

# Result of Proficiency Test and Comparison of Accuracy Using a European Spine Phantom among the Three Bone Densitometries

Ae Ja Park<sup>1</sup>, Jee-Hye Choi<sup>1</sup>, Hyun Kang<sup>2</sup>, Ki Jeong Park<sup>3</sup>, Ha Young Kim<sup>4</sup>, Seo Hwa Kim<sup>4</sup>, Deog-Yoon Kim<sup>5</sup>, Seung-Hwan Park<sup>6</sup>, Yong-Chan Ha<sup>6</sup>

<sup>1</sup>Department of Laboratory Medicine, Chung-Ang University College of Medicine, Seoul;

<sup>2</sup>Department of Anesthesiology, Chung-Ang University College of Medicine, Seoul;

<sup>3</sup>Departments of Radiology, Chung-Ang University College of Medicine, Seoul;

<sup>4</sup>Division of Endocrinology and Metabolism, Sanbon Medical Center, University of Wonkwang College of Medicine, Iksan;

<sup>5</sup>Department of Nuclear Medicine, College of Medicine, Kyung Hee University, Seoul;

<sup>6</sup>Department of Orthopedic Surgery, Chung-Ang University College of Medicine, Seoul, Korea

## Corresponding author

Yong-Chan Ha

Department of Orthopaedic Surgery,  
Chung-Ang University College of Medicine,  
102 Heukseok-ro, Dongjak-gu, Seoul 156-755,  
Korea

Tel: +82-2-6299-1577

Fax: +82-2-822-1710

E-mail: hayongch@naver.com

Received: February 6, 2015

Revised: May 28, 2015

Accepted: May 28, 2015

No potential conflict of interest relevant to this article was reported.

**Background:** Although dual energy X-ray absorptiometry (DXA) is known to standard equipment for bone mineral density (BMD) measurements. Different results of BMD measurement using a number of different types of devices are difficult to use clinical practice. The purpose of this study was to evaluate discrepancy and standardizations of DXA devices from three manufactures using a European Spine Phantom (ESP). **Methods:** We calculated the accuracy and precision of 36 DXA devices from three manufacturers (10 Hologic, 16 Lunar, and 10 Osteosys) using a ESP (semi-anthropomorphic). The ESP was measured 5 times on each equipment without repositioning. Accuracy was assessed by comparing BMD ( $\text{g}/\text{cm}^2$ ) values measured on each device with the actual value of the phantom. Precision was assessed by the coefficient of variation (CVsd). **Results:** Lunar devices were, on average, 22%, 8.3%, and 5% overestimation for low (L1) BMD values, medium (L2), and high (L3) BMD values. Hologic devices were, on average, 6% overestimation for L1 BMD, and 5% and 6.2% underestimation for L2 and L3 BMD values. Osteosys devices was, on average, 12.7% ( $0.063 \text{ g}/\text{cm}^2$ ), 6.3% ( $0.062 \text{ g}/\text{cm}^2$ ), and 5% ( $0.075 \text{ g}/\text{cm}^2$ ) underestimation for L1, L2, and L3, respectively. The mean CVsd for L1-L3 BMD were 0.01%, 0.78%, and 2.46% for Lunar, Hologic, and Osteosys devices respectively. **Conclusions:** The BMD comparison in this study demonstrates that BMD result of three different devices are significant different between three devices. Differences of BMD between three devices are necessary to BMD standardization.

**Key Words:** Bone density, Densitometry, Lumbar vertebrae, Reference standards

Copyright © 2015 The Korean Society for Bone and Mineral Research

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.



## INTRODUCTION

Osteoporosis is a disease of the bone metabolism characterized by loss of the bone mass and microarchitectural alterations which results in bone fragility and increased risk of fractures.[1] The representative diagnostic tool of osteoporosis is a measurement of bone mineral density (BMD) and biochemical markers.

