

## Predicting the Risk of Fracture at Any Site in the Skeleton: Are All Bone Mineral Density Measurement Sites Equally Effective?

G. M. Blake,<sup>1</sup> K. M. Knapp,<sup>2</sup> T. D. Spector,<sup>3</sup> I. Fogelman<sup>1</sup>

<sup>1</sup>Imaging Sciences, Guy's, King's and St Thomas' School of Medicine, Guy's Campus, London, United Kingdom

<sup>2</sup>School of Physics, University of Exeter, Exeter, United Kingdom

<sup>3</sup>Twin Research and Genetic Epidemiology Unit, St Thomas' Hospital, London, United Kingdom

Received: 20 July 2005 / Accepted: 30 August 2005 / Online publication: 5 December 2005

**Abstract.** The ability to assess a patient's risk of fracture is fundamental to the clinical role of bone densitometry. Fracture discrimination is quantified by the *relative risk* (RR), defined as the increased risk of fracture for a 1 standard deviation decrease in bone mineral density (BMD). The larger the value of RR, the more effective measurements are at identifying patients at risk of fracture. Epidemiological studies show that RR values for predicting the risk of any fracture are approximately the same for all BMD measurement sites. In this study, we show theoretically that this interesting observation is predictable and a consequence of two related observations: (1) that fracture prediction by BMD measurement sites distant from the fracture site is quantitatively explained by the correlation of BMD measurements and (2) that all correlation coefficients between distant BMD sites are comparable, with values in the range  $r = 0.55$ – $0.65$ . The first of these conditions (referred to as the *correlation hypothesis*) is important because it sets a lower limit on the RR values at distant BMD sites on the assumption that measurements at these sites contain no independent information about fracture risk over and above that provided by their correlation with the fracture site BMD. If the correlation hypothesis is true, the present study points to the importance of the correlation coefficient between BMD sites as a key index that is indicative of the ability of different types of measurement to predict fracture risk. If, on the contrary, the correlation hypothesis is not valid, there is scope to improve bone densitometry by further studies to better identify those measurements that do provide independent information about fracture risk and how best to integrate this information with existing techniques to improve decision making.

**Key words:** Bone densitometry — Correlation coefficient — Fracture risk prediction

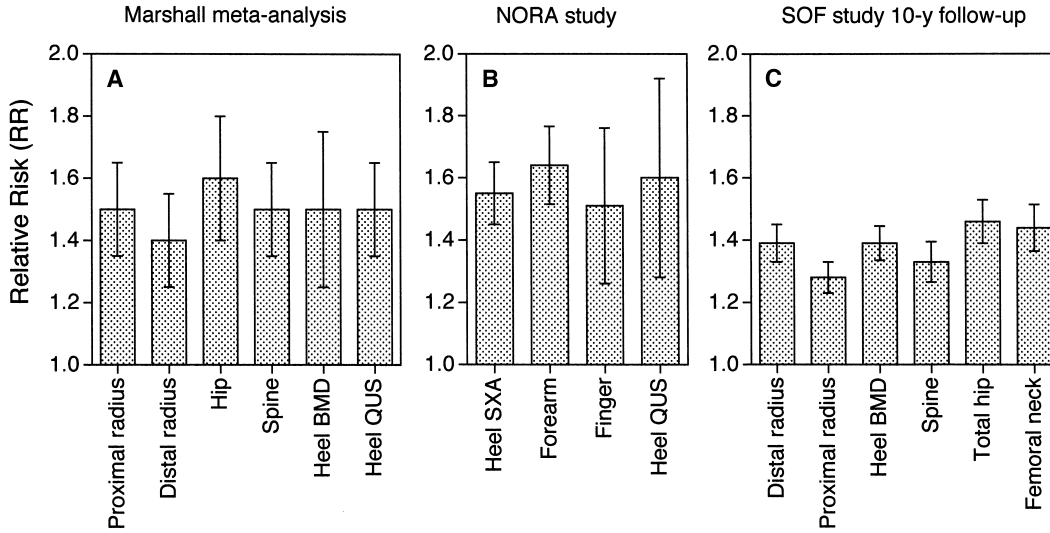
Osteoporosis is widely recognised as a serious problem in public health. Increased awareness of the impact of fractures on the lives of elderly people [1], the costs of

healthcare [2], and the availability of new treatments [3–5] have all contributed to a growing demand for bone densitometry services. Today, scans to measure bone mineral density (BMD) have an essential role in the evaluation of patients with osteoporosis [6]. Often, the preferred method of investigation is to measure hip and spine BMD using dual X-ray absorptiometry (DXA) [6–8]. The reasons for this choice include the fact that hip BMD is the best predictor of hip fracture risk [9, 10], the use of spine BMD for monitoring treatment [11], and the consensus that spine, hip, and forearm BMD results should be interpreted using the World Health Organisation (WHO) definition of osteoporosis of a T-score of  $-2.5$  or below [8, 12].

In addition to axial DXA systems that measure spine and hip BMD, a variety of devices for measuring sites in the peripheral skeleton are available [13, 14]. These include peripheral DXA (pDXA) systems for the forearm, heel, or hand and quantitative ultrasound (QUS) devices for measurements in the heel and other peripheral sites. Because osteoporosis is common and usually managed in primary care, there is a need for cheaper and more convenient methods of evaluating BMD than axial DXA, and it is widely believed that peripheral devices might fulfil this role [15, 16].

Fundamental to the clinical role of BMD scans is the ability to assess a patient's risk of fracture. The most reliable way of evaluating the effectiveness of any bone densitometry technique is through prospective studies of incident fractures [9, 10]. Such studies are analysed using a proportional hazards model in which the findings are expressed as the *relative risk* (RR), defined as the increased risk of fracture for a 1 standard deviation (SD) decrease in BMD [9]. The greater the value of RR, the more effective a technique is at identifying patients at risk of fracture [17].

