Osteoporosis Risk Assessment and Ethnicity:

Validation and Comparison of 2 Clinical Risk Stratification Instruments

Alvah R. Cass, MD, SM, Angela J. Shepherd, MD, Carol A. Carlson, BA
The University of Texas Medical Branch, Galveston, TX, USA.

BACKGROUND: Dual energy x-ray absorptiometry (DXA), coupled with early treatment, may reduce morbidity and mortality associated with osteoporosis. Clinical tools to enhance selection of women for DXA screening have not been developed or validated in an ethnically diverse population.

OBJECTIVE: To compare the performance of the osteoporosis risk assessment instrument (ORAI) and the simple calculated osteoporosis risk estimation (SCORE) instrument across 3 racial/ethnic groups to identify women who would benefit from DXA scans.

DESIGN: Blinded comparison of the instruments in a cross-sectional sample.

PARTICIPANTS: Two-hundred twenty-six postmenopausal women were recruited from a university-based family medicine clinic. Women with a prior diagnosis of osteoporosis or those taking bone active medications were excluded.

MEASUREMENTS: Participants completed a questionnaire that contained the ORAI and the SCORE questions; 203 completed a DXA scan.

RESULTS: The sensitivity and specificity for the ORAI (0.68, [0.49 to 0.88, 95% CI]; 0.66, [0.59 to 0.73, 95% CI]) and the SCORE instrument (0.54, [0.34 to 0.75, 95% CI]; 0.72, [0.65 to 0.78, 95% CI]) differed significantly from previous reports. Overall, the accuracy of the ORAI (66.5%) and SCORE instrument (70.0%) were similar (McNemar's test *P* value = .37). The accuracy between instruments differed significantly in African-American women (McNemar's test, *P* value <.001). In African Americans, the SCORE instrument correctly identified more women without osteoporosis, but missed 70% of those with osteoporosis.

CONCLUSIONS: The performance of the ORAI and SCORE instrument differed significantly from previous reports. Although both can reduce the use of DXA scans for screening for osteoporosis, lower sensitivities resulted in underrecognition of osteoporosis and may limit their clinical usefulness in an ethnically diverse population.

KEY WORDS: osteoporosis; diagnosis; women's health; cross-cultural medical issues.
DOI: 10.1111/j.1525-1497.2006.00459.x
J GEN INTERN MED 2006; 21:630–635.

O steoporosis, characterized by low bone mineral density (BMD) and loss of structural integrity, increases with age and is more prevalent in women than men. In the elderly, osteoporosis is a major risk factor for fractures of the wrist, vertebrae, and hip and accounts for substantial morbidity and mortality. Based on 1995 data, total direct medical expenditures for the treatment of osteoporotic fractures were estimated at \$13.8 billion.[1] Adjusting for inflation, expenditures exceeded \$17 billion in 2001 dollars.

Effective screening modalities for osteoporosis are available, as are effective prevention and treatment strategies.[2–12] However, these clinical options may be underutilized. In 2002, the U.S. Preventive Services Task Force (USPSTF) recommended that women in all racial/ethnic groups who are 65 years of age and older should be offered osteoporosis screening with dual energy x-ray absorptiometry (DXA)[13]; however, the National Osteoporosis Foundation (NOF) estimated that only 12% of women in this age group had been screened with DXA.[14] In 2003, a survey of patients receiving care in an academic setting reported that only 34% of white women who met NOF criteria, which are similar to the USPSTF screening recommendations, had received a DXA scan. Furthermore, only 8% of African-American women who met these criteria had received a DXA scan.[15] These findings demonstrate that DXA screening is underused, despite evidence that early recognition and treatment reduce osteoporotic fractures and the associated morbidity and mortality.

