

# The effect of weight and weight change on the long-term precision of spine and hip DXA measurements

R. Rajamanohara · J. Robinson · J. Rymer · R. Patel ·  
I. Fogelman · G. M. Blake

Received: 25 March 2010 / Accepted: 7 June 2010 / Published online: 11 August 2010  
© International Osteoporosis Foundation and National Osteoporosis Foundation 2010

## Abstract

**Summary** We examined the effect of weight and weight change on the long-term precision of spine and hip bone mineral density (BMD) in a group of 64 postmenopausal women studied over a 10-year period. Long-term precision errors were 50% larger than short-term errors. Over the range 50–90-kg weight was associated with a statistically significantly larger precision error when precision was expressed in BMD units, but not when expressed as the coefficient of variation (CV). Weight changes up to 5 kg had little effect on precision.

**Introduction** Reliable knowledge of the precision of bone mineral density (BMD) measurements is important for the interpretation of follow-up dual-energy X-ray absorptiometry (DXA) scans. In this study, we examined the effect of body weight and change in weight on the long-term precision of spine and hip BMD.

**Methods** The study population was a group of 64 postmenopausal women enrolled in a 16-year trial of tibolone. We analyzed the spine, femoral neck, and total hip BMD data acquired over a 10-year period on a Hologic QDR4500A densitometer using linear regression to examine the trend of BMD with time for each subject. Precision was expressed in BMD units ( $\text{g cm}^{-2}$ ) (standard error of the estimate, SEE) and also as the coefficient of variation (CV). **Results** The long-term precision errors were in BMD (CV) units: 0.018  $\text{g cm}^{-2}$  (1.9%) for spine, 0.017  $\text{g cm}^{-2}$  (2.3%) for femoral neck, and 0.016  $\text{g cm}^{-2}$  (1.7%) for total hip BMD. An inverse relationship between CV and BMD was found for the spine ( $P=0.003$ ) and total hip ( $P=0.043$ ) sites, but none between SEE and BMD. For spine BMD, there were statistically significant correlations between SEE and weight ( $P=0.025$ ) and body thickness ( $P=0.027$ ). For femoral neck BMD, there were correlations between SEE and weight ( $P=0.030$ ), body mass index (BMI) ( $P=0.023$ ) and thickness ( $P=0.021$ ), but no correlations for total hip BMD or when precision was expressed as the CV. When study subjects were grouped in quartiles according to weight, the spine BMD SEE increased from 0.014  $\text{g cm}^{-2}$  for women in the lowest quartile (46–62 kg) to 0.018  $\text{g cm}^{-2}$  for women in the highest quartile (80–105 kg) ( $P=0.008$ ). There was a trend for SEE to be greater in individuals with larger weight changes, although these tended to be the heavier subjects.

**Conclusions** From the study, we were able to come up with the following conclusions: (1) long-term precision errors were 50% larger than short-term errors, (2) over the range 50 to 90 kg (BMI: 20–35  $\text{kg m}^{-2}$ ), body weight had a small but statistically significant effect on precision expressed in BMD units, but not when expressed as the CV, and (3) weight changes up to 5 kg had little effect on precision. More studies of individuals >100 kg are required to fully

---

R. Rajamanohara · I. Fogelman · G. M. Blake  
Osteoporosis Research Unit, Division of Imaging Sciences,  
King's College London Guy's Campus,  
London, UK

J. Robinson · J. Rymer  
Menopause Research Unit, Division of Reproduction &  
Endocrinology, King's College London Guy's Campus,  
London, UK

R. Patel  
Academic Bone Densitometry Unit, Imperial College London,  
London, UK

G. M. Blake (✉)  
Osteoporosis Research Unit,  
1st Floor, Tower Wing Guy's Hospital,  
London SE1 9RT, UK  
e-mail: glen.blake@kcl.ac.uk

investigate the dependence of DXA scan precision on weight.

**Keywords** Body weight · Bone mineral density · Dual-energy X-ray absorptiometry · Long-term precision · Weight change

## Introduction

Dual-energy X-ray absorptiometry (DXA) is widely used to identify postmenopausal women at increased risk of fracture who can be advised to take preventive treatment for osteoporosis [1, 2]. Because DXA offers a safe, sensitive and precise method of measuring changes in bone mineral density (BMD), many patients treated for osteoporosis have follow-up scans to assess their response to therapy [1, 3–5].

Knowledge of the precision error of DXA measurements is required to decide whether a follow-up scan provides evidence of a clinically relevant change in BMD [6]. It is widely accepted that, given a single baseline measurement and a single follow-up measurement, the smallest BMD change that can be regarded as statistically significant at the 95% confidence level is 2.8 times the precision error [7]. Reliable knowledge of the precision error is therefore important for the clinical interpretation of follow-up DXA examinations in individual patients.

Although precision errors are important for interpreting DXA scan findings, it is a topic that continues to cause controversy [8–12]. One reason is that measuring precision is an inherently noisy process that can involve significant statistical errors, even for studies with quite large groups of subjects. For example, a widely adopted protocol for performing a precision study is to acquire duplicate measurements in a group of 30 subjects [3, 7]. However, based on this number, the 95% confidence interval in the final precision figure is still a relative error of  $\pm 30\%$ , and studies of over 200 subjects are required to reduce this error to less than  $\pm 10\%$  [8].

Because little is known about the clinical factors that can modify the precision error in an individual patient, the general practice is to adopt a single figure for precision and apply it to all subjects [3]. Body weight is one factor that might be expected to influence precision [12, 13]. Patients who are heavier or have a larger body mass index (BMI) will, in general, have a greater thickness of soft tissue to attenuate X-rays and, as weight increases, at some point the signal reaching the detectors will become sufficiently deficient in photons to result in a noisier and less precise measurement. There is a second reason why obesity might influence precision. DXA scans are affected by accuracy errors that arise from the heterogeneous distribution of adipose tissue external to bone and variations in marrow

composition within bone [14]. When patients with a normal BMI lie on the scanning table for a follow-up scan, the disposition of soft tissues will be similar to the baseline scan, and the accuracy error arising from the heterogeneous distribution of adipose tissue will be almost the same for every scan. This explains why precision errors in DXA are smaller than the accuracy errors [14]. However, in obese patients, the soft tissues may be distributed differently on follow-up examinations, causing scan-to-scan changes in the accuracy error that will have an adverse effect on precision [15]. Weight change is a second factor that might influence precision. In patients showing a sufficiently large weight change, the resulting change in the amount and distribution of adipose tissue may cause a change in the accuracy error that affects precision.

