Pulse-echo ultrasound method for detection of post-menopausal women with osteoporotic BMD

Janne P Karjalainen[1], Ossi Riekkinen[1] and Heikki Kröger[2]

[1] Bone Index Finland Ltd., Kuopio, Finland
[2] Department of Orthopaedics, Traumatology and Hand Surgery, Kuopio University Hospital, Kuopio, Finland.

Conflicts of Interest:

Janne P Karjalainen
Bone Index Finland Ltd., stock ownership

Ossi Riekkinen
Bone Index Finland Ltd., stock ownership

Heikki Kröger
Bone Index Finland Ltd., stock ownership

Submitted to Osteoporosis International

Correspondence:
Janne Karjalainen
Microkatu 1
70211 Kuopio
Finland
janne.p.karjalainen@iki.fi

Acknowledgements

Supported in part by a research grant from the Investigator Initiated Studies Program of Merck Sharp & Dohme Corp/MSD. The opinions expressed in this paper are those of the authors and do not necessarily represent those of Merck Sharp & Dohme Corp/MSD. The study was funded in part by Bone Index Finland Ltd.

Mini Abstract

We lack effective diagnostics of osteoporosis at the primary health care level. An ultrasound device was used to identify subject in the osteoporotic range as defined by DXA. A case finding strategy combining ultrasound results with DXA measurements for patients with intermediate ultrasound results is presented.

ABSTRACT

Purpose

We lack effective screening and diagnostics of osteoporosis at primary health care. In this study, a pulse-echo ultrasound (US) method is investigated for osteoporosis screening.

Methods

A total of 1091 Caucasian women (aged 50 – 80 years) were recruited for the study and measured with US in the tibia and radius. This method measures cortical thickness and provides an estimate of bone mineral density (BMD), Density Index (DI). BMD assessment of the hip was available for 988 women. A total of 888 women had one or more risk factors for osteoporosis (OP_{susp}) and 100 women were classified healthy. Previously determined thresholds for the DI were evaluated for assessment of efficacy of the technique to detect hip BMD at osteoporotic range (T-Score at or below -2.5).

Results

In the OP_{susp} group, the application of thresholds for the DI showed that approximately 32% of the subjects would require an additional DXA measurement. The multi-site US measurement based DI showed 93.7 % sensitivity and 81.6% specificity whereas the corresponding values for single site US measurement based DI were 84.7% and 82.0%, respectively. The ultrasound measurements showed a high negative predictive value 97.7% to 99.2% in every age decade examined (ages 50-59, 60-69, 70-79 years).

Conclusions

The study data demonstrate that a strategy of combining ultrasound measurement with added DXA measurements in cases with intermediate ultrasound results (about 30%) can be useful for identifying subjects at risk for a low bone mineral density in the osteoporotic range.

Keywords: Ultrasound, osteoporosis, screening, DXA, bone.

1. INTRODUCTION

In osteoporosis bones undergo several changes including porosity increase and thinning of cortical bone, deterioration of trabecular bone structure and changes in bone tissue composition. Osteoporosis is often diagnosed only after fractures have already occurred and it is estimated that 75% of the osteoporotic patients are not diagnosed. Diagnosis is typically based on bone mineral density (BMD) assessment at hip or spine by dual energy x-ray absorptiometry (DXA). According World Health Organization (WHO) statement, osteoporosis is present if the BMD reading is -2.5 standard deviations below young adult average, typically reported as T-Score. Low bone density is an important determinant of hip fracture, and one standard deviation decrease in femoral bone density increases the risk by a factor of two to three (1).

Hip fracture is the most serious outcome of osteoporosis in terms of mortality and morbidity (2, 3). During the first year after a hip fracture, over 24% of the patients at or over 65 years of age will die (3). The highest risk gradient for hip fractures has been shown to be with BMD measurements at proximal femur. Similarly for other locations, site-specific measurement at e.g. radius or spine show highest prediction for fractures at that location with DXA. However it has been shown that spine measurements especially in elderly, at age 65 or more, are compromised by vertebral fractures and spondylarthrosis changes, which may mislead diagnosis by showing higher BMD in analyses due to structural changes and artifacts. Moreover, the forearm site (radius) has been suggested to be used only when measurements at hip or spine are not possible or cannot be interpreted ((4), ISCD official position).