One of the important findings to emerge from fracture studies is that when it comes to the ability to predict the risk of fracture at any site in the skeleton, different types of BMD measurement appear to be equally



**Fig. 1.** Bar charts showing RR values and 95% CI for the prediction of any fracture by bone densitometry measurements at different skeletal sites. (A) Figures from the meta-analysis of Marshall et al. [9]. (B) Results from the NORA study [18]. (C)

Results from the SOF 10-year follow-up [10]. RR values are defined as the increased risk of fracture for a 1 SD decrease in BMD. SXA, single X-ray absorptiometry.

effective (Fig. 1). This is in contrast to the prediction of fracture risk at a specific site when the best measurement is one made at the fracture site itself [9, 10]. The fact that RR values for the prediction of any fracture are approximately the same for all BMD sites was first suggested by the meta-analysis of Marshall et al. [9], which found RR values of around 1.5 for spine, hip, forearm, and heel BMD as well as heel QUS (Fig. 1A). Subsequently, the National Osteoporosis Risk Assessment (NORA) study [18] reported a similar result for heel, forearm, and finger BMD as well as heel QUS (Fig. 1B). These findings were confirmed by the publication of the 10-year follow-up data for the Study of Osteoporotic Fractures (SOF) [10], which reported RR values of around 1.4 for six different BMD sites (Fig. 1C). With the longer follow-up time, the new SOF study data show slightly smaller RR values than earlier studies, although with the larger number of fractures the statistical errors are smaller. While some sites may perform slightly better or worse than others, overall the different RR values in Figure 1C are remarkably consistent.

In this study, we show that the observation that RR values for predicting any fracture are broadly equal for all BMD sites is a consequence of two related observations: (1) fracture prediction by BMD sites distant from the fracture site is quantitatively explained by the correlation of the BMD measurements [19] and (2) all correlation coefficients between distant BMD sites are comparable, with values in the range  $r = 0.55\text{--}0.65$  [20]. On this basis, we show that the findings presented in Figure 1 are predictable and have a simple explanation.

## Material and Methods

### Theoretical Analysis of Fracture Risk Prediction

We begin by considering the prediction of fracture risk at a single site  $i$  in the skeleton by a BMD measurement made at site  $j$ . It is convenient to express BMD measurements in terms of their Z-scores. Z-scores are calculated by taking the difference between the measured BMD and the mean BMD for healthy subjects matched for age, gender, and ethnic group and dividing by the population SD. We assume that for a group of subjects representative of the nonfracture population, the distribution of Z-score values for the measurement site  $j$  approximates a gaussian curve, with its peak at  $Z_j = 0$  and  $SD_j = 1.0$  [21]:

$$N(Z_j) = \frac{1}{\sqrt{2\pi}} \exp(-Z_j^2/2) \quad (1)$$

The proportional hazards model used to analyse fracture studies gives the following exponential equation between the fracture risk at site  $i$  and the Z-score at site  $j$ :

$$P_{ij}(Z_j) = P_i \exp(-\beta_{ij}^2/2) \bullet \exp(-\beta_{ij}Z_j) \quad (2)$$

In equation 2,  $P_i$  is the mean fracture risk at site  $i$  averaged over the whole population,  $\beta_{ij}$  is the logarithm of the RR for the measurement at site  $j$  to predict a fracture at site  $i$  [ $\beta_{ij} = \ln(RR_{ij})$ ], and the factor  $\exp(-\beta_{ij}^2/2)$  ensures that the total number of fractures at site  $i$  is independent of the value of  $\beta_{ij}$ . The Z-score distribution of the fracture population,  $N_i(Z_j)$ , is found by multiplying equation 1 by equation 2. After rearrangement, one obtains the following:

$$N_i(Z_j) = \frac{1}{\sqrt{2\pi}} P_i \exp[-(Z_j + \beta_{ij})^2/2] \quad (3)$$

Equation 3 shows that the Z-score distribution of the fracture population is a gaussian curve with the same SD as the nonfracture population and a mean Z-score of  $Z_j = -\beta_{ij}$ .

Equation 3 can be used to predict the risk of fracture at any site in the skeleton by adding the contribution of each fracture site. For fracture site  $i$ , the mean Z-score of the fracture cases at the measurement site  $j$  is  $Z_j = -\beta_{ij}$ . The mean Z-score for

all fracture cases is therefore equal to the sum over all sites of the mean Z-score  $-\beta_{ij}$  weighted by the fraction  $f$  of all fractures that occur at the site ( $f_i = P_i / \sum_i P_i$ ):

$$Z_{j,all} = -\sum_i f_i \beta_{ij} \quad (4)$$

It follows from equation 3 that the  $\beta$  value  $\beta_{j,all}$  for predicting any fracture from a BMD measurement at site  $j$  is found by substituting  $\beta_{j,all}$  for  $-Z_{j,all}$  in equation 4:

$$\beta_{j,all} = \sum_i f_i \beta_{ij} \quad (5)$$

In a previous study [19], we showed that data from fracture studies are consistent with the hypothesis that fracture prediction by BMD measurements made at sites distant from the fracture site is explained by the correlation between the BMD measurements at the two sites. We refer to this assumption as the *correlation hypothesis*. The correlation hypothesis is useful because it sets a lower limit on the RR value for fracture prediction by distant BMD sites on the null hypothesis that measurements at distant sites contain no information about fracture risk over and above that explained by their correlation with the BMD measurement at the fracture site. The correlation hypothesis predicts the following relationship between the  $\beta$  value  $\beta_{ii}$  for a BMD measurement at the fracture site, the  $\beta$  value  $\beta_{ij}$  for a distant BMD site, and the correlation coefficient  $r_{ij}$  between the Z-scores of the two BMD measurements [19]:

