

**Mihail A. Boyanov**

**Quality Assurance and Control  
in Dual-energy X-ray Absorptiometry**



**Sofia, 2013**  
**Central Medical Library**

Quality Assurance and Control in Dual-energy  
X-ray Absorptiometry

© Mihail A. Boyanov, 2013

First edition

ISBN 978-954-9318-27-2

Mailing address: Prof. Mihail Boyanov, MD, PhD, DSc

Clinic of Endocrinology and Metabolism,

Department Internal Medicine

Medical University

1 G. Sofiyski st., Sofia 1431, Bulgaria

Phone: +3592 9230 784, Fax: +3592 9230 779

E-mail: [mihailboyanov@yahoo.com](mailto:mihailboyanov@yahoo.com) or

[boyanov@alexandrovska-hospital.bg](mailto:boyanov@alexandrovska-hospital.bg)

## Foreword

The Dual Energy X-ray Absorptiometry (DXA) has become the gold standard in the diagnosis of osteoporosis and low bone mass. The wide spread use of DXA even in areas with previously difficult access to this technique precludes the need for proper education of the technologists and physicians performing and reading DXA studies. Quality assurance (QA) and control in each measurement methodology are essential as a basis for the credibility of the yielded results.

This book is focused on QA and control in DXA. It is conceived and written by one of the leading specialists in the field in our country, a recognized expert, scientist and educator. It contains original experiments and data reflecting his own experience. The work represents a synthesis of some basic science and statistics and more advanced concepts of QA in bone densitometry such as cross-calibration or use of more sophisticated techniques for precision calculations.

The level of scientific content is high. All the latest developments are presented and discussed. The text can easily be understood by both DXA technologists and physicians, as well as by all specialists reading DXA reports in their everyday clinical practice. It adds some new insights into the problems inherent to QA. Thus it could be of great help for all those dedicated to the field of bone densitometry and DXA.

The structure of the book is coherent and balanced. The tables are precise and informative. Well selected references are given at the end of the work.

I am sure that this excellent contemporary book will be well received and appreciated. I highly recommend it.

*Ljubomir Diankov*

*Professor of Radiology*



## CONTENTS

<b>Introduction and definitions</b> .....	7
<b>Instrument QA procedures</b> .....	12
<b>The Technologist QC</b> .....	24
<b>A few more practical problems</b> .....	30
<i>Manufacturer supplied versus own precision data</i> .....	30
<i>Long-term precision of DXA</i> .....	34
<i>Intra-observer versus inter-observer precision errors</i> .....	34
<i>Other factors affecting precision</i> .....	36
<i>Some more statistical considerations</i> .....	40
<b>Replacing or upgrading a bone densitometer.</b>	
<b>Cross-calibration of DXA devices from different manufacturers or different models of the same manufacturer</b> ....	42
<i>Upgrading central DXA equipment</i> .....	42
<i>Replacing central DXA equipment</i> .....	43
<i>QC and cross-calibration in the research setting</i> .....	49
<b>Conclusion</b> .....	50
<b>Acknowledgments</b> .....	50
<b>References</b> .....	51

## **Abbreviations**

BMC	– bone mineral content
BMD	– bone mineral density
CV	– coefficient of variation
DXA	– dual-energy X-ray absorptiometry
GE	– General Electric
ISCD	– International Society for Clinical Densitometry
LSC	– Least Significant Change
MTI	– Monitoring Time Interval
QA	– quality assurance
QC	– quality control
QCT	– quantitative computed tomography
sBMD	– standardized BMD
SD	– standard deviation

## Introduction and definitions

The concept of osteoporosis, as a quantifiable bone metabolic disease, relies on measurements of bone mineral density (BMD) [1]. In addition, BMD plays a crucial role in predicting fracture risk, as well as in monitoring antiresorptive treatment [2, 3]. Dual-energy X-ray absorptiometry (DXA) is the gold standard for BMD testing. It is of utmost importance that the measurements reflect the true value of BMD (i.e. the initial bone mineral content measurement should match the value of any repeated measurements). These two major requirements are translated into the concepts of *trueness* and *precision*. Together they contribute to the *accuracy* of the technique. Several years ago the term accuracy was used instead of trueness. Nowadays the term accuracy is used to depict the broader concept, which incorporates both trueness and precision.

The term **accuracy** was defined by the International Standard Organization (ISO 3534:3.11) as **the ability of a measurement to match the accepted reference value**. As such, accuracy can be affected by two types of error: systematic error (trueness error) and random error (precision error) [4-6].

**Trueness** (ISO 3534:3.12) is the **ability to measure a true value of BMC or BMD using densitometry**. When measuring bone mineral content (BMC), the measured value can be compared to the mineral content of ashed bone. Ideally, accuracy should be assessed on human cadavers but this is rarely done. However, if that is the case, the region of interest (ROI) is scanned; the pure bone in the ROI is then excised and incinerated in a special chamber. The remnants, consisting mainly of minerals, are weighed and compared to the BMC value measured by bone densitometry. This approach is very extensive

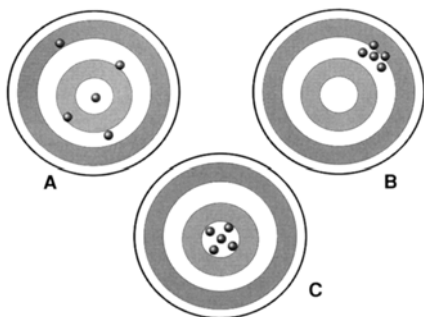
and may pose some problems. Another possible approach for BMC assessment is to measure equivalent material (e.g. hydroxyapatite with known calcium content) in artificial phantoms [7]. Testing trueness of the bone densitometry procedure is mainly the task of manufacturers. It is rarely done under clinical conditions. FDA 510K clearance requires an error of less than 10 % for all BMD devices. Most manufacturers in bone densitometry claim that the trueness of the bone mass measurement is lower than 5-6 % [4, 6, 7]. This is the so called systematic error inherent in the technique. The trueness affects the correct use of DXA for the diagnosis of low bone mass and osteoporosis.

**Precision** (ISO 3534:3.14) is the other attribute of quantitative measurement techniques. It refers **to the ability of the technique to reproduce the same numerical result** when the test is repeated in an identical fashion on the same object under the assumption that there is no biological change [6, 8]. It is also referred to as *reproducibility* or *repeatability* as it is based on serial measurement comparisons. Random errors result in scattered multiple measures. However, the average value should still be “true”. In other words, precision shows whether repeated measurements would yield approximately the same result irrespective of the true value (i.e. of the accuracy). Precision is important in bone densitometry monitoring (the so-called serial measurements). It reflects both the inherent precision error of the technique and the error introduced by incorrect or inconsistent patient positioning and scan analysis [6, 8].

The concepts of trueness and precision are very well illustrated if each measurement is regarded as a dart aiming to hit the same point, which represents the true value. Trueness reflects

how close the hits are to the true value (the central point), while precision reflects whether the hits are close to each other irrespective to their proximity to the central point (the true value). This concept is illustrated in Figure 1.

- A – good trueness, low precision
- B – low trueness, good precision
- C – good trueness and precision (the ideal measurement method)



**Figure 1.** The concept of trueness and precision is shown

The concept of precision was first developed in clinical chemistry. The idea of having some kind of quality control grew with the technical development of DXA [9-13]. It was practically introduced in bone densitometry due to the seminal work of C. Glüer et al. [14, 15]. In their work, Glüer et al. presented typical problems encountered during research studies and guidelines for their solutions. Detailed descriptions of the precision and trueness (accuracy) concepts can be found in a variety of sources [5, 6, 8, 16, 17]. The mathematical construct behind precision is explained in detail in a paper written by S. L. Bonnick et al. [8]. If an object is measured  $n$  times, the mean value of all measurements is regarded as the true value for each ob-

ject (true value = (measurement 1 + measurement 2 + .... + measurement  $n$ ) /  $n$ ), where  $n$  is the number of measurements. The differences between each measurement and the mean are then calculated (difference 1 = measurement 1 – mean value; difference 2 = measurement 2 – mean value; etc.). The differences are then squared, added and the total is divided by the number of measurements minus 1 ( $n-1$ ), where  $n-1$  represents the degrees of freedom. The square root is then taken. This value is the standard deviation (SD) for the set of  $n$  measurements. It is expressed in the units of measurement (e.g.,  $\text{g/cm}^2$  in the case of BMD) and represents the variability from the mean of all  $n$  measurements.

$$SD = \sqrt{\frac{\sum_{i=1}^n (X_i - X_{\text{mean}})^2}{n-1}} \quad (1)$$

where  $X_i$  is the value of the  $i$ -th consecutive measurement,  $X_{\text{mean}}$  is the mean value of all  $n$  measurements and  $n$  is the number of measurements.

In order to present the SD as a proportion of the mean, it is divided by the mean, multiplied by 100 and expressed as a percentage. This is known as the coefficient of variation (%CV):

$$\%CV = \frac{SD}{\text{Mean}} \times 100 \quad (2)$$

If  $m$  objects are scanned repeatedly, the average value, SD and %CV can be found for each set of repeated measurements (i.e. for each object). The precision value for all measurements can then be expressed as the root-mean-square SD (RMS-SD) or the root-mean-square CV (RMS-CV) of the calculated SD

and CV for each object. They are then calculated for the entire object group by squaring SD and CV for each patient, adding these values, dividing them by the number of patients and then taking the square root:

$$SD_{RMS} = \sqrt{\frac{\sum_{i=1}^n SD_i^2}{n}} \quad (3)$$

$$CV_{RMS} = \sqrt{\frac{\sum_{i=1}^n CV_i^2}{n}} \quad (4)$$

where  $n$  is the number of objects.

When the CV is presented as a percentage of the mean (%CV) its value depends on the mean value. In the case of BMD the average would differ substantially in measurements obtained from subjects with normal versus low BMD. One would expect that the higher the average BMD, the lower the %CV (where the CV is a proportion of the mean) [18]. In order to avoid this source of bias, absolute CV values (e.g. grams per square centimeter) are preferred. The RMS-SD or RMS-CV are respectively the appropriate expressions of the so-called short-term precision for bone densitometry facilities.

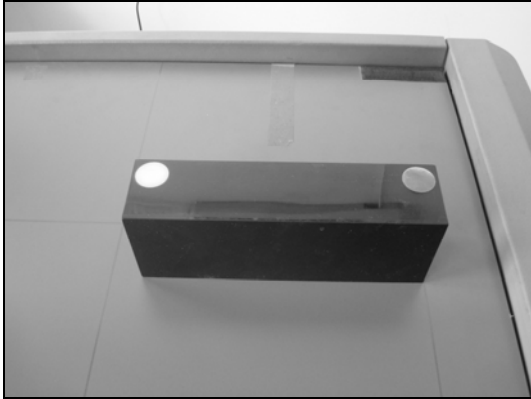
A precision study can be performed on a specific phantom (*in vitro* precision study) or on real patients (*in vivo* precision study). The first type of study tests for the reproducibility of the measurements and the precision error, which are generated solely by a machine. This error is due to the calibration and stability of the X-ray source, the detectors and the software. An *in vitro* precision study performed on a phantom is mandatory part of the

Instrument Quality Assurance and Control (Instrument QA, QC) [19, 20]. Precision testing on real patients additionally depends on proper patient positioning, scanning and data analysis. It represents more clearly the real clinical situation and is part of the Technologist's Quality Control.

### **Instrument QA procedures**

QA procedures need to be standard procedures in DXA facilities, in order to have a control for operator and machine variability [19-23]. **Instrument QA means, that the machine's performance is monitored over time.** One must keep in mind that DXA devices are X-ray based and must meet all radiation safety requirements. Periodical monitoring of the X-ray source, the scattered radiation and the patient and technologist exposure to radiation should be done. Radiation exposure and safety procedures in DXA facilities are regulated through established international and local standards.

Instrument QA procedures are performed according to the manufacturer's recommendations and include instrument calibration as well as testing of some software features. The *instrument calibration* is done differently by different manufacturers. GE and Norland bone densitometers require an external calibration scan with a special manufacturer-provided phantom (calibration standard) to be performed each day (Figure 2). This calibration standard tests the mechanical operation and calibration of the DXA machine. Hologic machines perform a continuous internal calibration with each pixel measurement by automatically using a rotating internal calibration wheel or drum.

