Effect of Precision on Longitudinal Follow-Up of Bone Mineral Density Measurements in Elderly Women and Men.

Janaka Lenora[1, 2], Kristina Åkesson[1, 2], and Paul Gerdhem[3, 4]

[1] Department of Clinical Sciences, Clinical and Molecular Osteoporosis Research Unit, Lund University, Malmö, Sweden

[2] Department of Orthopaedics, Malmö University Hospital, Lund University, Malmö, Sweden

[3] Department of Clinical Science, Intervention and Technology, Karolinska Institute, Karolinska University Hospital, Stockholm, Sweden

[4] Department of Orthopaedics, Karolinska University Hospital, Stockholm, Sweden

Abstract

Precision error of dual-energy X-ray absorptiometry exceeds the expected annual rate of bone loss in the elderly. The capacity to detect changes in areal bone mineral density (aBMD; g/cm(2)) over a 5-yr period was assessed. Six hundred ninety-one women, 75.2 (0.1)yr, from the Malmö OPRA-study, were measured using Lunar DPX-L (GE Lunar, Madison, WI), and 211 men, 74.7 (3.2)yr, from the Malmö Mr Os-study, were measured using Lunar Prodigy (GE Lunar) with follow-up 5yr later. Precision error was determined with 30 degrees of freedom. Least significant change (LSC, i.e., 2.77xprecision error) was calculated. Women's precision errors (g/cm(2)) for DPX-L were 0.028 (total hip [TH]) and 0.016 (lumbar spine [LS]), and for Prodigy, they were 0.009 (TH) and 0.039 (LS). In men, corresponding results for Prodigy were 0.014 and 0.031. In women, 41% and in men, 39% had aBMD changes exceeding the LSC at TH. Follow-up intervals (i.e., LSC/median rate of aBMD change) for both women and men were 8yr (TH) and 13yr (LS). Based on Prodigy precision data, follow-up intervals for women were 3 and 32yr at TH and LS. In summary, several years were needed to detect change. Only when a high rate of bone loss is suspected, a short follow-up time is possible, in elderly persons.

Key words: - Dual-energy x-ray absorptiometry, elderly individuals, least significant change, longitudinal follow-up, precision error.

Introduction

Dual-energy X-ray absorptiometry (DXA)-based bone mineral densitometry is the standard tool for diagnosing osteoporosis. Serial measurements of DXA are performed to monitor patients on treatment, in order to estimate the therapeutic response and to monitor high-risk individuals (1). A large decrease in areal bone mineral density (aBMD) in high-risk individuals is of clinical importance, signifying an increasing risk of fracture and supporting the indication for osteoporosis treatment. A significant increase in aBMD following treatment indicates the efficacy of the treatment. The ability to accurately detect the longitudinal change in aBMD depends on the precision or the reproducibility of the technique, which includes the inherent precision error of the technique and the error caused by inconsistent or incorrect positioning of the patient and incorrect analysis (2). Precision of DXA can be affected by factors such as the machine, technologist, the patient population and the skeletal region involved. Because of this, it is recommended that each densitometry lab should determine its own precision values for comparable populations and skeletal regions (1). Sometimes it is difficult to determine whether the observed change in aBMD is a true biological change or a random fluctuation in aBMD due to the precision error of the measurement procedure. To overcome this problem, a least significant change (LSC) or smallest detectable difference has been defined. For longitudinal follow-up studies, the 95% confidence LSC value is defined as 2.77 times the precision error and the 80% confidence LSC is defined as 1.81 times the precision error (2,3). To the best of our knowledge, the practical possibility of detecting an aBMD change exceeding the LSC in community-based longitudinal studies has not been reported. Based on the LSC and the median rate of aBMD change, the monitoring time interval of disease progression has been defined as the time period after which half of individuals with normal changes will show a measured change exceeding the LSC (3). In elderly women, aBMD decreases at a slower rate than in early post-menopausal women (4,5). The aBMD change is usually also reported as being low in elderly men (4,5). Consequently, it is not clear whether the commonly used follow-up period of 1–2 years is applicable to elderly individuals.

The purpose of this study was 1) to assess the ability to detect individuals with aBMD changes exceeding the LSC over a period of 5 years in two population-based cohorts of elderly individuals, and 2) to compare the follow-up intervals using two different DXA machines. Study subjects were recruited from two cohorts: elderly women from the Malmö-OPRA study (6) and elderly men from the Mr Os Malmö sample of the Mr Os Sweden study (7).

