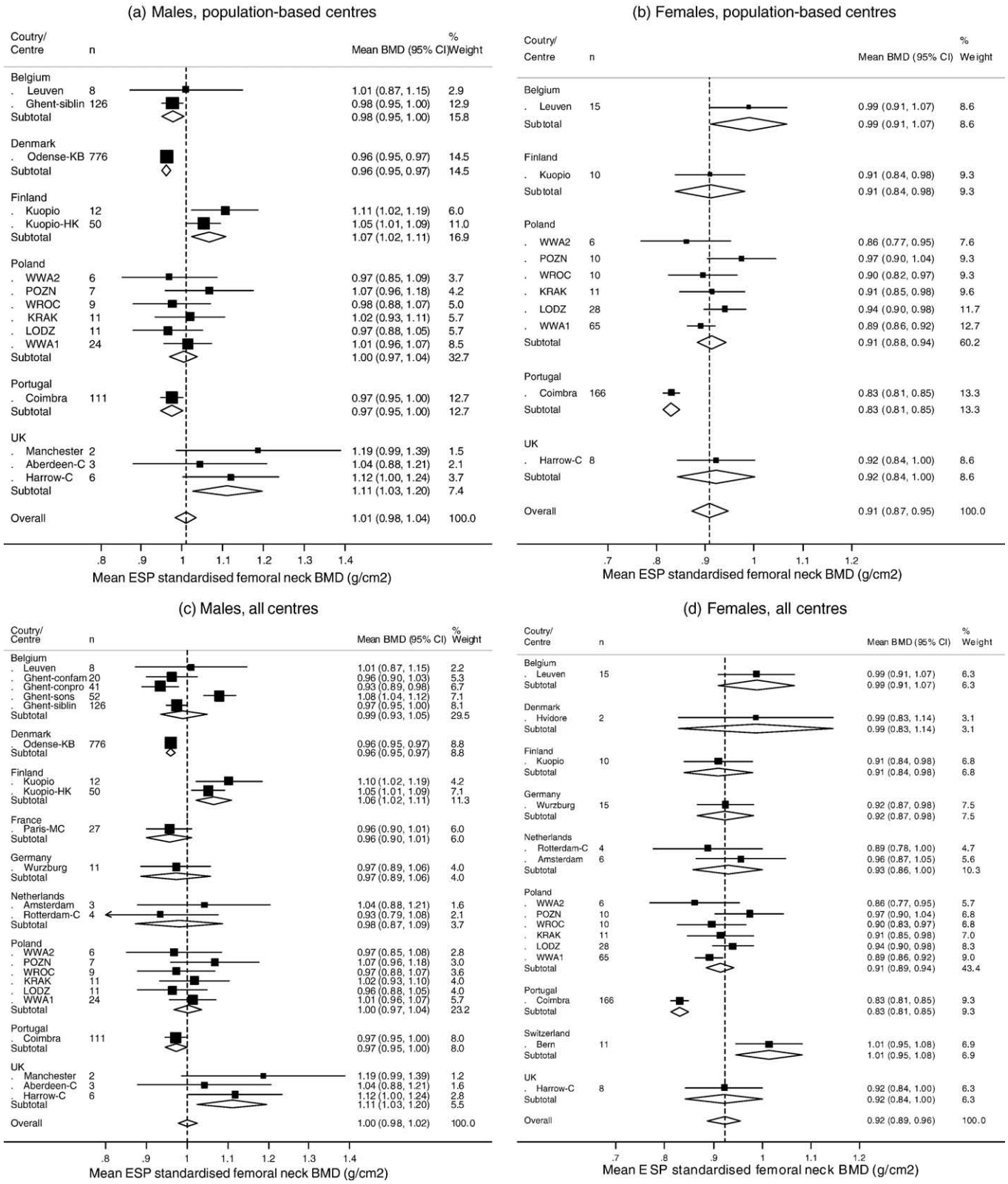

Fig. 1. ESP standardised femoral neck BMD means and 95% CIs for each country represented. Top: Population-based data for (a) males and (b) females; bottom: all data for (c) males and (d) females. Centre means were adjusted for men to age 26y, weight 80 kg, and height 181 cm and for women to age 25y, weight 59 kg, and height 164 cm; The numbers after each country indicate participant numbers. Country effects for population-based data: $p < 0.0001$ in men and $p = 0.004$ in women; for combined data $p < 0.0001$ for both sexes.