One of the most common methods of BMD measurements is dual energy X-ray

absorptiometry (DXA). It is a safe, accurate and precise technique. Various devices are available, and the values of measurement may differ among them, for technical reasons.[2, 3] Different results of BMD measurement using a number of different types of devices are difficult to use clinical practice when patients are followed on different machines.[4]

For appropriate treatment and study, we need quality control and calibration of the devices by measurement of phantoms.[5] The European Spine Phantom (ESP) had been developed as a universal standard for instruments measuring bone density. The ESP is composed of three semi-anthropomorphic hydroxyapatite vertebrae of varying densities surrounded by soft tissue equivalent plastic designed to resemble human bone and soft tissue when scanned on bone densitometers.[6] The ESP has been developed by an independent group under the auspices of the Comité d'Actions Concertées-BioMedical Engineering (COMAC-BME) organization,[7-9] for use with different types of DXA devices.[6] This phantom has been used for standardization of BMD results.[7] So far, comparison studies using world widely popular several bone densitometers were reported. However, there is no comparison study including regional bone densitometry. Therefore, the purpose of this study was to evaluate discrepancy and standardizations of DXA devices from three manufactures using a ESP.

## METHODS

Thirty-two centers equipped with DXA devices participated in this study. These centers were distributed throughout the country. Ten centers had 10 Hologic Discovery-W devices (Hologic Inc., Waltham, MA, USA). Ten centers had

10 Osteosys Dexam-T devices (OsteoSys, Seoul, Korea). Twelve centers had 16 Lunar Prodigy advance devices (GE Healthcare, Madison, WI, USA).

A single ESP (no. 126) was used to compare the results of different devices. The ESP is a semianthropomorphic phantom, comprising three vertebral-like structures of different sizes and densities. The three vertebrae represent low (L1), medium (L2), and high (L3) densities, with actual BMD values of 0.496 g/cm<sup>2</sup>, 0.990 g/cm<sup>2</sup>, and 1.499 g/cm<sup>2</sup>, respectively. According to Kolta et al.'s methods,[4] the phantom was scanned 5 times without repositioning on each device. BMD (g/cm<sup>2</sup>), bone mineral content (BMC, g) and area (cm<sup>2</sup>) were collected for each vertebra (L1, L2, L3) and for the three together (L1-L3). We compared the BMD and BMC results on different types of device using ANOVA and Tukey post hoc test. To assess the accuracy of measurement, individual BMD values observed on each device were compared with actual BMD and BMC values of the phantom. To assess precision we plotted the differences between each replicate measurement of BMD and the estimated true value for a particular manufacturer, and calculated the limits of agreement as defined by Bland and Altman.[7] The true value was estimated by the average of all replicate measurements for each manufacturer. In addition, the mean values of the coefficient of variation (CVsd), for each type of device, were calculated for BMD and BMC for each vertebra as well as for the three vertebrae together, using the root mean square average.[8,10]

## RESULTS

The average ( $\pm$  standard deviation [SD]) results of BMD

**Table 1.** Bone mineral density differences of three devices

	Hologic (n=50)	Lunar (n=80)	Osteosys (n=50)	P-value
L1	0.526 $\pm$ 0.012	0.605 $\pm$ 0.018 <sup>a)</sup>	0.433 $\pm$ 0.026 <sup>a,b)</sup>	<0.001
L2	0.940 $\pm$ 0.023	1.072 $\pm$ 0.0314 <sup>a)</sup>	0.928 $\pm$ 0.042 <sup>b)</sup>	<0.001
L3	1.406 $\pm$ 0.021	1.574 $\pm$ 0.062 <sup>a)</sup>	1.424 $\pm$ 0.097 <sup>b)</sup>	<0.001
L1-L3	0.965 $\pm$ 0.018	1.119 $\pm$ 0.038 <sup>a)</sup>	0.925 $\pm$ 0.045 <sup>a,b)</sup>	<0.001
L1 (BMC)	4.450 $\pm$ 0.222	4.992 $\pm$ 0.204 <sup>a)</sup>	4.514 $\pm$ 0.355 <sup>b)</sup>	<0.001
L2 (BMC)	10.155 $\pm$ 0.322	10.684 $\pm$ 0.277 <sup>a)</sup>	9.486 $\pm$ 0.509 <sup>a,b)</sup>	<0.001
L3 (BMC)	15.738 $\pm$ 0.310	16.387 $\pm$ 0.452 <sup>a)</sup>	14.513 $\pm$ 0.909 <sup>a,b)</sup>	<0.001
L1-L3 (BMC)	30.343 $\pm$ 0.681	31.675 $\pm$ 1.346 <sup>a)</sup>	28.513 $\pm$ 1.530 <sup>a,b)</sup>	<0.001

Data were presented as mean  $\pm$  standard deviation, and compared using ANOVA and Tukey post hoc test.