Although body weight and weight change are factors that theoretically might affect the precision of spine and hip DXA measurements, relatively little is known quantitatively about what weight or weight change are required before clinically significant changes in the precision error occur [12, 13]. In the present study, we examined annual BMD measurements made during the OD14 study [16–18], a clinical trial of tibolone in postmenopausal women. Tibolone is a synthetic compound with weak hormonal properties that exerts beneficial effects on bone, as well as on climacteric symptoms, but does not stimulate the endometrium [19]. A retrospective analysis was performed on spine and hip BMD data from the OD14 trial with the aim of examining the effect of weight and weight change on the long-term precision error.

## Methods

### Subjects

The subjects were healthy postmenopausal women enrolled in an open-label study of tibolone. All were 6–36 months since their last menstrual period. A detailed description of the subjects and the entry criteria has been published previously [16–18]. Subjects were given the choice of receiving either oral tibolone 2.5 mg/day (Livial<sup>®</sup>; NV Organon, Oss, The Netherlands) or no treatment (controls). As far as possible, the women in the two treatment arms were matched for age and time elapsed since menopause. One hundred and ten subjects (59 tibolone, 51 controls) were enrolled over an 18-month period starting in November 1988, of whom 41 (21 tibolone, 20 controls) completed the study during 2005–2006. All the women gave written consents to participate in the study. The protocol was approved by Guy's Hospital Ethics Committee, and the study was carried out according to Good Clinical Practice and in accordance with the declaration of Helsinki.

## DXA measurements

Spine and hip DXA scans were performed at baseline, 6 months, and then annually for up to 16 years. During the course of the study, four different Hologic DXA scanners were used (Hologic, Bedford, MA) [20]. Daily quality control was performed using a Hologic spine phantom. Each new densitometer was cross-calibrated in vitro with its predecessor using the spine phantom, and the results checked with an in vivo cross-calibration study [21–23]. To evaluate the BMD changes over the full 16-year period of the study, the BMD data were corrected using the relevant in vivo cross-calibration factors. For the study of the effect of weight and weight change on long-term precision, the analysis was restricted to the data obtained on a single scanner, a QDR-4500A, between 1995 and 2006. All scans were performed in the array mode. Long-term precision of the daily scans of the spine phantom was 0.49%.

The QDR-4500A BMD data used for the precision study were taken from the original scan analyses performed by the DXA technologist during the period 1995 to 2006. Scan files were recovered from the archive, and patients' height, weight, lumbar spine (L1–L4), femoral neck, and total hip BMD recorded for each visit, along with the DXA measurements of body thickness for the spine and hip regions of interest (ROI). Figures for body mass index (BMI) were determined from height and weight ( $BMI = Wt(kg)/Ht(m)^2$ ). Scan images were evaluated visually for factors that might adversely affect precision, including the presence of scoliosis and degenerative changes on spine scans, and whether manual intervention by the technologist to delete bone from the pelvis (ischium) in order to make space for the femoral neck ROI was necessary on the hip scans [24].

## Phantom study of DXA measured body thickness

Scan printouts from Hologic DXA systems include an estimate of the body thickness measured through the soft tissue reference area that is derived from the X-ray transmission factor and calibrated using the daily scan of the spine phantom. The thickness measurements (labeled TH on scan printouts) are given in units of inches, so all figures were multiplied by 2.54 to convert them into centimeters. To evaluate the reliability of this measurement, phantom scans were made through a water bath with increasing depths of water from 8–30 cm and the DXA measurements of thickness compared with the known depths of water.

## Evaluation of long-term precision

In precision studies based on repeated BMD measurements obtained over a short-time period (<1 or 2 months), the

short-term precision error can be expressed either as the standard deviation (SD) of the measurements (precision expressed in BMD units of  $g\ cm^{-2}$ ), or as the coefficient of variation (CV) obtained by writing the SD as a percentage of the mean BMD (precision expressed in %) [7–9]. If for one particular individual a series of  $N$  BMD measurements are available for the evaluation of their precision error, the statistical error in the precision is determined by the number of degrees of freedom ( $df$ ) given by  $df = N - 1$  [7, 8].

For precision studies made over longer time intervals, true BMD changes may occur, and a different mathematical approach is required [8, 20]. If it is assumed that the changes approximate to a linear change in BMD with time, then any random variations about the best fitting straight line can be quantified using linear regression analysis. When the BMD results for any one individual are plotted against time, then the variability about the regression line is given by the standard error of the estimate (SEE). The long-term precision error can be expressed either as the SEE (in BMD units), or as the CV obtained by writing the SEE as a percentage of the mean BMD. If linear regression analysis is performed, then two free parameters are fitted to the data (the intercept and slope), and the number degrees of freedom is reduced to  $df = N - 2$  [8].

The individual long-term precision error was evaluated in all the women in the study who had at least three visits for BMD measurements performed on the QDR4500A instrument, so there were sufficient data points for linear regression analysis. To preserve the maximum possible overall number of degrees of freedom (and hence statistical power), the statistical significance ( $P$  value) of the slope of the linear regression line was used to determine whether the SEE or SD was used to estimate precision. Slopes with a  $P$  value < 0.40 were judged to show a sufficient trend with time that a linear regression fit was required, and in these subjects, the SEE was used for the precision figure. Slopes with a  $P$  value > 0.40 were interpreted as consistent with no trend of BMD with time, and the SD was used for the precision figure. In this way, the maximum statistical power was kept for the analysis of the effect of weight and weight change on precision.

## Data analysis

Overall values for long-term precision at the spine, femoral neck, and total hip were obtained by taking the root mean square (RMS) precision figure for the individual patients weighted by their individual  $df$  figure [8]. Precision results were expressed in both BMD units (based on the SEE or SD figures, as appropriate in the individual; henceforth, we shall refer to this as the SEE precision figure), and in percentage units calculated from the RMS CV. For spine BMD, the precision was also calculated separately for the

subjects with and without scoliosis and/or degenerative disease. Similarly, for femoral neck and total hip BMD, precision was calculated separately for subjects in whom the technologist did or did not have to paint out bone in the pelvis to fit the femoral neck ROI. F-tests [25] were performed to establish whether scoliosis and degenerative disease, or painting out bone in the pelvis, were factors that predicted poorer precision. The effect of the treatment arm on precision (tibolone vs. controls) was similarly assessed. Statistical significance of the F-test was evaluated using a *P* value of less than 0.05.