Please cite this article as: Kaptoge S, et al, Geographical variation in DXA bone mineral density in young European men and women. Results from the Network in Europe on male osteoporosis (NEMO) study, Bone (2008), doi:10.1016/j.bone.2008.04.001

the adjusted mean BMD levels for the population-based centres in both genders (Figs. 1 & 2 and Supplementary Table 1). The % weight column in Figs. 1 and 2 shows the weight as a proportion assigned to each centre/country by the random effects meta-analysis method when calculating the overall mean value, and it accounts for both the precision of the mean value for each centre as well as estimated between centre/country

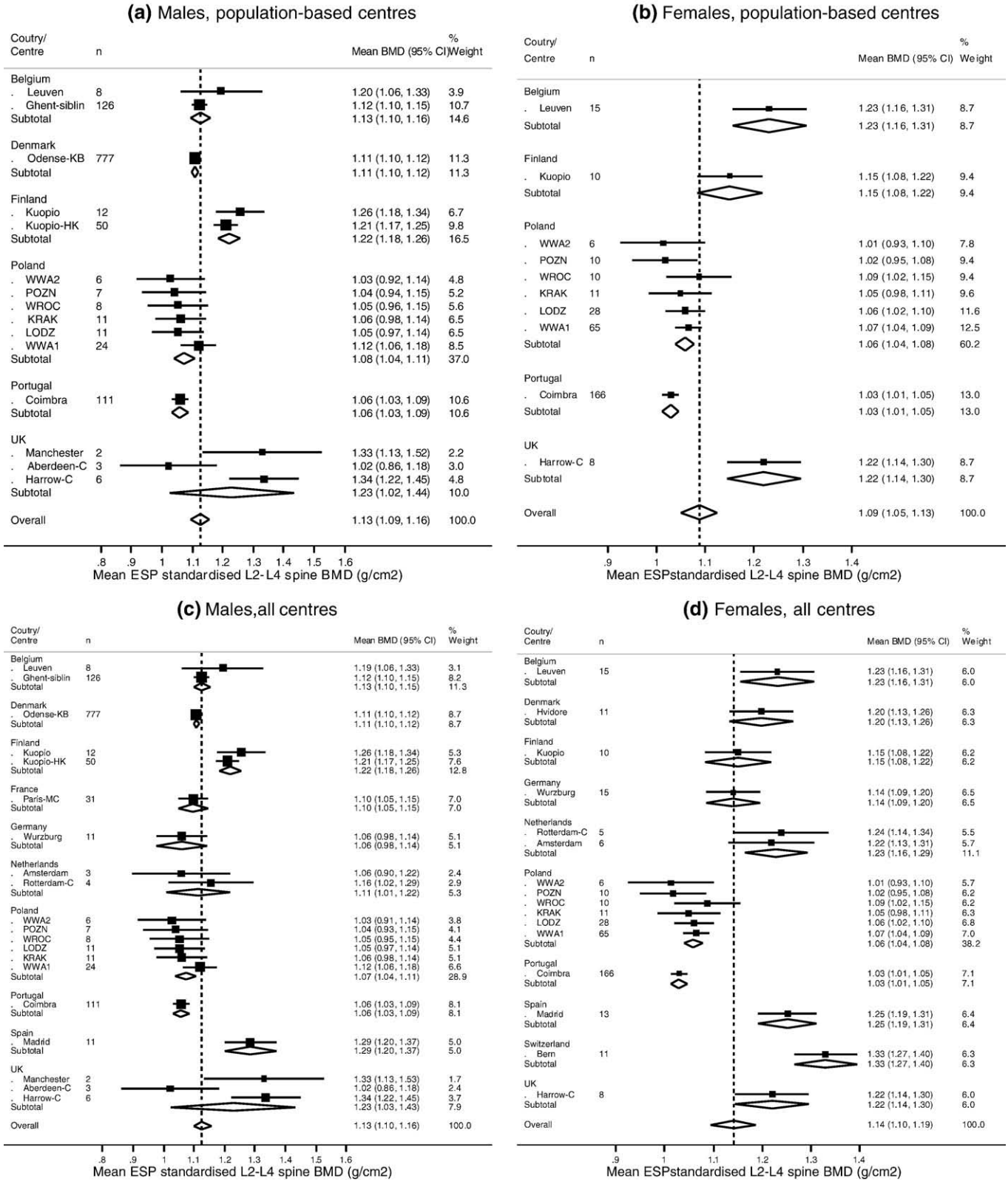

Fig. 2. ESP standardised L2-L4 BMD means and 95% CIs for each country represented. Top: Population-based data for (a) males and (b) females; bottom: all data for (c) males and (d) females; (the right shift in the overall mean value for females in the all data plot was due to the statistically significant effect of recruitment method – see text). Centre means were adjusted for men to age 26y, weight 80 kg, and height 181 cm and for women to age 25y, weight 59 kg, and height 164 cm; The numbers after each country indicate participant numbers. Country effects for all data shown: $p<0.0001$.

Please cite this article as: Kaptoge S, et al, Geographical variation in DXA bone mineral density in young European men and women. Results from the Network in Europe on male osteoporosis (NEMO) study, Bone (2008), doi:10.1016/j.bone.2008.04.001

Table 3
Summary of pooled mean value and statistics quantifying heterogeneity between centres and between countries in universally standardised young normal population-based BMD values at the femoral neck, trochanter and L2-L4 spine in males and females adjusted for age, weight, and height