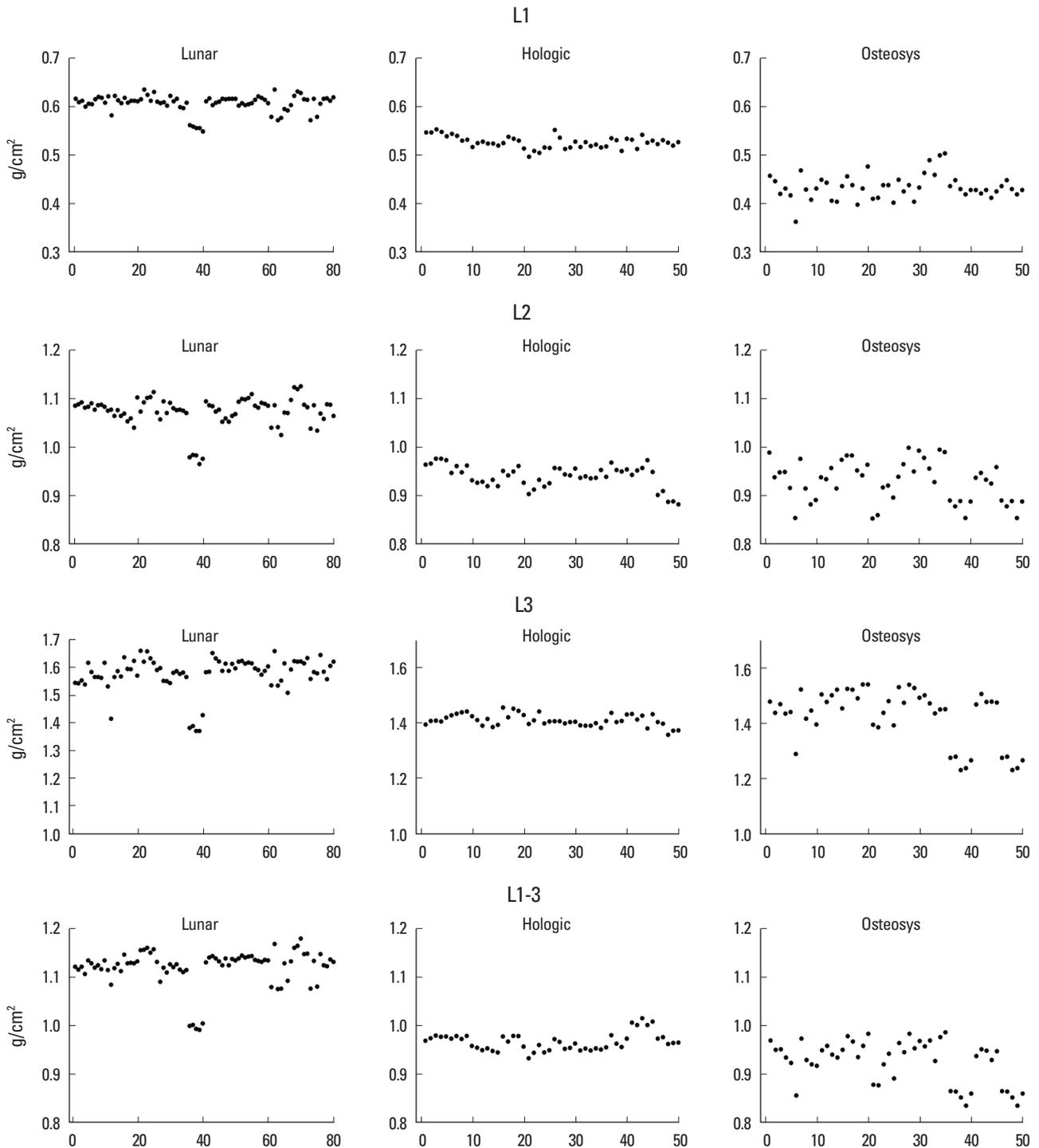
<sup>a)</sup>P<0.05 compared with Hologic. <sup>b)</sup>P<0.05 compared between Lunar and Osteosys.

