Osteoporos Int (2000) 11:203–210

Osteoporosis International

Original Article

Hand Ultrasound for Osteoporosis Screening in Postmenopausal Women

C. L. Benitez[1], D. L. Schneider[2], E. Barrett-Connor[3] and D. J. Sartoris[4]

[1]School of Medicine, [2]Department of Medicine, [3]Department of Family and Preventive Medicine and [4]Department of Radiology, University of California–San Diego, La Jolla, California, USA

Abstract. There is a need for low-cost screening methods to detect low bone mass (osteopenia or osteoporosis) in postmenopausal women. The utility of quantitative ultrasonography (QUS) of the hand was assessed for osteoporosis screening using the WHO criteria. Bone mineral density (BMD) was measured in 206 postmenopausal Mexican-American women at the total hip and lumbar spine by dual-energy X-ray absorptiometry (DXA). The amplitude-dependent speed of sound (AD-SoS) was measured in the phalanges by QUS. Subjects identified by DXA as having osteopenia or osteoporosis had significantly lower AD-SoS values in comparison with normals. Estrogen users had significantly higher spine and hip BMD and AD-SoS values compared with non-estrogen users. The areas under the receiver operating characteristic (ROC) curves (AUC) for AD-SoS to screen for osteoporosis (T-score $\leqslant -2.5$) at the spine or hip were 0.73 for all subjects, 0.74 for estrogen users and 0.68 for non-estrogen users. The AUC for non-estrogen users to screen for osteopenia (T-score -1 to -2.5) was 0.77. Performance comparisons of AD-SoS with SCORE (a risk factor questionnaire) and body weight showed AUC values of 0.73, 0.69 and 0.65, respectively. QUS was the superior screening test when considering both the AUC and the shape of the ROC curves. For non-estrogen users, the group at higher risk for osteoporosis, QUS correctly identified 31% as normal, and 62% as having low bone mass and needing DXA referral; and the remaining 7% were false negatives. These data suggest phalangeal QUS can be effectively used for screening osteoporosis in postmenopausal women.

Correspondence and offprint requests to: Diane L. Schneider, MD, MSc, UC San Diego, Department of Medicine, 9500 Gilman Drive, La Jolla, CA 92093-0638, USA. Tel: +1 (858) 534 6046. Fax: +1 (858) 534 5939. e-mail: dlschneider@ucsd.edu

Keywords: Bone density; Osteoporosis; Phalanx; Screening; Ultrasound

Introduction

Osteoporosis is a disease characterized by low bone mass and microarchitectural deterioration of bone tissue, leading to enhanced bone fragility and a consequent increase in fracture risk [1]. Prospective studies have shown bone mineral density (BMD) to be a strong predictor of fracture risk [2–4]. For osteoporotic hip fractures, the BMD assessment has the highest predictive value when done at the site of fracture risk [5]. Other studies have shown that pre-existing vertebral fractures and BMD predict future vertebral fractures [6]. Dual-energy X-ray absorptiometry (DXA) is the current gold standard for diagnosis of osteoporosis through a quantitative measurement of BMD. However, DXA is not adequate for population screening purposes due to cost constraints and availability. Consequently, there is a need for low-cost screening methods to detect low bone mass (osteopenia or osteoporosis) in postmenopausal women. If low bone mass is recognized, subsequent referral to DXA for a definitive diagnosis would be recommended. These pre-screening methods may serve as more efficient case-finders and subsequently increase the diagnostic utility of DXA scans. Substantial long-term savings on the treatment of osteoporotic bone fractures would become apparent as more people with low bone mass begin to be treated early.

Screening methods for low bone mass include medical history questionnaires to ascertain risk factors such as age, weight, ethnicity, hormonal factors, nutrition, medications, immobility and existence of diseases. Edelstein et al. [7] and Dawson-Hughes et al. [8] have

identified body weight as a strong and consistent marker of BMD at weight-bearing sites for people without fractures. Mazess et al. [9,10] have studied the prediction of spine and hip BMD from body weight. Recently, SCORE (Bone Measurement Institute, PA), a simplified patient questionnaire composed of six questions, was developed to screen for low bone mass in the femoral neck and optimize the use of DXA [11]. It uses six risk factors (age, weight, ethnicity, presence of rheumatoid arthritis, estrogen status and history of fractures) found to have the strongest association with femoral neck BMD using regression modeling.

Quantitative ultrasound (QUS) is now a rapidly emerging technology in the field of densitometry with a largely proven potential to assess bone fracture risk using heel measurements [12,13]. Multiple anatomic sites have been studied, most commonly the calcaneus, followed by the phalanges, radius, tibia and patella. All QUS sites investigated have a relatively small amount of soft tissue around bone. QUS measures the average speed of sound in bone and soft tissue. Therefore, the sites are restricted to those with thin layers of soft tissue to ensure that the measurements remain sensitive to the status of the bone. As a screening tool, QUS has low cost, portability, short scanning time, and does not involve ionizing radiation. Phalangeal QUS has been shown to be an effective method in the identification of osteoporosis[14–16] and fracture risk in cross-sectional and longitudinal studies [14,17].