	Using data from all population-based centres						Excluding centres with <10 participants				
Grouping factor / BMD region	Number of groups	Random effects pooled mean (95% CI)	Between-group SD τ [a]	Q statistic χ^2 (df)	p	I^2 (95% CI)	Random effects pooled Mean (95% CI)	Between-group SD τ [a]	Q statistic χ^2 (df)	P	I^2 (95% CI)
By centre											
Male											
Femoral neck	15	1014 (992, 1036)	29	48 (14)	<0.0001	71% (50%, 83%)	1007 (983, 1031)	27	33 (7)	<0.0001	79% (58%, 89%)
Trochanter	15	867 (846, 889)	29	50 (14)	<0.0001	72% (53%, 83%)	871 (846, 896)	30	41 (7)	<0.0001	83% (68%, 91%)
L2-L4 spine	15	1111 (1085, 1137)	39	84 (14)	<0.0001	83% (74%, 89%)	1110 (1082, 1138)	35	56 (7)	<0.0001	88% (78%, 93%)
Female											
Femoral neck	10	930 (897, 962)	45	55 (9)	<0.0001	84% (71%, 91%)	928 (882, 974)	48	50 (4)	<0.0001	92% (84%, 96%)
Trochanter	10	746 (729, 764)	18	19 (9)	0.022	53% (5%, 77%)	735 (721, 748)	8	6 (4)	0.235	28% (0%, 72%)
L2-L4 spine	10	1080 (1049, 1111)	43	59 (9)	<0.0001	85% (74%, 91%)	1070 (1035, 1106)	36	31 (4)	<0.0001	87% (72%, 94%)
By country											
Male											
Femoral neck	6	1005 (978, 1032)	29	35 (5)	<0.0001	86% (71%, 93%)	1001 (975, 1028)	27	31 (4)	<0.0001	87% (72%, 94%)
Trochanter	6	871 (843, 898)	30	40 (5)	<0.0001	87% (75%, 94%)	870 (843, 897)	28	34 (4)	<0.0001	88% (76%, 95%)
L2-L4 spine	6	1111 (1078, 1144)	37	61 (5)	<0.0001	92% (85%, 96%)	1107 (1074, 1139)	35	53 (4)	<0.0001	92% (85%, 96%)
Female											
Femoral neck	5	926 (879, 972)	47	46 (4)	<0.0001	91% (83%, 96%)	922 (861, 982)	51	46 (2)	<0.0001	96% (91%, 98%)
Trochanter	5	757 (733, 782)	21	15 (4)	0.004	74% (35%, 90%)	738 (718, 757)	13	5 (2)	0.068	63% (0%, 89%)
L2-L4 spine	5	1110 (1064, 1156)	47	56 (4)	<0.0001	93% (86%, 96%)	1082 (1033, 1130)	40	30 (2)	<0.0001	93% (84%, 97%)