BMC, bone mineral content.

and BMC of L1, L2, L3, and L1-L3 for three devices was described in Table 1. There was significant difference between BMD values on three devices (Table 1).

BMD values of L1, L2, and L3 on three types of device were different from their actual values (Fig. 1). For Lunar devices

there was, on average, 22% (0.109 g/cm<sup>2</sup>), 8.3% (0.082 g/cm<sup>2</sup>), and 5% (0.075 g/cm<sup>2</sup>) overestimation for low (L1) BMD values, medium (L2), and high (L3) BMD values. For Hologic devices there was, on average, 6% (0.03 g/cm<sup>2</sup>) overestimation for low (L1) BMD values, and 5% (0.05 g/cm<sup>2</sup>) and



**Fig. 1.** Individual bone mineral density (BMD) values of and Lunar (n=16 devices, 80 measurements), Hologic (n=10 devices, 50 measurements), and Osteosys (n=10 devices, 50 measurements) devices.

**Table 2.** Coefficients of variation (CVsd, %) of the bone mineral density and bone mineral content L1, L2, L3, and L1 to L3 for Hologic, Lunar, and Osteosys devices

	Hologic (n=50)	Lunar (n=80)	Osteosys (n=50)
L1	1.72	0.01	5.24
L2	1.05	0.35	3.03
L3	1.00	0.01	15.06
L1-L3	0.78	0.01	2.46
L1 (BMC)	2.66	0.02	6.20
L2 (BMC)	1.07	0.01	3.42
L3 (BMC)	0.90	0.01	1.89
L1-L3 (BMC)	0.90	0.03	2.43

BMC, bone mineral content.

6.2% (0.093 g/cm<sup>2</sup>) underestimation for medium (L2) and high (L3) BMD values. For Osteosys devices there was, on average, 12.7% (0.063 g/cm<sup>2</sup>), 6.3% (0.062 g/cm<sup>2</sup>), and 5% (0.075 g/cm<sup>2</sup>) underestimation for L1, L2, and L3, respectively.

The limits of agreement, as defined by Bland and Altman, for Hologic devices were  $\pm 0.024$ ,  $\pm 0.045$ ,  $\pm 0.042$ , and  $\pm 0.026$  g/cm<sup>2</sup> for L1, L2, and L3 respectively. For Lunar devices these limits were  $\pm 0.035$ ,  $\pm 0.062$ ,  $\pm 0.122$  g/cm<sup>2</sup> for L1, L2, and L3. For Osteosys devices these limits were  $\pm 0.050$ ,  $\pm 0.083$ ,  $\pm 0.190$  g/cm<sup>2</sup> for L1, L2, and L3.

Differences in extreme results between devices from the same manufacturer were on average 1.4%, 2.8%, and 5.2% for L1-L3 BMD, 1.6%, 7.7%, and 4.9% for L1-L3 BMC for Hologic, Lunar, and Osteosys devices respectively. These differences reached up to 3.95%, 3.77%, and 10.8% for low (L1) BMD on Hologic, Lunar, and Osteosys devices respectively. By comparing these extreme results with the mean values for devices from the same manufacturer, they ranged from -2.3% to +5.0%, -11.5 to +3.6%, and -9.8% to +6.5% on Hologic, Lunar, and Osteosys devices respectively. The mean CVsd values are given in Table 2. For L1-L3 BMD they were 0.01%, 0.78%, and 2.46% for Lunar, Hologic, and Osteosys devices respectively.

## DISCUSSION

In the *in vitro* ESP study, the BMD comparison shows that BMD result of three different devices are significant different between three devices. However, the accuracy and precision of three devices are moderately satisfaction, as are

the limits of agreement among devices.