The common parameters measured by different manufacturers include the speed of sound (SoS) or ultrasound transmission velocity (UTV) and the frequency-dependent attenuation or broadband ultrasound attenuation (BUA) of the sound beam as it traverses bone tissue. Attenuation of the QUS signal is a result of energy dissipation from the sound wave by absorption and scattering in the soft tissues, the bone and marrow. SoS and attenuation are both correlated with bone density and strength. Therefore, healthy bone will have a higher SoS and attenuation whereas osteoporotic bone will have a lower SoS and attenuation. It has been shown that both SoS and BUA are parameters that can predict fracture risk independently of BMD [14,18,19]. These two parameters reflect the overall strength of bone. Strength is a function of density, elasticity and microstructure. The latter is influenced by bone trabecular connectivity, separation and orientation [20]. The relationship of how SoS and BUA predict these bone mechanical properties needs to be explored further.

The aim of this study was to assess the precision and accuracy of phalangeal QUS as a screening tool for the identification of low bone mass (osteopenia and osteoporosis) defined by BMD at the lumbar spine or total hip. In order to identify high-risk groups for referral to DXA, receiver operator characteristic (ROC) analysis was used to study the sensitivity and specificity of the test and to select appropriate QUS threshold values that minimized false negative results. QUS was also compared with other low-cost screening methods: SCORE and body weight.

Subjects and Methods

Subjects

The study included 206 Mexican-American postmenopausal women from San Diego County who volunteered to participate in a skeletal health study. Exclusion criteria included diseases or conditions known to affect bone health, including Paget's disease, rheumatoid arthritis, long-term immobilization, chronic kidney disease and medications (fluoride, calcitonin, bisphosphonates, corticosteroids). Appropriate informed consent forms were obtained from every subject. All participants were mailed a Spanish or English medical questionnaire depending on their native language. This included questions on reproductive and gynecologic history, smoking, alcohol, exercise habits and use of estrogen. Exercise was defined as physical activity for 20 min duration at least three times a week. Regular use of alcohol was defined as three or more drinks per week. Height and weight were measured to calculate the body mass index (BMI). Self-administered questionnaires were reviewed with the participant in clinic and the clinic bilingual staff obtained additional history.

Ultrasound

Phalangeal ultrasonography was carried out using a DBM Sonic 1200 (Igea, Italy). This instrument measures amplitude-dependent speed of sound (AD-SoS) transmission through the metaphysis of the proximal phalanx. The system uses 16 mm diameter QUS transducers with a frequency of 1.25 MHz. The transducers are mounted on a high-precision caliper (0.01 mm) that measures the thickness of the finger. The ultrasound signals are attenuated as they cross the soft and bone tissues in a lateromedial direction. Since osteoporotic bone produces a smaller beam arrival-amplitude than normal bone, it will not trigger a transducer response for the first portion of the signal. However, a response will occur as the received signal amplitude increases. This amplitude-dependent characteristic magnifies the differences between normal and osteoporotic patients. QUS measurements were performed on fingers two to five of both dominant and nondominant hands. For each hand, an average speed of sound was calculated for the four fingers. The probes were positioned over the metaphysis on the mediolateral phalangeal surfaces using the phalanx condyle as a reference point. The total scanning time with analysis on each hand took no more than 5 min. Measurements were performed on both hands to reduce individual variability error.

The device automatically advises the operator during the measurement when the BMI value of the subject is out of the range of 19–32 kg/m^2. The amount of soft tissue around the phalanx is related to the BMI of the subject and influences the AD-SoS measurement. Therefore, two subjects were excluded from the study because they had a BMI greater than 32 kg/m^2.

To quantify the reproducibility of this instrument, the inter-operator and intra-operator coefficients of variation (CV) were calculated. The inter-operator CV was obtained by having three different operators perform at least five sequential measurements on the same subject. The intra-operator CV was obtained by having the same operator perform at least five independent measurements on three subjects. Hand QUS had good reproducibility with an intra-operator CV of 0.61% and an inter-operator CV of 0.74%

Densitometry

Certified bone densitometry technologists measured BMD (Hologic QDR 2000, Bedford, MA) at the hip and lumbar spine (L2–4). BMD results are reported as *T*-scores compared with young adult normative data for women. At the hip site, the NHANES non-Hispanic white women were the reference group and for the lumbar spine the manufacturer's reference database was used. BMD (in g/cm^2), is based on the mineral content of a region of interest divided by the area of that region. The women were divided into three groups based on their *T*-scores: normal (BMD > -1 SD), osteopenic (BMD between -1 and -2.5 SD) and osteoporotic (BMD $\leqslant -2.5$ SD) for lumbar spine and total hip measurements independently. Each individual was classified as osteopenic or osteoporotic based on both lumbar spine and total hip measurements (positive for the disease if either of these two sites met the WHO BMD criteria).

Data Management and Analysis

To ascertain whether the AD-SoS values could identify normal, osteopenic and osteoporotic groups, and subjects on current estrogen status, the AD-SoS mean and 95% confidence intervals were calculated for each group. Comparisons between groups were made using an unpaired two-tailed Student's *t*-test. The Pearson correlation coefficients were calculated for weight, BMI, AD-SoS, lumbar spine BMD and proximal femoral BMD.