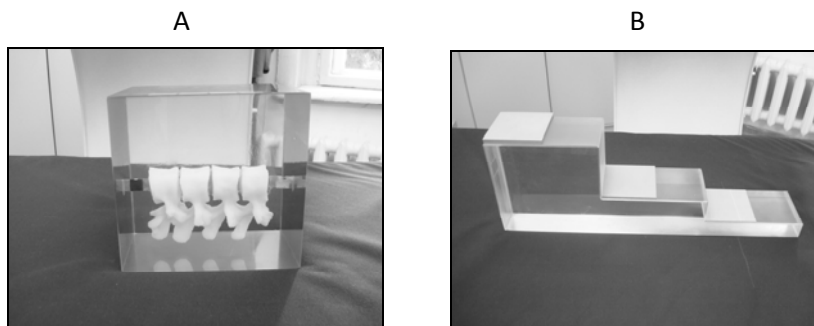


**Figure 2.** GE Lunar external calibration standard

A second type of phantom is used to: (1) test the machine performance over time by detecting changes in BMD values, and (2) monitor short-term and long-term precision *in vitro*. Those phantoms with known BMC, area and BMD usually mimic a region of the human skeleton. Some of them resemble lumbar vertebrae or the hip and are termed *anthropomorphic phantoms*. Their aim is to test machine function within a range of bone densities (replicating normal, osteopenic and osteoporotic bone) as well within a range of soft tissue densities. The phantoms test edge detection (the ability to distinguish bone from soft tissue) and allow for the assessment of bones that have a non-uniform density throughout the tissue length. Bone is imitated by aluminium or hydroxyapatite of different densities, while soft tissue is simulated by embedded “bone equivalents” into epoxide, resin or other materials.

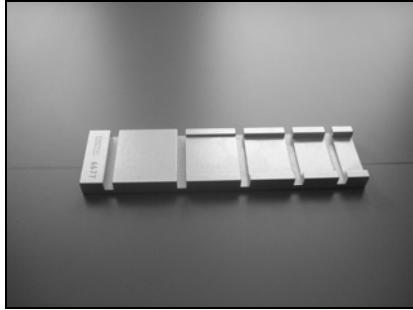
Scanning of an *anthropomorphic phantom* is required for all systems. The Hologic anthropomorphic phantom (see Figure 3 A) consists of four anatomically correct vertebrae with similar den-

sities and areas, consisting of epoxy-resin to simulate 60 % fat. This phantom is often used with devices from other manufacturers. The lack of a BMD value range, however, makes it less suitable for cross-calibration purposes than the newly developed European Spine Phantom and Bona Fide Phantom (see below). Hologic has developed an anthropomorphic hip phantom as well, but it is considerably less used in clinical medicine and remains reserved solely for research purposes. The use of an original step phantom for the calibration of whole body measurements in Hologic machines is also possible, where the phantom should be scanned at least once weekly (see Figure 3 B).



**Figure 3.** Hologic calibration phantoms. (A) Hologic anthropomorphic spine phantom. (B) Whole body scanning calibration phantom

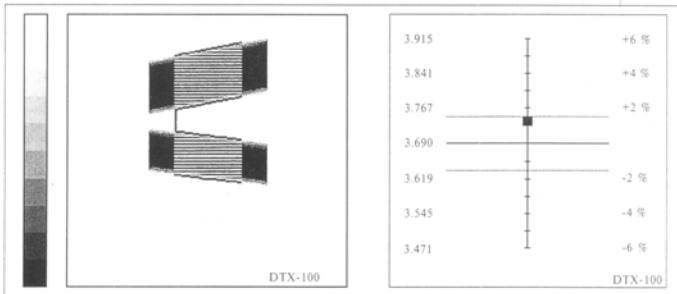
The Lunar Spine Phantom represents a rectangular aluminum bar mimicking all vertebrae at the level of L1-L4, as well as parts of the Th12 and L5 vertebrae (see Figure 4). Each vertebra has a different area and density. To simulate soft tissue, the phantom is submerged in water or it is embedded in an epoxy-resin block, which is the case in the newly developed version of the phantom.



**Figure 4.** The Lunar spine phantom

All Norland scanners are calibrated against a primary AP spine phantom, composed of hydroxyapatite equivalent materials, and a primary body composition phantom.

Peripheral DXA devices are calibrated using manufacturer-provided phantoms. The DTX-100, for instance, uses an aluminum phantom mimicking two rectangles for the radius and ulna (Figure 5).



**Figure 5.** Phantom scan on the DTX-100 forearm bone densitometer. Note that the phantom value is at the upper limit of the 1.5 % range

The European Spine Phantom was developed in an attempt to produce the “perfect” phantom. It is a semi-anthropomorphic phantom with three levels of BMD values intended for the

standardization of any central bone densitometer. It is rather expensive and is used primarily for research purposes [24].

The Bone Fide Spine Phantom is a calcium hydroxyapatite step phantom embedded within an acrylic block. The acrylic mimics soft tissue, while the step wedges represent a range of BMD values similar to those measured in patients.

The forearm presents more difficulties for standardization as different devices measure different ROIs: ultradistal, distal, middistal, ¼ site and 33% proximal site. A European semi-anthropomorphic phantom for the cross-calibration of peripheral bone densitometers was developed [25]. A universal standardization approach was also tested but the results were not encouraging [26, 27].

In short, QC phantoms test both machine parameters and software features of edge detection and analysis. The purpose of this type of QA scanning is to detect any machine calibration failures or drifts occurring over a period of time, as well as changes resulting from maintenance and repair of the DXA machine. The phantom should be re-scanned multiple times (10-20 times) after the instrument installation to define the mean values of the measured BMD, BMC and area. These mean values are further used when monitoring *in vitro* precision (based on daily phantom measurements). The anthropometric phantom has to be scanned on each working day and the phantom BMD, BMC, and area should be plotted on graphs based on the, so-called, Shewhart control charts. Shewhart is the name of the scientist who developed the QC procedures in analytical chemistry [28].

There are different ways to monitor machine stability over time using QC charts. One simple way to do this is by visual inspection and the application of the, so-called, *Shewhart*

*rules*. They give simple recommendations about the kind of BMD, BMC or area changes that should be closely monitored and, which would require instrument re-calibration or servicing. Firstly, it is necessary to establish a baseline value and control limits. The baseline value represents the mean of 10 consecutive phantom measurements on the same day or the mean of 15 to 25 measurements on consecutive days. The standard deviation (SD) for the set of scans is then calculated. Control limits are expressed either as a percentage of the mean or in terms of SD. Daily scan values are then plotted on a graph depicting the baseline value and the control limits. If any of the Shewhart rules are violated, a warning message is issued requiring specific actions (for details see Table 1).

<b>The phantom measurement is falling outside range</b>	<b>What should be done</b>
1 phantom value exceeds the average by $\pm 1\%$	Visual inspection of the QC chart
1 phantom value exceeds the average by $\pm 1.5\%$	Instrument re-calibration should be performed*
2 consecutive phantom values exceed the average by $\pm 1\%$ range	Instrument re-calibration should be performed
A difference between 2 consecutive measurements $> 2\%$ (and if both are outside the $\pm 1\%$ range)	Instrument re-calibration should be performed
4 consecutive phantom values on one side and $> +0.5\%$ or $\leq 0.5\%$	Instrument re-calibration should be performed
10 consecutive measurements fall on one side of the average	Instrument re-calibration should be performed
* The instrument re-calibration requires the obtainment of 5 consecutive phantom measurements with repositioning. If the mean value differs by more than 1% from the manufacturer-provided value (usually registered on the phantom), the user will be prompted to maintenance service the instrument!	

**Table 1.** Shewhart rules in QC. The left column shows possible aberrations, while the right column shows the required actions [6]

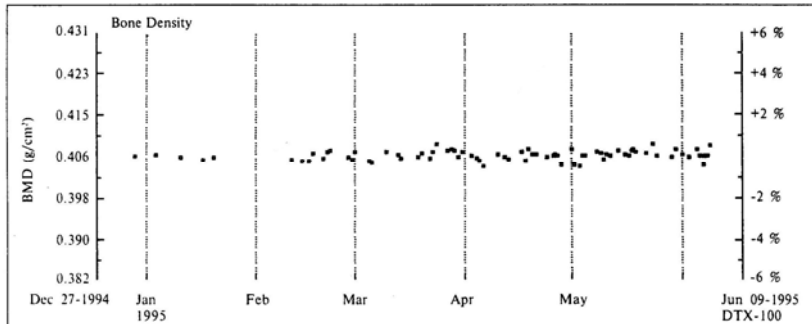
Usually, 1 SD is close to 0.5 % of the average phantom value. Therefore 1 SD = 0.5 %, 2SD = 1 %, 3 SD = 1.5 %, 4 SD = 2 %. The Shewhart rules are, therefore, known as the 3 SD or 1.5 % rule, the 2 SD or 1 % twice rule, the range of 4 SD or range of 2 % rule, the four  $\pm$  1 SD or four  $\pm$  0.5 % rule, and the mean  $\times$  10 rule. The violation of any of these rules must be regarded as a warning for machine malfunction.

However, the Shewhart rules tend to produce a high false alarm rate. This is due to the fact that even DXA machines in perfect condition may, from time to time, violate the Shewhart rules. In the work by Lu et al. this was the case with every 39<sup>th</sup> scan [29]. A, so-called, filter may be used to reduce the occurrence of false alarms [29]. This filter requires that the mean of 10 phantom scans is calculated after a violation of the Shewhart rules has occurred. The violation is confirmed if the difference from the baseline average value exceeds 1 SD. Alternatively, one can set the 3 SD rule as the crucial rule, whose violation requires the immediate application of the other rules.

The Shewhart rules are easy to use and can detect both sudden changes in machine calibration (i.e. a shift), as well as a slowly growing bias in machine performance (i.e. a drift). The shift might be due to recent maintenance, changes in the X-ray tube or detectors, or relocations (and incorrect machine calibration), as well as to sudden machine malfunction. A drift (slow changes) might be due to an aging X-ray tube or detectors, to change in the power supply or to slowly changing room conditions (e.g. temperature, humidity).

Figures 6 and 7 show the QC charts of the DTX-100 forearm bone densitometer (see Figure 6) and the Hologic QDR 4500

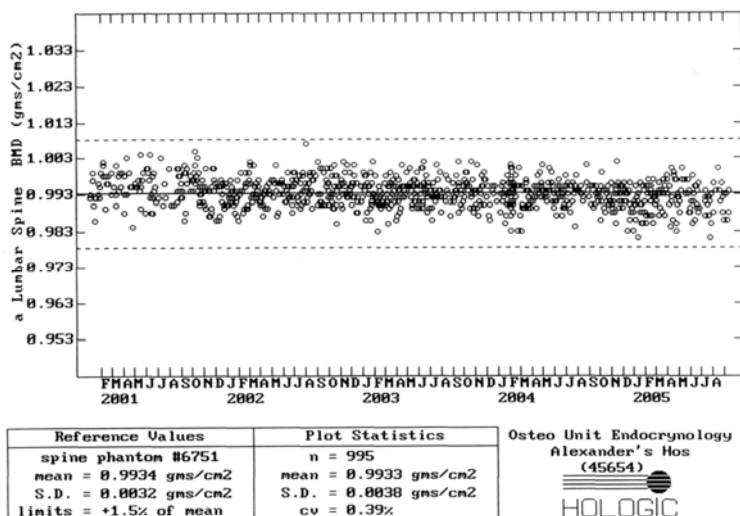
A device (see Figure 7) at the Bone Metabolic Unit. Figure 8 illustrates the same DTX-100 device (as in Figure 6) during a prolonged malfunction period (servicing had been postponed).



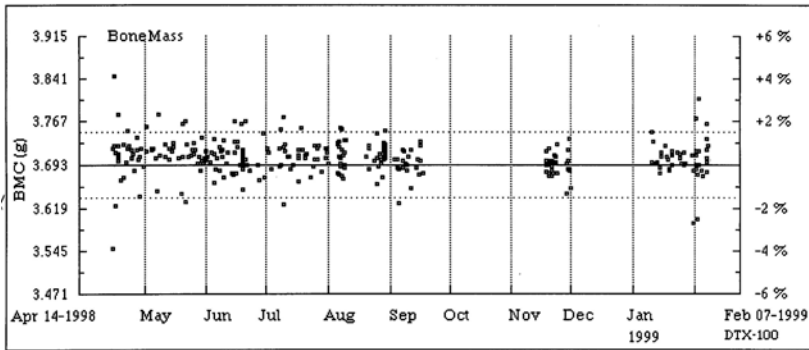
**Figure 6.** Phantom statistics (BMD chart) of a forearm single X-ray absorptiometry device, the DTX-100 (Osteometer Meditech, USA). All BMD values are within the  $\pm 1.5\%$  values. Proper machine functioning is observed.