$$\beta_{ij} = r_{ij} \beta_{ii} \quad (6)$$

This basic prediction must be modified to allow for the effects of the random BMD measurement errors caused by inhomogeneities in soft tissue composition [22]. The effect of random errors in the distant BMD measurement is to decrease  $\beta_{ij}$  and  $r_{ij}$  by equal factors so that the underlying relationship in equation 6 is unchanged [19]. In contrast, the effect of the errors in the BMD measurements at the fracture site is to decrease  $r_{ij}$  and  $\beta_{ii}$  by equal factors such that the relationship for the distant BMD sites becomes  $\beta_{ij} = r_{ij} s_i \beta_{ii}$ , where  $s_i = \sigma^2 / (\sigma^2 - \epsilon^2)$  is the ratio of the population variance  $\sigma^2$  with and without the contribution  $\epsilon^2$  of the random measurement errors [19]. After substitution in equation 5, we derive the following relationship:

$$\beta_{j,all} = \sum_{i \neq j} f_i r_{ij} s_i \beta_{ii} + f_j \beta_{jj} \quad (7)$$

where the first term represents the sum over the fractures occurring at all the distant skeletal sites and the second term, the contribution of fractures occurring at the BMD measurement site itself.

In practice, all correlation coefficients between BMD measurements made at distant sites in the skeleton have similar values, with figures in the range  $r = 0.55-0.65$  [20]. Substituting the mean value  $\bar{r}$  for  $r_{ij}$  in equation 7 and taking the appropriate part of the contribution from fractures at the measurement site into the sum over all sites, we finally obtain:

$$\beta_{j,all} = \bar{r} \left( \sum_i f_i s_i \beta_{ii} \right) + (1 - \bar{r} s_j) f_j \beta_{jj} \quad (8)$$

The random errors in BMD due to inhomogeneities in soft tissue composition are typically about 50% of the total population SD [22]. It follows that the value of the factor  $s_i$  in equation 8 is  $1/(1 - 0.5^2) = 1.33$ . Assuming  $\bar{r} = 0.6$ , this gives a value of 0.2 for the coefficient  $(1 - \bar{r} s_j)$  in the second term in equation 8.

#### Calculation of RR for All Fractures

The estimate of  $\beta_{j,all}$  given in equation 8 is dominated by the first term, which is a sum over all fracture sites  $i$  whose value is

independent of the particular choice of BMD measurement site  $j$ . It follows that, with the exception of a small additional contribution from those fractures that occur at the measurement site itself expressed in the second term in equation 8, the RR value for all fractures is constant across all BMD sites. To evaluate this common RR value, we used data from the SOF listing the frequency of different types of fracture and their RR values [23]. We also used published SOF data [20] to examine the assumption of a common value of the correlation coefficient.

There is an important proviso to the inference above that the  $\beta$  value for all fractures can be calculated using equation 5. Unlike the contributions from the individual fracture sites, the modelled curve of fracture risk against Z-score for all fractures is not a single exponential. To evaluate the consequences of this, the SOF data were used to calculate the Z-score distribution of the entire fracture population. The fracture risk curve for all fractures was fitted by a single exponential between the Z-score limits representing the 5th and 95th centiles of all fracture cases and the resulting value of RR compared with the prediction from equation 5.

#### Value of the Correlation Coefficient between Distant BMD Sites

The explanation of the constant RR value for all fractures across different BMD sites depends on the assumption that the correlation coefficient between BMD measurements made at distant sites in the skeleton is approximately the same for all pairs of sites. While there is substantial evidence to support this assumption [20, 24], it is restricted to comparisons between a small number of sites such as the spine, hip, heel, and forearm. To examine the evidence for a larger number of sites, we examined BMD measurements obtained from spine, hip, forearm, and total body DXA scans in a study of 422 healthy white female twins aged 21–80 years [25]. The scans were performed on a QDR4500 (Hologic, Bedford, MA) and the data examined to determine values of the Z-score correlation coefficient between the measurements at the different BMD sites including all the subregions of the total body scans (lumbar spine, thoracic spine, pelvis, head, ribs, arms, and legs). Since reference data were not available for the total body subregions, Z-scores were calculated using the 10-year running mean of the BMD data.

## Results

Data on the incidence of fractures at different sites, the  $\beta$  values and 95% confidence intervals (CIs) taken from the SOF study 2-year follow-up [23] were substituted in equation 5 to calculate RR values for prediction of any fracture for measurements of BMD at the distal radius, proximal radius, and calcaneus. The results (and 95% CIs) were as follows: distal radius BMD = 1.54 (1.44–1.66), proximal radius BMD = 1.44 (1.35–1.54), and calcaneus BMD = 1.55 (1.44–1.66) (Table 1). As explained above, equation 5 does not allow for the fact that the dependence of fracture risk on Z-score for all fractures is no longer exactly a single exponential. RR values were therefore recalculated by substituting the same data in equation 2, adding the contributions from all the different fracture sites and fitting the resulting curves of total fracture risk against Z-score with a single exponential between the 5th and 95th centiles for all incident fractures. The three curves showed a close fit to a single exponential ( $r^2 = 0.9981, 0.9988, \text{ and } 0.9986$

**Table 1.** An example of the use of equation 5 to calculate RR values for calcaneus, proximal radius, and distal radius BMD for prediction of any osteoporotic fracture<sup>a</sup>