Materials and Methods

Participants

The Malmö-OPRA study is an ongoing population-based study on risk factors for osteoporosis and fracture in elderly women. The participants were recruited from the population registers of ZZZ. Details of the study have been published elsewhere (6,8). In brief, 1,044 women attended at baseline, and 995 had a DXA measurement at at least one site. At the 5-year follow-up, 691 had a second aBMD measurement performed at at least one site. Of the 691 women in this study 82 were treated with hormone replacement therapy or bisphosphonates at any time between baseline

and the 5 year follow-up. According to the self-assessment questionnaire, women had sustained a fracture at least once at any time before the baseline investigation.

The Mr Os Sweden study is an international study on risk factors for osteoporosis and fracture in elderly men. The men in the Mr Os Malmö cohort of the Mr Os Sweden study were recruited from the population registers of Malmö. This report contains the subset of men who were recruited during the first 13 months of the study. Of the 316 men with a baseline measurement, 211 men returned for the 5-year follow-up measurement. Of the two at any time between baseline and the 5 year follow-up. Fracture data for the men was not available when this study was performed.

Bone density measurements

In women, aBMD measurements were performed for the total body, the total hip, the femoral neck and the lumbar spine (L2-L4) by using a Lunar DPX-L scanner (Lunar DPX-L; GE Lunar, Madison, WI, USA), which uses the pencil beam technique. Scan analysis at baseline was done with software versions 1.33 and 1.35. Follow-up scans at five years were done with software version 4.7b, with the exception of hip scans, which were all analysed with software version 4.7b. Men were measured at the same regions using a Lunar Prodigy scanner (Lunar Prodigy; GE Lunar, Madison, WI, USA), which uses the fan beam technique. Software version 2.05 was used for scan analysis at baseline and at five years. DXA scans were performed and analysed by experienced personnel according to the guidelines provided by the manufacturers. Longitudinal stability of the measurements was established by daily calibrations using a Hologic spine phantom (Hologic Inc., Bedford, MA, USA).

DXA precision measurements

Reproducibility assessments were performed by the same technician, using consecutive volunteer samples from the two cohorts. Between measurements, each individual stepped out of the machine, stood up on the floor, lay down on the machine again and was repositioned. Coefficient of variation (CV%) and absolute precision error (root mean square standard deviation; RMS-SD) were calculated as suggested by Gluer (9). Precision of the Lunar DPX-L was calculated by duplicate measurements of 30 elderly women (80–85 years of age) taking part in the Malmö-OPRA study, and 31 measurements of the spine phantom. Precision for the Lunar prodigy was calculated by duplicate measurement of 30 elderly men taking part in the Mr Os Malmö study and triplicate measurement of 15 85-year-old women from the Malmö-OPRA study, and 31 measurements of the spine phantom. For estimation of precision with the phantom, the phantom was taken out of the machine for a while and then placed on the machine again.

Statistics

In the longitudinal follow-up, aBMD change between baseline and 5-year scans (5-year aBMD minus baseline scan aBMD) were calculated. The 95% confidence LSC was calculated as 2.77

times the absolute precision error, and 80% confidence LSC as 1.81 times the absolute precision error. Individuals with aBMD change beyond the 95% confidence LSC and the 80% confidence LSC were identified. Monitoring time intervals were also calculated for both follow-up cohorts using the 95% confidence LSC (95% LSC/median annual rate of change in aBMD).

All steps of the study were approved by the ethical review committee of AAA University in accordance with the Declaration of Helsinki. Informed consent was obtained from each of the participants prior to the study.

Results

Precision measurements

Age, aBMD values, absolute precision error and CV% of the individuals who attended the precision measurements are summarised in Table 1. Precision errors for the phantom (L2–L4) were 0.017 g/cm^2 (CV = 1.4%) for the Lunar DPX-L scanner and 0.006 g/cm^2 (0.6%) for the Lunar Prodigy scanner. Precision errors for total body aBMD ranged between 0.009 and 0.010 g/cm^2; for the lumbar spine, aBMD ranged between 0.016 and 0.039 g/cm^2; for the total hip, aBMD ranged between 0.009 and 0.028 g/cm^2 and for the femoral neck, aBMD ranged between 0.011 and 0.030 g/cm^2. Lunar DPX-L total hip and femoral neck absolute precision errors for elderly women were greater than corresponding values for Lunar Prodigy, both for elderly women and elderly men (Table 1).