According to the International Society for Clinical Densitometry (ISCD) recommendations, use of other methods than axial DXA i.e. peripheral DXA or ultrasound, the detection of osteoporotic BMD should be based on determination of 90% sensitivity and

specificity thresholds against axial DXA (5-7). Recently, new ultrasound based approach targeted to primary care has been suggested to be used as an aid in osteoporosis diagnostics (8, 9). In this approach, the cortical thickness in the radius and tibia is measured, and an estimate called density index (DI), for proximal femur BMD as measured by DXA, is reported. The thresholds for detection of osteoporotic BMD were suggested in accordance to ISCD recommendations. The results were promising for the technique being suitable for locating individuals with osteoporotic BMD and suggested good performance when applied in osteoporosis management pathways with fracture risk calculator tool (FRAX®).

The aim of the present study is to investigate the novel technique and suggested thresholds for detection of osteoporotic BMD in Caucasian female population.

2. METHODS

2.1. Subjects

A total of 1091 women were recruited for the study. The subjects were tested at six study sites in different cities in Finland. Five of the sites were units that belonged to a nationwide healthcare service provider (Terveystalo Ltd., at Jyväskylä, Mikkeli, Tampere, Lahti and Kouvola, Finland) and one was university hospital (Kuopio University Hospital, Kuopio, Finland). Patients were included into the study from the daily clinical patient flow whenever there was a free time slot and the patient was willing to participate in the study. From the population of 1091, the data for 23 subjects was lost due to corrupted database leaving 988 women with valid US data. Reproducibility was assessed with 85 subjects, of which 11 were measured also with DXA and included in other groups based on exclusion and inclusion criteria.

The subjects (n = 988) that were measured both with ultrasound and DXA, were divided into two groups based on the presence of risk factors. A total of 888 women had one or more risk factors for osteoporosis (OP_{susp}) and 100 women were classified healthy. The subjects in the OP_{susp} group had been referred to the DXA examination by treating physician due suspected osteoporosis. From OP_{susp} subjects 171 women had received osteoporosis medication and they were classified as a *Treated* group. The inclusion and exclusion criteria for subjects in different groups are shown in Tables 1 and 2. The study was approved by the local ethical committee, and informed written consent was obtained from each subject (Kuopio University Hospital Ethical Committee, 75/2013).

2.2. Ultrasound Measurements

Ultrasound measurements were conducted by using Bindex ultrasound device (model BI-100, Bone Index Finland Ltd., Kuopio, Finland, Software v.2.0). Ultrasound measurements were done by trained study nurses. The device consists of a pulser unit plugged into the USB port of a laptop and a focused ultrasound probe (3.0 MHz nominal center frequency). US measurements were conducted at 1/3 of the length of the radius from the distal head and 1/3 of the length of the tibia from the proximal and distal heads, respectively. Five parameters were collected including cortical thickness (Ct.Th) at the distal radius ($Ct.Th_{rad}$) and at the proximal ($Ct.Th_{prox}$) and distal ($Ct.Th_{dist}$) tibia and density indices based on measurement of all three sites (DI_3) and single site measurement (DI_1) at proximal tibia. The method for cortical thickness measurement has been described earlier in detail (8, 9). Five repetitions were made at each location. One subject had no measurement at the distal tibia and therefore DI_3 could not be calculated.

At each study site, reproducibility assessment was made. A total of 85 subjects participated reproducibility assessment for repeated measurements, relocating the measurement sites (sites were marked with a water soluble marker pen and wiped clear before the next repeated measurement). Root mean coefficient of variation (CV_{rms}) was calculated for DI according to the method earlier described by Gluer et al. (10). Devices were calibrated by measuring same phantom twice, prior first subject and after last subject visit at each site. Average difference between the devices was 0.41% at first assessment. Between the phantom measurements prior first and after last subject visit the average difference was 0.32%. These differences were considered to be negligible and thus were not accounted for in analyses.

Previously published thresholds for osteoporosis were applied for DI_3 (upper 0.876, lower 0.803) and DI_1 (upper 0.844, lower 0.779) (9).

2.3. DXA Measurements

Axial DXA (Lunar Prodigy, GE Healthcare Ltd, Pollards Wood, UK) measurements of BMD were conducted along the guidelines of the manufacturer. BMD values were recorded for the femoral neck (BMD_{neck}) and total hip (BMD_{total}). The subject was considered osteoporotic if the T-score at either the femoral neck or total hip was -2.5 or less. Finnish reference thresholds for BMD at osteoporotic range were 0.684 g/cm^2 at femoral neck and 0.708 g/cm^2 at total hip. At all study sites the same manufacturer and model of DXA device was used. Cross-calibration was performed twice, before the first patient visit and after last patient visit, by measuring the spine phantom (DPA/QDR-1, Hologic Inc., Waltham, MA, USA) 10 times with each device (11). The average difference between the first and second phantom measurement was 0.38%. The difference in the average BMD of phantom measurements between the study sites was 0.42% and 0.29% prior first patient and after last patient visit, respectively. These differences were considered to be negligible and thus were not accounted for in analyses.