Another way to monitor QC is by using *CUSUM charts* [30]. They are more sophisticated and suitable for professional purposes. The CUSUM charts reflect the random variation in phantom measurements. The magnitude and direction of the variation should remain relatively constant compared to the average mean. The principle underlying the CUSUM charts is in that the differences between an actual phantom measurement and the baseline average value (from the 10 consecutive scans on the same day or the 15-25 measurements on different days) should be randomly scattered around the zero line. This means that the cumulative sum of all of the differences should be close to zero if there is no bias introduced. To build the CUSUM graph, each sequential value of the cumulative

sum is plotted against time. If the machine is functioning properly, the values should be scattered in a horizontal pattern around the 0 value. Otherwise, there will be a rising or falling pattern. Although the CUSUM charts can be inspected visually, there are other ways to detect subtle changes. One of them is by determining the slopes of the arms on the V-mask. Alternatively, tabular CUSUM can be used, in which upper and lower control limits are introduced. The CUSUM charts are less intuitive than the Shewhart graph and require the use of a software program. The main advantage of the CUSUM charts is in their higher sensitivity. Good examples for CUSUM charts can be found in ref. 6.



**Figure 7.** Phantom statistics (BMD chart) of the Hologic QDR 4500 A device (Hologic Inc., Bedford, USA). The machine is functioning within the pre-specified range of BMD



**Figure 8.** Phantom statistics of the DTX-100 forearm bone densitometer during a period of prolonged malfunction. BMC values are scattered around the mean and some of them exceed the acceptable variation range.

The visual examination of QC charts has been compared to process-control charts (CUSUM charts) [31]. The regular Shewhart chart, with its associated sensitizing rules, has a high false alarm rate. Introducing an additional filter reduces the false alarm rate and overall sensitivity. The CUSUM method offers the best sensitivity/specificity ratio and can estimate when (i.e.the date) a change in DXA scanner performance occurred [31]. In another study, the ROC analysis using the Shewhart multi-rule charts was shown to be as effective, if not better, than the one done using the CUSUM chart. However, the stringency of the rules must be tightened to obtain an optimal performance [32, 33].

An interesting approach is to use, not only the baseline phantom value (from multiple scans), but to also gradually incorporate the phantom values during periods of proper machine functioning. This is called “the moving average” [17].

A detailed description of how *in vitro* QC charts are to be handled can be found in the Canadian ISCD standards for technologists [34]. Hologic software supplies these charts automatically, Norland-Cooper Surgical maintains BMD charts only, while GE Lunar operators need to create and maintain these charts manually. Practically all manufacturers provide sophisticated software programs for calculating phantom value variations and displaying drifts in machine performance.

A phantom log of all phantom scans should be maintained. Service logs should also be maintained (including printouts of error messages), as should all service, government, and radiation survey reports. After any preventive maintenance or service, the phantom should be scanned 10 times without moving the phantom. If the mean BMD exceeds the mean of the daily phantom scans that were performed during the 10 days prior to maintenance or servicing, the machine should be recalibrated and a new mean should be established based on new 10 scan set of the phantom. In this scenario, a 1% change in the readings is considered within normal limits following a software change.

A practical question, one begs to ask is who can ensure that the QC at a particular bone densitometry center is performed correctly. Large imaging facilities, such as BioImaging or Synarc, require source information from participating DXA units and review actual QC data. Centralized control is usually performed in multicenter clinical trials [24, 35-37]. Alternatively, one can have centralized control for all bone densitometry units in a particular country – an approach that might be im-

plemented in the future. It is very difficult to decide who should perform this centralized control in this case. The QC procedures could also be made mandatory through the National Standards/Guidelines for Clinical Densitometry in a given country. In this case, all health care funding bodies will have the right to check for the implementation of the QC procedures before any payment for bone densitometry occurs.

The concept of centralized quality control of bone densitometry was applied in a national health survey (NHANES III) using three mobile examination centers [37]. A small percentage (3.5%) of all scans was rejected, but 33% of the remaining required reanalysis at the QC center. Precision in spine and hip phantom was below 1 %. *In vivo* precision for the femur neck BMD was 3.2% (CV) and was found to be 5.1 % for Ward's triangle BMD [34]. A few similar trials can be found in the literature [35, 36]. In Bulgaria, a National Osteoporosis Program 2006-2010 was implemented. During this program, an epidemiology study on the prevalence of osteoporosis using the DXA technology was performed [38]. Six DXA centers were subjected to a centralized review of QA and QC procedures and the results were published elsewhere [38]. The QC study was done on 5 pencil beam DXA densitometers – one Hologic QDR 1000, one Hologic QDR 1500, two Lunar DPX-IQ and one DPX-L, as well as one fan-beam Hologic QDR 4500 A. Table 2 shows *in vitro* precision errors of the machines that occurred while scanning the Hologic anthropomorphic phantom 20 times consecutively.

DXA Unit	Nr 1	Nr 2	Nr 3	Nr 4	Nr 5	Nr 6
SD (g/cm <sup>2</sup> )	0.0034	0.0039	0.004	0.007	0.011	0.004
%CV (%)	0.334	0.307	0.377	0.524	0.848	0.428

**Table 2.** Short-term *in vitro* precision errors of the six DXA centers participating in the National Osteoporosis Program in Bulgaria 2006-2010 [39]

The data from Table 2 show that DXA machine number 5 differed significantly from all others. The data from the *in vivo* precision study are presented in the text below (see also Table 7). This example displays the importance of reviewing QC logs and charts at a particular DXA center as part of site certification and accreditation.

## The Technologist QC

At least 30 degrees of freedom are needed for a precision study to be valid [40, 41]. Thirty degrees of freedom ensure that the upper limit for the 95% confidence interval (CI) of the calculated precision value does not exceed 34%. Table 3 represents the possible combinations of objects (i.e. patients) and scans per patient to achieve 30 degrees of freedom.

No. of patients (objects)	No. of scans per patient
1	31
10	4
15	3
30	2

**Table 3.** Number of patients and scans per patient ensuring at least 30 degrees of freedom in a short-term precision study

In clinical practice, 30 patients are scanned twice over 2 weeks to 1 month. This is called a **short-term precision study**. In **long-term precision studies** the objects are followed over an extended period of time (> 1 month). Long-term precision studies are easily performed with phantoms as their mineral content does not change over time. This type of precision error is called long-term precision *in vitro*. It reflects the calibration and proper functioning of the bone densitometer. In the case of precision studies on real patients, a major confounding factor is the BMD change over time. This logistic difficulty requires the use of linear regression to account for this biological change and the calculation of *standard error of the estimate* (SEE) in order to accurately express precision. The long-term precision errors are expected to exceed the short-term precision estimates. In clinical practice, *in vivo* short-term precision studies are usually performed, where RMS-SD and RMS-CV (in absolute values or %) are calculated.

If the *in vivo* precision error at a particular site is known, the magnitude of BMD change measured indicating real biological change at that site can be determined. This is called the **least significant change** (LSC). One has to choose which level of statistical confidence will be used for calculating LSC. According to the level of statistical confidence, a specific coefficient, as shown in Table 4, should be used to calculate the LSC as follows:

$$LSC = \sqrt{2} \times \text{coefficient} \times \text{precision error} \quad (5),$$

where the coefficient can be found in Table 4 according to the desired level of statistical significance and precision is expressed as either RMS-CV or RMS-SD.

Level of confidence	Coefficient value
99	2.58
95	1.96
90	1.65
80	1.28

**Table 4.** Coefficients used in the calculation of LSC according to the desired level of statistical significance

Therefore, if a 95% significance is required the  $LSC^{95} = 1.44 \times 1.96 \times \text{precision} = 2.77 \times \text{precision}$  (6). For 80% confidence, the  $LSC^{80} = 1.44 \times 1.28 \times \text{precision} = 1.84 \times \text{precision}$  (7). Usually the 95% confidence level is used and the precision error has to be multiplied by 2.77 to obtain the LSC.

The International Society for Clinical Densitometry has provided a useful tool for the calculation of %CV and the LSC at [www.iscd.org](http://www.iscd.org). The user must insert the absolute values of consecutive measurements (30 patients measured twice or 15 patients measured trice) and the values for %CV, SD, RMS-SD, RMS-CV and LSC are then calculated automatically.

In 2005, the International Society for Clinical Densitometry (ISCD) issued standards for the minimum acceptable precision for an individual technologist [2, 42]:

- Lumbar spine: 1.9 % (LSC = 5.3 %)
- Total hip: 1.8 % (LSC = 5.0 %)
- Femoral neck: 2.5 % (LSC = 6.9 %)

If a technologist's precision is worse than these values, re-training is required [2, 41].

In order to detect real changes in BMD, a serial scan should be repeated when the biological change is expected to exceed LSC. The **time interval needed** depends on both the expected rate of change per year and the LSC:

$$\text{Time interval} = \text{LSC} / \text{expected rate of change per year} \quad (6)$$

The greater the LSC and the smallest the annual change, the longer the time interval should be. A simple approach in defining the interrelationship between precision error and anticipated rate of change, in establishing the time interval for serial measurements needed to exceed the LSC, is given in Table 5.

Precision as %CV	Change per year (%)	Time interval (yr)
1.0	1	2.77
	3	0.92
	5	0.55
1.5	1	4.16
	3	1.39
	5	0.83
2.0	1	5.54
	3	1.85
	5	1.11

**Table 5.** Interrelationship between precision and anticipated rate of change in determining the time interval needed to exceed LSC [8]

For example, the change in BMD is rapid in corticosteroid-induced osteoporosis, and repeated measurements can be performed within 9-12 months. In the usual osteoporosis patient treated with antiresorptive drugs, the biological changes in BMD could be detected after at least 2-3 years. However, it is safer to obtain the first BMD measurement early after ther-

apy initiation as non-responders to therapy can be identified earlier. Typically, the first repeat scan might be performed a year after treatment has started. If BMD stability (i.e. lack of significant change) or improvement are noted, the next scans should be performed at longer time intervals (2-3 years) [3]. Repeated measurements should be performed on the same device and, preferably, by the same operator as outlined in the cross-calibration section [3, 42].

A White Paper on precision assessment and radiation safety for DXA was published [22]. A number of important questions were exhausted: the definition and importance of accuracy and precision, methods of DXA precision assessment, common factors affecting precision and others. According to this paper “clinical DXA precision is influenced by a combination of short- and long-term variability of the scanner, patient motion, body habitus, and operator dependent factors such as patient positioning and scan analysis. Patient and operator related sources of variability are more important than the scanner variability itself with operator related factors having the most influence on the overall precision of DXA measurements.” The White Paper strongly recommended “that each technologist and bone density facility conduct a precision assessment. Technologists should perform a precision assessment after they have scanned at least 100 patients. Precision assessment does not need to be repeated as long as there is no reason to believe that there has been a change in the technologist’s level of competence or in the machine software and hardware” [3, 21]. Using manufacturer-provided precision data was not endorsed. The question of defining a mean precision error for a DXA center with

multiple technologists was solved with the averaging of their individual precision errors. This is applicable if the confidence intervals of the precision errors for each technologist overlap, indicating that the small inter-technologist differences might not represent true differences [3]. As the White Paper also reviewed Radiation Safety, a short paragraph indicated that for the need of precision studies no special consent form would be required from the patient, but he/she should be informed about the purpose of the duplicate scans and given the opportunity to refuse participation.

Standards and guidelines for technologists performing central DXA were published in 2007 as a part of the ISCD Canadian standards [34]. A number of important practical points in DXA precision assessment were discussed:

1. Precision studies in a particular center should be done on patients who are typical for that center. Translated into clinical terms, this means that all kinds of subjects – with normal, low or osteoporotic BMD should be included in the precision study. The influence of the absolute BMD value on the precision value is discussed further in this chapter.
2. Same-day precision is generally better than different day precision, with intra-observer variability being lower than inter-observer variability. To reflect clinical reality in the best way, a repeat scan a few days after the first scan can be advocated, while even having the two scans performed by different technologists.
3. A monitoring time interval (MTI) is defined as the time required to achieve the LSC given an assumed rate of BMD change for that patient and the particular skeletal site.