Longitudinal follow-up

The mean age of elderly women at baseline was 75.2 ± 0.1 years and mean aBMDs (g/cm^2) was 1.008 ± 0.093 for total body, 0.857 ± 0.147 for total hip, 0.766 ± 0.138 for femoral neck and 0.987 ± 0.190 for lumbar spine. Based on the femoral neck aBMD at baseline, 219 women (30.5%) were osteoporotic, 378 women (52.6%) were osteopenic and 121 women (16.9%) were considered normal.

After 5.0 ± 0.1 years, mean aBMD of the total body, the total hip and the femoral neck were lower than baseline figures (femoral neck and total hip: $p < 0.001$; total body: $p = 0.054$) while mean aBMD of the lumbar spine was slightly higher than at baseline ($p = 0.098$) (Table 2). The percentage of women who had a change in aBMD that exceeded the 95% confidence LSC ranged from 39% to 48% in different regions. The corresponding percentage range was 51–63 % when the 80% confidence LSC was used. However, when Lunar Prodigy precision values for elderly women were applied for calculation of LSC, the percentage of women who had a change in aBMD that exceeded the 95% confidence LSC was 44% for total body, 76% for the femoral neck, 78% for total hip and 19% for the lumbar spine.

The mean age of elderly men at baseline was 74.7 ± 3.20 years and mean aBMDs (g/cm^2) were 1.187 ± 0.097 for total body, 0.982 ± 0.138 for total hip, 0.898 ± 0.133 for femoral neck and 1.240 ± 0.190 for lumbar spine. Based on the femoral neck aBMD at baseline, 28 men (13.5%) were osteoporotic, 100 men (48.3%) were osteopenic and 79 men (36.2%) were considered

normal. At the follow-up (mean 5.1 ± 0.1 years), mean aBMD for total body, total hip and femoral neck were lower than at baseline ($p < 0.001$) while mean aBMD for the lumbar spine was higher ($p < 0.001$). The percentage of men with a change in aBMD exceeding the LSC at the different regions ranged from 25% to 39% at the 95% confidence level (Table 2) and from 45% to 55% when the 80% confidence level was used.

When osteoporotic, osteopenic and normal individuals were analyzed separately, there were no substantial differences in the proportions of women or men who had a change in aBMD of more than the LSC (data not shown).

Calculated interval between DXA measurements

Based on the findings above, the time gap between DXA measurements to reach a level where at least half of the individuals had a change in aBMD exceeding 95% LSC was calculated (Table 3). The calculated minimum follow-up intervals for the total hip were 8.1 years in both elderly women and men. For the femoral neck region, the time intervals were 7.8 years for women and 20.6 years for men. The lumbar spine follow-up intervals were 13.0 years for women and 12.6 years for men. The intervals for the total body were 13.2 years for women and 9.6 years for men.

When the Lunar Prodigy precision values for women and the median rate of aBMD changes obtained by Lunar DPX-L scanner were used for calculation of follow-up interval with DXA, the interval was 2.9 years for total hip and femoral neck, 11.9 years for total body and total hip, and 31.8 years for lumbar spine (Table 3).

A separate calculation without the that were using hormone replacement therapy or bisphosphonates did not substantially alter the results (data not shown).

Discussion

In this study of two population-based cohorts of elderly women and men, the number of individuals with aBMD changes greater than the least significant change during a 5-year period was low. The estimated interval for meaningful reassessment was long, based on the precision error values of the cohorts and rate of aBMD change. These findings have clear clinical implications for the follow-up of osteopenic individuals who are at a risk of developing osteoporosis later in life. The usual recommendation is to follow up osteoporotic patients on treatment for about 1–2 years using aBMD measurements (1), and in practice individuals at risk of developing osteoporosis are followed with DXA using similar time intervals. However, the findings in this study raise the question of whether such an interval may be too short, especially for elderly individuals, unless a high rate of bone loss or gain in aBMD is expected. Factors associated with fast bone loss (and osteoporosis), such as long-term immobilization, glucocorticoid therapy, primary hyperparathyreoidism and high bone turnover should of course be taken into account in the clinical judgment concerning the need of a remeasurement (10,11). Except for remeasurement of aBMD, measurements of bone turnover may add information on rate of bone loss (12).