Fracture site	Number of women (%)	Calcaneus RR (95% CI)	Proximal radius RR (95% CI)	Distal radius RR (95% CI)
Wrist	171 (19.3%)	1.76 (1.49–2.09)	1.63 (1.39–1.90)	1.82 (1.53–2.16)
Foot	91 (10.3%)	1.39 (1.11–1.75)	1.29 (1.04–1.59)	1.43 (1.14–1.80)
Ankle	81 (9.2%)	1.11 (0.88–1.39)	1.14 (0.90–1.43)	1.11 (0.88–1.39)
Humerus	79 (8.9%)	1.85 (1.44–2.38)	1.96 (1.56–2.47)	2.04 (1.57–2.65)
Hip	78 (8.8%)	1.70 (1.32–2.19)	1.42 (1.12–1.80)	1.46 (1.13–1.88)
Vertebra	67 (7.6%)	1.70 (1.30–2.22)	1.40 (1.09–1.80)	1.51 (1.15–1.98)
Rib	62 (7.0%)	1.53 (1.16–2.01)	1.54 (1.19–2.00)	1.80 (1.35–2.41)
Toe	56 (6.3%)	1.39 (1.04–1.85)	1.39 (1.06–1.82)	1.54 (1.15–2.05)
Elbow	43 (4.9%)	1.25 (0.90–1.72)	1.18 (0.87–1.61)	1.15 (0.84–1.59)
Finger	29 (3.3%)	1.15 (0.78–1.70)	1.08 (0.74–1.59)	1.22 (0.83–1.79)
Leg	27 (3.1%)	1.73 (1.12–2.67)	1.52 (1.03–2.25)	1.97 (1.27–3.05)
Patella	27 (3.1%)	1.77 (1.14–2.75)	1.32 (0.89–1.97)	1.25 (0.82–1.91)
Pelvis	26 (2.9%)	2.44 (1.55–3.84)	1.61 (1.08–2.40)	1.30 (0.86–1.97)
Hand	22 (2.5%)	1.63 (1.00–2.63)	1.69 (1.10–2.62)	1.87 (1.15–1.68)
Face	15 (1.7%)	1.10 (0.64–1.88)	0.99 (0.58–1.69)	0.99 (0.59–1.68)
Clavicle	10 (1.1%)	1.10 (0.57–2.11)	1.96 (1.03–3.72)	2.63 (1.23–5.62)
<b>Any site</b>	<b>884 (100%)</b>	<b>1.55 (1.44–1.66)</b>	<b>1.44 (1.35–1.54)</b>	<b>1.54 (1.44–1.66)</b>

<sup>a</sup> Data for RR values at individual fracture sites taken from the SOF 2-year follow-up [23]

for distal radius, proximal radius, and calcaneus, respectively), and the RR values agreed well with those calculated from equation 5 (distal radius BMD 1.54, proximal radius BMD 1.44, calcaneus BMD 1.54).

If it is accepted that equation 5 gives a reasonably accurate estimate of the RR values for the prediction of any fracture, we have shown that two further steps are required to explain the observation in Figure 1 that RR values for the prediction of any fracture are approximately constant for different BMD measurement sites. These are the correlation hypothesis and the constant value of the correlation coefficient between distant BMD sites. The evidence for the correlation hypothesis has been presented previously [19] and will not be discussed in detail again here. Instead, we concentrate on the second requirement, the constant value of the BMD correlation coefficient.

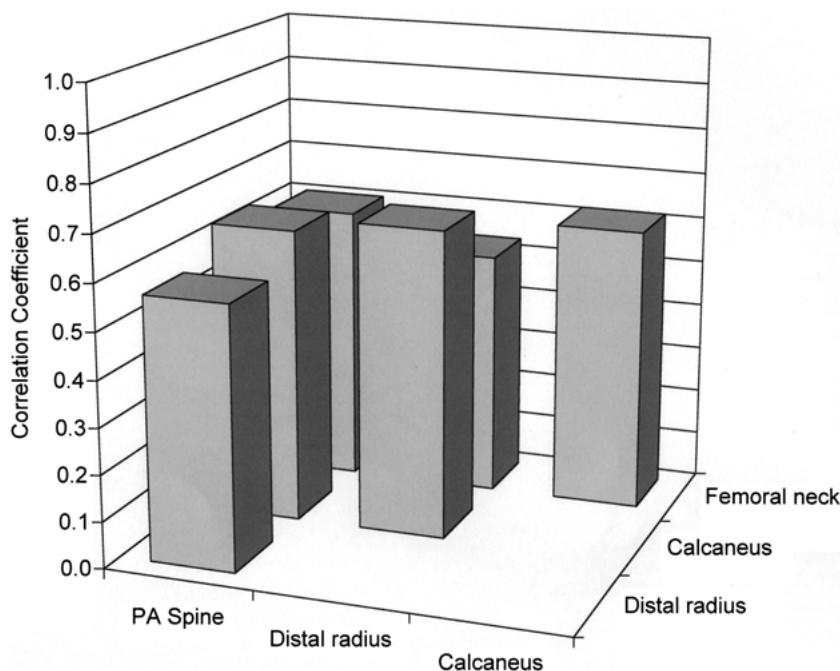
Statistically, some of the most accurate data for correlation coefficients between different BMD sites are those published for over 7,600 postmenopausal white women enrolled in the SOF [20]. It is notable that the SOF correlation coefficients between the more distant BMD sites in the skeleton (spine, hip, forearm, and heel) fall within a relatively narrow range ( $r = 0.53–0.67$ , Fig. 2). In contrast, the correlation coefficients for adjacent sites (e.g., the different regions of interest in the hip) are consistently larger ( $r = 0.76–0.90$ ) [20].

Data for a wider range of sites across the skeleton were provided by the twin study. However, since subjects in the latter study had a wide range of ages (21–81 years) compared with the SOF (age 65 years and over), the effect of the age difference on the resulting correlation coefficients was examined by comparing the results for the twin and SOF studies for the measurement sites

they had in common (spine, femoral neck, total hip, and distal radius) (Fig. 3). The results showed that, provided the Z-score correlation coefficient data were used for the twin study, the effects of the wide age range were minimised and the two sets of correlation coefficients were consistent within the limits set by the statistical errors.

