Osteoporos Int (1998) 8:1–3

Osteoporosis International

Editorial

Standardization of BMD Measurements

C. A. Formica

Regional Bone Center, Helen Hayes Hospital, West Haverstraw, and College of Physicians and Surgeons of Columbia University, New York, NY, USA

The diagnosis of osteoporosis is usually based on the WHO diagnostic criteria defining persons as normal, osteopenic or osteoporotic depending on their standardized score relative to a young adult population mean (*T*-score) for bone mass and the presence or absence of fractures. The accurate diagnostic classification of patients is dependent on many factors, in particular the accurate derivation of the reference population mean and standard deviation used to calculate the *T*-score, and the comparability of bone mass measurements across different manufacturers. Discrepancies between patients classified according to the WHO or other diagnostic criteria using different bone densitometers have been previously documented [1–4]. Recently, several investigators have reported concerns regarding the use of the Hologic reference database for the femoral neck for the diagnostic classification of patients [3]. Any discrepancies between manufacturers may have significant impact in clinical therapeutic efficacy trials and in the clinical diagnosis and management of patients, as the same individual measured on one instrument may be classified differently if measured on another instrument.

The International Committee for Standards in Bone Measurement was established to address these and other issues concerned with the standardization of DXA results. The committee has already presented its recommendations for standardization of spine BMD measurements, and, in the previous volume of *Osteoporosis International*, have presented their recommendations for the standardization for proximal femur bone mineral density (BMD) measurements [5,6]. It is without doubt a difficult task to achieve universal consensus regarding standardization given the diverse opinions from academia and industry. However, the Committee has done extremely well to bring forward the definition and standardization of proximal femur BMD measurements. The Committee will promote the Standardized Total Femur BMD (sBMD) in place of the currently widely employed manufacturer-specific Femoral Neck BMD.

In anticipation of the widespread application of standardized BMD measurements, it is both timely and appropriate to suggest caution before standardized measurements are openly endorsed. Standardized BMD measurements will play an important role in prospective clinical trials where patient data obtained using densitometers from different manufacturers are pooled. However, in the clinical setting, standardized BMD measurements may complicate patient diagnosis and monitoring.

It is in the clinical setting where the greatest problems concerning sBMD measurements may be manifested. The concept of sBMD measurements, particularly with the manufacturers using this variable as their default measure, may encourage the use of multiple scanners in the clinical setting. It must still be emphasized that any baseline and follow-up measurements should be performed on the same instrument to avoid misdiagnosis or patient mismanagement. The concerns arguing against the widespread clinical use of sBMD will be discussed and are not mutually exclusive. There is considerable overlap between factors that influence the derivation of sBMD, and each may have an additive effect that significantly increases the error associated with sBMD. It is important to bear in mind that the International Committee for Standards in Bone Measurement has not set out to bring about agreement in the definition of regions, or their detection algorithms, but has sought to achieve agreement for the diagnosis of osteoporosis.

Correspondence and offprint requests to: Carmelo A. Formica, Regional Bone Center, Helen Hayes Hospital, Route 9W, West Haverstraw, NY 10993, USA. Tel: +1 (914) 947–3000. Fax: +1 (914) 947–2485.

The International Committee for Standards in Bone Measurement has recommended the use of the Total Femur region as the region of interest for femur evaluations. This region has been reported by several studies and has been shown to be equally diagnostic and more precise than the Femoral Neck region. Furthermore, only about 45% of hip fractures occur at the femoral neck, and therefore the Total Femur region may provide a theoretical advantage over the Femoral Neck region. Unfortunately, the Total Femur region as defined by one manufacturer is not equivalent to that defined by another, and significant differences in the skeletal localization and size of the regions exists. Furthermore, different manufacturers employ different edge detection algorithms to separate bone from soft tissue. Bärenholdt et al. [7] using bone mineral equivalent standards demonstrated significant erroneous increases in calculated area with increasing mass using a Hologic QDR 4500 and QDR 2000 bone densitometer. These differences in proprietary algorithms and definitions contribute to lack of comparability and measurement variability between instruments.

The standard error of the estimate for standardized BMD is approximately 3%. Therefore, the intra-individual variability is expected to be between $\pm$ 3–6%. An error of this magnitude would not be expected to influence patient diagnosis greatly. However, if the sBMD were utilized for patient monitoring then significant errors could arise when determining clinically significant changes. The objective for developing a sBMD was for universal diagnostic standardization, yet the Committee has suggested implementing standardized units by default on new systems. It must be stressed that for accurate clinical monitoring the same instrument must be used for all patient follow-up measurements. It would be undesirable to allow clinicians to develop the perception that universal sBMD is also acceptable for patient monitoring and determining treatment efficacy.