The ISCD Canadian standards also introduced the concept of “quality assurance procedures”. Quality assurance (QA) controls for both operator and machine variability, thus consisting of two distinct parts: Instrument Quality Control and Technologist Quality Control. Instrument QA is achieved by monitoring the machine’s performance over time (as described below), while the Technologist QA displays the reproducible positioning, performing and analysis of patient scans (as described above) [34].

### **A few more practical problems**

The concept of precision is very important in real life as it is the foundation of DXA credibility as a precise technology. It can also be applied to various DXA applications such as lateral spine, morphometric X-ray absorptiometry, hip structural analysis or body composition analysis [43-48].

Some real life developments and research in the field of QC and QA in central DXA are described below.

#### *Manufacturer supplied versus own precision data*

The ISCD does not recommend using the manufacturer supplied precision data [42]. Instead, each technologist must establish his own set of precision data. The conditions reproduced by the manufacturer could be regarded as “ideal”, while the conditions in real life are far from ideal. Table 6 summarizes the manufacturer supplied precision data derived from ref. 49.

Manufacturer	DXA model	%CV BMD AP spine	%CV BMD total hip	%CV BMD whole body	In vitro QC procedures
Hologic, Inc., Bedford, MA, USA	QDR 4500® A	< 1.0 %	< 1.0 %	< 1.0 %	Self-calibrating with Hologic Automatic Internal Reference system and automated quality control program
	QDR 4500® C	< 1.0 %	< 1.0 %		
	QDR 4500® SL	< 1.0 %	< 1.0 %	< 1.0 %	
	QDR 4500® W	< 1.0 %	< 1.0 %	< 1.0 %	
	Delphi™	< 1.0 %	< 1.0 %	< 1.0 %	
	Discovery™	< 1.0 %	< 1.0 %	< 1.0 %	
	Explorer™	< 1.0 %	< 1.0 %		
	DPX Bravo®	< 1.0 %	< 1.0 %		
	DPX Duo®	< 1.0 %	< 1.0 %		
	DPX-IQ™	0.5 %	1.0 %	0.5 %	
GE Healthcare Madison, WI, USA	DPX MD™	1.0 %	0.7 %	0.5 %	Block phantom and aluminum spine phantom
	DPX-NT™	1.0 %	1.0 %	1.0 %	Automatic test program
	Prodigy™	1.0 %	< 1.0 %	< 1.0 %	Automated QA program
	Expert®-XL	< 1.0 %	< 1.0 %	< 1.0 %	Automatic test program with QA trending
					Internal hydroxyapatite; Automated QA program with spine phantom
	iDXA™	1.1 %	0.7 %	0.5 %	Automated 6-point calibration and quality assurance with QA trending
	Excell™	1.0 %	1.2 %		Automated with 77-step calibration standard and QC phantom
	Excell™plus	1.0 %	1.2 %		Automatic with supplied calibration standard and QC phantom
	XR-46™	1.0 %	1.2 %	1.0 %	Automatic with supplied calibration standard and QC phantom
	XR-600™	1.0 %	1.4 %		Automatic with supplied calibration standard and QC phantom; automated Shewhart Chart analysis
	XR-800™	1.0 %	1.4 %	0.8 %	

**Table 6.** Precision data for the FDA-approved densitometry devices [49]

The precision data from the literature are so abundant that only a few illustrative sets will be cited below. A study tested 3 GE Healthcare Prodigy and 3 Hologic Delphi devices and found that *in vivo* averaged %CV for the AP spine, total hip and femoral neck to be 1.0 %, 0.9 % and 1.5 % respectively for the Prodigy and 1.2 %, 1.3 % and 1.9 % respectively for the Hologic devices [50]. This study suggested that short-term BMD precision errors were skeletal-site and manufacturer specific.

The introduction of fan-beam technology required precision studies. In such a study, the precision error of fan-beam DXA of the spine, hip and forearm was calculated [51]. The CV at the lumbar spine was 1.1 %, at the total hip – 1.2 %, at the total forearm – 0.5 % [51]. The authors concluded that fan-beam DXA of central sites and the forearm were as precise as scans performed in slower scanning modes. When using the Compare function, they found a significant decrease in precision at the forearm and no improvement at the spine or hip [51]. However, the use of the Compare function whenever possible is advocated.

With the introduction of new DXA devices into clinical practice, precision studies are usually performed. Very often they are compared to older models from the same manufacturer, as was the case with GE Lunar Prodigy and iDXA [52]. A new pencil-beam bone densitometer (the DMS Stratos) was assessed and the RMS-CV values were found to be 1.22%, 1.38%, 2.11% and 0.86% for the lumbar spine (L1-L4), lumbar spine (L2-L4), femoral neck and total hip respectively [53]. Another research group tested the first cone-beam DXA system (the DMS Lexxos) and found a precision of 1.3 % for total hip, 2.0 % for femoral neck and 2.3 % for spine BMD [54].

In our own precision study, we tested the short-term *in vivo* precision of our Hologic QDR 4500 A bone densitometer. The results were published elsewhere [55] and were within the pre-specified ISCD limits (see Table 7).

	Precision %CV	Least significant change (LSC) – absolute value	Least significant change (LSC) – in %
BMD L <sub>1</sub> -L <sub>4</sub> (g/cm <sup>2</sup> )	1.35 %	0.026 g/cm <sup>2</sup>	3.57 %
Total hip BMD (g/cm <sup>2</sup> )	0.95 %	0.019 g/cm <sup>2</sup>	2.52 %
Femoral neck BMD (g/cm <sup>2</sup> )	1.29 %	0.023 g/cm <sup>2</sup>	3.71 %
Trochanter BMD (g/cm <sup>2</sup> )	1.31 %	0.019 g/cm <sup>2</sup>	3.59 %
Ward's BMD (g/cm <sup>2</sup> )	3.14 %	0.036 g/cm <sup>2</sup>	8.76 %

**Table 7.** Short-term *in vivo* precision of the Hologic QDR 4500 A at the Bone Metabolic Unit of the University Hospital Alexandrovska, Sofia [55]

During the implementation of the National Osteoporosis Program 2006-2010 in Bulgaria, six DXA centers were subjected to a centralized review of their precision data (see Table 8). One of them showed severe deviations from the ISCD recommendations. The data generated by this center were not included in the final analysis.

DXA Center	Nr 1	Nr 2	Nr 3	Nr 4	Nr 5	Nr 6
%CV	0.604	0.982	1.880	3.589	0.440	1.622
LSC (%)	1.670	2.720	5.210	9.940	1.220	4.493
LSC-I <sup>a</sup> (g/cm <sup>2</sup> )	0.017	0.027	0.052	0.094	0.012	0.045
LSC-II <sup>b</sup> (g/cm <sup>2</sup> )	0.015	0.024	0.047	0.096	0.013	0.043

a calculated manually

b calculated using the ISCD calculation tool

**Table 8.** The *in vivo* %CVs for the lumbar spine measurements of six DXA centers, which participated in the epidemiology study of the National Osteoporosis Program, are shown [39].

### *Long-term precision of DXA*

The long-term precision of DXA has been tested in a few studies. It was assessed over 7 years in forty postmenopausal women [56]. Using trimmed data (untrimmed data in brackets) the calculated %CV was found to be 1.12 % (1.65 %) for the AP spine, 1.32 % (1.57 %) for the total hip, and 2.21 % (2.48 %) for the femoral neck. The long-term precision of central DXA was studied in 64 postmenopausal women over a period of 10 years [57]. The long term precision errors expressed in SD or %CV were 0.018 g/cm<sup>2</sup> (1.9 %) for the spine, 0.016 g/cm<sup>2</sup> (1.7 %) for the total hip and 0.017 g/cm<sup>2</sup> (2.3 %) for the femoral neck BMD. In another study, using two Hologic DXA scanners, the long-term precision %CV was found to be 2.4 % for lumbar spine, 2.3 % for total hip and 2.7 % for femoral neck [58]. These values were almost twice as high when compared to those obtained from the short-term precision data: 1.3 % (AP spine), 1.2 % (total hip) and 1.4 % (femoral neck). The expected long-term *in vivo* precision was compared to expected annual bone loss rates in men and women in a subset from the Malmö OPRA-study, where several years were needed to detect changes (e.g. LSC/median range of BMD change) [59].

### *Intra-observer versus inter-observer precision errors*

The introduction of a second technologist is usually expected to increase the precision error as both technologists may position patients and analyze the scans in a different manner.

The *in vivo* intrarater and interrater precision of measuring BMD using supine lateral DXA was assessed in two cohorts – a younger (with high BMD) and an older one (with presumably

lower BMD) [43]. Intrarater precision was better in the younger cohort – it ranged in %CV from 0.50-3.68 % for the young cohort, and 1.46 -5.30 % for the older cohort. Interrater precision was lower with %CV of 1.11-2.35 % for the younger cohort and 1.85 – 4.32 % for the older cohort.

In our own precision study, we tested the *in vivo* precision of a single X-ray forearm bone densitometer – the DTX-100 device (Osteometer Meditech, USA) and compared it to previous publications [60, 61]. On this specific device the distal region of interest, ROI, begins at the 8 mm separation point between radius and ulna and then continues proximally for a distance of 24 mm. It contains around 60-70 % cortical bone and has a more cylindrical and, thus, reproducible shape. The ultradistal ROI extends from the radial endplate proximally to the 8 mm point and is built primarily by trabecular bone (60-70 %). The radial endplate shows much more anatomical variations and, thus, results in less precise measurements. Over the duration of the study, the short-term *in vitro* precision %CV was 0.58 %. The short-term *in vivo* precision %CVs for a single or two different technologists, as well as the long-term *in vivo* precision %CVs (same patients re-scanned 3 months later) are shown in Table 9.

	Distal site	Ultradistal site
intra-observer short-term %CV	1.65%	2.03%
inter-observer short-term %CV	1.86%	2.54%
intra-observer long-term %CV	1.92 %	2.13 %

**Table 9.** The *in vivo* intra-observer and inter-observer precision %CV on the DTX-100 forearm single X-ray device [60, 61]

Our results demonstrate that the magnitude of the precision error varies with the anatomical site, as well as with the introduction of a second technologist. That is why the ISCD recommends the use of an average value for all technologists performing DXA at a particular DXA site [3].

#### *Other factors affecting precision*

A number of non-modifiable factors are affecting precision: type and model of the equipment used, site of measurement, software version, range of BMD values, soft tissue thickness (especially in body composition studies) and others. Among the modifiable factors are the proper training of technologists and good patient compliance [62].

W. Leslie investigated the *factors affecting short-term BMD precision* assessment during the Manitoba Bone Density Program [63]. There were no significant differences in precision errors when scan-pairs were acquired by a single technologist or two technologists. Scanner mode (pencil-beam versus fan-beam), age, height and bone area also did not affect the precision error. Weight, BMI and T-scores showed inconsistent differences. The single factor affecting precision significantly was acquiring the scan-pairs on different days. The author concluded that the most widely used procedure for performing a BMD precision assessment (same technologist on the same day) might systematically underestimate precision error and would lead to over-categorization of change in many patients [63]. A similar finding was published in a study where a subgroup of patients were rescanned 3 to 10 days later, and a second study where patients were rescanned 188 days later on average [64]. The short-term precision for the left femoral

neck with immediate rescanning was 0.007 g/cm<sup>2</sup>, 0.017 g/cm<sup>2</sup> with rescanning within days and 0.024 g/cm<sup>2</sup> with rescanning within months. This study underlined the fact that short-term precision of immediate replicate scans might underestimate the error over time [64].

*Different measurement sites yield different precision errors.* The AP spine and the total hip usually show the best precision: AP spine ≈ Total hip > Femoral neck > Wards. The expected biological changes at the spine (mostly trabecular bone) are, however, greater than those at the total hip (mixed bone) and that is why it is the primary recommended site for serial bone density measurements [2, 3]. The total hip (either single or dual femur) might also be used for monitoring. No other site should be used. The forearm site has excellent precision but is rather unresponsive to treatment. The lateral spine is less precise than the AP spine, hip and total body DXA. This hypothesis was proven in a study by GM Blake et al. [65]. The results were normalized for the ratio of treatment effect/precision. The authors concluded that, although the treatment effect was larger for lateral rather than for PA spine BMD, this advantage was offset by the greater precision errors [65].