For the analyses the results of the left leg were used. If the measurement of the left leg was not possible the results from the right leg were applied for DXA and ultrasound. A total of 4 subjects did not have valid BMD measurement at the hip on either side, leaving 984 subjects with femoral neck and 983 subjects with total hip DXA result. T-Score values reported here refer to femoral neck or total hip values, whichever showed smaller value in T-Score.

2.4. Statistical analyses

The Pearson's correlation analysis was applied for normally distributed parameters. The Shapiro-Wilk's test and visual evaluation of histograms were used to evaluate normality of the parameters distributions and Leneve's test was applied for evaluation of homogeneity of variances. The Student's *t*-test was used to compare normally distributed parameter values

between healthy (T-Score >-2.5) and osteoporotic (T-Score <= -2.5) subjects. The one-way ANOVA was used for normally distributed variables when comparing subject's characteristics or US and DXA data between *Healthy*, OP_{susp} or *Treated* groups. For multiple comparisons Tukey's or Games-Howell post hoc tests were applied depending whether or not equal variances were assumed. The sensitivity and specificity were calculated with the following equations

$$Sensitivity = \frac{true\ positives}{true\ positives + false\ negatives}$$

$$Specificity = \frac{true\ negatives}{true\ negatives + false\ positives}$$

where *false negatives* are subjects with OP above higher threshold and *true positives* are subjects with OP below the higher threshold. The *false positives* are healthy subjects below the lower threshold and *true negatives* are healthy subjects with DI value above lower threshold. Statistical analyses were conducted with SPSS software version 23 (SSPS Inc., Chicago, IL, USA).

3. RESULTS

In the study population with both US and DXA data (*n*=988) 114 women (11.5%) had osteoporotic (T-Score <= -2.5) value either at the femoral neck (*n*=97) or at the total hip (*n*=67).

In the OP_{susp} group (*n* = 888), using the previously determined thresholds for the DI_3, 32.5% of the subjects would have been referred to DXA investigation. The approach showed 93.7% sensitivity and 81.6% specificity to detect osteoporosis based on DXA. With single-site cortical thickness assessment (DI_1), 31.6% of subjects would have required a BMD measurement by DXA, however, the sensitivity was also lower at 84.7% whereas the specificity was similar at 82.0%. The subject classification by DI_1 and DXA are shown in

table 6. Both parameters (DI_1 and DI_3) showed very high negative predictive values (97.4% and 98.8%, respectively) but lower positive predictive values (40.2% and 41.9%, respectively). The average parameter values in different groups classified by DI are reported in the Table 3.

In the *healthy* group (n =100) three subjects were found to have osteoporotic BMD reading at the hip. The density indices showed high sensitivity (100%) and specificity (90.7-93.8%) for osteoporosis. The negative predictive value was 100% for both indices (DI1 and DI3). The DI_3 showed statistically significantly higher values in the *healthy* when compared to OP_{susp} or *Treated* group ($p < 0.01$) (Table 4.). The DI_1 showed also significantly higher values in *Healthy* and OP_{susp} when compared to *Treated* group ($p < 0.01$), however, no difference was observed between *Healthy* and OP_{susp} groups. Significant differences between the groups was detected in age, the average age of subjects seemed to be highest in the *Treated* group. There was also significant difference observed in height and weight of subjects in the *Treated* group when compared to *Healthy* or OP_{susp} groups, however no statistically significant difference was observed between *Healthy* and OP_{susp} groups.

In the *treated* group (n =171), a total of 28 subjects were observed with osteoporotic BMD at the hip. The density indices showed high sensitivity (96.4-100.0%) and negative predictive (99.0-100.0%) value with DXA based osteoporosis diagnosis.

When analyzing OP_{susp} group divided in three age decades, an increase of osteoporosis prevalence from 8.0%, 10.9% up to 16.9% was noted in age groups of 50 to 59, 60 to 69 and 70 to 79 years, respectively. The sensitivity and specificity of US method with DXA osteoporosis diagnosis seemed to change with age as shown in table 5. The negative predictive value was high in all age groups for both DI_1 and DI_3.