Results from the twin study for the Z-score correlation coefficient between the different subregions in the total body scans (Fig. 4A) and between the total body subregions and PA spine, hip, and forearm BMD (Fig. 4B) were plotted for comparison with the SOF data in Figure 2. Results for head BMD were systematically low compared with the other correlation coefficients and varied from  $r = 0.38$  (head vs. femoral neck BMD) to  $r = 0.53$  [head vs. posteroanterior (PA) spine BMD]. However, when the head site was excluded, the values between the other six total body subregions (arms, ribs, thoracic spine, lumbar spine, pelvis, and legs) varied between  $r = 0.51$  and  $0.69$  (Fig. 4A), in good agreement with the SOF data ( $r = 0.53–0.67$ , Fig. 2). When the spine, hip, and forearm BMD Z-scores were compared with the total body subregions, the correlation coefficients were  $r = 0.49–0.74$  (Fig. 4B, excluding the high value  $r = 0.90$  between the PA spine BMD and the equivalent lumbar spine region on the total body scans).

Values of the correlation coefficients were smoothed by plotting the mean value of  $r$  between each total body subregion and each of the other six subregions (Fig. 5A). The mean values of the correlation coefficient between the PA spine, femoral neck, total hip, and distal forearm and the seven total body subregions are also shown in the same figure. As noted above, the lowest

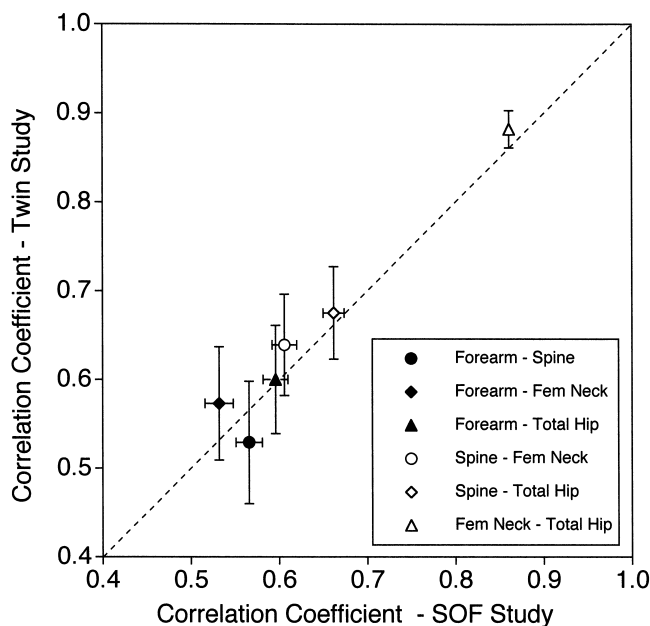


**Fig. 2.** Values of the correlation coefficient between BMD measurements in four distant regions of interest in the skeleton (hip, spine, forearm, and heel) for 7,600 white women aged 65 years and over enrolled in the SOF [20].

value was for head BMD ( $r = 0.47$ ), while for other regions the mean varied between  $r = 0.53$  and  $0.61$  for the total body subregions and  $r = 0.53$  and  $0.66$  for the PA spine, hip, and forearm BMD. There is a trend in the data shown in Figure 4A for correlation coefficients between adjacent total body subregions to be slightly larger than those between more distant sites. Subregions were therefore ranked in the order arms, ribs, T spine, L spine, pelvis, and legs; and values of the mean correlation coefficient for adjacent sites (i.e., pelvis and legs), sites twice removed (i.e., lumbar spine and legs), etc. were plotted (Fig. 5B). Values decreased from  $r = 0.63$  for adjacent sites to  $r = 0.55$  for the most distant pairs.