Precision data are affected by the *range of BMD values if they are expressed in percent of the mean*. This is very well illustrated in a number of studies [18, 66]. One of these studies highlighted that calculated CV shows a progressive fall in value as BMD rises. Therefore, the SD should be used to calculate significant BMD changes [66]. Another study investigated the fractions of patients whose BMD changes exceeded LSC in the Manitoba Bone Density program [18]. Spectrum bias was observed when BMD monitoring was based upon relative change

(in percent) rather than absolute measurements ( $\text{g}/\text{cm}^2$ ). Therefore, it advocated the categorization of change in BMD based upon absolute values [18]. The precision error dependence on the absolute value is why patients, representative of the population being usually referred to a DXA center, should be used in precision studies.

The *effect of bone area on short-term BMD measurement error* was investigated during the Manitoba Bone Density Program [67]. Differences in bone area exceeding the 2% rule were common and also caused greater BMD measurement errors. This finding was corroborated in another study showing that greater interpretable bone area improved precision [68].

*Software advances may additionally improve precision.* This was the case with the Hologic Apex v2.0 analysis software in comparison with the Hologic Delphi v11.2 software. A study found that the precision of Apex and Delphi was respectively 1.0 % and 1.2 % (L1-L4 spine), 1.1 % and 1.3 % (total hip), 1.6% and 1.9% (femoral neck) and 0.7% and 0.9% (dual total femur) [69]. The same was true for the Encore software used by GE Healthcare. Software advances also proved useful in reducing the precision errors in body composition studies. In a study comparing two versions of software (Hologic V8.1a and V8.21), the reproducibility of DXA in obese women was tested [70]. The CVs for fat and lean weight and BMC were 1.2%, 1.1%, and 1.7% respectively for the new software V8.21, compared to 1.3%, 1.3%, and 2.1% respectively for the older V8.1a [70].

The *effect of femoral rotation on hip BMD* was tested in a number of older and more recent studies [71-73]. External rotation by 5 to 10 degrees was able to change BMD by more

than 5 % [71]. In another study, the authors investigated the effect of subtle positioning flaws on hip BMD measurements [72]. Overall, there was no significant difference between all right or all left hips. In hip pairs, in which one hip was positioned flawlessly and the contra-lateral hip was flawed (vertically, rotationally or both), the measured BMD of the latter hip was not predictably greater than the former. The conclusion was that subtle positioning flaws do not generate predictable changes in measured BMD at any hip ROI [72]. The effect of femur rotation on BMD precision was further confirmed in a methodological study of the distal femur following total knee arthroplasty [73].

The *negative effect of increasing weight on short-term precision* of DXA was tested on a GE Lunar Prodigy scanner [74]. The %CVs for the lumbar spine, total hip and total body increased by approximately 1/3 with increasing BMI [74]. The effect of weight and weight change on the long-term precision of spine and hip DXA measurements was studied in 64 postmenopausal women over a period of 10 years [57]. The authors concluded that body weight had a small effect on precision expressed in absolute values of BMD, but not on %CV, and this was true only for weight changes of over 5 kg. Another study simulated change in body fatness and registered some changes in Hologic QDR 4500A whole body and central DXA bone measures [75]. The authors concluded that, on average, simulated weight change minimally impacted bone measures. However, individual variability in measurement error was noteworthy [75].

The effect of weight is more pronounced in studies of body composition. A review on the trueness and precision of DXA

in the assessment of body composition has been summarized in a number of studies showing trueness errors of DXA % fat estimates between -5.3 % and +2.9 % [76]. For the Hologic QDR 4500A, the %CV for lean and fat mass of 1.3 % was found, for the GE Lunar Prodigy it was found to be 0.7 – 1.0 % for lean and 1.2 % for fat mass, and for the GE Lunar iDXA it was found to be 0.4 – 0.5 % and 0.7 – 0.8 % respectively [76].

### *Some more statistical considerations*

The *sample size* required for bone density precision assessment is a matter of discussion. A few authors have hypothesized that a sample size of 30 might be too small for a robust precision estimate, and compared it to a pool of approximately 200 replicate scans [41]. The pooled spine RMS-SD was found to be 0.017 and pooled hip RMS-SD was found to be 0.009 g/cm<sup>2</sup>. Sample sizes of 30 gave a range of RMS-SD estimates from 0.012 to 0.021 for the spine and from 0.008 to 0.012 g/cm<sup>2</sup> for the hip. The conclusion was that a sample size of 30 is insufficient to reliably characterize precision error or change during clinical monitoring [41].

The *validity of the LSC* can be undermined by a long-term gradual shift in the machine performance. A research group used a specially designed phantom to monitor machine stability over time [77]. Two Hologic scanners were observed to cause clinically insignificant BMD shifts (maximum of 0.34%), whereas two GE/Lunar scanners revealed higher BMD shifts (1.5% and 2.1%). As a result, using the LSC calculations based only on short-term *in vivo* precision studies might not be

valid in case of poorer long-term machine stability [77]. In a similar study, when testing the long-term performance of a 6-year old bone densitometer, the manufacturer's QC failed to detect a 2% shift in the phantom BMD values, which required regular measurements of the Lunar aluminum phantom in addition to the daily QC measurement of the tissue equivalent block [78].

In addition to LSC, the measurement error can be detected using *Bland and Altman's plots* [79]. They are also used in comparisons of different technologies measuring similar tissues [80]. A good description of Bland and Altman's plots can be found in ref. [81]. Precision expressed using this method gives an absolute estimate of random measurement error, which is known as SDD. If there are two observations for each subject, the standard deviation of the differences ( $SD_{diff}$ ) estimates within measurement variability. Most disagreements between measurements are expected to be in the so called "limits of agreement" – between  $-1.96 \times SD_{diff}$  and  $+ 1.96 \times SD_{diff}$  (95 % CI limits of agreement). A test is capable to detect a difference of at least that established by the magnitude of the limits of agreement. The authors recommend the use of the SDD additionally to, or instead of, the LSC. In their study, the SDD was  $\pm 0.0218 \text{ g/cm}^2$  when both femurs were measured, and  $\pm 0.0339 \text{ g/cm}^2$  when only one femur was measured. Thus, they advocated the use of the dual femur feature [81].

Alternative approaches to longitudinal studies have also been proposed. To find outliers more easily of the primary outcome, a *Bayesian method* was proposed for calculating a prediction interval that can incorporate external process information [40].

The authors believed that this approach outperformed the least squares method. Another research group evaluated *the empirical Monte Carlo simulation method* for estimating the significance of an observed change in BMD, which simultaneously considers the magnitude of the change [82]. Using this approach they identified a progressive increase in the ability to detect BMD change using larger precision study sample sizes [82].

### **Replacing or upgrading a bone densitometer. Cross-calibration of DXA devices from different manufacturers or different models of the same manufacturer**

The industry is constantly developing new bone densitometry devices. So it is not uncommon for a Bone Unit to upgrade an old device or replace it with a new, more sophisticated one. The ISCD has published guidelines for how to proceed in different situations of system upgrade or change [2, 3, 83]. The topic is also extensively presented in the literature.

#### *Upgrading central DXA equipment*

In the case of **software upgrade**, the manufacturer has to be asked as to whether the new software has implemented changes in the edge-detection algorithms or the calculation of BMD, T- and Z-scores (e.g. by introducing a new reference database). The first possibility can be tested by scanning the phantom 10 times and establishing a new baseline BMD. The difference to the old phantom baseline BMD should not exceed 1 %. The BMD, T- and Z-scores calculations can be

tested by reanalyzing scans done with the old software (without changing the region of interest). The numerical results should not change. Unintentional consequences can, however, occur with software changes (84). An example is a study of body composition analysis, in which measurements by a GE Lunar DPX-L machine differed between two (standard and extended) analysis modes [85].

In the case of **hardware changes or upgrades**, a technologist should perform 10 phantom scans with repositioning before and after the hardware change. If a difference of more than 1% in BMD is found, the manufacturer should be contacted for servicing.

### *Replacing central DXA equipment*

The transfer of patients to a **completely new device for measurements** usually creates a set of specific problems. They are exhausted in detail in the publications by the ISCD, as well as those by other authors [2, 3, 83, 86-88]. Different DXA systems from the same manufacturer may use different acquisition methods, software and reference databases. Interdevice variability is expected to be  $\pm 2\%$ . Different manufacturers use different technologies to produce and collimate the X-ray beam. They also use different calibration methods, edge-detection software, regions of interest, and normative databases. As a result, the interdevice variability is expected to be  $\pm 5-7\%$ . Interdevice variability is considered too high for practical purposes and data are not interchangeable. **Cross-calibration** is needed in these cases.

When replacing a bone densitometer, there are three possible outcomes:

1. The new device is from the same manufacturer and model, but exhibits some technical and program improvements.
2. The new device is from the same manufacturer but the technology differs substantially. A good example is the transition from dual-photon to X-ray absorptiometry (from DPA to DXA). The DPA technology is rather old-fashioned and this type of transition is very rarely seen. That is why it will not be discussed further. Another example is the replacement of a pencil-beam with a fan-beam device.
3. The new device is from a different manufacturer

In the first scenario cross-calibration should be performed by scanning the phantom 10 times with repositioning on the new device. A new baseline BMD and intra-system LSC is established. If there is more than 1% difference in phantom BMD between the two devices, the manufacturer has to be contacted for servicing. If the difference cannot be corrected, cross-calibration by scanning patients is mandatory.

In the second (e.g. replacing pencil-beam with fan-beam technology from the same manufacturer) and third scenarios (when changing to a system made by a different manufacturer) cross-calibration is needed. Cross-calibration is an approach used to overcome differences between different DXA machines by creating a regression line based on BMD values of the same patients measured using both devices. This process is described in detail in the Cross-calibration and Minimum Precision Standards for DXA: The 2005 ISCD Official Positions [42].

One approach to cross-calibration is to scan 30 subjects, who represent the facility's patient population, on the initial system and then scan them twice on the new system within 60 days (lumbar spine and hip are measured). The subjects should be scanned on both machines preferably on the same day. This approach creates the inconvenience of: (1) keeping both machines at the DXA Center and (2) of multiple scanning for the patients. The data collected are used in regression analyses to generate a slope and intercept for each region of interest. The ISCD has developed a specific Cross-Calibration Tool ([www.iscd.org](http://www.iscd.org)). The user has to simply insert his BMD data from the two machines and all of the calculations are done automatically. Using this tool, the average BMD relationship (from the regression line) and LSC for the new and old machine can be calculated.

Inter-system LSC for two densitometers of different technology is expected to be 2 to 3 times higher than that of one single machine. Some authors have suggested using the root-mean-square error (RMSE) twice for the LSC between measures, for the measured change to reach statistical significance [89]. The inter-system LSC incorporates three sources of variation: (1) the LSC of the initial machine, (2) the LSC of the new machine and (3) the SEE of the regression equations. The newly established LSC on the new system, which is derived from the duplicate scans, should be used further.

Others have advocated for the use of a general LSC [90, 91]. This approach requires 30 patients to be scanned on the old system for a precision study, 30 patients to be scanned on the new system for the same purpose, and 30 patients to be scanned on both machines in order to build a regression line.

If two machines are not cross-calibrated, any direct comparisons are prone to inherent errors, as described above, and are

not allowed. Thus, it is not possible to quantitatively compare BMD from different machines without cross-calibration.

Extensive work has been done in the field of cross-calibration. Most of the older studies comparing devices from different manufacturers were performed on Lunar, Norland and Hologic devices. The different machines were calibrated differently and BMD in  $\text{g}/\text{cm}^2$  was not identical (i.e. universal) on any one machine. It was readily seen that Hologic machines usually provide lower values for BMD [88, 92]. On the other hand, even machines from the same manufacturer did not produce identical results [93-96]. Two ways were proposed to partially correct for this problem: (1) to cross-calibrate different machines (as described above); or (2) to introduce a universal kind of BMD (called “standardized” BMD or sBMD), which might be calculated from the raw data of the machine.