The average reproducibility (CV_{rms} for six operators) of the DI_1 and DI_3 measurements were 3.2% (range 1.4-4.3%) and 3.3% (range 0.9-5.3%), respectively. Both

DI_1 and DI_3 showed statistically significant differences between the osteoporotic (T-Score ≤ -2.5 at the hip or femoral neck) and healthy group (T-Score > -2.5) ($p < 0.01$).

4. DISCUSSION

The performance of the DI in osteoporosis detection was evaluated in a large group of clinically relevant patient population in five private health care centers and one public university hospital. Based on the data reported here, the new ultrasound method could diminish the need for DXA referrals as only approximately 32% of the patients were recommended for DXA investigation, if the method were applied in accordance to ISCD and NOS recommendations (5, 6, 12).

The study population consisted mainly of those under osteoporosis suspicion (90%) *i.e.* had at least one risk factor for osteoporotic fracture and was referred to DXA examination by physician. Yet, the prevalence of osteoporosis was surprisingly only 11.2%. This may be partly explained by healthy selection bias at private sector, as subjects may have asked a referral for a DXA scan. On the other hand, 17.3% of the subjects had received treatment for osteoporosis. Patients with vertebral fractures or osteoporosis in the spine were not evaluated in the present study, and this may explain the discrepancy between the number of osteoporotic and treated subjects. In accordance to the Finnish osteoporosis guidelines, some of the subjects may have received treatment also based on high FRAX risk without BMD information.

The *healthy* group was included in the study to investigate differences in the US parameters in comparison to subjects “at risk” or to those that have received treatment for osteoporosis. The data suggests that the DI values as well as the $Ct.Th_{prox}$ are lower in *treated* group when compared to *healthy*. One should note that the groups were not perfectly balanced by number or subject characteristics, which may have an effect to this result. It should be also noted that all subjects in OP_{susp} group were referred to DXA examination, which can be considered as certain pre-selection since physician may have considered also other risk factors that were not controlled in the present study.

The presented ultrasound method with previously published thresholds applied, showed high sensitivity (85 to 93%) and specificity (82%) for detection of osteoporotic BMD. This finding is in line with the predicted performance of thresholds by Blake et al. (7) *i.e.* when at least 70 healthy and 70 osteoporotic patients are used for development of the thresholds, there is 95% confidence that the true sensitivity and specificity does not fall below 80%. The sensitivity and specificity of the method seemed to be associated with age. The results reported here included a small number of osteoporotic patients and therefore the estimates for positive predictive value are unreliable and may be underestimated. For the same reason, the age dependency of sensitivity and specificity may be partly caused by the domination of healthy subject in the population. Nonetheless, as the average T-score values of the patients under the lower threshold of DI were low, according to the NOS guideline the use of the triage approach is unlikely to have any significant effect on the efficacy of treatment (4). One should note that reported sensitivity and specificity is achieved only when subjects between the thresholds by ultrasound will be examined by axial DXA.

The present results are in line with those reported previously in US population (13). The reported sensitivity and specificity in the present study were higher, which can be mostly explained by the challenging population in the study conducted in the US as large portion (41%) of the study population had T-score near -2.5 *i.e* differences between the femoral neck and /or total hip BMD measures and the osteoporosis cut-point values were less than the precision errors of used densitometers. In the present study, the average reproducibility of the ultrasound method was approximately 3%, which is in line with CVs ranging from 1.2% to 3.4% reported in previous studies (9, 13). As with the axial DXA (reproducibility approximately 1.0% and 2.0% at spine and femoral neck, respectively(14, 15)), the effect of higher reproducibility may decrease the classification performance of patients with the proposed method (16). Hence, the reproducibility can affect the sensitivity and specificity in detection of subject with osteoporotic BMD. It should be also noted that in the present study

application of the method aimed at detecting osteoporotic areal BMD, and did not assess other options, like finding subjects at high risk of fracture nor evaluating cases with e.g. vertebral fractures. Naturally, if the aim is to identify only subjects with areal BMD in osteoporotic range, the method may miss those patients who have a high fracture risk but normal or osteopenic areal BMD.

The threshold approach has been also investigated in other devices based on the use of either X-rays or ultrasound at peripheral locations (5-7). The 90% sensitivity and specificity thresholds have been suggested for peripheral DXA devices, where the percentage of those needing additional DXA examination varied from 39% to 50% (7). However, no studies to our knowledge are available confirming and evaluating these findings. For one calcaneal ultrasound device the same approach has been applied and 56% need for additional DXA testing was reported (6). For axial transmission ultrasound technique, the number is higher at 60% - 75% (17). Our findings in the present study compare favorably with only 32% of the subjects in need for additional DXA. In addition, presently the heel ultrasound and axial ultrasound devices provides only T-score classifications (not above mentioned 90% sensitivity/specificity thresholds). Therefore, most of patients with osteoporotic BMD are classified to be healthy (for -2.4 T-score value with heel ultrasound the sensitivity is 22% (17)). By using these ultrasound devices most of the patients who should be treated do not receive antiosteoporosis treatment.