In general, it is not clear whether precision is better expressed in BMD units as the SEE, or by the CV [9]. To examine this question, scatter plots were drawn of the spine, femoral neck, and total hip precision figures for individual subjects, expressed in both BMD units and as the CV, against the mean BMD. Trends for precision to vary with BMD were evaluated using the Spearman rank correlation coefficient [25].

The question of whether larger body weight, BMI, or DXA measured body thickness were associated with poorer precision expressed either as the SEE or the CV was examined by plotting the spine, femoral neck, and total hip precision data for individual subjects against their weight, BMI, and thickness, and evaluating the statistical significance of the trends using the Spearman rank correlation coefficient. Statistical significance was assessed using a one-tailed test with a *P* value of less than 0.05 to evaluate the hypothesis that larger body weight, BMI, or body thickness had an adverse effect on precision.

For the spine BMD data, the variation of precision with body weight was also examined by dividing women into quartiles according to their weight, and calculating the RMS SEE for the subjects in each quartile. A similar analysis was performed by dividing subjects into quartiles according to the SD of their change in weight during the study. Finally, to separate the effects of weight and weight change on precision, the women were divided into four

groups according to whether their body weight and body weight SD were above or below the median figures for the whole study.

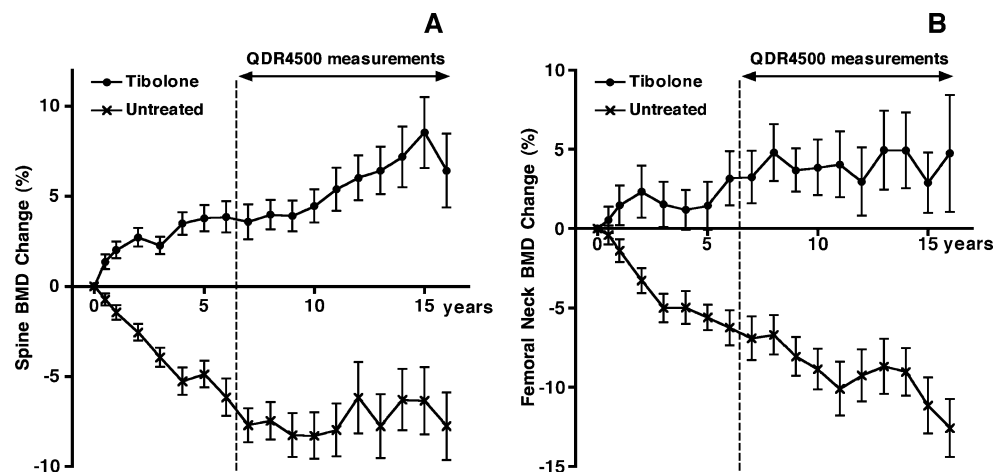
## Results

Figure 1 shows the mean percentage changes in lumbar spine and femoral neck BMD over the entire 16-year period of the study. The DXA data used for the assessment of long-term precision were all obtained on a single densitometer over the final 10 years of the study (years 7 to 16) (Fig. 1). A total of 64 subjects (33 tibolone, 31 controls) had at least three DXA visits during this 10-year period, and between them had a total of 513 spine BMD examinations. One subject had a hip replacement, leaving 63 subjects with a total of 504 hip BMD examinations. The BMD data were taken from the original scan analysis performed by the DXA technologist during the period 1995 to 2006. Demographic data for the 64 subjects, including age, body weight, BMI, DXA measured body thicknesses at the spine and hip, and spine and hip T-scores are listed in Table 1.

Body thickness figures obtained from the DXA scan analysis ranged from 14.1–24.9 cm at the spine and 12.2–19.9 cm at the hip (Table 1). Results of the phantom study of DXA-measured thickness plotted against water depth are shown in Fig. 2a. The DXA-derived thickness varied from 7.8 cm at 8.5-cm water depth to 26.6 cm at 29.4-cm water depth. Figure 2b shows the scatter plots of DXA-measured thickness at the spine and hip against body weight. Pearson correlation coefficients between body weight, BMI, and body thickness were all around  $r=0.9$  (Table 2).

The RMS SEE (RMS CV) for the 64 subjects with spine BMD data was  $0.018 \text{ g cm}^{-2}$  (1.9%). For the 63 subjects with hip data, the figures were  $0.017 \text{ g cm}^{-2}$  (2.3%) for femoral neck BMD and  $0.016 \text{ g cm}^{-2}$  (1.7%) for total hip BMD. In four subjects with scoliosis and/or degenerative changes, the spine BMD precision results were  $0.031 \text{ g cm}^{-2}$

**Fig. 1** Percentage change (mean  $\pm$  SEM) from baseline for **a** lumbar spine BMD (L1–L4), **b** femoral neck BMD in postmenopausal treated with tibolone 2.5 mg daily for 16 years compared with an untreated control group. The BMD measurements from years 7 to 16 were all made on the same Hologic QDR4500A densitometer



**Table 1** Demographic data for study subjects ( $N=64$ )

	Mean	Range
Age (y) <sup>a</sup>	58.8	50.9–65.4
Body Weight (kg) <sup>b</sup>	70.5	46.4–105.0
BMI (kg m <sup>-2</sup> ) <sup>b</sup>	26.9	19.5–40.9
Spine ROI body thickness (cm) <sup>b</sup>	19.2	14.1–24.9
Hip ROI body thickness (cm) <sup>b</sup>	15.7	12.2–19.9
Spine T-Score (L1–L4) <sup>b</sup>	-0.39	-3.48–3.87
Femoral neck T-score <sup>b</sup>	-0.70	-2.44–1.87
Total hip T-score <sup>b</sup>	-0.23	-2.29–2.62