ROC analysis comparing AD-SoS with lumbar spine or total hip BMD was done to assess the accuracy of phalangeal QUS compared with DXA scans. Areas under the curve (AUC) and their standard errors were obtained. ROC curves to assess AD-SoS performance versus BMD for estrogen and non-estrogen users were plotted and compared. ROC analysis was used to select AD-SoS thresholds that had a high sensitivity yet maintained a moderate specificity. AD-SoS threshold values were selected from portions of the ROC curves optimizing for a high sensitivity with a specificity of at least 50%. The selected AD-SoS threshold values were used to obtain the percentage of subjects identified as abnormal who would be referred to DXA (true positives plus false positives), the percentage of those correctly identified as being normal (true negatives) and the percentage of those with osteoporosis who were missed (false negatives). For screening purposes the number of false negatives should be smaller than that of the false positives.

The performances of AD-SoS, SCORE and body weight were compared with ROC curves that illustrated their accuracy in identifying osteoporosis at either spine or hip for all patients. Although the screening accuracy of SCORE was proven for the femoral neck BMD, Lydick et al. [11] observed that SCORE also performed well when compared with spine BMD. *p* values were calculated for the AUC using the method of Hanley and McNeil [21,22]. Data were managed and analyzed with Microsoft Excel 95 and the add-in statistical function package Analyse-It (University of Leeds, UK). Analyse-It uses the ROC methodology formulated by Hanley and McNeil [21]. The screening methods were then combined to explore the possibility of improving the sensitivity and specificity attained by QUS alone for the risk non-estrogen group.

Results

A total of 206 postmenopausal women of Mexican-American heritage residing in San Diego County participated; their average age was 63 years. As shown in Table 1, the women stratified by estrogen status had similar habits: relatively small percentages were smokers or regularly used alcohol; and more than half exercised over three times a week. However, estrogen users had a significantly lower BMI even though body weight was not significantly different. Estrogen users, whose mean duration of use was 11.2 years (8.5–13.5, 95% CI), had significantly higher BMD and AD-SoS than non-estrogen users.

As displayed in Tables 2 and 3, the women were classified by the WHO criteria for BMD at the lumbar spine and total hip. By diagnostic categories at the lumbar spine, Table 2 shows there were no significant differences among the groups in terms of age, years since menopause and BMI. Women with normal BMD were heavier than women with osteoporosis at the spine. Both hip BMD and AD-SoS identified low bone mass at the spine. Similarly, Table 3 shows that both spine BMD and AD-SoS identified low bone mass at the hip. It was also seen that women with osteoporosis at the hip were older and leaner compared with those with osteoporosis at the spine.

Correlation coefficients for BMD and AD-SoS between different anatomic sites are shown in Table 4. Total hip and lumbar spine BMD had a correlation of 0.697. AD-SoS correlated better with lumbar spine BMD (0.415) than with any of the femoral sites. The correlations between AD-SoS versus femoral neck BMD and total hip BMD were almost identical (0.311 and 0.316, respectively). Both weight and BMI were negatively correlated with AD-SoS. The best correlation (0.939) was seen between right and left hand AD-SoS,

Table 1. Sample characteristics, bone density at spine and hip, and phalangeal ultrasound measurements for all subjects and by estrogen use

	All subjects (n = 206)	Current estrogen users (n = 74)	Non-estrogen users (n = 132)
Age (years)	63.3 (62.3–64.3)	62.4 (60.8–63.9)	63.8 (62.5–65.1)
Years since menopause	18.1 (16.8–19.4)	17.8 (15.7–19.9)	18.2 (16.5–19.9)
Body weight (kg)	69.3 (67.7–70.8)	67.4 (65.1–69.8)	70.3 (68.2–72.3)
Body mass index (kg/m^2)	28.8 (28.2–29.5)	27.7 (26.9–28.6)**	29.4 (28.6–30.3)
L2–4 spine BMD (g/m^2)	0.912 (0.890–0.934)	0.977 (0.941–1.013)***	0.876 (0.850–0.902)
Total hip BMD (g/m^2)	0.813 (0.796–0.831)	0.855 (0.827–0.882)**	0.790 (0.769–0.811)
Left hand AD–SoS (m/s)	1977 (1966–1988)	2014 (1997–2032)***	1957 (1944–1969)
Current smoking	6.6%	5.3%	7.4%
Alcohol use ($\geqslant$3 drinks per week)	5.2%	6.6%	4.4%
Exercise ($\geqslant$3 times per week)	63.75%	69.7%	60.3%

Values are the mean (95% CI) or percent.
Compared with non–estrogen users: **p<0.001; ***p<0.0001.