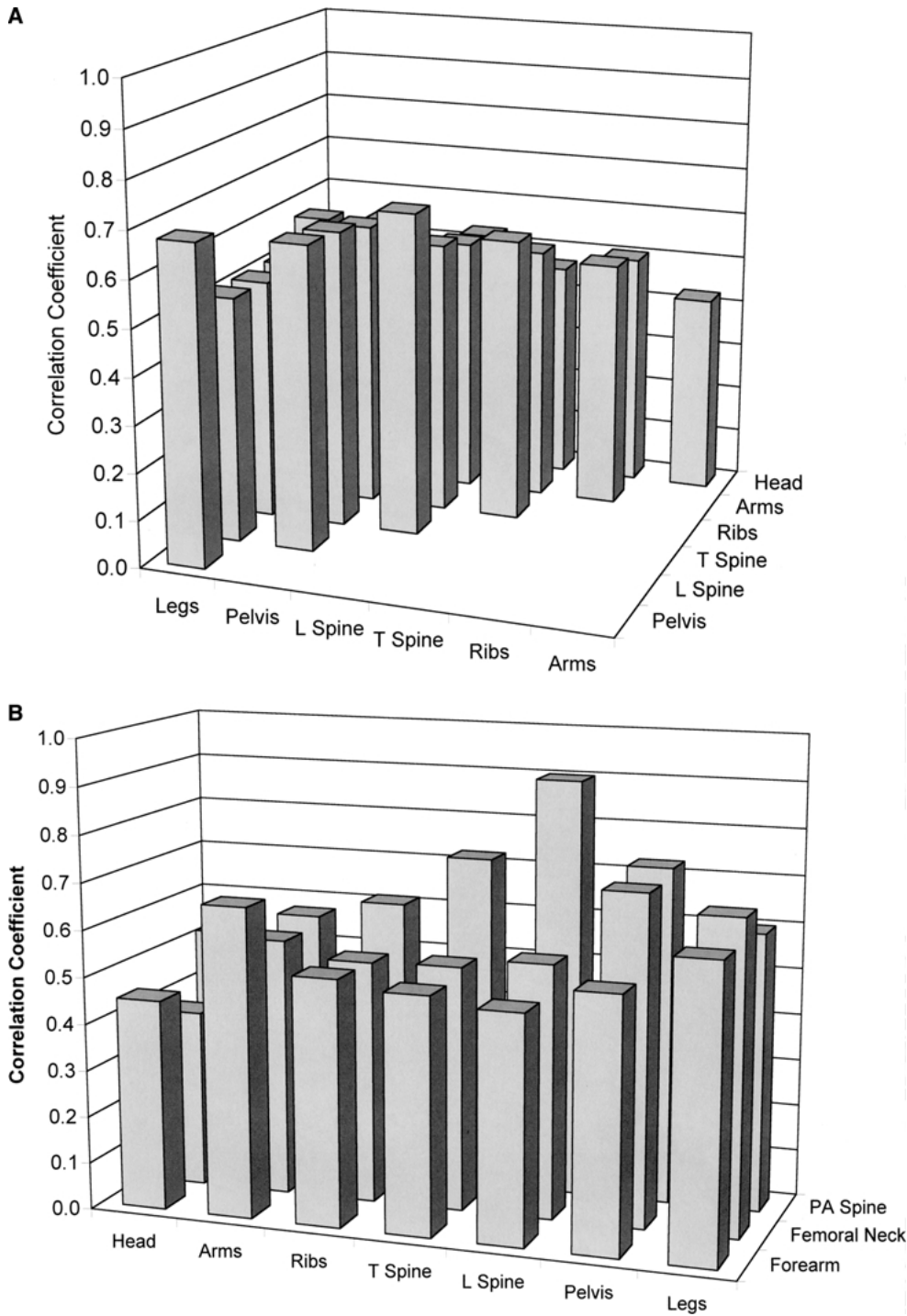
**Discussion**

The ability to assess a patient’s risk of fracture is fundamental to the clinical role of bone densitometry. The most reliable way of establishing the effectiveness of any type of BMD measurement is through prospective studies of incident fractures. Epidemiological studies verify fracture discrimination by showing that the RR value is statistically significantly greater than 1.0. The statistical quality of the available data was improved by the publication of the SOF 10-year follow-up study [10]. With several hundred fractures recorded at each of the principal sites (hip, wrist, and spine) and over 2,000 fractures at all sites, the errors in the RR values are smaller than those of previous studies. The new data confirm with a high degree of statistical significance that hip BMD is the most effective predictor of hip fracture risk. They also allow for the first time evaluation of the quantitative contribution of the correlation between



**Fig. 3.** Comparison of correlation coefficient values between the same four regions of interest in the skeleton (PA spine, femoral neck, total hip, and distal forearm) for the SOF population [20] and the twin study population [25]. For the twin study population, the Z-score correlation coefficient was plotted to reduce the effect of the wide age range of the subjects. The dashed line shows the line of identity.

BMD measurements to fracture prediction by measurement sites distant to the fracture site. After allowing for the effect of random errors arising from variations in soft tissue composition on the estimates of RR and measurements of the correlation coefficients between different BMD sites, the SOF 10-year data for hip, wrist,



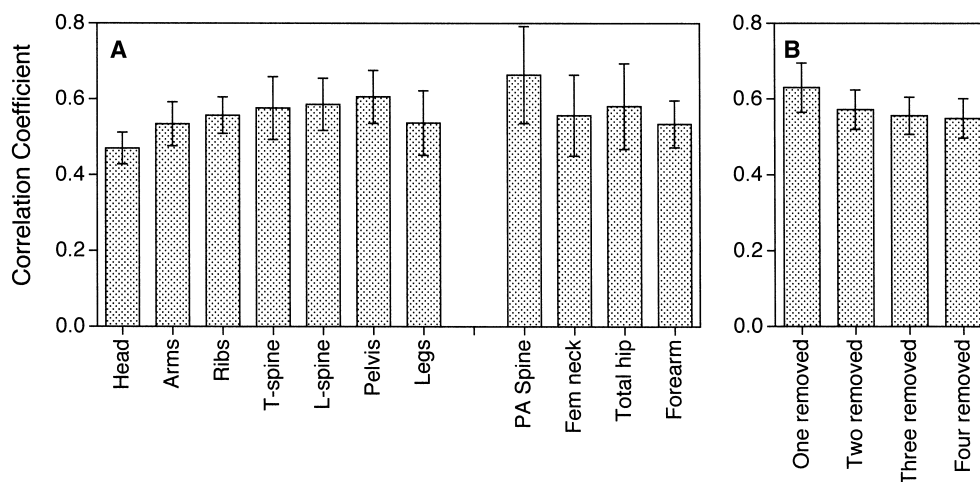
**Fig. 4.** (A) Values of the Z-score correlation coefficient between BMD measurements in the seven subregions of the total body DXA scans (head, arms, ribs, thoracic spine, lumbar spine, pelvis, and legs) for the twin study population. (B) Values of the Z-score correlation coefficient between BMD measurements in the PA spine, femoral neck, and distal forearm and the same seven subregions of the total body scans for the twin study population.

and spine fractures were shown to be consistent with the predictions of the correlation hypothesis [19].

The SOF 10-year data confirm with improved statistical accuracy earlier reports that RR values for the prediction of any fracture are approximately constant for different BMD measurement sites (Fig. 1C). The average RR value of 1.4 is rather smaller than the figures from the Marshall et al. meta-analysis [9] and the NORA study [18] and reflects the longer follow-up

period. Although there is evidence in the data that there may be small but real differences between some BMD sites, it is the relative constancy of the RR values that is the most striking feature of Figure 1C.

The present study has shown that the constant RR values at different BMD sites for predicting any fracture can be explained if (1) the correlation hypothesis is valid and (2) the correlation coefficients between distant BMD sites are equal. Although both these propositions are



**Fig. 5.** (A) Mean values of the Z-score correlation coefficient between each total body subregion and the remaining six subregions. Mean values of the correlation coefficient between the PA spine, femoral neck, total hip, and distal forearm and the seven total body subregions are also shown. (B) Mean

values of the Z-score correlation coefficient between adjacent (i.e., once removed) subregions on the total body DXA scans, twice removed, etc. For this analysis, the total body subregions were placed in the order arms, ribs, thoracic spine, lumbar spine, pelvis, and legs and the head was excluded.

approximations, they are useful because they provide insight into some of the limitations that constrain the clinical effectiveness of bone densitometry.

