Introduction

Bone densitometry is frequently used to determine an individual's fracture risk at a particular point in time but may also be used to assess changes in bone mineral density (BMD) due to therapy, aging or disease. Since the magnitude of any such changes is likely to be small, perhaps in the region of a few percent per annum, it is essential that the measurements obtained are as precise as possible.

Dual Energy X-ray Absorptiometry (DXA) is currently the most widely used technique for the assessment of skeletal status in both the clinical and research setting and therefore will be the focus for this discussion on Quality Assurance. However the principles described will be equally applicable to other methods of assessing skeletal status such as Single Energy X-ray Absorptiometry (SXA), peripheral DXA (pDXA), Quantitative Computed Tomography (QCT), peripheral QCT (pQCT) and ultrasound.
Precision

The precision of a technique is the ability of that technique to reproduce the same result whereas the accuracy of the technique describes the difference between the "true" value and the measured value.

There are two main factors that contribute to the precision of BMD measurements, instrument-induced errors and operator-induced errors, although factors such as changes in an individual's weight, degenerative changes particularly in the spine and unexplained biological variation may also contribute to the precision in longitudinal studies.

Precision of Instruments

The ideal method for monitoring the performance of bone densitometry instruments is by the regular scanning of a phantom, preferably daily but at least weekly, containing a stable amount of bone mineral or equivalent and assessing the ability of the instrument to reproduce the BMD of the phantom. The BMD data can then be evaluated using statistical quality control (QC) procedures. QC procedures that have been adapted and applied to bone densitometry include the multi-rule Shewhart method, the Cusum technique and simple visual inspection of the data.\textsuperscript{1,3-6}

One of the major benefits of instigating a QC protocol involving the regular scanning of a phantom is that it may allow the early identification of changes in instrument performance prior to instrument malfunction, although of course not all malfunctions will affect BMD data. The prevention or prediction of instrument malfunction is particularly useful in the clinical trial setting where the timing of a BMD measurement may be important. Phantom measurements are also useful in determining if changes in BMD data have occurred following changes in instrument components or software (figure 1). The meticulous recording of any changes to the instrument will assist in the identification and attribution of any changes seen in the BMD data. A further benefit of performing regular scans of phantoms is that it will allow the retrospective adjustment of any patient data acquired during periods affected by change in instrument performance.\textsuperscript{7,8}

The short-term precision of BMD measurements using DXA has been assessed in a number
of studies using a variety of phantoms. A short-term precision (expressed as a coefficient of variation) of approximately 0.5% for spine phantom measurements and 1-4% for femur phantoms has been reported. The longer-term precision over three years has also been reported to be good, approximately 0.5% for spine phantoms. However, small but significant changes in instrument performance can occur during longitudinal use. These changes can take the form of drifts or jumps in the BMD data, and may be clinically relevant when they are comparable to the expected changes in BMD. The daily calibration procedures required by some DXA instruments, prior to patient scanning, may not detect small but potentially clinically significant changes.

Types of Phantoms
A variety of phantoms have been used to assess instrument performance. The Hologic spine phantom is supplied with each Hologic DXA instrument (Hologic Inc, Waltham, MA, USA) for use as a QC phantom. The phantom is composed of four semi-anthropomorphic (ie. with some resemblance to a human spine) hydroxyapatite (HA) vertebrae of a single density level (approximately 1.05 g/cm² when scanned on Hologic instruments) and therefore calibration of the other levels is not checked. A further potential disadvantage of the Hologic spine phantom is that the material surrounding the spine is not soft-tissue equivalent and this may influence the soft tissue compensation made by instruments from different manufacturers.

Hologic also manufacture a HA step spine phantom which provides a range of BMD’s (0.6 - 1.6 g/cm²) but it is not anthropomorphic and thus does not provide a check of the edge detection algorithms. An HA hip phantom is also manufactured by Hologic with the configuration of a proximal femur and some internal structures to reflect the varying densities of the trochanter and Ward’s triangle and this has been used in some precision studies.