The first approach (the cross-calibration) is the “gold standard” today. It can create difficulties because of the great variety between different DXA devices. An attempt was made in the past to overcome this diversity by introducing conversion formulas based on different manufacturers. A number of studies developed conversion formulas for the QDR 1000, the Lunar DPX and the Norland XR-26 [97-99]. With today’s variety, with at least 15 central DXA devices, each machine should be cross-calibrated against each other. That is why the International Committee for Standards in Bone Measurement supported a study, in which 100 women underwent PA spine and proximal femur measurements on 3 devices – the Hologic QDR-2000, the Norland XR-26 Mark II and the Lunar DPX-L [100]. Equations for the conversion of PA lumbar spine BMD from device to device were published. In the same study, the European Spine Phantom was

scanned on each of the three machines. The data were used to develop formulas for conversion of raw BMD data into sBMD, which would not differ by more than 3-5 % on the different machines. These formulas were approved in 1994 [101]. The sBMD is expressed in  $\text{mg}/\text{cm}^2$ , not in  $\text{g}/\text{cm}^2$ , to be readily distinguished by the reader. Two years later, formulas for the total femur (total hip) were approved [102]. The formulas for sBMD are presented below [101, 102]:

$$\text{sBMD}_{\text{SPINE}} = 1000(1.0761 \times \text{Norland XR-26 BMD}_{\text{SPINE}})$$

$$\text{sBMD}_{\text{SPINE}} = 1000(0.9522 \times \text{Lunar DPX-L BMD}_{\text{SPINE}})$$

$$\text{sBMD}_{\text{SPINE}} = 1000(1.0755 \times \text{Hologic QDR-2000 BMD}_{\text{SPINE}})$$

$$\text{sBMD}_{\text{TOTAL FEMUR}} = 1000 [(1.008 \times \text{Hologic BMD}_{\text{TOTAL FEMUR}}) + 0.006]$$

$$\text{sBMD}_{\text{TOTAL FEMUR}} = 1000[(0.979 \times \text{Lunar BMD}_{\text{TOTAL FEMUR}}) - 0.031]$$

$$\text{sBMD}_{\text{TOTAL FEMUR}} = 1000[(1.012 \times \text{Norland BMD}_{\text{TOTAL FEMUR}}) + 0.026]$$

A few years later Lu et al. [103] developed equations for sBMD of the femoral neck, trochanter and Ward's area. The formulas for the femoral neck are presented below:

$$\text{sBMD}_{\text{FEMORAL NECK}} = 1000 [(1.087 \times \text{Hologic BMD}_{\text{FEMORAL NECK}}) + 0.019]$$

$$\text{sBMD}_{\text{FEMORAL NECK}} = 1000 [(0.939 \times \text{Lunar BMD}_{\text{FEMORAL NECK}}) - 0.023]$$

$$\text{sBMD}_{\text{FEMORAL NECK}} = 1000 [(0.985 \times \text{Norland BMD}_{\text{FEMORAL NECK}}) + 0.006]$$

sBMD is calculated automatically using the software and is part of the printout of all state-of-the-art DXA machines. However, one should keep in mind that: (1) the conversion equations used data from older technologies; and (2) the mean error introduced by using sBMD can reach 5 %. It is superimposed on the precision error of 1-2 % and, therefore, has a deteriorating ability to detect early changes in BMD. That is why *sBMD should not be used for individual patients* [2-4, 42].

The European Spine Phantom was used for cross-calibration in the NOREPOS study and was a valid substitute when assessing the agreement between scanners of the same model, not between different models by the same manufacturer [104]. Therefore, cross-calibration of specific DXA devices, while creating associated regression lines and Bland-Altman plots, is the only valid way of comparing BMD data between devices.

In a comparative study, cross-calibration was done between densitometers of different manufacturers [88]. As expected, BMD measured on the GE Lunar machine was about 15 % higher than that on the Hologic. Furthermore, using sBMD to cross-calibrate gave a mean difference of 3 % at the lumbar spine. The *in vivo* calibration gave better agreement than using sBMD [88].

An example for the impact of DXA technology on cross-calibration results is the comparison of pencil-beam and fan-beam technology. The effect of bone density, scan mode and tissue depth on spine measurements was investigated in an *in vivo* and phantom study using GE Lunar MD and Prodigy densitometers [105]. This study demonstrated that the choice of scan mode was the most important factor for the Prodigy and for subjects who were thin, obese or who had a low BMD [105]. Another publication discussed the cross-calibration and variability of the GE Lunar DPX-L and Prodigy bone densitometers [106]. GE Lunar iDXA and Prodigy bone densitometers have been cross-calibrated in a number of studies, but two of them demonstrated very well the technique of cross-calibration [52, 107]. In one of these studies, BMD was measured in 3 groups, each consisting of 30 subjects, by using a cross-over design and both a GE Lunar iDXA and a Prodigy [107]. Cross-calibration equations decreased the systematic errors between the Prodigy and the iDXA by 0.4 % at the

spine, 0.8 % at the femoral neck, and 0.1 % at the total femur. In another study, the greatest mean BMD difference was only 0.007 g/cm<sup>2</sup> and was, therefore, of no clinical consequence [52].

The forearm site has been the subject of extensive research focused on trying to find a way of standardization. Due to the great variety of devices, technologies and forearm regions of interest, the attempt of universal forearm standardization has obviously failed [26, 108, 109]. No real conversion from device to device is reliable.

### *QC and cross-calibration in the research setting*

The increasing efforts to fill all of the gaps in the field of DXA QC and cross-calibration have led to the inclusion of different population groups. A candidate for testing the limits of DXA technology is the pediatric population [110, 111]. Children and adolescents have smaller bones, whose bone density is rather low. Excellent machine performance is needed to overcome these difficulties. Special pediatric phantoms have been developed [112]. Furthermore, new phantoms for testing body composition analysis and morphometric X-ray absorptiometry have also been introduced [113-116].

Another interesting direction of research is in testing DXA performance in small animals such as mice, rats, piglets and others [117-119].

The concept of QC might also be explored in more advanced DXA applications, such as morphometric X-ray absorptiometry or hip structural analysis, as stated earlier [120]. It is also used in other methods, which contribute to the quantification of bone mass, such as QCT or pQCT [121].

## **Conclusion**

DXA technology is now believed to provide a robust basis for the diagnosis of osteoporosis, the assessment of fracture risk and monitoring of bone mass changes over time. The diagnostic classification in osteoporosis is based on arbitrary T-scores, which are highly dependent on the precision of the measurement technique [122, 123]. Low precision can yield misleading results, especially when combined with the deleterious effect of low accuracy [124]. Thus, both accuracy and precision are crucial for the validity of the DXA technique. There are attempts that are aimed to continue to improve our understanding of both types of error and to reduce their impact on the every day clinical practice in bone densitometry [125-127]. Proper QA and QC in bone densitometry has become a standard tool in clinical practice, as well as a mandatory criterion for a DXA site certification and accreditation [128, 129].

## **Acknowledgments**

The author would like to thank Dr. Margarita Blajeva for improving the English language of this manuscript

## References

- [1] Kanis JA, Melton LJ, III, Christiansen C, Johnston CC, Khaltsev N. The diagnosis of osteoporosis. *Journal of Bone and Mineral Research* 1994;9(8) 1137-1141.
- [2] Binkley N, Bilezikian JP, Kendler DL, Leib ES, Lewiecki EM, Petak SM. Official Positions of the International Society for Clinical Densitometry and Executive Summary of the 2005 ISCD Position Development Conference. *Journal of Clinical Densitometry* 2006;9(1) 4-14.
- [3] Baim S, Binkley N., Bilezikian JP, Kendler DL, Hans DB, Lewiecki EM, Silverman S. Official Positions of the International Society for Clinical Densitometry and Executive Summary of the 2007 ISCD Position Development Conference. *Journal of Clinical Densitometry* 2008;11(1) 75-91.
- [4] Blake GM, Wahner HW, Fogelman IF. Assessment of instrument performance: precision, installation of new equipment and radiation dose. In: Blake GM, Wahner HW, Fogelman IF (ed.). *The evaluation of osteoporosis. Dual-energy X-ray absorptiometry and ultrasound in clinical practice*. London: Martin Dunitz; 1999, p147-172.
- [5] Kolta S, Roux C. Contrôle de qualité en densitométrie osseuse. *Journal de Radiologie* 1999;80(1) 93-98.
- [6] Bonnick SL. Chapter 4: Quality control. In: Bonnick SL (ed.) *Bone densitometry in clinical practice- application and interpretation*, 3rd ed. Totowa NJ: Humana Press; 2010, p105-126.
- [7] Dunn WL, Wahner HW. Instrument evaluation and routine quality control procedures. In: Wahner HW, Fogelman I (eds.) *The evaluation of osteoporosis: Dual energy X-ray absorptiometry in clinical practice*. Martin Dunitz, London 1994, p49-74.
- [8] Bonnick SL, Johnston CC, Jr., Kleerekoper M, Lindsay R, Miller P, Sherwood L, Siris E. Importance of precision in bone density measurements. *Journal of Clinical Densitometry* 2001;4(2) 105-110.
- [9] Kelly TL, Slovik DM, Neer RM. Calibration and standardization of bone mineral densitometers. *Journal of Bone and Mineral Research* 1989;4(5) 663-669.
- [10] Lilley J, Walters BG, Heath DA, Drolc Z. In vivo and in vitro precision for bone density measured by dual-energy X-ray absorption. *Osteoporosis International* 1991;1(3) 141-146.

- [11] Orwoll ES, Oviatt SK, Biddle JA. Precision of dual-energy x-ray absorptiometry: development of quality control rules and their application in longitudinal studies. *Journal of Bone and Mineral Research* 1993;8(6) 693-699.
- [12] Kelly T, Grane G, Baran D. Single X-ray Absorptiometry of the Forearm: Precision, Corelation, and Reference Data. *Calcified Tissue International* 1994;54(2) 212-218.
- [13] Bassett CA. Quality assurance for bone densitometry research studies. *Osteoporosis International* 1994;4(5) 283 (comment).
- [14] Glüer CC, Faulkner KG, Estilo MJ, Engelke K, Rosin J, Genant HK. Quality assurance for bone densitometry research studies: concept and impact. *Osteoporosis International* 1993;3(5) 227-235.
- [15] Glüer CC, Blake GM, Lu Y, Blunt BA, Jergas M, Genant HK. Accurate assessment of precision errors: How to measure the reproducibility of bone densitometry techniques. *Osteoporosis International* 1995;5(3) 262-270.
- [16] Miller CG. Instrument quality control and phantoms for DXA scanners. In: *The evaluation of osteoporosis: dual energy X-ray absorptiometry in clinical practice*. Wahner HW, Fogelman I (eds.) London: Martin Dunitz; 1994, p173-93.
- [17] Njeh CF, Hans D. Instrument Evaluation. In: *The Physical measurement of Bone*. Langton CM, Njeh CF (eds.) IoP, 2004, p91-122.
- [18] Leslie WD. The importance of spectrum bias on bone density monitoring in clinical practice. *Bone* 2006;39(2) 361-368.
- [19] Damilakis J, Guglielmi G. Quality assurance and dosimetry in bone densitometry. *Radiology Clinics of North America* 2010;48(3) 629-640.
- [20] Sheahan NF, Dowling A, O'Reilly G, Malone JF. Commissioning and quality assurance protocol for dual energy X-ray absorptiometry (DEXA) systems. *Radiation Protection Dosimetry* 2005;117(1-3) 288-290.
- [21] Baim S, Wilson CR, Lewiecki EM, Luckey MM, Downs RW Jr, Lentle BC. Precision assessment and radiation safety for dual-energy X-ray absorptiometry. *Journal of Clinical Densitometry* 2005;8(4) 371-378.
- [22] Hawkinson J, Timins J, Angelo D, Shaw M, Takata R, Harshaw F. Technical white paper: bone densitometry. *Journal of the American College of Radiology* 2007;4(5) 320-327.