The evidence for the correlation hypothesis has been discussed elsewhere [19] and will not be repeated here. Instead, we concentrate on the values of the correlation coefficient between different BMD measurement sites. Values for distant sites in the skeleton, such as the hip, spine, forearm, and heel, lie in the range  $r = 0.55$ – $0.65$  (Fig. 2), while correlations between adjacent sites are rather higher. The SOF data plotted in Figure 2 were included here because of their small statistical errors due to the large number of subjects [20]. For the data plotted in Figure 2, the 95% CIs in the correlation coefficients are approximately  $\pm 0.015$ . With such small errors, there are undoubtedly real differences in the correlation coefficients between the different sites in Figure 2. Nevertheless, as with the RR data in Figure 1C, it is the relative constancy of the correlation coefficient that is the most striking feature of Figure 2.

To examine the values of the correlation coefficient over a wider range of skeletal sites, total body BMD data from a study of white female twins [25] were analysed. There were not enough older twins in the same age range as the SOF subjects to allow an analysis with sufficiently small statistical errors. However, by analysing the twin study data over the whole age range to evaluate Z-score rather than T-score correlation coefficients, the effects of the wide age range were minimised (Fig. 3). We were thus able to show that the correlation coefficients between the different total body subregions (Fig. 4A) and between the PA spine, hip, and forearm BMD and the total body subregions (Fig. 4B) lay in the same range as the SOF data in Figure 2. There are some

provisos to this general conclusion. Correlations including head BMD were lower than the general trend (Fig. 5A) and, as with the SOF data, there was a trend for adjacent sites to correlate slightly better than more distant sites (Fig. 5B).

The present study has important limitations. The three propositions we have attempted to relate are only approximations: (1) the constancy of the RR values at different BMD sites for predicting fracture risk at any skeletal site; (2) the explanation of RR values for predicting fracture risk at distant sites by BMD correlation; and (3) the constancy of the correlation coefficient between distant BMD sites. As stated previously [19], it is important that the correlation hypothesis is examined using raw fracture data from the SOF. In addition, we were not able to examine the probable variation of both the RR values and the BMD correlation coefficients with the age of the subjects. Other limitations include the assumptions that the distribution of Z-score values follows a gaussian function and that the variation of fracture risk with Z-score is exponential, which are also approximations. Finally, we have assumed that total body DXA scans provide subregion BMD results that are accurate enough to obtain reliable values of the correlation coefficients.

If the arguments provided in the present study are valid, they point to the importance of the correlation coefficient between different BMD measurement sites as a key index that is predictive of the ability of different techniques to identify patients at risk of fracture. For example, it is the narrow range of values of the correlation coefficient between BMD measurements at different sites that leads to the prediction that such measurements should have comparable power to predict

the risk of any fracture. Conversely, types of measurement that correlate significantly better or worse with hip and spine BMD than those discussed here would be expected to perform better or worse in the prediction of fractures at distant sites. Consistent with this expectation, some types of ultrasound measurement that correlate poorly with central BMD have also been shown to be less effective at predicting fracture risk [26, 27]. For the same reason, we expect that the ability of bone densitometry techniques to predict fracture risk could be improved by the development of new methods that correlate better with BMD in the central skeleton. This might be achieved, for example, by novel methods that attempt to reduce the accuracy errors associated with the variations in soft tissue composition [28].

If, on the contrary, the correlation hypothesis is not valid, the clinical effectiveness of bone densitometry could be improved through further studies to more reliably identify those techniques that are capable of providing independent information about fracture risk and how best this information might be integrated with existing methods. The best-known counterexample to the correlation hypothesis is the use of heel ultrasound measurements to predict hip fracture risk. Two reports have presented evidence that, after correction for femoral neck BMD, heel QUS is an independent predictor of hip fracture risk [29, 30]. However, more recent QUS data with larger numbers of fracture cases show smaller RR values that appear to be consistent with the correlation hypothesis [31].

In summary, we have shown that the well-known observation that RR values for predicting the risk of any fracture are comparable for different BMD measurement sites has a simple explanation. If valid, this explanation provides further evidence to support the correlation hypothesis. If the correlation hypothesis is true, it points to the correlation coefficient between different bone densitometry techniques as a useful guide to their likely effectiveness at predicting fracture risk. Bone densitometry techniques can be improved either by developing novel methods that correlate better with existing measurements or by studies to more reliably identify methods that provide independent information about fracture risk.

## References

1. Cooper C, Campion G, Melton LJ (1992) Hip fractures in the elderly: a world-wide projection. *Osteoporosis Int* 2:285–289
2. Ray NF, Chan JK, Thamer M, Melton LJ (1997) Medical expenditures for the treatment of osteoporotic fractures in the United States in 1995: report from the National Osteoporosis Foundation. *J Bone Miner Res* 12:24–35
3. Black DM, Cummings SR, Karpf DB, Cauley JA, Thompson DE, Nevitt MC, Bauer DC, Genant HK, Haskell WL, Marcus R, Ott SM, Torner JC, Quandt SA, Reiss TF, Ensrud KE (1996) Randomised trial of the effect

of alendronate on risk of fracture in women with existing vertebral fractures. *Lancet* 348:1535–1541