[a]This estimate of the between-group standard deviation (τ) provides a means of calculating the expected range of mean values across centres/countries, for example, the pooled mean femoral neck sBMD in males was 1014 mg/cm^2 and as shown above τ=29, from which the 95% range of mean BMD values for individual centres would be estimated to be 1014±1.96*29 or 957 to 1071 mg/cm^2.

heterogeneity summarised in Supplementary Table 1. The estimate of the between centre/country standard deviation (τ) provides a means of calculating the expected range of mean values across centres/countries, for example, from Fig. 1 (a) the pooled mean femoral neck BMD in males was 1.01 g/cm^2 and Supplementary Table 1 shows that τ=0.035, from which the 95% range of mean BMD values for individual centres would be estimated to be 1.01±1.96*0.035 or 0.94 to 1.08 g/cm^2. There was moderate to high heterogeneity between centres, with the I^2 statistic in males being 70% at the femoral neck and 83% at the L2-L4 spine. In females the I^2 estimates were 84% at the femoral neck and 53% for L2-L4 spine. Similar levels of heterogeneity were observed for trochanter BMD with the I^2 statistic being 72% in men and 91% in females (Supplementary Table 1). There was no significant heterogeneity in BMD levels for centres within the same country at the femoral neck and trochanter sites in either sex (P>0.05). L2-L4 spine BMD appeared to be heterogeneous in the 3 UK centres in men, but there were only 11 male UK subjects (χ^2_2=10.7, P=0.005, I^2=81% (42%, 94%)). These estimates of between centre and country differences in mean BMD levels were unchanged in further sensitivity analysis that excluded centres that provided data from <10 participants (Supplementary Table 1).

Supplementary analyses using universally standardised BMD values (sBMD) gave similar conclusions to the analyses based on ESP-standardised BMD values. Thus the presence of large between-centre differences in mean BMD values persisted (Table 2, and Supplementary Figs. 2 and 3). The sBMD values were slightly higher in magnitude than the ESP-standardised BMD values (Table 3 in comparison with Supplementary Table 1, or Figs. 1 and 2 in comparison with Supplementary Figs. 1 and 2). The between-centre heterogeneity

Table 4
Comparison of European femoral neck and trochanter BMD data with US NHANES III data

Sex/ BMD region	European Data (current study)			US NHANES III data			European vs. US	
	n	Mean BMD g/cm^2 (95% CI)	SD	n	Mean BMD g/cm^2 (95% CI)	SD	P for difference in means	P for difference in SD
Male								
Femoral neck	1163	0.912 (0.904, 0.920)	0.132	382	0.93 (0.92, 0.94)	0.137	0.022	0.363
Trochanter	1163	0.793 (0.786, 0.800)	0.124	382	0.78 (0.77, 0.79)	0.118	0.072	0.244
L2-L4 spine	1162	1.027 (1.020, 1.034)	0.123	-	-	-	-	-
Female								
Femoral neck	330	0.826 (0.813, 0.838)	0.115	409	0.86 (0.85, 0.87)	0.120	<0.0001	0.420
Trochanter	330	0.670 (0.660, 0.680)	0.093	409	0.71 (0.70, 0.72)	0.099	<0.0001	0.236
L2-L4 spine	330	0.983 (0.971, 0.994)	0.107	-	-	-	-	-

To enable direct comparison with the NHANES III BMD data measured using Hologic QDR 1000 densitometer, the ESP standardised European data were first converted to Hologic QDR 1000 equivalents using Cross-calibration parameters for one such densitometer in the current study.

Please cite this article as: Kaptoge S, et al, Geographical variation in DXA bone mineral density in young European men and women. Results from the Network in Europe on male osteoporosis (NEMO) study, Bone (2008), doi:10.1016/j.bone.2008.04.001

statistics in these models using sBMD values were identical to those described above for the models using ESP-standardised BMD values.