The Lunar aluminium spine phantom is supplied with each Lunar DXA instrument (Lunar Corp., Madison, WI, USA) for use as a QC phantom. The aluminium spine ranges in thickness to provide four complete vertebrae with a range of densities (approximately 0.9 to 1.4 g/cm² when scanned on Lunar instruments). The aluminium spine has to be scanned with some soft tissue-equivalent and this is usually accomplished by placing the phantom in a water bath containing 15 cm depth of water. Disadvantages of this phantom are that it is not made of HA, it is not anthropomorphic and that even the lowest density level represented (approximately 0.9 g/cm²) does not test the calibration at the lowest levels of BMD seen clinically. Lunar also manufacturer an aluminium hip phantom with a step to simulate the Ward’s triangle region. As with the spine phantom it needs to be scanned with soft tissue equivalent and is not anthropomorphic.

Various cadaveric phantoms have also been used for assessing the performance of DXA instruments. The advantage of using this type of phantom is that it is as anthropomorphic as possible ex vivo. The disadvantages are that without ashing the phantom the true bone mineral content is not known and that a range of phantoms would be required to assess the linearity at all levels of BMD.

Spine and forearm phantoms have been developed, under the auspices of an EU organisation, COMAC-BME (Committee d’Actions Concertes - Bio Medical Engineering), for use with bone densitometry instruments. These are now known as the European Spine Phantom (ESP) and the European Forearm Phantom (EFP).

The ESP (figure 2) is composed of three semi-anthropomorphic vertebrae representing low, medium and high bone mineral content (BMC), BMD and wall thicknesses respectively. The ESP is composed of epoxy-resin based plastics with various additional constituents to achieve water and bone equivalent solid materials. The HA densities were selected to represent the range of values found clinically and thus the area densities (including the intervertebral spaces) of 0.5, 1.0 and 1.5 g/cm² HA are provided when the phantom is scanned in the posteroanterior mode on DXA instruments. The ESP may therefore, be a suitable phantom for QC purposes as it is relatively anthropomorphic, provides a range of
density levels of known BMC and is composed of true soft tissue equivalent. The ESP and the EFP have been used as standardisation phantoms to cross-calibrate bone densitometry instruments since there are differences not only between bone densitometry instruments from different manufacturers but there may also be small differences between instruments from the same manufacturer. These inter-instrument differences require that repeat scans on an individual should ideally be performed using the same instrument that was used for the baseline scans.

**Figure 2**

**Precision in subjects**
The reported precision of BMD measurements in subjects is usually higher than that of phantoms due to variations in positioning, patient movement, the heterogeneity of the tissues measured and the more complex anatomy that the edge-detection algorithms are required to assess. The reported precision of measurements of the proximal femur (approximately 1-4%) is generally poorer than that of the spine (approximately 1-2%) and the reasons for this include variation in leg positioning, smaller size of the region-of-interest (ROI) at the femoral neck and lower BMD values. Variable positioning of the foot or leg even within the recommended range of 0° to 20° inward rotation can result in a steady change in the BMD values obtained.

Other sources of operator error include using the incorrect scan speed or current during the acquisition of data, or errors during the analysis of data where there may be variation in the size or position of the ROI selected for analysis. There may be also be variation between operators; in one study a precision of 3.7% for femoral neck measurements was reported when site radiographers performed the scan analysis compared with 2.1% when the analysis was performed by a trained operator. BMD measurements performed on different days (1-4 weeks apart) are reported to have a higher variance than those performed on the same day and this discrepancy was found to be most pronounced in post-menopausal women compared to pre-menopausal women. Other studies have also found lower precision in older individuals or those with a low BMD. Precision has also been found to be compromised in those with a greater body thickness.
Therefore, it is important for each centre performing bone densitometry measurements to determine their own "in-house" precision values that are applicable to the population to be scanned (e.g. by performing repeat measurements of volunteers with repositioning between scans). Once a centre has determined their precision values it will enable the least significant change that may be observed in patients to be calculated. Operator-induced errors may be reduced by measurements being performed by a small number of dedicated and highly trained individuals who follow standard operating procedures for the positioning of patients and for the acquisition and analysis of scan data. Errors in the positioning of ROIs in follow-up scans may be minimised by use of the "compare" facility which is available on most DXA instruments. This facility allows the simultaneous viewing and superimposing of serial scans enabling comparable ROIs to be analysed.

Summary

Precision errors in bone densitometry measurements are attributable in the most part to instrument errors and operator errors. The regular scanning of a phantom will enable the performance of an instrument to be monitored and thus for instrument errors to be quantified and/or reduced. Furthermore, phantom data may be used to correct BMD data, where necessary, at the end of a study. Operator errors may be reduced by scans being performed by a small number of highly trained operators and by the use of standardised protocols for the acquisition and analysis of scans.
References