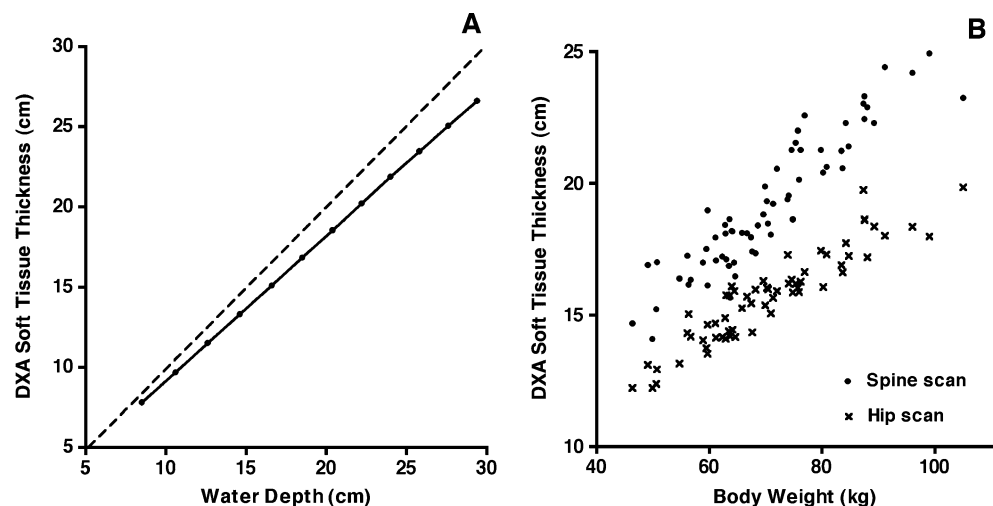
BMI body mass index

<sup>a</sup> Age at first DXA examination on QDR4500 system in 1995

<sup>b</sup> Mean of all DXA examinations on QDR4500 system (1995–2006)

(2.9%) compared with 0.017 g cm<sup>-2</sup> (1.8%) in the other 60 subjects ( $F=3.34$ ,  $P<0.0001$ ). In 21 subjects with hip scans that required the DXA technologist to paint out bone in the pelvis during scan analysis, the femoral neck BMD precision results were 0.018 g cm<sup>-2</sup> (2.5%) compared with 0.017 g cm<sup>-2</sup> (2.2%) in the other 42 subjects ( $F=1.19$ ,  $P=0.12$ ). For total hip BMD, the corresponding figures were 0.015 g cm<sup>-2</sup> (1.8%) compared with 0.016 g cm<sup>-2</sup> (1.7%) ( $F=1.03$ ,  $P=0.42$ ). For hip BMD data, the fact of having to paint out bone in the pelvis did not significantly affect the precision of the BMD results, and the subsequent analysis was performed using the data for all 63 subjects. However, scoliosis and degenerative changes in the spine were a predictor of poor spine BMD precision and therefore, two analyses were performed, the first for all 64 subjects, and the second for the 60 subjects without scoliosis or degenerative disease. When the effect of treatment arm (tibolone vs. controls) on precision was examined, there was no significant difference for the spine ( $F=1.08$ ,  $P=0.30$ ) or total hip BMD SEE ( $F=1.05$ ,  $P=0.38$ ), but a possible small difference was found for femoral neck BMD ( $F=1.35$ ;  $P=0.020$ ).

**Fig. 2 a** Results of a phantom experiment in which DXA scan measurements of body thickness were compared with the measured depth of water in a water bath. The *broken line* is the line of identity. **b** Scatter plot of DXA measurements of body thicknesses for the spine and hip regions of interest plotted against body weight for the 64 subjects in the present study



When subjects' individual precision was plotted against BMD, the results for spine BMD showed no trend of SEE with BMD (Fig. 3a) but there was an inverse relationship between CV and BMD (Fig. 3b). The Spearman correlation coefficient between individual subjects' spine CV figure and their spine BMD was statistically significant for both the whole study group ( $N=64$ ;  $R=-0.296$ ;  $P=0.009$ ) and the selected group excluding the four subjects with scoliosis and degenerative disease ( $N=60$ ;  $R=-0.349$ ;  $P=0.003$ ) (Table 3). In contrast, the spine precision data expressed in BMD units showed no correlation with BMD (Table 3). The choice between the SEE and CV descriptions of precision was less clear cut for the hip sites (Table 3). For femoral neck BMD, neither choice could be excluded, while for total hip BMD, there was a marginally significant trend for the CV to vary inversely with BMD ( $R=-0.218$ ;  $P=0.043$ ) similar to the relationship found for the spine.

The scatter plots of precision against BMD (Fig. 3) show a wide variation in the figures for individual subjects. Given that the data points were derived from, on average, eight annual BMD measurements, an important question is to what extent this variation is explained by random statistical fluctuations. For each BMD site, the individual SEE figures were normalized to the RMS value and plotted as a histogram to show the distribution with ratios that typically varied between 0.2 and 1.8 (Fig. 4). The solid curves in Fig. 4 are the predicted histograms based on the number of subjects with different degrees of freedom ( $df$  ranges between 1 and 9) and the assumption that there is a single precision figure applicable to all subjects. The solid curves therefore show the expected spread of data purely due to statistical fluctuations assuming there were no real differences in precision between subjects.

When the Spearman correlation coefficients between body weight, BMI, and DXA-derived body thickness with spine precision expressed in BMD units were examined in the group of 60 subjects that excluded those with scoliosis,

**Table 2** Pearson correlation coefficients between subjects' body weight, BMI, and the soft tissue thicknesses at the lumbar spine and proximal femur measured on the DXA scans ( $N=64$ )

	BMI	Spine ROI body thickness	Hip ROI body thickness
Body weight	0.924	0.925	0.930
BMI	–	0.906	0.923
Spine ROI body thickness	–	–	0.888

BMI body mass index

the relationships were statistically significant for body weight ( $R=+0.254$ ;  $P=0.025$ ) and body thickness ( $R=+0.249$ ;  $P=0.027$ ), but not BMI ( $R=+0.189$ ;  $P=0.074$ ) (Table 4). For the femoral neck, the trend was statistically significant for body weight, BMI, and thickness. However, there was no relationship for the total hip (Table 4). When precision data were expressed as the CV, none of the trends were statistically significant (Table 4).