Table 2. WHO criteria for lumbar spine BMD for all subjects

	Normal (n = 62)	Osteopenic (n = 96)	Osteoporotic (n = 48)
L2–4 spine T-score:	$T \geqslant -1$	$T = -1$ to -2.5	$T \leqslant -2.5$
L2–4 spine BMD (g/m^2)	1.104 (1.073–1.135)	0.880 (0.870–0.889)***	0.728 (0.713–0.744)***
Age (years)	60.8 (60.8–64.4)	63.3 (61.8–64.7)	64.1 (62.1–66.2)
Years since menopause	16.9 (14.6–19.2)	18.3 (16.3–20.3)	19.2 (16.5–21.)
Body weight (kg)	72.6 (69.8–75.3)	68.9 (66.6–71.2)*	65.7 (62.8–68.6)**
BMI (kg/m^2)	29.6 (28.6–30.6)	28.7 (27.7–29.7)	28.1 (26.6–29.7)
Left-hand AD-SoS (m/s)	2019 (2000–2038)	1975 (1959–1990)**	1929 (1912–1945)***

Values are the mean (95% CI).
Compared with the normal group: *p< 0.05; **p<0.001; ***p<0.0001.

Table 3. WHO criteria for total hip BMD for all subjects

	Normal (n = 73)	Osteopenic (n = 103)	Osteoporotic (n = 30)
Total hip T-score:	$T \geqslant -1$	$T = -1$ to -2.5	$T \leqslant -2.5$
Total hip BMD (g/m^2)	0.943 (0.926–0.960)	0.781 (0.771–0.790)***	0.610 (0.594–0.626)***
Age (years)	60.8 (59.4–62.2)	63.4 (62.0–64.8)*	69.0 (66.5–71.4)***
Years since menopause	15.3 (13.3–17.3)	18.2 (16.3–20.0)*	24.4 (21.4–27.5)***
Body weight (kg)	74.2 (71.6–76.8)	67.6 (65.5–69.7)***	63.0 (60.3–65.8)***
BMI (kg/m^2)	30.8 (29.7–31.8)	27.9 (27.0–28.8)***	27.3 (26.0–28.6)**
Left-hand AD-SoS (m/s)	2000 (1982–2019)	1974 (1959–1989)*	1933 (1911–1955)***

Values are the mean (95% CI).
Compared with the normal group: *p< 0.05; **p<0.001; ***p<0.0001.

Table 4. Correlation coefficients between body weight, BMI, hip and lumbar spine BMD, and left- and right-hand AD-SoS

	BMI (kg/m^2)	L2–4 spine BMD (g/m^2)	Total hip BMD (g/m^2)	Femoral neck BMD (g/m^2)	Left-hand AD-SoS (m/s)	Right-hand AD-SoS (m/s)
Body weight (kg)	0.868	0.215	0.364	0.352	−0.208	
BMI (kg/m^2)		0.123	0.308	0.291	−0.259	
L2–4 spine BMD			0.697	0.603	0.415	
Total hip BMD				0.873	0.316	
Femoral neck BMD					0.311	
Left-hand AD-SoS						0.939
Right-hand AD-SoS						

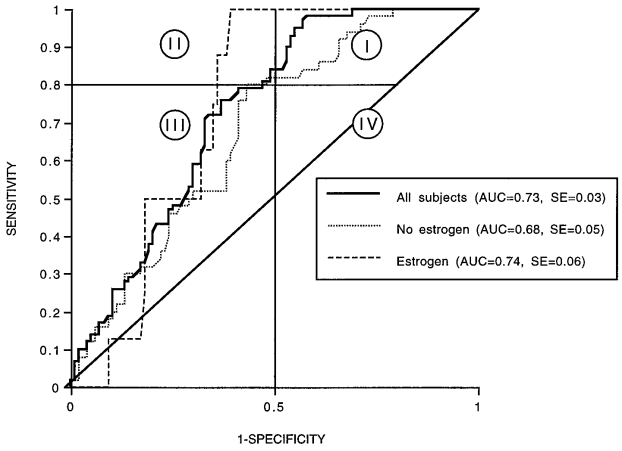

Fig. 1. AD-SoS ROC curves to screen for hip or spine osteoporosis for all subjects and, separately, for estrogen users and non-users. AD-SoS threshold values were chosen from quadrant II (sensitivity >80% and specificity >50%).

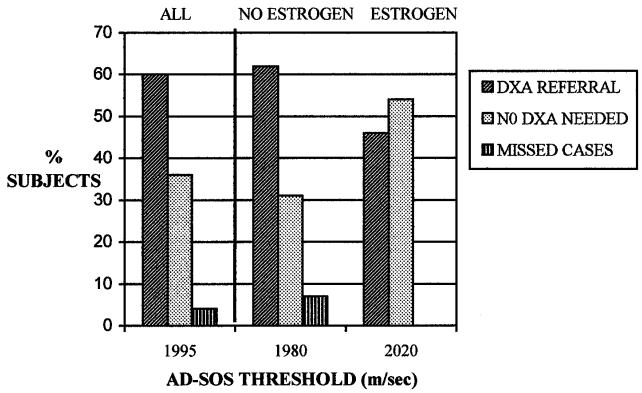

Fig. 2. DXA referrals with selected AD-SoS threshold values for osteoporosis screening at spine or hip for all subjects, estrogen users and non-users. *DXA referral*, true positives + false positives; *No DXA needed*, true negatives; *Missed cases*, false negatives.

demonstrating symmetry between the dominant and nondominant hand. For clinical use, only one hand (dominant or nondominant) with measurements on four fingers is necessary for a reliable hand QUS assessment.