Leslie reported results from a mixed clinical cohort with a mean age of 56 years (13). After a mean follow-up of 1.75 years, 27% had a change in aBMD exceeding the 95% confidence LSC

at the spine and the corresponding change was 40% for the total hip. Precision for hip data in that study was similar to precision for data from the women measured on Lunar Prodigy in the present study, while precision for the spine was better. The study individuals consisted of a clinical cohort and they were closer to menopausal age. They were therefore more likely to have a higher rate of bone loss than the elderly women in the Malmö-OPRA study. To the best of our knowledge, only one previous study has reported precision errors for individuals older than 70 years (14); Ravaud *et al.* reported that elderly women have a greater precision error (RMS-SD) for the lumbar spine and total hip than early postmenopausal women. The reason for higher precision errors in elderly individuals could be due to the degenerative changes and compression fractures that are common in this age group. According to our own experience, 25-year-old women examined at our research laboratory with Lunar Prodigy have similar precision errors to those of elderly men and women in the hip and the total body, but lower absolute precision error for the spine (unpublished data). On the other hand, in a study comparing CV% of 60 young individuals (30 men and 30 women aged 28.2 ± 5.5 years), 102 post-menopausal women (age 58 ± 7.0 years), and 60 patients (age 47.1 ± 15.1 years) who had chronic rheumatic disease, El Maghraoui *et al.* concluded that total hip and lumbar spine precision error does not depend on age, aBMD or disease conditions (15).

The precision error of DXA measurements varies depending on the equipment used, the populations of individuals included and different skeletal regions measured. When the precision errors of elderly women were compared for the two machines, the proximal femur CV% and absolute precision error of the pencil beam (Lunar DPX-L) machine were greater than the precision values calculated for the fan beam machine (Lunar Prodigy). Part of this difference can be explained by technical differences, which is evidenced by the higher precision error, absolute and CV%, from the phantom measurements with the Lunar DPX-L machine when compared to the Lunar Prodigy machine. Other contributing factors may be the different leg-positioning aids used in the two machines, since the angle of leg rotation affects the aBMD value and the precision error (16). However, some of the CV% values obtained in our study are higher than the ISCD-recommended levels (1). Although the precision measurements were performed by the same technician, who was well-trained and had several years of experience, and repeated measurements were performed on the same day, this could be due to the age-related skeletal deformities and physical difficulties in positioning the individuals. One earlier study has reported that the Lunar Prodigy system has lower precision error values than the Hologic Delphi system, both of which are fan beam systems (17). On the other hand, some authors have shown that precision error does not depend on the machine type, pencil beam or fan beam (18,19), or whether the same technician or different technicians performed the test, but it depends on whether repeated measurements were performed on the same day or on different days (18). These studies were performed using younger individuals than were used in the current study (18,19).

When the time interval required between DXA measurements in order to detect significant changes was calculated using the precision error values and median rate of aBMD changes, we found that the time to reassessment should be around 8 years for total hip and 13 years for the lumbar spine. The extremely long follow-up interval required for the lumbar spine is probably caused by the masking of bone loss by vertebral compression fractures, osteophytic changes of the spine and aortic calcifications, as evidenced by the increase in lumbar spine aBMD in both men and women. Consequently, the evaluation of spinal aBMD measurement must always

consider the quality of the scan and unless it is possible to evaluate each vertebra, spine DXA should not be used for longitudinal assessment in elderly individuals. It is therefore advisable not to use lumbar spine aBMD changes in elderly individuals for follow-up.

For calculation of CV% values, we used samples of 30 degrees of freedom (30 individuals measured twice, 15 individuals measured 3 times and 31 phantom measurements) as recommended by others (2,9). Although Leslie and Moayyeri have reported that sample sizes of 30 degrees of freedom may not be sufficient to reliably estimate the precession error (20), this study was initiated earlier and the recommendations of Gluer *et al.* (9) were used. The aim of this study was especially to evaluate the precision of DXA in elderly individuals, and the strength of our study is the inclusion of two elderly population-based cohorts of both women and men, and the long follow-up and high degree of reassessment rate.