4. McClung MR, Geusens P, Miller PD, Zippel H, Bensen WG, Roux C, Adami S, Fogelman I, Diamond T, Eastell R, Meunier PJ, Reginster J-Y (2001) Effect of risedronate treatment on hip fracture risk in elderly women. *N Engl J Med* 344:333–340
5. Neer RM, Arnaud CD, Zanchetta JR, Prince R, Gaich GA, Reginster J-Y, Hodsman AB, Eriksen EF, Ish-Shalom S, Genant HK, Wang O, Mitlak BH (2001) Effect of recombinant human parathyroid hormone (1–34) fragment on spine and non-spine fractures and bone mineral density in postmenopausal osteoporosis. *N Engl J Med* 344:1434–1441
6. Kanis JA, Delmas P, Burckhardt P, Cooper C, Torgerson D, on behalf of the European Foundation for Osteoporosis, Bone Disease (1997) Guidelines for diagnosis and treatment of osteoporosis. *Osteoporosis Int* 7:390–406
7. Royal College of Physicians (1999) Osteoporosis: clinical guidelines for prevention and treatment. RCP, London
8. National Osteoporosis Society (2002) Position statement on the reporting of dual X-ray absorptiometry (DXA) bone mineral density scans. Bath, UK
9. Marshall D, Johnell O, Wedel H (1996) Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *Br Med J* 312: 1254–1259
10. Stone KL, Seeley DG, Lui L-Y, Cauley JA, Ensrud K, Browner WS, Nevitt MC, Cummings SR (2003) BMD at multiple sites and risk of fracture of multiple types: long-term results from the Study of Osteoporotic Fractures. *J Bone Miner Res* 18:1947–1954
11. Eastell R (1998) Treatment of postmenopausal osteoporosis. *N Engl J Med* 338:736–746
12. World Health Organization (1994) Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. WHO Technical Report Series 843. WHO, Geneva
13. Genant HK, Engelke K, Fuerst T, Gluer C-C, Grampp S, Harris ST, Jergas M, Lang T, Lu Y, Majumdar S, Mathur A, Takada M (1996) Noninvasive assessment of bone mineral and structure: state of the art. *J Bone Miner Res* 11:707–730
14. Gluer C-C, Jergas M, Hans D (1997) Peripheral measurement techniques for the assessment of osteoporosis. *Semin Nucl Med* 27:229–247
15. Jones T, Davie M (1998) Bone mineral density at distal forearm can identify patients with osteoporosis at spine or femoral neck. *Br J Rheumatol* 37:539–543
16. Fordham JN, Chinn DJ, Kumar N (2000) Identification of women with reduced bone density at the lumbar spine and femoral neck using BMD at the os calcis. *Osteoporosis Int* 11:797–802
17. Blake GM, Fogelman I (2001) Peripheral or central densitometry: does it matter which technique we use? *J Clin Densitom* 4:83–96
18. Miller PD, Siris ES, Barrett-Connor E, Faulkner KG, Wehren LE, Abbott TA, Chen Y-T, Berger ML, Santora AC, Sherwood LM (2002) Prediction of fracture risk in postmenopausal white women with peripheral bone densitometry: evidence from the National Osteoporosis Risk Assessment. *J Bone Miner Res* 17:2222–2230
19. Blake GM, Fogelman I (2005) Fracture prediction by bone density measurements at sites other than the fracture site: the contribution of BMD correlation. *Calcif Tissue Int* 76:249–255
20. Lu Y, Genant HK, Shepherd J, Zhao S, Mathur A, Fuerst TP, Cummings SR (2001) Classification of osteoporosis based on bone mineral densities. *J Bone Miner Res* 16:901–910
21. Patel R, Blake GM, Fogelman I (2004) An evaluation of the National Osteoporosis Society position statement on the use of peripheral X-ray absorptiometry. *Osteoporosis Int* 15:497–504

22. Svendsen OL, Hassager C, Skodt V, Christiansen C (1995) Impact of soft tissue on in-vivo accuracy of bone mineral measurements in the spine, hip and forearm: a human cadaver study. *J Bone Miner Res* 10:868–873
23. Seeley DG, Browner WS, Nevitt MC, Genant HK, Scott JC, Cummings SR (1991) Which fractures are associated with low appendicular bone mass in elderly women? *Ann Intern Med* 115:837–842
24. Grampp S, Genant HK, Mathur A, Lang P, Jergas M, Takada M, Gluer C-C, Lu Y, Chavez M (1997) Comparisons of non-invasive bone mineral measurements in assessing age-related bone loss, fracture discrimination and diagnostic classification. *J Bone Miner Res* 12: 697–711
25. Knapp KM, Andrew T, MacGregor AJ, Blake GM, Fogelman I, Spector TD (2003) An investigation of the unique and shared gene effects on speed of sound and bone density using axial transmission quantitative ultrasound and dual X-ray absorptiometry in twins. *J Bone Miner Res* 18:1525–1530
26. Knapp KM, Blake GM, Spector TD, Fogelman I (2001) Multisite quantitative ultrasound: precision, age and menopause related changes, fracture discrimination and T-score equivalence with dual X-ray absorptiometry. *Osteoporosis Int* 12:456–464
27. Clowes JA, Peel NFA, Eastell R (2002) Fractures of the distal forearm, humerus, hip and vertebrae: association with peripheral and axial densitometry measurements. *J Bone Miner Res* 17:S186
28. Kullenberg R, Falch JA (2003) Prevalence of osteoporosis using bone mineral measurements at the calcaneus by dual X-ray and laser (DXL). *Osteoporosis Int* 14:823–827
29. Bauer DC, Gluer C-C, Cauley JA, Vogt TM, Ensrud KE, Genant HK, Black DM (1997) Broadband ultrasonic attenuation predicts fractures strongly and independently of densitometry in older women. *Arch Intern Med* 157:629–634
30. Hans D, Dargent-Molina P, Schott AM, Sebert JL, Cormier C, Kotzki PO, Delmas PD, Pouilles JM, Breart G, Meunier PJ (1996) Ultrasonographic heel measurements to predict hip fracture in elderly women: the EPIDOS prospective study. *Lancet* 348:511–514
31. Blake GM, Fogelman I (2004) Bone densitometry and fracture risk prediction. *Eur J Nucl Med* 31:785–786