When subjects were divided into quartiles according to their body weight, there was a trend for the spine BMD RMS SEE precision error to increase with weight, with statistically significant differences between the lowest quartile (46–62 kg) and each of the three highest quartiles (63–69 kg; 70–79 kg; 80–105 kg), as assessed by the F-test (Fig. 5a). A second plot obtained by dividing subjects into quartiles according to the standard deviation of their change in weight during the study (SD quartiles: 0.8–1.5 kg; 1.7–2.3 kg; 2.4–3.0 kg; 3.2–10.8 kg) was similar in appearance to Fig. 5a, although the associated  $P$  values were all less significant.

Inferring whether weight, or weight change SD, or both affect precision is confused by the correlation between them (Spearman rank correlation coefficient  $R=0.43$ ;  $P=0.0004$ ). To resolve this issue, the subjects were divided into four groups according to whether they were above or below the median values for body weight and weight change SD

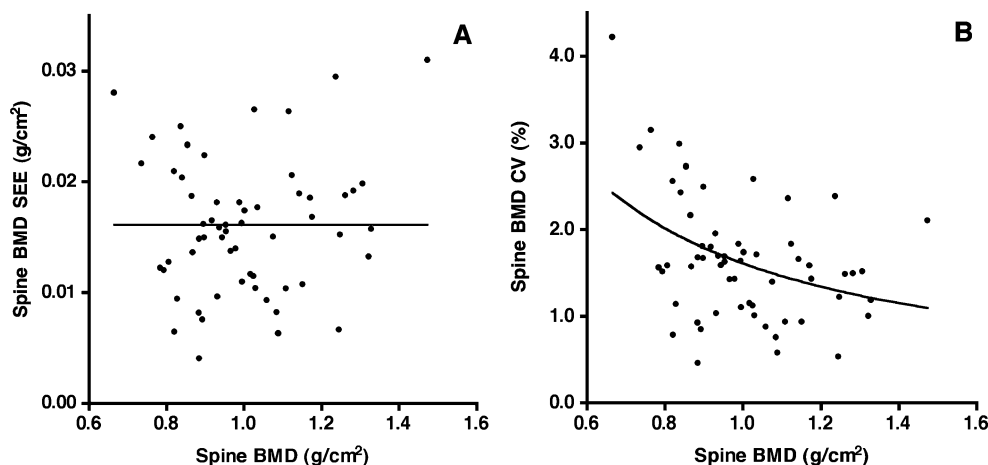
(Fig. 5b). The effect of weight change SD on the SEE precision error (Group 2 vs. Group 1; Group 4 vs. Group 3) was considerably smaller than the effect of body weight (Group 3 vs. Group 1; Group 4 vs. Group 2). However, none of the differences in RMS SEE in Fig. 5b were statistically significant apart from that between Group 4 and Group 1 ( $F=1.44$ ;  $P=0.018$ ). When subjects in Group 2 were pooled with those in Group 1, and those in Group 4 pooled with those in Group 3 to evaluate whether subjects with body weight above the median (70 kg) had the same SEE precision error as those below the median, a statistically significant difference was found ( $F=1.39$ ;  $P=0.013$ ). However, when subjects in Group 3 were pooled with those in Group 1, and those in Group 4 pooled with those in Group 2 to evaluate whether subjects with weight change SD above the median (2.4 kg) had different SEE precision error to those below the median, no significant difference was found ( $F=1.23$ ;  $P=0.082$ ). A similar analysis of the CV precision errors in Fig. 5b showed no difference due to either weight ( $P=0.31$ ) or weight change ( $P=0.50$ ).

## Discussion

The data used for this precision study were the spine and hip BMD measurements from years 7 to 16 of the OD14 study (Fig. 1). By year 7, the women had been postmenopausal for 7–10 years, so the phase of rapid bone loss during the early years of menopause was over. During the final 10 years of the OD14 study, mean BMD was either constant over time (spine BMD in untreated women; femoral neck BMD in tibolone group), or was increasing or decreasing linearly with time (spine BMD in tibolone group; femoral neck BMD in untreated women).

The overall long-term precision error evaluated over an average period of 8 years was in BMD (percentage) units  $0.018 \text{ g cm}^{-2}$  (1.9%) for spine BMD,  $0.017 \text{ g cm}^{-2}$  (2.3%)

**Fig. 3** Scatter plots of the individual spine BMD precision measurements for 60 women in the OD14 study plotted against spine BMD for **a** precision expressed in BMD units. The horizontal line is the mean SEE figure, **b** precision expressed as the coefficient of variation (CV). The curved line is the inverse relationship between CV and BMD corresponding to the horizontal line in (A)



**Table 3** Spearman correlation coefficients between subjects' spine, femoral neck, and total hip BMD and their individual precision measurements expressed as the standard error of the estimate (SEE) or coefficient of variation (CV)

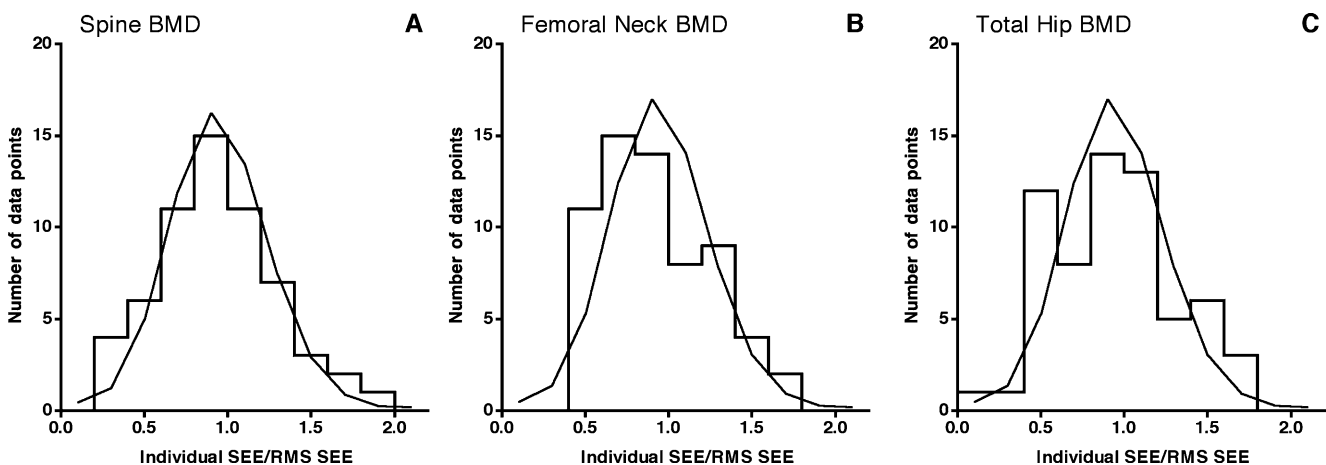
	<i>N</i>	SEE (g cm <sup>-2</sup> ) vs. BMD	CV (%) vs. BMD
Spine BMD (all)	64	<i>R</i> =0.038 <i>P</i> =0.383	<i>R</i> =-0.296 <i>P</i> =0.009
Spine BMD (excluding scoliotics)	60	<i>R</i> =0.007 <i>P</i> =0.476	<i>R</i> =-0.349 <i>P</i> =0.003
Femoral neck BMD (all)	63	<i>R</i> =0.184 <i>P</i> =0.075	<i>R</i> =-0.188 <i>P</i> =0.071
Total hip BMD (all)	63	<i>R</i> =0.138 <i>P</i> =0.140	<i>R</i> =-0.218 <i>P</i> =0.043