Table 4 shows the femur data for the neck and trochanter regions as mean and SDs for use in deriving T-scores. These, when averaged across all centres, were compared statistically with the NHANES III population-based US data. For women it can be seen that mean values at both sites were significantly lower; but at all sites and in both sexes the SDs were very similar. Lumbar Spine L2-4 BMD was not studied in NHANES III. Thus, when it is necessary to adjust T-scores for differences between the young normative values in the US and a European country of interest, it should only be necessary to adjust the T-score by adding or subtracting the fraction of a SD unit represented by the difference between mean BMD values in the two countries. The expected 95% range for such adjustments in individual European centres can be estimated from the current findings based on the estimated standard deviation of the between centre/country heterogeneity (τ). For example, based on the between centre/country heterogeneity estimate (τ) in femoral neck BMD values in males being 0.035 g/cm^2 and the pooled mean being 0.91 g/cm^2, the 95% range of the mean femoral neck BMD values for individual European centres would be estimated to be [0.91 ± 1.96*0.035] or 0.84 to 0.99 g/cm^2. Hence the 95% range of the adjustments to be subtracted or added to the male femoral neck T-scores calculated based on US reference values (mean=0.93 g/cm^2, SD=0.137 g/cm^2) in individual European centres would be expected to be: (0.93 – 0.99)/0.137 to (0.93 – 0.84)/0.137 or -0.44 to +0.66 SD's.

Supplementary Fig. 3 shows the cross-sectional age-related changes in ESP cross-calibrated DXA BMD at the femoral neck, trochanter and L2-L4 spine in the young normal males from population-based and non population-based centres. There was a modest but statistically significant decline in femoral neck and trochanter BMD and an increase in L2-L4 spine BMD between the ages 20 and 30 years. Supplementary Fig. 4 shows the cross-sectional age-related changes for the young normal females.

Discussion

This is the first large population-based study across a continent to describe young normal values for BMD of the spine. It therefore extends and complements the NHANES III study [4] of femoral BMD. It was also of great interest to contrast the femoral BMD results obtained in the US and Europe because of the world-wide use of US data as a normative standard. The key message is that while mean European values for BMD in the femur were sometimes different to those in the US, the SDs were indistinguishable. This is of crucial importance when it is desired to use T-scores because the T-score is calculated from the difference between the measured value and the mean young normal value divided by the young normal SD. In consequence, when it is necessary to adjust T-scores for differences between the young normative values in the US and a European country of interest, it should only be necessary to adjust the T-score by adding or subtracting the fraction of a SD unit represented by the difference between means BMD values in the two countries.

The small but significant differences observed in our study between countries lead us to conclude that for public health or therapeutic purposes it might be necessary to collect more extensive young normal data. However this would only be of major importance where treatment decisions are made on the basis of a fixed T-score threshold as in the UK [12]. In previous work it has been observed that there are quite substantial between centre differences in BMD in older populations of men and women [6] and that rates of bone loss or gain differ between centres [13]. The between country and between centre differences in mean BMDs described in this paper will allow the study of whether these differences are apparently growing over time (implying an impact of ecology on the evolution of BMD) or whether they are constant across the range of adult ages in men and women.

This study has limitations. It was undertaken over a long time-span (15 years), using 3 brands of DXA equipment that differed in design and performance, where cross calibration can only be imperfect. Quality control issues are important in DXA studies, but all the staff responsible for collecting data were experienced operators and obtained their scan data according to the scanner manufacturers' recommended procedures. This was a cross-sectional study, so issues relating specifically to longitudinal studies and the measurement of small differences in BMD over time in the same subjects do not apply. Moreover, collecting young normal data has proved challenging in the past and our data is likely to be the best available European data-set for some time to come. Our data regarding the effects of recruitment procedure suggest that asking for volunteers, at least in young women, will not always provide a fully representative sample of BMD values in the normal population. We were limited by the availability of submitted data in making inferences on country-specific adjustments that may be necessary to correct T-scores calculated based on a US reference population, since this would necessarily require a nationally representative sampling scheme to be employed, whereas only a few centres represented each country in our study with small numbers in some. That the between-centre differences persisted despite the application of universally standardised BMD values may imply that it is unlikely that such differences were due to variations in DXA equipment, although we cannot exclude the possibility that differences in placement of the region of interest may be a contributing factor. For public health purposes it is safer to assume that these differences are real.

In conclusion, we have provided young normal data for both sexes for the spine for the first time and a European database for femoral BMD values that should allow more confident interpretation of clinical BMD measurements in Europeans. The origins of the fairly modest differences between European populations in mean spine and femoral BMD values, not addressed in this study, might reflect differences in genetic background or environmental exposures.