- [23] Lewiecki EM, Binkley N, Petak SM. DXA quality matters. *Journal of Clinical Densitometry* 2006;9(4) 388-392.
- [24] Kolta S, Ravaud P, Fechtenbaum J, Dougados M, Roux C. Accuracy and precision of 62 bone densitometers using a European Spine Phantom. *Osteoporosis International* 1999;10(1) 14-19.
- [25] Pearson J, Ruegsegger P, Dequeker J, Henley M, Bright J, Reeve J, Kalender W, Felsenberg D, Laval-Jeantet AM, Adams JE, et al. European semi-anthropomorphic phantom for the cross-calibration of peripheral bone densitometers: assessment of precision accuracy and stability. *Bone Mineral* 1994;27(2) 109-120.
- [26] Shepherd JA, Cheng XG, Lu Y, Njeh C, Toschke J, Engelke K, Grigorian M, Genant HK Universal standardization of forearm bone densitometry. *Journal of Bone and Mineral Research* 2002;17(4) 734-745
- [27] Prevrhal S, Lu Y, Genant HK, Toschke JO, Shepherd JA. Towards standardization of dual X-ray absorptiometry (DXA) at the forearm: a common region of interest (ROI) improves the comparability among DXA devices. *Calcified Tissue International* 2005;76(5) 348-354.
- [28] Shewhart WA. On the Measurement of a Physical Quantity Whose Magnitude is Influenced by Primary Causes beyond the Control of the Observer and on the Method of Determining the Relation between Two Such Quantities. *Procedures National Academy of Science USA* 1922;8(8) 248-251.
- [29] Lu Y, Mathur AK, Blunt BA, Gluer CC, Will AS, Fuerst TP, Jergas MD, Andriano KN, Cummings SR, Genant HK. Dual X-ray absorptiometry quality control: comparison of visual examination and process-control charts. *Journal of Bone and Mineral Research* 1998;11(5) 626-637.
- [30] Westgard JO, Groth T, Aronsson T, Falk H, de Verdier CH. Performance characteristics of rules for internal quality control: probabilities for false rejection and error detection. *Clinical Chemistry* 1977;23(10) 1857-1867.
- [31] Westgard JO, Groth T, Aronsson T, de Verdier CH. Combined Shewhart-cusum control chart for improved quality control in clinical chemistry. *Clinical Chemistry* 1977;23(10) 1881-1187.
- [32] Garland SW, Lees B, Stevenson JC. DXA longitudinal quality control: a comparison of inbuilt quality assurance, visual inspection, multi-rule

- Shewhart charts and Cusum analysis. *Osteoporosis International* 1997;7(3) 231-7.
- [33] Pearson D, Cawte SA. Long-term quality control of DXA: a comparison of Shewart rules and Cusum charts. *Osteoporosis International* 1997;7(4) 338-343.
- [34] Khan AA, Colquhoun A, Hanley DA, Jankowski LG, Josse RG, Kendler DL, Lentle B, Leslie WD, Lewiecki EM, O'Neill E, Robertson S, Syed ZA, Tanner SB, Webster D. Standards and guidelines for technologists performing central dual-energy X-ray absorptiometry. *Journal of Clinical Densitometry* 2007;10(2) 189-195.
- [35] Faulkner KG, McClung MR. Quality control of DXA instruments in multicenter trials. *Osteoporosis International* 1995;5(4) 218-227.
- [36] Khan KM, Henzell SL, Broderick C, Prince RL, Saul A, Lomman J, Wark JD. Instrument performance in bone density testing at five Australian centres. *Australian and New Zealand Journal of Medicine* 1997;27(5) 526-530.
- [37] Wahner HW, Looker A, Dunn WL, Walters LC, Hauser MF, Novak C. Quality control of bone densitometry in a national health survey (NHANES III) using three mobile examination centers. *Journal of Bone and Mineral Research* 1994;9(6) 951-960.
- [38] Borissova A-M, Rashkov R, Boyanov M, Shinkov A, Popivanov P, Temelkova N, Vlahov J, Gavrilova M. Femoral neck bone mineral density and 10-year absolute fracture risk in a national representative sample of Bulgarian women aged 50 years and older. *Archives of Osteoporosis* 2011;6(1-2) 189-195.
- [39] Slavchev A, Boyanov M. Impact of precision on long-term patient monitoring during osteoporosis treatment. In: Dossel O, Schlegel WC. (eds.), WC 2009, IFMBE Proceedings 2009;25(7) 352-355.
- [40] Ambrosius WT, Hui SL. A quality control measure for longitudinal studies with continuous outcomes. *Statistics in Medicine* 2000;19(10) 1339-1362.
- [41] Leslie WD, Moayerri A. Minimum sample size requirements for bone density precision assessment produce inconsistency in clinical monitoring. *Osteoporosis International* 2006;17(11) 1673-1680.

- [42] Shepherd JA, Lu Y, Wilson K, Fuerst T, Genant H, Hangartner TN, Wilson C, Hans D, Leib ES; International Society for Clinical Densitometry Committee on Standards of Bone Measurement. Cross-calibration and minimum precision standards for Dual-Energy X-ray absorptiometry: The 2005 ISCD Official Positions. *Journal of Clinical Densitometry* 2006;9(1) 31-36.
- [43] Briggs AM, Wark JD, Kantor S, Teh R, Greig AM, Fazzalari NL, Bennell KL. In vivo intrarater and interrater precision of measuring apparent bone mineral density in vertebral subregions using supine lateral dual-energy X-ray absorptiometry. *Journal of Clinical Densitometry* 2005;8(3) 314-319.
- [44] Ferrar L, Jiang G, Eastell R. Short-term precision for morphometric X-ray absorptiometry. *Osteoporosis International* 2001;12(9) 710-715.
- [45] Ferrar L, Jiang G, Eastell R. Longitudinal evaluation of morphometric X-ray absorptiometry for the identification of vertebral deformities. *Osteoporos. Int.* 2001;12(8);661-671.
- [46] Hind K, Oldroyd B, Truscott JG. In vivo precision of the GE Lunar iDXA densitometer for the measurement of total body composition and fat distribution in adults. *European Journal of Clinical Nutrition* 2011;65(1) 140-142.
- [47] Oldroyd B, Smith AH, Truscott JG. Cross-calibration of GE/Lunar pencil and fan-beam dual energy densitometers – bone mineral density and body composition studies. *European Journal of Clinical Nutrition* 2003;57(8) 977-987.
- [48] Khoo BC, Beck TJ, Qiao QH, Parakh P, Semanick L, Prince RL, Singer KP, Price RI. In vivo short-term precision of hip structure analysis variables in comparison with bone mineral density using paired dual-energy X-ray absorptiometry scans from multi-center clinical trials. *Bone* 2005;37(1) 112-121.
- [49] Bonnick SL. Chapter 15: FDA-approved bone densitometry devices. In: Bonnick SL (ed.) *Bone densitometry in clinical practice- application and interpretation*, 3rd ed. Totowa NJ: Humana Press; 2010, p371-425.
- [50] Shepherd JA, Fan B, Lu Y, Lewiecki EM, Miller PD, Genant HK. Comparison of BMD precision for Prodigy and Delphi spine and femur scans. *Osteoporosis International* 2006;17(9) 1303-1308.

- [51] Henzell, S., S. Dhaliwal, R. Pontifex, Gill F, Price R, Retallack R, Prince R. Precision error of fan-beam dual X-ray absorptiometry scans at the spine, hip, and forearm. *Journal of Clinical Densitometry* 2000;3(4) 359-364.
- [52] Krueger D, Vallarta-Ast N, Checovich M, Gemar D, Binkley N. BMD measurement and precision: a comparison of GE Lunar Prodigy and iDXA Densitometers. *Journal of Clinical Densitometry* 2012;15(1) 21-25.
- [53] Nalda E, Mahadea KK, Dematei C, Kotzki PO, Pouget JP, Boudousq V. Assessment of the Stratos: a new pencil-beam bone densitometer: dosimetry, precision, and cross calibration *Journal of Clinical Densitometry* 2011;14(4) 395-406.
- [54] Blake GM, Knapp KM, Fogelman I. Dual X-ray absorptiometry: clinical evaluation of a new cone-beam system. *Calcified Tissue International* 2005;76(2) 113-120.
- [55] Boyanov M, Shinkov A, Nestorova R. Quantitativer Ultraschall am Kalkaneus und am distalen Radius: eine Vergleichsstudie mit der Zwei-Spektren Röntgen-Absorptiometrie der Lendenwirbelsäule und des proximalen Femurs. *Deutsche Medizinische Wochenschrift* 2007;132(16) 869-873 (article in German).
- [56] Patel R, Blake GM, Rymer J, Fogelman I. Long-term precision of DXA scanning assessed over seven years in forty postmenopausal women. *Osteoporosis International* 2000;11(1) 68-75.
- [57] Rajamanohara R, Robinson J, Rymer J, Patel R, Fogelman I, Blake GM. The effect of weight and weight change on the long-term precision of spine and hip DXA measurements. *Osteoporosis International* 2011;22(5) 1503-1512.
- [58] Tothill P, Hannan WJ. Precision and accuracy of measuring changes in bone mineral density by dual-energy X-ray absorptiometry. *Osteoporosis International* 2007;18(11) 1515-1523.
- [59] Lenora J, Akesson K, Gerdhem P. Effect of precision on longitudinal follow-up of bone mineral density measurements in elderly women and men. *J. Clin. Densitom.* 2010;13(4);407-412.
- [60] Boyanov M, Prodanova A, Popivanov P. Reproducibility of forearm Single X-ray absorptiometry. *Rentgenologia i Radiologia* 1999;38(2) 15-19 (article in Bulgarian).

- [61] Boyanov M. Quality control and standardization of forearm X-ray bone densitometry. *Rentgenologia i Radiologia* 2000;39(2);95-100 (article in Bulgarian).
- [62] Frimeth J, Galiano E, Webster D. Some physical and clinical factors influencing the measurement of precision errors, least significant change, and bone mineral density in dual-energy X-ray absorptiometry. *Journal of Clinical Densitometry* 2010;13(1) 29-35.
- [63] Leslie WD, for the Manitoba Bone Density Program. Factors affecting short-term bone density precision assessment and the effect on patient monitoring. *Journal of Bone and Mineral Research* 2008;23(2) 199-204.
- [64] Kiebzak GM, Morgan SL. Long-term versus short-term precision of dual-energy X-ray absorptiometry scans and the impact of interpreting change in bone mineral density at follow-up. *Journal of Clinical Densitometry* 2011;14(2) 108-115.
- [65] Blake GM, Herd RJ, Fogelman IF. A longitudinal study of supine lateral DXA of the lumbar spine: a comparison with posteroanterior spine, hip and total-body DXA. *Osteoporosis International* 1996;6(6) 462-470.
- [66] Wong JC, Griffiths MR. Precision of bone densitometry measurements: when is change true change and does it vary across bone density values? *Australasian Radiology* 2003;47(3) 236-239.
- [67] Leslie WD, for the Manitoba Bone Density Program. The impact of bone area on short-term bone density precision. *Journal of Clinical Densitometry* 2006;9(2) 150-153.
- [68] Blank RD, Malone DG, Christian RC, Vallarta-Ast NL, Krueger DC, Drezner MK, Binkley NC, Hansen KE. Patient variables impact lumbar spine dual-energy X-ray absorptiometry precision. *Osteoporosis International* 2006;17(5) 768-774.
- [69] Fan B, Lewiecki EM, Sherman M, Lu Y, Miller PD, Genant HK, Shepherd JA. Improved precision with Hologic Apex software. *Osteoporosis International* 2008;19(11) 1597-1602.
- [70] Cordero-MacIntyre ZR, Peters W, Libanati CR, España RC, Abila SO, Howell WH, Lohman TG. Reproducibility of DXA in obese women. *Journal of Clinical Densitometry* 2002;5(1) 35-44.

- [71] Lekamwasam S, Lenora RSJ. The effect of leg rotation on hip bone mineral density measurements. *Journal of Clinical Densitometry* 2003;6(4) 331-336.
- [72] McKiernan F, Washington W. Effect of subtle positioning flaws on measured bone mineral density of the hip. *Journal of Clinical Densitometry* 2005;8(3) 330-334.
- [73] Therbo M, Petersen MM, Schrøder HM, Nielsen PK, Zerahn B, Lund B. The precision and influence of rotation for measurements of bone mineral density of the distal femur following total knee arthroplasty: a methodological study using DXA. *Acta Orthopaedica Scandinavica* 2003;74(6) 677-682.
- [74] Knapp KM, Welsman JR, Hopkins SJ, Fogelman I, Blake GM. Obesity increases precision errors in dual-energy X-ray absorptiometry measurements. *Journal of Clinical Densitometry* 2012;15(3) 315-319.
- [75] Evans EM, Mojtahedi MC, Kessinger RB, Misic MM. Simulated change in body fatness affects Hologic QDR 4500A whole body and central DXA bone measures. *Journal of Clinical Densitometry* 2006;9(3) 315-322.
- [76] Toombs RJ, Duche G, Shepherd JA, DeSouza MJ. The impact of recent technological advances on the trueness and precision of DXA to assess body composition. *Obesity* 2012;20(1) 30-39.
- [77] Hangartner TN. A study of long-term precision of dual-energy X-ray absorptiometry bone densitometers and implications for the validity of the least-significant-change calculation. *Osteoporosis International* 2007;18(4) 513-523.
- [78] Wells J, Ryan PJ. The long-term performance of DXA bone densitometers. *British Journal of Radiology* 2000;73(871) 737-739.
- [79] Feldmann U, Schneider B, Klinkers H, Haeckel R. A multivariate approach for the biometric comparison of analytical methods in clinical chemistry. *Journal of Clinical Chemistry and Clinical Biochemistry* 1981;19(3) 121-137.
- [80] Boneva Z, Boyanov M. Body composition analysis by leg-to-leg bio-electrical impedance and by dual-energy X-ray absorptiometry in non-obese and obese individuals. *Diabetes Obesity Metabolism* 2008;10(10) 1012-1018.