for femoral neck BMD, and 0.016 g cm<sup>-2</sup> (1.7%) for total hip BMD. For the spine and total hip sites, these figures are around 50% larger than the short-term precision errors of 1.2–1.3% reported for this generation of Hologic equipment [26]. The evaluation of long-term precision is more difficult than short-term precision because over a time period of several years, real biological changes are likely to occur. The assumption of a linear BMD change with time is an idealized case, and if non-linear changes occurred, then the long-term precision error will be overestimated. As far as could be ascertained from the BMD plots of individual subjects, the assumption of linear change was reasonable for the majority of women [20].

Although there are difficulties with the accurate evaluation of the long-term precision error, it should be recognized that the reliable evaluation of short-term precision is also problematic. As long ago as 1974, Holland and Whitehead commented that: “too frequently precision is determined under optimal conditions and produces an unrealistically optimistic view of what is likely to be achieved in routine practice” [27]. In recent years, a number of studies have suggested that the widely adopted practice of performing precision assessments with the same technologist performing duplicate examinations with simple

repositioning of the patient between scans underestimates the true precision error [11, 20, 28]. Our finding of a long-term precision figure 50% greater than short-term precision is consistent with the results of previous studies whenever the scans being compared were performed on different days [11, 12, 20].

Leslie drew attention to the fact that, despite advice from organizations such as ISCD [3] that precision data are best expressed in BMD units, in practice, many clinicians still prefer to use the coefficient of variation, and to interpret clinical BMD changes in terms of the percentage change from baseline [9]. We examined this question by plotting the precision data for individual OD14 subjects, expressed in either BMD units or as the CV, against BMD (Fig. 3). For spine BMD, there was a statistically significant inverse relationship between CV and BMD, but no relationship between SEE and BMD. For femoral neck BMD, it was not possible to distinguish between the two alternative descriptions of the precision error (Table 3), while for the total hip site, the findings were similar to the spine, but were less statistically significant. The present findings therefore tend to support the conclusions of Leslie that precision should be expressed in BMD units [9], although our ability to discriminate between the two methods of reporting the



**Fig. 4** Histograms showing the distribution of individual precision results for women in the OD14 study for **a** spine BMD, **b**: femoral neck BMD, **c** total hip BMD. Solid curves show the predicted distributions based on the number of degrees of freedom for each

subject and the assumption that each subject has the same true precision. Figures on the horizontal axis show the ratio of individual precision to the root mean square (RMS) figure for all the subjects

**Table 4** Spearman correlation coefficients between subjects' weight, body mass index (BMI), and tissue thickness (thickness) and their individual precision measurements expressed as the standard error of the estimate (SEE) or coefficient of variation (CV)

	N	SEE (g cm <sup>-2</sup> )		CV (%)			
		Weight (kg)		Thickness (cm)			
		Weight (kg)	BMI (kg m <sup>-2</sup> )	Thickness (cm)	Weight (kg)	BMI (kg m <sup>-2</sup> )	Thickness (cm)
Spine BMD (all)	64	R=0.200 P=0.056	R=0.128 P=0.156	R=0.192 P=0.064	R=0.106 P=0.202	R=0.042 P=0.371	R=-0.109 P=0.197
Spine BMD (excluding scoliotics)	60	R=0.254 P=0.025	R=0.189 P=0.074	R=0.249 P=0.027	R=0.158 P=0.116	R=0.100 P=0.222	R=-0.163 P=0.106
Femoral neck BMD (all)	63	R=0.238 P=0.030	R=0.252 P=0.023	R=0.258 P=0.021	R=0.085 P=0.253	R=0.092 P=0.237	R=-0.092 P=0.237
Total hip BMD (all)	63	R=0.007 P=0.480	R=-0.02 P=0.437	R=0.040 P=0.379	R=-0.11 P=0.197	R=-0.14 P=0.140	R=-0.083 P=0.259

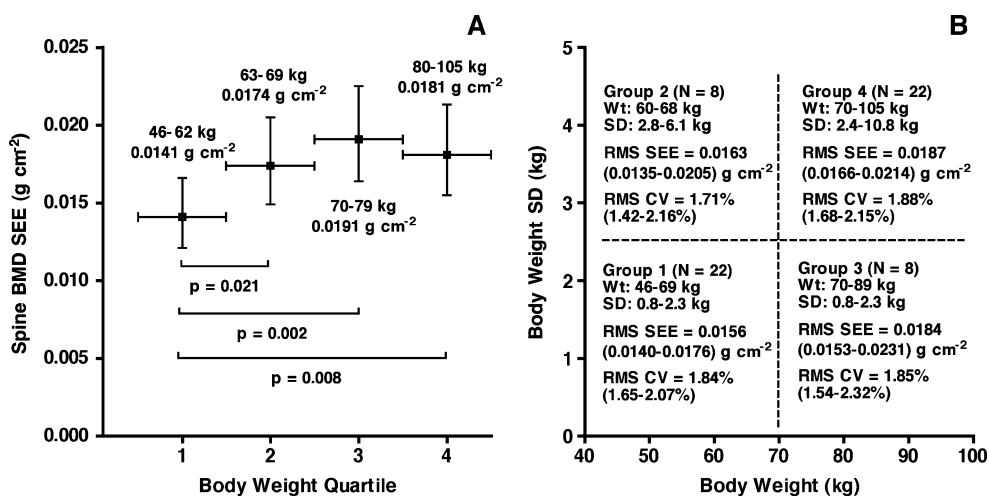
precision error was limited by the scatter of the results for individual subjects shown in Fig. 3. Given the limited number of repeated BMD measurements per subject, the scatter in individual results was consistent with the expected random statistical variations (Fig. 4).