Figure 1 shows the AD-SoS ROC curve for all subjects and, separately, the curves for the non-estrogen users and estrogen users for screening osteoporosis at either the spine or hip combined. The AUC values were 0.73, 0.68 and 0.74, respectively. AUC differences between estrogen users and non-users were not statistically significant. However, the curves separate for higher sensitivities, with the estrogen users showing a better specificity than non-users. The vertical and horizontal lines show arbitrary lower limits chosen for the sensitivity (80%) and specificity (50%) used to define phalangeal QUS threshold values. These lines divide the chart into four quadrants. Quadrant II represents the area that maximizes both sensitivity and specificity. AD-SoS threshold values were chosen from points in the curves over quadrant II. The threshold values, their ranges, and corresponding sensitivity and specificity were: 1980 (1765–2118) m/s, with 82% and 50% for non-estrogen users: 1995 (1765–2200) m/s, with 84% and 50% for all subjects; and 2020 (1839–2200) m/s, with 100% and 61% for estrogen users.

Figure 2 shows the DXA referral percentages when screening for spine or hip osteoporosis in all subjects, non-estrogen users and estrogen users after applying the selected AD-SoS threshold values from the ROC analysis. DXA referrals were lower for the estrogen users as well as the missed cases. Conversely, DXA referrals were higher for non-estrogen users with only a slight increase in missed cases. For all subjects, 60% would be referred to DXA, 36% were appropriately screened as normal and 4% were missed cases. For non-estrogen users, 62% would be referred to DXA, 31% were appropriately screened as normal and 7% were missed cases. For estrogen users, 46% would be referred to DXA, 54% were appropriately screened as normal and no cases were missed.

For screening purposes, more attention was given to the higher-risk group for fractures: the non-estrogen users. Table 5 is a numerical tabulation of the ROC curve to screen for spine or hip osteopenia and osteoporosis for non-estrogen users. It shows the AD-SoS threshold values with their corresponding sensitivity and specificity. Table 6 shows the AD-SoS performance in screening non-estrogen users for osteoporosis as well as osteopenia at the lumbar spine or total hip. The AUC for osteoporosis (T-score $\leqslant -2.5$ SD) was smaller than the AUC for osteopenia (T-score $\leqslant -1$ SD). As the BMD cutoff was changed from osteoporosis to osteopenia, the AD-SoS threshold required to maintain an adequate sensitivity and specificity increased. DXA referrals together with missed cases increased for osteopenia.

Figure 3 shows the ROC curves for AD-SoS, body weight and the SCORE questionnaire to screen for osteoporosis at the spine or hip for all subjects. The methods were compared using the AUC values as well as the shape of the curves. The AUC difference of AD-SoS and body weight had a p value of 0.07. The AUC of AD-SoS compared with SCORE had a p value of 0.40,

Table 5. Sensitivity and specificity for non-estrogen users (n = 132) for various threshold values of AD-SoS for osteopenia (T-score $\leqslant -1$) and osteoporosis (T-score $\leqslant -2.5$) at spine or hip

AD-SoS threshold (m/s)	T-score $\leqslant -1$		T-score $\leqslant -2.5$	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
1980	69	82	82	50
1990	73	65	84	41
2000	77	53	86	35
2010	83	53	94	32
2020	87	47	98	27
2030	88	41	98	23
2040	90	41	98	22

Table 6. AD-SoS performance for non-estrogen users (n = 132) for osteopenia (T-score $\leqslant -1$) and osteoporosis (T-score $\leqslant -2.5$) at spine or hip

	n	AUC	SE (AUC)	AD-SoS threshold	DXA referral (%)	No DXA (%)	Missed cases (%)
T-score $\leqslant -1$	115	0.770	0.057	2010	103 (78)	9 (7)	20 (15)
T-score $\leqslant -2.5$	50	0.682	0.046	1980	82 (62)	41 (31)	9 (7)

AUC, area under the curve.

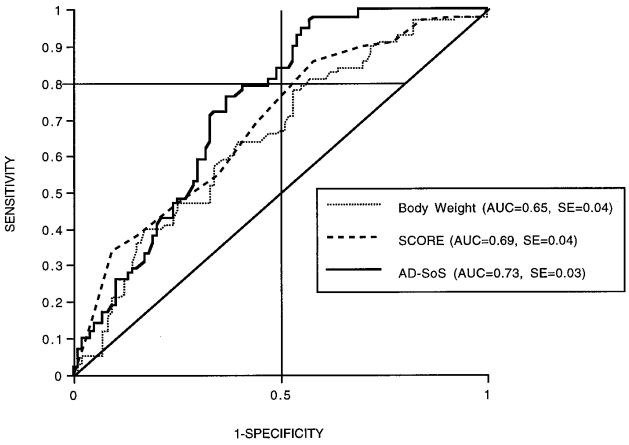

Fig. 3. ROC curve comparison between AD-SoS, SCORE and body weight to screen for osteoporosis at the spine or hip for all subjects.