The underlying basis for sBMD is that all machines from the same manufacturer provide identical results. While substantial efforts are undertaken by the manufacturers to ensure inter-instrument comparability, significant differences can be seen in the clinical environment. Figure 1 shows the results for the European Spine Phantom (ESP) measured on 17 different Lunar DPX-L machines across the USA and the resulting probability plot. The variability from the minimum to the maximum is about 5%, for identical instruments from the same manufacturer. Gaither et al. [8] using both Hologic and Lunar instruments have also demonstrated that, in the worst case, differences of 5.0% were observed between systems from the same manufacturer.

The Committee should also enforce the fact that the sBMD measurements are derived on pencil beam systems, and that application of these equations to combined pencil and fan beam systems could lead to considerable variability and further complicate the use of sBMD in the clinical setting. Figure 2 shows the results of the ESP measured 10 times on pencil and fan beam systems from the same manufacturer. Differences of 6–8% are observed, particularly when comparing the Hologic QDR 1500W (pencil beam) and the Hologic QDR 4500W (fan beam) systems. What are the clinical implications of such variability between measurements? If we assume a 65-year-old women with a *true* sBMD of 809 mg/cm^2 (*T*-score = -1.2 SD) then her measured sBMD could range from 716 mg/cm^2 (*T*-score = -1.6 SD) to 857 mg/cm^2 (*T*-score = -0.8 SD). This variability in derived sBMD could not only affect patient diagnosis, but may have more significant implications for fracture prediction and patient management.

The International Committee for Standards in Bone Measurement has also recommended as the reference database for the Total Femur region, the bone density data collected in phases 1 and 2 of the third National Health and Nutrition Survey (NHANES III). There is little doubt that the NHANES III reference data are the most applicable for the US population. Unfortunately, the NHANES III data are only available for the femur.

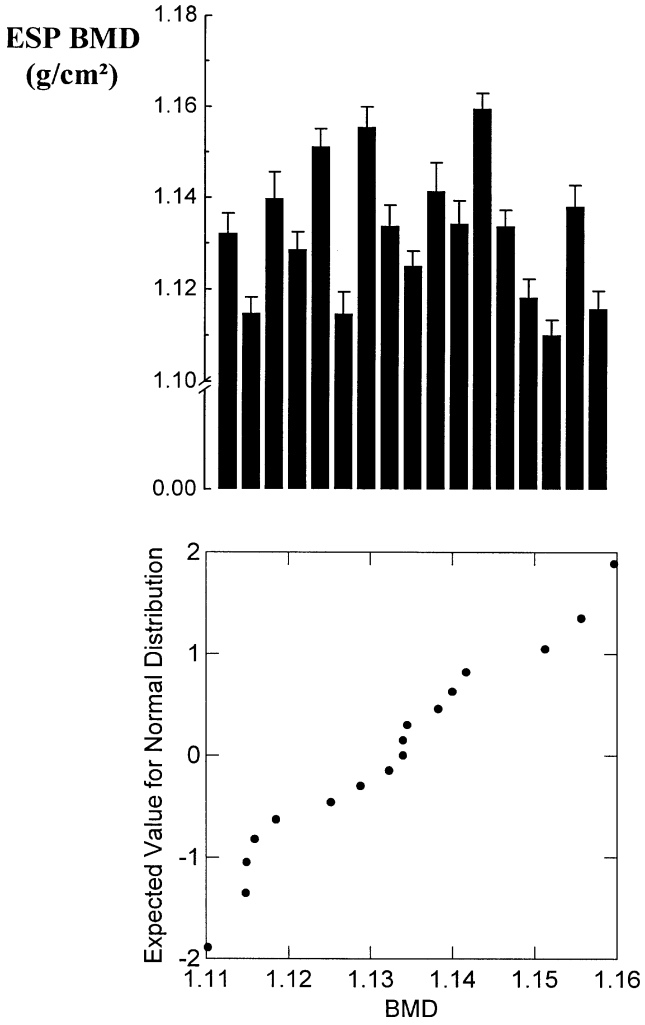

Fig. 1. Inter-machine variability determined by measuring the European Spine Phantom (ESP) 10 times on 17 different Lunar DPX-L DXA machines across the USA and the resulting probability plot.