The effect of weight, BMI, and body thickness on precision was examined for both SEE and CV (Table 4). We considered the effect of body thickness, as well as weight and BMI on precision, since a thickness figure is routinely included on scan printouts from Hologic densitometers. Although not intended as an accurate measurement of true body thickness (for example, it is not corrected for the effects of beam hardening on linearity), it does provide a reliable measure of the attenuation of the X-ray beam through the patient's body. For this reason, we hypothesized that it might provide a better predictor of the effect of body size on precision than body weight or BMI. Phantom studies made with a water bath showed that the reported thickness figure was around 10% less than the measured water depth (Fig. 2a). Correlation coefficients between weight, BMI, and thickness measured at the spine and hip were all around  $r=0.9$  (Table 2).

Weight and body thickness were statistically significantly correlated with the spine SEE precision error for the group of 60 subjects that excluded four subjects with poor individual figures due to scoliosis or degenerative disease (Table 4). For femoral neck BMD, weight, BMI, and thickness all correlated significantly with SEE. Only for the total hip site was there no trend for the individual precision figure to be poorer for larger subjects. Welsman et al. reported similar findings, with subjects with higher BMI having poorer precision at the spine and femoral neck sites, but no effect at the total hip [13]. These findings can be explained by the greater body thickness at the spine compared with the hip (Table 1), making the spine BMD precision error more susceptible to the influence of tissue thickness. The greater effect of body size on the precision of femoral neck BMD measurements compared with the total hip site can be explained by the effect of a fat panniculus on BMD measurements at the proximal femur in overweight subjects [15]. DXA scan images often show the edge of the fat panniculus lying over the femoral neck ROI, leaving the majority of the total hip ROI unaffected.

In contrast to the SEE precision error, when precision was expressed as the CV, none of the correlations with body size were statistically significant, and in every instant, the Spearman correlation coefficient was smaller (less positive) than when precision was expressed in BMD units. It is well established that patients' individual BMD measurements at the spine and hip correlate with their body weight [29]. Hence, the findings in Table 4 can be explained by the trend for heavier subjects to have a higher BMD, which partly offsets the poorer precision in BMD





**Fig. 5** **a** Spine BMD RMS SEE precision results for OD14-study women divided into quartiles according to their body weight. Figures against each group show the range of body weights and the root mean square (RMS) precision figure. Vertical error bars show the 95% confidence intervals calculated from the number of degrees of freedom (approximately 100) for each group. **b** Spine BMD precision results for OD14-study women divided into four groups according to whether

their body weight and weight change standard deviation (SD) were greater or less than the median values. For each group, the figure shows the number of subjects *N*, the range of weights and weight change SDs, the RMS SEE and the RMS CV precision figures, and their 95% confidence intervals calculated from the number of degrees of freedom

units in these individuals when precision was expressed as the CV.

To better characterize the effect of body size on precision, the OD14 study women were divided into four quartiles according to their body weight, and the RMS SEE for spine BMD plotted for each quartile (Fig. 5a). The results showed an increase in the precision error from 0.014 g cm<sup>-2</sup> for women in the lowest quartile (46–62 kg) to 0.018 g cm<sup>-2</sup> for women in the highest quartile (80–105 kg). Weight change, as well as body weight itself, is likely to be a factor influencing the precision of BMD measurements. However, because the heavier women tended to be the ones who experienced the greatest changes in weight, it was difficult to separate the two effects. When the OD14-study-women were divided into four groups according to whether their body weight or weight change were above or below the respective median values, the data showed that weight, but not weight change had a significant effect on the SEE.

The present study has a number of limitations. We have already mentioned that the derivation of the long-term precision error requires the assumption of a linear BMD change with time, and that if non-linear changes occur, then the true precision error will be overestimated. The statistical power of the study was limited by the amount of BMD data available. With an average of eight measurements for each subject, the individual precision figures for each woman were subject to significant statistical errors that limited our ability to investigate the correlations with BMD and the effect of weight and the other body-size variables on precision. When individual precision data were pooled into

larger groups, the statistical power of the study was limited by the total of 513 spine BMD and 504 hip BMD examinations. After allowing for the degrees of freedom lost in calculating precision, there remained a total of 398 *df* for the calculation of spine and femoral neck BMD precision, and 389 *df* for the total hip site. Thus, when women were divided in quartiles according to body weight, there were approximately 100 *df* for each quartile, implying a 95% confidence figure of  $\pm 15\%$  for the relative error. Finally, it is clear from the results shown in Fig. 5 that the effects of weight and weight change on the precision error expressed in BMD units were relatively modest over the ranges studied, and there was no effect when precision was expressed in terms of the coefficient of variation. These modest effects may be because of a lack of severely overweight subjects or individuals showing very large weight changes. In the present study, there were only four women with a body weight >90 kg or a BMI >35 kg m<sup>-2</sup>, and only three women with a weight change SD of >5 kg. Therefore, we were unable to examine the effect of body weight on precision in more obese subjects (BMI >40 kg m<sup>-2</sup>), or the effect of weight changes of 10 kg or more on BMD, when there may be considerably larger effects. An important strength of the present study was that we used the BMD data from the original scan analysis performed during the period 1995 to 2006, so our results reflect what was really achieved at the time.

In conclusion, we found that long-term precision errors were around 50% larger than short-term errors. For spine and total hip BMD, it was more appropriate to express the precision errors in BMD units rather than as the coefficient

of variation. When precision was expressed in BMD units, weight had a small but statistically significant effect on the precision of spine and femoral neck BMD over the range 50–90 kg (BMI: 20–35 kg m<sup>-2</sup>), but none for total hip BMD. When precision was expressed as the coefficient of variation, the relationships with weight were no longer statistically significant, probably because of the correlation between body weight and BMD at weight-bearing sites. Weight changes up to 5 kg had little effect on precision, but the relationship was difficult to investigate because subjects with the biggest weight changes also tended to be the heaviest. More studies of individuals >100 kg are required to fully investigate the dependence of DXA scan precision on body weight.