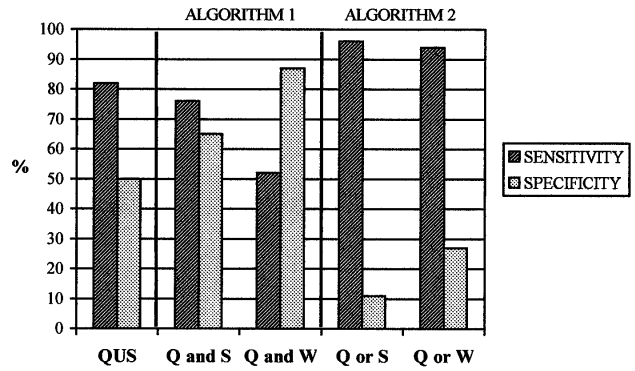

Fig. 4. Performance comparison using sensitivity and specificity for combinations of QUS with SCORE and body weight to screen for osteoporosis at the spine or hip on all subjects versus QUS alone. *Q*, QUS; *S*, SCORE; *W*, body weight. In algorithm 1 ('and'), both methods classify the subject as having osteoporosis at either spine or hip. In algorithm 2 ('or'), a positive diagnosis of osteoporosis at either spine or hip is required from either method.

whereas SCORE compared with body weight had a p value of 0.23. The most apparent differences between the curves were seen in the upper portion. For a specificity of 50%, AD-SoS had a sensitivity of 84%, SCORE had a sensitivity of 76% and body weight had sensitivity of 68%.

Fig. 4 shows the effect on sensitivity and specificity of combining QUS with SCORE, and QUS with body weight for all patients. In this study, QUS alone had a sensitivity of 82% and specificity of 50%. The methods were combined in two ways. In the first algorithm, the cases were detected only if they were identified by both methods. Fewer cases were detected here. The sensitivity worsened and the specificity improved. In the second algorithm, cases of osteoporosis at the spine or hip were detected if they were identified by either method. This algorithm detected more cases. The sensitivity was improved but the specificity worsened. None of these two algorithms achieved a sensitivity of more than 80% with specificity of 50% or above. The closest approximation was that of QUS and SCORE using the first algorithm with a sensitivity of 76% and specificity of 65%.

Discussion

In this cross-sectional study, it was shown that phalangeal QUS was able to identify Mexican-American women with osteopenia and osteoporosis at both the lumbar spine and total hip. As expected, women with osteoporosis at the hip were older and leaner than women with osteoporosis at the spine. The first site to show changes in BMD after menopause is the spine, while hip changes become more apparent after age 65 years [23,24]. The approach of using QUS to screen for osteoporosis at either spine or hip combined and not only at one site is considered important to identify low bone mass at any site regardless of age. Previous epidemiologic studies have reported osteoporosis prevalence based on the existence of low bone mass at any one site (lumbar spine, total hip or midradius) [25].

Both BMD and AD-SoS were significantly different in women who did or did not use estrogen. The average duration of estrogen use in this population was over 11 years. It has been previously reported by Ettinger et al. [26] and Felson et al. [27] that women who used estrogen for more than 7 years had significantly higher BMD. Looking at p values, AD-SoS was similar to spine BMD but better than total hip BMD in differentiating subjects who used estrogen. Presumably, the population of interest for low BMD mass screening is represented by the non-estrogen group, who could benefit from any type of bone-strengthening treatment. Based on the results shown on Fig. 2, a greater proportion of patients would be referred to DXA from the non-estrogen group compared with estrogen users, in which the majority is classified as not needing a DXA scan.

The best correlation of phalangeal QUS and DXA was that for the spine. One of the reasons for different correlations between sites is probably bone composition. Buckwalter et al. [28] reported that the phalangeal metaphysis has a greater proportion of cancellous bone

than cortical bone. This proportion makes the metaphysis resemble more a vertebral body than the total hip in young adults. Since early osteoporosis is more evident in cancellous bone, the phalanx seems to be an adequate site to study systemic bone mass changes. Another advantage of the phalanx is that it is a non-weight-bearing site that is less influenced by remodeling changes that could hide early bone mass loss. Kleerekoper et al. [29] have shown the phalanx to have the largest age-related variations in bone mass (by radiographic absorptiometry) among several peripheral and axial measurement sites.

In ROC curve analysis, the AUC represents the probability that a randomly chosen subject with osteoporosis defined by BMD at the spine or hip ranked with greater suspicion than a randomly chosen subject with no osteoporosis at either of these two sites [21]. Swetts [30] suggests that areas of 0.5 and 0.7 indicate low test accuracy, 0.7 to 0.9 moderate accuracy, and greater than 0.9 high accuracy. However, the AUC does not provide all the information necessary to assess the validity of the diagnostic test [31]. Therefore, this study also used the rule-out decision threshold achieved by setting minimum values for sensitivity and specificity. The rule-out threshold (a test with a high sensitivity) is used when the disease is serious and should not be missed, the disease is treatable, and false positive results do not have serious psychological or economic consequences for the patient [32]. The shape of the ROC curve, specifically the upper portion (high sensitivity), is important to validate performance in this study in addition to the AUC. The importance of using both of these elements is seen in the comparison of AD-SoS with SCORE and body weight. Even though having similar total AUC values, a clear difference is seen in the upper-portion of the curves (Fig. 3). The three curves separate, with AD-SoS having the highest specificities for sensitivity values above 50%. SCORE and body weight performed worse than QUS for higher sensitivity values. When combining methods to improve the accuracy for screening osteoporosis in the non-estrogen group, it was seen that either sensitivity or specificity could be improved but not both at the same time. One parameter improved at the expense of the other. Combining tests was not useful to optimize both sensitivity and specificity. AD-SoS alone remained the best screening method, with a sensitivity of 82% and specificity above 50%.