To conclude, the results in the present study are in line with the previous findings for the performance of examined ultrasound technique in the detection of osteoporotic hip BMD. The study data demonstrate that a strategy of combining ultrasound measurement with added DXA measurements in cases with intermediate ultrasound results (about 30%) can be useful for identifying subjects at risk for a low bone mineral density in the osteoporotic range.

References

1. Kanis JA. Diagnosis of osteoporosis and assessment of fracture risk. Lancet. 2002;359(9321):1929-36.

2. Center JR, Nguyen TV, Schneider D, Sambrook PN, Eisman JA. Mortality after all major types of osteoporotic fracture in men and women: an observational study. Lancet. 1999;353(9156):878-82.

3. Keene GS, Parker MJ, Pryor GA. Mortality and morbidity after hip fractures. BMJ. 1993;307(6914):1248-50.

4. National Osteoporosis Foundation. Clinician's Guide to Prevention and Treatment of Osteoporosis. Washington, DC, National Osteoporosis Foundation. 2013.

5. Hans DB, Shepherd JA, Schwartz EN, Reid DM, Blake GM, Fordham JN, et al. Peripheral dual-energy X-ray absorptiometry in the management of osteoporosis: the 2007 ISCD Official Positions. J Clin Densitom. 2008;11(1):188-206.

6. Krieg MA, Barkmann R, Gonnelli S, Stewart A, Bauer DC, Barquero L, et al. Quantitative Ultrasound in the Management of Osteoporosis: The 2007 ISCD Official Positions. J Clin Densitom. 2008;11(1):163-187.

7. Blake GM, Chinn DJ, Steel SA, Patel R, Panayiotou E, Thorpe J, et al. A list of device-specific thresholds for the clinical interpretation of peripheral x-ray absorptiometry examinations. Osteoporos Int. 2005;16:2149-2156.

8. Karjalainen J, Riekkinen O, Töyräs J, Hakulinen M, Kröger H, Rikkonen T, et al. Multi-site bone ultrasound measurements in elderly women with and without previous hip fractures. Osteoporos Int. 2012;23(4):1287.

9. Karjalainen J, Riekkinen O, Töyräs J, Jurvelin JaK,H. New method for point-of-care osteoporosis screening and diagnostics. Osteoporos Int. 2016(27(3)):971-7.

10. Gluer CC, Blake GM, Lu Y, Blunt A, Jergas M, Genant HK. Accurate Assessment of Precision Errors: How to Measure the Reproducibility of Bone Densitometry Techniques. Osteoporos Int. 1995;5:262.

11. Shepherd JA, Lu Y, Wilson K, Fuerst T, Genant HK, Hangartner TN, et al. Cross-Calibration and Minimum Precision Standards for Dual-Energy X-ray Absorptiometry: The 2005 ISCD Official Positions. J Clin Densitom. 2006;9(1):31-36.

12. Patel R, Blake GM, Fordham JN, McCrea D, Ryan P. Peripheral X-ray absorptiometry in the management of osteoporosis. National Osteoporosis Society - Practical Guides. 2011.

13. Schousboe JT, Riekkinen O, Karjalainen JP. Prediction of Hip Osteoporosis by DXA Using a Novel Pulse-Echo Ultrasound Device. Osteoporos Int. 2016 Aug 4.

14. Shepherd JA, Fan B, Lu Y, Lewiecki EM, Miller P, Genant HK. Comparison of BMD precision for Prodigy and Delphi spine and femur scans. Osteoporos Int. 2006;17(9):1303-8.

15. Henzell S, Dhaliwal S, Pontiflex R, Gill F, Price R, Retallack R, et al. Precision error of fan-beam dual X-ray absorptiometry scans at the spine, hip, and forearm. J Clin Densitom. 2000;3(4):359-64.

16. 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. J Clin Densitom. 2007;10(3):239-243.

17. Clowes J, Peel NFA, Eastell R. Device-specific thresholds to diagnose osteoporosis at the proximal femur: an approach to interpreting peripheral bone measurements in clinical practice. Osteoporos Int. 2006;17:1293.