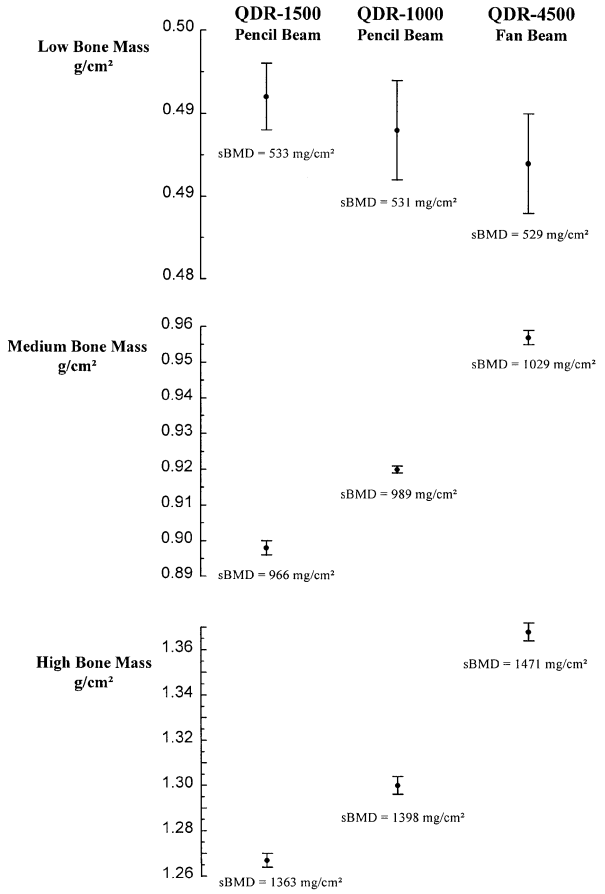

Fig. 2. European Spine Phantom measured 10 times using three different Hologic DXA Systems: QDR-1500W (pencil beam), QDR-1000 (pencil beam) and QDR-4500W (fan beam).

Other population-based reference data such as those currently being evaluated by the European Foundation for Osteoporosis are needed. With the Committee recommending the incorporation of sBMD as the default measure for both the lumbar spine and femur, concern arises as to which reference database for the lumbar spine manufacturers will utilize. Concerns addressing the validity of manufacturer-supplied reference databases have been well documented. Therefore, it would perhaps be prudent before the widespread application of sBMD that considerably more effort be directed toward the definition and establishment of reference data specific to geographic regions.

The use of sBMD will have important applications, particularly in the area of clinical research. In the correct situation the application of sBMD may prove valuable; however, caution is suggested before full endorsement is given for sBMD in clinical practice. Concerns regarding region-specific reference ranges, intra- and inter-manufacturer differences in hardware and software algorithms, and the possible encouragement of allowing patients to be measured on different machines for longitudinal monitoring questions the validity of sBMD in clinical practice. All these factors influence measurement accuracy and variability, which will have significant implications for patient management. It is important to remember that the International Committee for Standards in Bone Measurement has sought to achieve agreement for the diagnosis of osteoporosis, and therfore their recommendations are not appropriate for patient monitoring and management. The debate concerning sBMD measurements will undoubtedly continue, and I encourage the Committee to continue and hasten its excellent work. However, adequate consensus derived from large-scale debate is necessary before any endorsements are given that may influence clinical practice.

References

1. Pocock NA, Sambrook PN, Nguyen T, Kelly P, Freund J, Eisman JA. Assessment of spinal and femoral bone density by dual x-ray absorptiometry: comparison of Lunar and Hologic instruments. J Bone Miner Res 1992;7:1081–4.
2. Greenspan SL, Maitland-Ramsey L, Myers E. Classification of osteoporosis in the elderly is dependent on site-specific analysis. Calcif Tissue Int 1996;58:409–14.
3. Faulkner KG, Roberts LA, McClung MR. Discrepancies in normative data between Lunar and Hologic DXA systems. Osteoporos Int 1996;6:432–6.
4. Manfredonia DJ, Weiss SR. Bolognese MA, Abbott TA, Steven GJ, Bolognese CJ, et al. The prevalence of osteoporosis and osteopenia among women: an analysis of two convenience populations. J Bone Miner Res 1996;11(Suppl 1):S236.
5. Steiger P for the Committee for Standards in DXA. Standardization of spine BMD measurements [letter]. J Bone Miner Res 1995;10:1602–3.
6. Hanson J for the International Committee for Standards in Bone Measurement. Letter to the editor. Osteoporos Int 1997;7:500–1.
7. Bärenholdt O, Kolthoff N, Nielsen SP. Accuracy errors in bone densitometry due to edge detection algorithms [abstract]. World Congress on Osteoporosis, 1996.
8. Gaither KW, Faulkner KG, Ostrem EC, Backman JK, Avecilla LS, McClung MR. Variations in calibration among like-manufacturer DXA systems. J Bone Miner Res 1996;(Suppl 1):S119.