The use of QUS to screen for BMD at different sites has been studied previously by several groups [33–35]. Faulkner et al. [33] used linear regression analysis to predict BMD at different sites from SoS and BUA measurements at the heel. Herd et al. [34] and Young et al. [35] used ROC analysis to assess low bone mass as defined by BMD at the spine and femoral neck independently. Herd et al. [34] investigated three QUS parameters at the heel: velocity of sound, BUA and a combination of both. In all studies, the QUS correlation coefficients and ROC analysis showed results similar to this study, but the interpretation was somewhat different. Both a high sensitivity and high specificity were set as necessary conditions for screening. In this paper, we propose that sensitivity is the critical value to be maximized rather than specificity because QUS would eventually depend on DXA to reach a definitive diagnosis. In addition, we looked not only at the AUC of the ROC curve to make a judgment about the performance of the test, but also at the shape of the upper portion of the curve that gave additional information for screening. Our analysis was taken one step further by showing the number of women screened in and out from DXA together with missed cases of osteoporosis. In the future, QUS may provide the clinician with additional information, such as bone strength, and be used directly to predict fracture risk. However at present, it is important to determine how QUS could be used in combination with DXA.

This is the first phalangeal QUS study of Mexican-American women. In 1995, Looker et al. [36] reported NHANES III data on total hip BMD for women aged 50 years or older. There was a 16% prevalence of osteoporosis and 36% prevalence of osteopenia in Mexican-Americans compared with a 21% prevalence of osteoporosis and 39% prevalence of osteopenia in non-Hispanic white women. The present study had a similar prevalence of osteoporosis (15%) and higher prevalence of osteopenia (50%) at the total hip. Only a few QUS studies have provided a multi-ethnic comparison of SoS or BUA. Agren et al. [37] showed that values for the attenuation of the sound wave (BUA) were similar for a given age in Caucasian, Greek and Spanish women.

A limitation of the phalangeal ultrasound device is an excessive amount of soft tissue around the fingers. As a result of this, two subjects who had extremely thick fingers were excluded from the study. Soft tissue fat and edema between skin and bone usually cause a slower transmission of the QUS beam [38]. The study participants had a larger variation of weight in comparison with height. Therefore, higher body weight generally indicated a higher BMI or a more obese patient. These patients had more finger soft tissue and slower SOS. The negative correlation factors of body weight and BMI with AD-SoS are attributed to this finding. This concept also shows why QUS is limited to peripheral sites where there is less soft tissue around bone.

This work was a cross-sectional study on a postmenopausal population that needs to be further validated with larger groups. However, promising results were seen in the application of phalangeal QUS as a prescreening tool for DXA. QUS was able to identify groups of women with osteopenia and osteoporosis at both lumbar spine and total hip and distinguish women who used estrogen. A different approach on how to use and interpret the results of ROC analysis applied to QUS to screen for osteoporosis was also presented. When phalangeal QUS was compared with other screening methods that used risk factors only, it showed a better performance in identifying high-risk groups. These

findings confirm that phalangeal QUS could be useful for screening osteoporosis in postmenopausal women.

Acknowledgements. We would like to acknowledge our subject recruiter, Elva Benitez, together with our DXA technologists, Diane Claflin and Deidre Price, for the work and dedication they put into this project. We also thank Jaap Samson and Riccardo Isani, from Igea, Italy, and Deborah Morton for their technical support. Funding sources: The Sam and Rose Stein Institute for Research on Aging; The University of California–San Diego, Academic Geriatric Resource Program; and National Institute of Health, Summer Research Training Grant (5T35HLO7495).