Acknowledgments

We thank all the skilled densitometry technicians for their expertise and the subjects who gave their time to establish normal BMD databases. This study was funded by the EU's FP5 programme (Ref NEMO C1PDCT925102, ERBC1PDCT 930105 & 940229).

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bone.2008.04.001.

References

[1] Boonen S, Kaufman JM, Goemaere S, Bouillon R, Vanderschueren D. The diagnosis and treatment of male osteoporosis: Defining, assessing, and preventing skeletal fragility in men. Eur J Intern Med 2007;18:6–17.
[2] Dequeker J, Pearson J, Reeve J, Henley M, Bright J, Felsenberg D, et al. Dual X-ray absorptiometry: Cross-calibration and normative reference ranges for the spine: results of a European Community Concerted Action. Bone 1995;17:247–54.
[3] Pearson J, Ruegsegger P, Dequeker J, Henley M, Bright J, Reeve J, et al. European semi-anthropomorphic phantom for the cross-calibration of peripheral bone densitometers: assessment of precision accuracy and stability. Bone Miner 1994;27:109–20.
[4] Looker AC, Orwoll ES, Johnston Jr CC, Lindsay RL, Wahner HW, Dunn WL, et al. Prevalence of low femoral bone density in older US adults from NHANES III. J Bone Miner Res 1997;12:1761–8.
[5] Kalender WA, Felsenberg D, Genant HK, Fischer M, Dequeker J, Reeve J. The European Spine Phantom - a tool for standardization and quality control in spinal bone mineral measurements by DXA and QCT. Eur J Radiol 1995;20:83–92.
[6] Lunt M, Felsenberg D, Adams J, Benevolenskaya L, Cannata J, Dequeker J, et al. Population-based geographic variations in DXA bone density in Europe: the EVOS study. Osteoporosis Int 1997;7:175–89.
[7] Genant HK, Grampp S, Glüer CC, Faulkner KG, Jergas M, Engelke K, et al. Universal standardisation for dual X-ray absorptiometry: patient and phantom cross-calibrations. J Bone Miner Res 1994;9:1503–14.
[8] Hui SL, Gao S, Zhou XH, Johnston Jr CC, Lu Y, Gluer CC, et al. Universal standardization of bone density measurements: a method with optimal properties for calibration among several instruments. J Bone Miner Res 1997;12:1463–70.

[9] Finkelstein JS, Butler JP, Cleary RL, Neer RM. Comparison of four methods for cross-calibrating dual-energy X-ray absorptiometers to eliminate systematic errors when upgrading equipment. J Bone Miner Res 1994;9:1945–52.
[10] 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. Osteoporos Int 1995;5:174–84.
[11] Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539–58.
[12] National Institute for Clinical Excellence (NICE). Osteoporosis: assessment of fracture risk and the prevention of osteoporotic fractures in individuals at high risk; 2007. http://www.nice.org.uk/page.aspx?o=guidelines.inprogress.osteoporosis.
[13] Kaptoge S, Reid DM, Scheidt-Nave C, Poor G, Pols HA, Khaw KT, et al. Geographic and other determinants of BMD change in European men and women at the hip and spine. A population-based study from the Network in Europe for Male Osteoporosis (NEMO). Bone 2007;40:662–73.
[14] Dequeker J, Reeve J, Pearson J, Bright J, Felsenberg D, Kalender W, et al. Multicentre European COMAC-BME study on the standardisation of bone densitometry procedures. Technol Health Care 1993;1:127–31.
[15] Kroger H, Laitinen K. Bone mineral density measured by dual-energy X-ray absorptiometry in normal men. Eur J Clin Invest 1992;22:454–60.
[16] Høiberg M, Nielsen TL, Wraae K, Abrahamsen B, Hagen C, Andersen M, et al. Population-based reference values for bone mineral density in young Danish men. Osteoporos Int 2007;18:1504–14.

Please cite this article as: Kaptoge S, et al, Geographical variation in DXA bone mineral density in young European men and women. Results from the Network in Europe on male osteoporosis (NEMO) study, Bone (2008), doi:10.1016/j.bone.2008.04.001