Although difference of measurement results between Lunar and Hologic devices is well established, BMD results of Osteosys devices in this study are also different from two other devices.[3,4] These findings are corresponded with other studies. They suggested that the reason of these differences were the difference in edge detection algorithms and devices for scan method.[6] For example, Hologic devices use a fixed threshold of 0.2 g/cm<sup>2</sup> and this excludes less of the transverse processes and Lunar devices use a different algorithm incorporating the first derivative of pixels at bone edge, thus eliminating more or less low-density bone.[9] In addition, Fan beam scanners do not measure area directly as do pencil beam scanners. Osteosys device used pencil beam as scan methods. Other two devices used Fan-beam system to measure BMD.[11,12]

According to the agreement between devices, large measurement errors in L3 were observed. In the usual range of BMD (0.5 to 1.0 g/cm<sup>2</sup>) the limits of agreement were similar for three manufacturers.[4] Although the mean CVsd values of BMD for Lunar and Hologic devices are comparable to those found by other authors,[2,4,6] the mean CVsd values of BMD for Osteosys device is slightly wider than other two devices. To minimize range of CVsd, more accurate calibration and quality control are mandatory.

In this study, BMD data of three devices are difficult to generalize BMD results among three devices. So far, cross-calibration formulae for Lunar and Hologic devices have been developed.[2,10] This is particularly useful for epidemiologic studies and therapeutic trials that deal with groups of patients.[13,14] Further cross calibration studies for three devices are required to evaluate a comparison of BMD measurement.

This study has a limitation. This study is *in vitro* ESP study and sample size of this study is not enough to generalize the results. Further study is necessary to calculate cross-calibration formula for three different devices.

In conclusion, the BMD comparison in this study demonstrates that BMD result of three different devices are significant different between three devices. However, the accuracy and precision of three devices are moderately satisfaction. Differences of BMD between three devices are necessary to BMD standardization.

## ACKNOWLEDGMENT

This research was supported by Ministry of Trade, Industry and Energy through Standard Reference Data Program.

## REFERENCES

1. Genant HK, Cooper C, Poor G, et al. Interim report and recommendations of the World Health Organization Task Force for Osteoporosis. *Osteoporos Int* 1999;10:259-64.
2. Genant HK, Grampp S, Gluer CC, et al. Universal standardization for dual x-ray absorptiometry: patient and phantom cross-calibration results. *J Bone Miner Res* 1994;9:1503-14.
3. Pocock NA, Sambrook PN, Nguyen T, et al. Assessment of spinal and femoral bone density by dual X-ray absorptiometry: comparison of lunar and hologic instruments. *J Bone Miner Res* 1992;7:1081-4.
4. Kolta S, Ravaud P, Fechtenbaum J, et al. Accuracy and precision of 62 bone densitometers using a European Spine Phantom. *Osteoporos Int* 1999;10:14-9.
5. Trevisan C, Gandolini GG, Sibilla P, et al. Bone mass measurement by DXA: influence of analysis procedures and interunit variation. *J Bone Miner Res* 1992;7:1373-82.
6. Lees B, Garland SW, Walton C, et al. Evaluation of the European Spine Phantom in a multi-centre clinical trial. *Osteoporos Int* 1997;7:570-4.
7. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-10.
8. Gluer CC, Blake G, Lu Y, et al. Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. *Osteoporos Int* 1995;5:262-70.
9. Mazess RB, Trempe JA, Bisek JP, et al. Calibration of dual-energy x-ray absorptiometry for bone density. *J Bone Miner Res* 1991;6:799-806.
10. Hui SL, Gao S, Zhou XH, et al. Universal standardization of bone density measurements: a method with optimal properties for calibration among several instruments. *J Bone Miner Res* 1997;12:1463-70.
11. Ward KA, Ashby RL, Roberts SA, et al. UK reference data for the Hologic QDR Discovery dual-energy x ray absorptiometry scanner in healthy children and young adults aged 6-17 years. *Arch Dis Child* 2007;92:53-9.
12. 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. *Eur J Clin Nutr* 2003;57:977-87.
13. The Writing Group for the ISCD Position Development Conference. Technical standardization for dual-energy x-ray absorptiometry. *J Clin Densitom* 2004;7:27-36.
14. Reid DM, Mackay I, Wilkinson S, et al. Cross-calibration of dual-energy X-ray densitometers for a large, multi-center genetic study of osteoporosis. *Osteoporos Int* 2006;17:125-32.