References

1. Consensus development conference. Prophylaxis and treatment of osteoporosis. Am J Med 1991;90:107–10.
2. Melton LJ, Atkinsin EJ, O'Fallon WM, et al. Long-term fracture prediction by bone mineral assessed at different skeletal sites. J Bone Miner Res 1993;8:1227–33.
3. Black DM, Cummings SR, Genant HK, et al. Axial and appendicular bone density predict fractures in older women. J Bone Miner Res 1992;7:633–8.
4. Lips P. Epidemiology and predictors of fractures associated with osteoporosis. Am J Med 1997;103:S3–8.
5. Cummings SR, Black DM, Nevitt MC, et al. Bone density at various sites for prediction of hip fractures. The Study of Osteoporosis Fractures Research Group. Lancet 1993;341:72–5.
6. Ross PD, Davis JW, Epstein RS, et al. Pre-existing fractures and bone mass predict vertebral fracture incidence in women. Ann Intern Med 1991;114:919–23.
7. Edelstein SL, Barrett-Connor E. Relation between body size and bone mineral density in elderly men and women. Am J Epidemiol 1993;138:160–9.
8. Dawson-Hughes B, Shipp C, Sadowski L, et al. Bone density of the radius, spine, and hip in relation to percent of ideal body weight in postmenopausal women. Calcif Tissue Int 1987;40:310-4.
9. Mazess RB, Barden HS, Ettinger M. Radial and spinal bone mineral density in a patient population. Arthritis Rheum 1988;31:891–7.
10. Mazess RB, Barden HS, Dimka PJ, et al. Influence of age and body weight on spine and femur bone mineral density in QUS white men. J. Bone Miner Res 1990;5:645–52.
11. Lydick E, Cook K, Turpin J, et al. Development and validation of a simple questionnaire to facilitate identification of women likely to have low bone density. Am J Managed Care 1998;4:37–48.
12. Njeh CF, Boivin CM, Langton CM. The role of ultrasound in the assessment of osteoporosis: a review. Osteoporos Int 1997;7:7–22.
13. Genant HK, Engelke K, Fuerst T, et al. Noninvasive assessment of bone mineral and structure: state of the art [review]. J Bone Miner Res 1996;11:707–30.
14. Mele R, Masci G, Ventura V, et al. Three-year longitudinal study with quantitative ultrasound at the hand phalanx in a female population. Osteoporos Int 1997;7:550–7.
15. Ventura V, Mauloni M, Mura M, et al. Ultrasound velocity changes at the proximal phalanges of the hand in pre-, peri-, and postmenopausal women. Osteoporos Int 1996;6:368–75.
16. Sili Scavalli A, Marini M, Spadaro A, et al. Ultrasound transmission velocity of the proximal phalanxes of the non-dominant hand in the study of osteoporosis. Clin Rheumatol 1997;16:396–403.
17. Reginster JY, Dethor M, Pirenne H, Dewe W, Albert A. Reproducibility and diagnostic sensitivity of ultrasonometry of the phalanges to assess osteoporosis. Int J Gynaecol Obstet 1998;63:21–8.
18. Bauer DC, Gluer CC, Cauley JA, et al. Broadband ultrasound attenuation predicts fractures strongly and independently of densitometry in older women: a prospective study. Arch Intern Med 1997;157:629–34.
19. Hans D, Dargent-Molina P, Schott AM, et al. Ultrasonographic heel measurements to predict hip fracture in elderly women: the EPIDOS prospective study. Lancet 1996;348:511–4.
20. Gluer CC, Wu CY, Jergas M, et al. Three quantitative ultrasound parameters reflect bone structure. Calcif Tissue Int 1994;55:46–52.
21. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36.
22. Hanley JA, McNeil BJ. The method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148:839–43.
23. Sartoris DJ, Lenchik L. Current concepts in osteoporosis. AJR 1997;168:905–11.
24. Baran DT, Faulkner KG, Genant HK. Diagnosis and management of osteoporosis: guidelines for the utilization of bone densitometry. Calcif Tissue Int 1997;61:433–40.
25. World Health Organization. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Technical report series 843. Geneva: WHO, 1994:1–129.
26. Ettinger B, Genant HK, Cann CE. Long-term estrogen replacement therapy prevents bone loss and fractures. Ann Intern Med 1985;102:319–24.
27. Felson DT, Zhang Y, Hannan MT. The effect of postmenopausal estrogen therapy on bone density in elderly women. N Engl J Med 1993;329:1141–6.
28. Buckwalter JA, Glimcher MJ, Cooper RR, et al. Bone biology. I. Structure, blood supply, cells, matrix, and mineralization. J Bone Joint Surg Am 1985;77:1256–74.
29. Kleerekoper M, Nelson DA, Flynn MJ, et al. Comparison of radiographic absorptiometry with dual-energy X-ray absorptiometry and quantitative computed tomography in normal older white and black women. J Bone Miner Res 1994;9:1745–9.
30. Swetts JA. Measuring the accuracy of diagnostic systems. Science 1988;240:1285–93.
31. Henderson AR. Assessing test accuracy and its clinical consequences: a primer for receiver operating characteristic curve analysis. Ann Clin Biochem 1993;30:521–39.
32. Galen RS, Gambino SR. Beyond normality: the predictive value and efficiency of medical diagnoses. New York: Wiley, 1975.
33. Faulkner KG, McClung MR, Coleman LJ. Quantitative ultrasound of the heel: correlation with densitometric measurements at different skeletal sites. Osteoporos Int 1994;4:42–7.
34. Herd RJ, Blake GM, Miller CG. The ultrasonic assessment of osteopenia as defined by dual X-ray absorptiometry. Br J Radiol 1994;67:631–5.
35. Young H, Howey S, Purdie DW. Broadband ultrasound attenuation compared with dual-energy X-ray absorptiometry in screening for postmenopausal low bone density. Osteoporos Int 1993;3:160–4.
36. Looker AC, Johnston CC, Wahner HW. Prevalence of low femoral bone density in older US women from NHANES III. J Bone Miner Res 1995;10:796–802.
37. Agren M, Karellas A, Leahy D, et al. Ultrasound attenuation of the calcaneus: a sensitive and specific discriminator of osteopenia in postmenopausal women. Calcif Tissue Int 1991;48:240–4.
38. Johansen A, Stone MD. The effect of ankle oedema on bone ultrasound assessment at the heel. Osteoporos Int 1997;7:44–7.

Received for publication 2 April 1998
Accepted in revised form 27 July 1999