- [81] El Maghraoui A, Achemlal L, Bezza A. Monitoring of dual-energy X-ray absorptiometry measurement in clinical practice. *Journal of Clinical Densitometry* 2006;9(3) 281-286.
- [82] Leslie WD, Moayyeri A, Sadatsafavi M, Wang L. A new approach for quantifying change and test precision in bone densitometry. *Journal of Clinical Densitometry* 2007;10(4) 365-369.
- [83] Writing Group for the ISCD Position Development Conference. Technical standardization for dual-energy X-ray absorptiometry. *Journal of Clinical Densitometry* 2004;7(1) 27-36.
- [84] Binkley N, Kiebzak GM, Lewiecki EM, Krueger D, Gangnon RE, Miller PD, Shepherd JA, Drezner MK. Recalculation of the NHANES Database SD improves T-Score agreement and reduces osteoporosis prevalence. *Journal of Bone and Mineral Research* 2005;20(2) 195-201.
- [85] Nakata Y, Tanaka K, Mizuki T, Yoshida T. Body composition measurements by dual-energy X-ray absorptiometry differ between two analysis modes. *Journal of Clinical Densitometry* 2004;7(4) 443-447.
- [86] Ormica C. Standardization of BMD measurements. *Osteoporosis International* 1998;8(1) 1-3.
- [87] Bonnick SL. Chapter 6: Bone density data among technologies and manufacturers. In: Bonnick SL (ed.) *Bone densitometry in clinical practice- application and interpretation*, 3rd ed. Totowa NJ: Humana Press; 2010, p141-161.
- [88] Pearson D, Horton B, Green DJ. Cross calibration of DXA as part of an equipment replacement program. *Journal of Clinical Densitometry* 2006;9(3) 287-294.
- [89] Blake GM, Harrison EJ, Adams JE. Dual X-ray absorptiometry: cross-calibration of a new fan-beam system. *Calcified Tissue International* 2004;75(1) 7-14.
- [90] Shepherd JA, Lu Y. A generalized least significant change for individuals measured on different DXA systems. *Journal of Clinical Densitometry* 2007;10(3) 249-258.
- [91] Shepherd JA, Morgan SL, Lu Y. Comparing BMD results between two similar DXA systems using the generalized least significant change. *Journal of Clinical Densitometry* 2008;11(2) 237-242.

- [92] Hübsch P, Schneider B, Seidl G, Kalchhauser G, Klaushofer K, Popovic R. [DEXA (dual-energy x-ray absorptiometry) and DPA (dual photon absorptiometry) in densitometry of the femoral neck: correlation of the measurements of three commercially available instruments]. *Aktuelle Radiologie* 1991;1(4) 169-73. (Article in German).
- [93] Abrahamsen B, Gram J, Hansen TB, Beck-Nielsen H. Cross calibration of QDR-2000 and QDR-1000 dual-energy X-ray absorptiometry densitometers for bone mineral and soft tissue measurements. *Bone* 1995;16(3) 385-390.
- [94] Barthe N, Braillon P, Ducassou D, Basse-Cathalinat B. Comparison of two Hologic DXA systems (QDR 1000 and QDR 4500 A). *British Journal of Radiology* 1997;70(835) 728-739.
- [95] Saarelainen J, Honkanen R, Vanninen E, Kröger H, Tuppurainen M, Niskanen L, Jurvelin JS. Cross-calibration of Lunar DPX-IQ and DPX dual-energy X-ray absorptiometry densitometers for bone mineral measurements in women *Journal of Clinical Densitometry* 2005;8(3) 320-329.
- [96] Choi D, Kim DY, Han CS, Kim S, Bok HS, Huh W, Ko JW, Hong SH. Measurements of bone mineral density in the lumbar spine and proximal femur using Lunar Prodigy and the new pencil-beam dual-energy X-ray absorptiometry. *Skeletal Radiology* 2010;39(11) 1109-1116.
- [97] Gundry CR, Miller CW, Ramos E, Moscona A, Stein JA, Mazess RB, Sartoris DJ, Resnick D. Dual-energy radiographic absorptiometry of the lumbar spine: clinical experience with two different systems. *Radiology*. 1990;174(2) 539-41.
- [98] Lai KC, Goodsitt MM, Murano R, Chesnut CH. A comparison of two dual-energy X-ray absorptiometry systems for spinal bone mineral measurement. *Calcified Tissue International* 1992;50(3) 203-208.
- [99] 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. *Journal of Bone and Mineral Research* 1992;7(9) 1081-1084.

- [100] Genant HK, Grampp S, Gluer CC, Faulkner KG, Jergas M, Engelke K, Hagiwara S, Van Kuijk C. Universal standardization for dual X-ray absorptiometry: patient and phantom cross-calibration results. *Journal of Bone and Mineral Research* 1994;9(10) 1503-1514.
- [101] Steiger P. Standardization of spine BMD measurements. *Journal of Bone and Mineral Research* 1995;10(10) 1602-1603.
- [102] Hanson J. Standardization of femur BMD. *Journal of Bone and Mineral Research* 1997;12(8) 1316-1317.
- [103] Lu Y, Fuerst T, Hui S, Genant HK. Standardization of bone mineral density at femoral neck, trochanter, and Ward's triangle. *Osteoporosis International* 2001;12(6) 438-444.
- [104] Omsland TK, Emaus N, Gjesdal CG, Falch JA, Tell GS, Forsen L, Berntsen GK, Meyer HE. In vivo and in vitro comparison of densitometers in the NOREPOS study. *Journal of Clinical Densitometry* 2008;11(2) 276-282.
- [105] Laskey MA, Murgatroyd PR, Prentice A. Comparison of narrow-angle fan beam and pencil beam densitometers. In vivo and Phantom study of the effect of bone density, scan mode and tissue depth on spine measurements. *Journal of Clinical Densitometry* 2004;7(3) 341-348.
- [106] Gagnon DR, McLean RR, Hannan MT, Cupples LA, Hogan M, Kiel DP. Cross-calibration and comparison of variability of two bone densitometers in research setting: the Framingham experience. *Journal of Clinical Densitometry* 2010;13(2) 210-218.
- [107] Choi YJ, Lee BJ, Lim HC, Chung Y-S. Cross-Calibration of iDXA and Prodigy on Spine and Femurs Scans in Korean Adults. *Journal of Clinical Densitometry* 2009;12(4) 450-455.
- [108] Hagiwara S, Engelke K, Yang SO, Dhillon MS, Guglielmi G, Nelson DL, Genant HK. Dual x-ray absorptiometry forearm software: accuracy and intermachine relationship. *Journal of Bone and Mineral Research* 1994;9(9) 1425-1427.
- [109] Forsén L, Bernsten GK, Meyer HE, Tell GS, Fønnebø V, NOREPOS Core Research Group. Differences in precision in bone mineral density measured by SXA and DXA: the NOREPOS study. *European Journal of Epidemiology* 2008;23(9) 615-624.

- [110] Litaker MS, Barbeau P, Humphries MC, Gutin B. Comparison of Hologic QDR-1000/W and 4500W DXA scanners in 13- to 18-year olds. *Obesity Research* 2003;11(12) 1545-1552.
- [111] Leonard CM, Roza MA, Barr RD, Webber CE. Reproducibility of DXA measurements of bone mineral density and body composition in children. *Pediatric Radiology* 2009;39(2) 148-154.
- [112] Picaud JC, Duboeuf F, Vey-Marty V, Delams P, Claris O, Salle BL, Rigo J. First all-solid pediatric phantom for dual-energy X-ray absorptiometry measurements in infants. *Journal of Clinical Densitometry* 2003;6(1) 17-24.
- [113] Diessel E, Shepherd JA, Fuerst T, Gonzalez M, Genant HK, Carew B, Looker A. Comparison of two phantoms for body composition with dual-energy X-ray absorptiometry. *Annals of the New York Academy of Science* 2000;904 107-110.
- [114] Ellis KJ, Shypailo RS, Steinberg FM, Lewis RD, Young RL, Wong WW. Reproducibility of fan-beam DXA measurements in adults and phantoms. *Journal of Clinical Densitometry* 2004;7(4) 413-418.
- [115] Rea JA, Blake GM, Fogelman I. Development of a phantom for morphometric X-ray absorptiometry. *British Journal of Radiology* 2001;74(880) 341-350.
- [116] Steel SA, Thorpe JA, Walker R, Howey S, Langton CM. Development and evaluation of a phantom for morphometric X-ray absorptiometry. *Osteoporosis International* 1999;9(1) 38-44.
- [117] Iida-Klein A, Lu SS, Yokoyama K, Dempster DW, Nieves JW, Lindsay R. Precision, accuracy and reproducibility of dual-energy X-ray absorptiometry measurements in mice in vivo. *Journal of Clinical Densitometry* 2003;6(1) 25-34.
- [118] Horton JA, Murray GM, Spadaro JA, Margulies BS, Allen MJ, Damron TA. Precision and accuracy of DXA and pQCT for densitometry of the rat femur. *Journal of Clinical Densitometry* 2003;6(4) 381-390.
- [119] Fink C, Cooper HJ, Huebner JL, Guilak F, Kraus VB. Precision and accuracy of a transportable dual-energy X-ray absorptiometry unit for bone mineral measurements in guinea pigs. *Calcified Tissue International* 2002;70(3) 164-169.

- [120] Hind K, Oldroyd B, Prajapati A, Rhodes L. In vivo precision of dual-energy X-ray absorptiometry-derived hip structure analysis in adults. *Journal of Clinical Densitometry* 2012;15(3) 302-307.
- [121] Engelke K, Libanati C, Liu Y, Wang H, Austin M, Fuerst T, Stampa B, Timm W, Genant HK. Quantitative computed tomography (QCT) of the forearm using general purpose spiral whole-body CT scanners: accuracy, precision and comparison with dual-energy X-ray absorptiometry (DXA). *Bone* 2009;45(1) 110-118.
- [122] Blake GM, Shepherd JA. Effect of random BMD measurement error on diagnostic classification using T-scores. *Journal of Clinical Densitometry* 2007;10(4) 415-416.
- [123] Kiebzak GM, Faulkner KG, Wacker W, Hamdy R, Seier E, Watts NB. Effect of precision error on T-scores and the diagnostic classification of bone status. *Journal of Clinical Densitometry* 2007;10(3) 239-243.
- [124] Kiebzak GM, Faulkner KG, Wacker W, Hamdy R, Watts NB. Accuracy error versus precision error and the effect on T-scores. *Journal of Clinical Densitometry* 2007;10(4) 417-419.
- [125] Bonnick SL. Monitoring changes in bone density. *Womens Health (London Engl)*. 2008;4(1) 89-97.
- [126] Engelke K, Glüer CC. Quality performance measures in bone densitometry: part 1: errors and diagnosis. *Osteoporosis International* 2006;17(9) 1283-1289.
- [127] Probyn S, Clarys JP, Wallace J, Scafoglieri A, Reilly T. Quality control, accuracy, and prediction capacity of dual energy X-ray absorptiometry variables and data acquisition. *Journal of Physiological Anthropology* 2008;27(6) 317-323.
- [128] Avcilla LS. Professional certification and site accreditation in bone densitometry. *Journal of Clinical Densitometry* 1998;1(1) 81-89.
- [129] Bennett HS, Dienstfrey A, Hudson LT, Oreskovic T, Fuerst T, Shepherd J. Standards and measurements for assessing bone health. *Journal of Clinical Densitometry* 2006;9(4) 399-405.