We conclude that precision has an impact on the shortest interval that should be advocated when doing repeated scans, when following changes in aBMD. In these population-based cohorts of elderly people, several years were required to detect a significant change between different measurements. The newer Lunar Prodigy machine had better precision than the older DPX-L machine. The lumbar spine is likely to be unreliable for the detection of bone loss in the elderly. When deciding on the interval between follow-up DXA measurements, precision error and the expected bone loss should be taken into account. A short follow-up is suitable only when a high degree of change in bone density is anticipated.

References

1. International Society of Clinical Densitometry (ISCD) Official Positions 2007, pp www.iscd.org/Visitors/pdfs/ISCD2007OfficialPositions-Combined-AdultandPediatric.pdf.
2. 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**(2)**:**105-10.
3. Gluer CC 1999 Monitoring skeletal changes by radiological techniques. J Bone Miner Res **14**(11)**:**1952-62.
4. Emaus N, Berntsen GK, Joakimsen R, Fonnebo V 2006 Longitudinal changes in forearm bone mineral density in women and men aged 45-84 years: the Tromso Study, a population-based study. Am J Epidemiol **163**(5)**:**441-9.
5. Jones G, Nguyen T, Sambrook P, Kelly PJ, Eisman JA 1994 Progressive loss of bone in the femoral neck in elderly people: longitudinal findings from the Dubbo osteoporosis epidemiology study. BMJ **309**(6956)**:**691-5.
6. Lenora J, Ivaska KK, Obrant KJ, Gerdhem P 2007 Prediction of bone loss using biochemical markers of bone turnover. Osteoporos Int **18**(9)**:**1297-305.
7. Mellstrom D, Johnell O, Ljunggren O, Eriksson AL, Lorentzon M, Mallmin H, Holmberg A, Redlund-Johnell I, Orwoll E, Ohlsson C 2006 Free testosterone is an independent predictor of BMD and prevalent fractures in elderly men: MrOS Sweden. J Bone Miner Res **21**(4)**:**529-35.
8. Gerdhem P, Akesson K 2007 Rates of fracture in participants and non-participants in the Osteoporosis Prospective Risk Assessment study. J Bone Joint Surg Br **89**(12)**:**1627-31.
9. 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**(4)**:**262-70.

10. Espallargues M, Sampietro-Colom L, Estrada MD, Sola M, del Rio L, Setoain J, Granados A 2001 Identifying bone-mass-related risk factors for fracture to guide bone densitometry measurements: a systematic review of the literature. Osteoporos Int **12**(10)**:**811-22.
11. Kanis JA 2002 Diagnosis of osteoporosis and assessment of fracture risk. Lancet **359**(9321)**:**1929-36.
12. Ivaska KK, Lenora J, Gerdhem P, Akesson K, Vaananen HK, Obrant KJ 2008 Serial assessment of serum bone metabolism markers identifies women with the highest rate of bone loss and osteoporosis risk. J Clin Endocrinol Metab **93**(7)**:**2622-32.
13. Leslie WD 2006 The importance of spectrum bias on bone density monitoring in clinical practice. Bone **39**(2)**:**361-8.
14. Ravaud P, Reny JL, Giraudeau B, Porcher R, Dougados M, Roux C 1999 Individual smallest detectable difference in bone mineral density measurements. J Bone Miner Res **14**(8)**:**1449-56.
15. El Maghraoui A, Do Santos Zounon AA, Jroundi I, Nouijai A, Ghazi M, Achemlal L, Bezza A, Tazi MA, Abouqual R 2005 Reproducibility of bone mineral density measurements using dual X-ray absorptiometry in daily clinical practice. Osteoporos Int **16**(12)**:**1742-8.
16. Lekamwasam S, Lenora RS 2003 Effect of leg rotation on hip bone mineral density measurements. J Clin Densitom **6**(4)**:**331-6.
17. 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**(9)**:**1303-8.
18. Leslie WD 2008 Factors affecting short-term bone density precision assessment and the effect on patient monitoring. J Bone Miner Res **23**(2)**:**199-204.
19. Omsland TK, Emaus N, Gjesdal CG, Falch JA, Tell GS, Forsen L, Berntsen GK, Meyer HE 2008 In vivo and in vitro comparison of densitometers in the NOREPOS study. J Clin Densitom **11**(2)**:**276-82.
20. Leslie WD, Moayyeri A 2006 Minimum sample size requirements for bone density precision assessment produce inconsistency in clinical monitoring. Osteoporos Int **17**(11)**:**1673-80.