**Conflicts of interest** None.

## References

- National Osteoporosis Foundation (2010) Clinician's Guide to prevention and treatment of osteoporosis. National Osteoporosis Foundation, Washington DC, USA. Available at: [http://www.nof.org/professionals/Clinicians\\_Guide.htm](http://www.nof.org/professionals/Clinicians_Guide.htm). Accessed 23 Feb 2010
- Compston J, Cooper A, Cooper C, Francis R, Kanis JA, Marsh D, McCloskey EV, Reid DM, Selby P, Wilkins M; National Osteoporosis Guideline Group (NOGG) (2009) Guidelines for the diagnosis and management of osteoporosis in postmenopausal women and men from the age of 50 years in the UK. *Maturitas* 62:105–108
- International Society for Clinical Densitometry (ISCD). Official positions 2007. Available at: <http://www.iscd.org/Visitors/positions/OP-Index.cfm>. Accessed 23 Feb 2010
- Bell KJ, Hayen A, Macaskill P, Irwig L, Craig JC, Ensrud K, Bauer DC (2009) Value of routine monitoring of bone mineral density after starting bisphosphonate treatment: secondary analysis of trial data. *BMJ* 338:b2266
- Watts NB, Lewiecki EM, Bonnick SL, Laster AJ, Binkley N, Blank RD, Geusens PP, Miller PD, Petak SM, Recker RR, Saag KG, Schousboe J, Siris ES, Bilezikian JP (2009) Clinical value of monitoring BMD in patients treated with bisphosphonates for osteoporosis. *J Bone Miner Res* 24:1643–1646
- Gluer CC (1999) Monitoring skeletal changes by radiological techniques. *J Bone Miner Res* 14:1952–1962
- Bonnick SL, Johnston CC Jr, Kleerekoper M, Lindsay R, Miller P, Sherwood L, Siris E (2001) Importance of precision in bone density measurements. *J Clin Densitom* 4:105–110
- Gluer CC, Blake G, Lu Y, Blunt BA, Jergas M, Genant HK (1995) Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. *Osteoporos Int* 5:262–270
- Leslie WD (2006) The importance of spectrum bias on bone density monitoring in clinical practice. *Bone* 39:361–368
- Shepherd JA, Morgan SL, Lu Y (2008) Comparing BMD results between two similar DXA systems using the generalized least significant change. *J Clin Densitom* 11:237–242
- Leslie WD (2008) Factors affecting short-term bone density precision assessment and the effect on patient monitoring. *J Bone Miner Res* 23:199–204
- Sadatsafavi M, Moayyeri A, Wang L, Leslie WD (2008) Heteroscedastic regression analysis of factors affecting BMD monitoring. *J Bone Miner Res* 23:1842–1849
- Welsman J, Knapp K, MacLeod K, Blake G (2009) Obesity increases precision errors in dual energy x-ray absorptiometry measurements. *Osteoporos Int* 20(Suppl 4):S267–S268
- Blake GM, Fogelman I (2008) How important are BMD accuracy errors for the clinical interpretation of DXA scans? *J Bone Miner Res* 23:457–462
- Binkley N, Krueger D, Vallarta-Ast N (2003) An overlying fat panniculus affects femur bone mass measurement. *J Clin Densitom* 6:199–204
- Rymer J, Chapman MG, Fogelman I (1994) Effect of tibolone on postmenopausal bone loss. *Osteoporos Int* 4:314–319
- Rymer J, Robinson J, Fogelman I (2001) Effects of 8 years of treatment with tibolone 2.5 mg daily on postmenopausal bone loss. *Osteoporos Int* 12:478–483
- Rymer J, Robinson J, Fogelman I (2002) Ten years of treatment with tibolone 2.5 mg daily: effects on bone loss in postmenopausal women. *Climacteric* 5:390–398
- Cummings SR, Ettinger B, Delmas PD, Kenemans P, Stathopoulos V, Verweij P, Mol-Arts M, Kloosterboer L, Mosca L, Christiansen C, Bilezikian J, Kerzberg EM, Johnson S, Zanchetta J, Grobbee DE, Seifert W, Eastell R, Trial Investigators LIFT (2008) The effects of tibolone in older postmenopausal women. *N Engl J Med* 359:697–708
- Patel R, Blake GM, Rymer J, Fogelman I (2000) Long-term precision of DXA scanning assessed over seven years in forty postmenopausal women. *Osteoporos Int* 11:68–75
- Blake GM, Parker JC, Buxton FMA, Fogelman I (1993) Dual x-ray absorptiometry: a comparison between fan beam and pencil beam scans. *Brit J Radiol* 66:902–906
- Blake GM, Fogelman I (1995) Replacing dual x-ray absorptiometry scanners: cross-calibration of a new multidetector array system. *J Bone Miner Res* 10(Suppl 1):S267
- Blake GM (1996) Replacing DXA scanners: cross-calibration with phantoms may be misleading. *Calcif Tissue Int* 59:1–5
- QDR for Windows XP Reference manual (2004) Hologic Inc, Bedford, MA, pp 9.21–9.23.
- Altman DG (1991) Practical statistics for medical research. Chapman & Hall, London
- Shepherd JA, Fan B, Lu Y, Lewiecki EM, Miller P, Genant HK (2006) Comparison of BMD precision for Prodigy and Delphi spine and femur scans. *Osteoporos Int* 17:1303–1308
- Holland WW, Whitehead TP (1974) Value of new laboratory tests in diagnosis and treatment. *Lancet* 304:391–394
- Hangartner TN (2007) A study of the long-term precision of dual-energy X-ray absorptiometry bone densitometers and implications for the validity of the least-significant change calculation. *Osteoporos Int* 18:513–523
- Noon E, Singh S, Cuzick J, Spector TD, Williams FM, Frost ML, Howell A, Harvie M, Eastell R, Coleman RE, Fogelman I, Blake GM (2010) Significant differences in UK and US female bone density reference ranges. *Osteoporos Int*. 2010 Jan 9. [Epub ahead of print].