Several clinical risk stratification or screening instruments have been developed to identifying women who would most benefit from measurement of BMD by DXA scan[16–22] to diagnose osteoporosis. These instruments were developed in primarily white populations and have not been validated in multiethnic populations. In light of the 2002 USPSTF recommendations, these instruments may be useful in identifying postmenopausal women under 65 years of age at increased risk of unrecognized osteoporosis. Their role in women over 65 is debatable. This study compares the simple calculated osteoporosis risk estimation (SCORE) instrument[16–22] and the osteoporosis risk assessment instrument (ORAI)[16–22] across 3 racial/ethnic groups. We selected these instruments because the initial development and validation studies were methodologically sound and afforded an opportunity to compare 1 instrument that included race/ethnicity (SCORE) to another that did not (ORAI). In this study, we compared the operating characteristics of these instruments in postmenopausal women to identify women for *screening* versus *not screening* for osteoporosis with a DXA scan. We also compared our results with results from previous reports. Finally, we compared the accuracy of the screening decisions based on results of both instruments. We report findings for the entire sample and for each of the 3 racial/ethnic groups.

METHODS

Design

We designed a cross-sectional study of primary screening for osteoporosis that was conducted as a comparison of 2 clinical

risk stratification instruments designed to assign postmeno-pausal women to 1 of 2 groups: (1) those likely to have os-teoporosis and therefore most likely to benefit from DXA screening and (2) those unlikely to have osteoporosis and therefore least likely to benefit from DXA screening. The Human Subjects Institutional Review Board approved this study. All subjects signed a written informed consent.

Participants

We enrolled a sample of postmenopausal women, 45 years of age and older, receiving usual care at a university-based family practice clinic, which is a combined faculty and resident practice. Women, including non-Hispanic white, African American, and Hispanic participants, were recruited during a regularly scheduled visit. Since this study focused on detecting osteoporosis in women without a prior diagnosis of osteoporosis, we excluded women who previously had been diagnosed with osteoporosis. In addition, we excluded women who were taking bone active medication (e.g., bisphosphonates, calcitonin, etc.) for osteoporosis or osteopenia because of potential effects on BMD; women who had other bone disease (e.g., Paget's disease, hip replacement surgery) that could interfere with interpretation of the DXA scans; and women who exceeded the weight limit of the DXA scanner.

Measures and Instruments

Each participant completed a survey that included demographic data, medical history, and risk factors identified by each of the clinical risk stratification instruments under consideration. Women subsequently underwent measurement of BMD by DXA, which we used as the reference standard to classify women as normal, osteopenic, or osteoporotic. Based on the World Health Organization definitions,[23] results from DXA scans of total hip and total lumbar spine were designated as follows: normal bone density (T score ≥ -1.0); osteopenia ($-1.0 > T$ score > -2.5); or osteoporosis (T score ≤ -2.5). Participants' BMD was classified based on the lower T score for *either* the total hip *or* the total lumbar spine. Bone mineral density and corresponding T scores were based on reference standards provided by the Hologic DXA scanners (Hologic, Inc., Bedford, MA) used in this study. All but 4 DXA scans were performed on the same Hologic 1000 QDR 4500A machine in the General Clinical Research Center. The other 4 were performed on a similar machine, a Hologic 1000 QDR 4500W in the radiology department.

Osteoporosis Risk Assessment Instrument. The ORAI[20] was developed in a large cohort of predominately white women in Canada and relies on age, weight, and estrogen replacement therapy to classify women into *screen* and *do not screen* categories. The instrument was validated in a second sample of Canadian women.[20,24] Women with a score of 9 points or greater are referred for DXA.

Simple Calculated Osteoporosis Risk Evaluation. The SCORE instrument also was developed in a predominantly white population.[17] In addition to age, weight, and estrogen replacement therapy, the SCORE instrument includes race/ethnicity, history of rheumatoid arthritis, and history of nontraumatic fractures after age 45 to classify women into *screen* and *do not screen* categories. Women with a score of 6 points or greater

Table 1. Computation Rules for the ORAI and SCORE Instrument

Guideline/Rule	Points
*Osteoporosis Risk Assessment Instrument (ORAI)**	
Points are given for:	
Age	
≥ 75	15
65 to 74	9
55 to 64	5
Weight	
< 60 kg (132 lbs)	9
60 to 69 kg (132 to 154 lbs)	3
≥ 70 kg (154 lbs)	0
Estrogen use	
Not current	2
Current	0
Simple Calculated Osteoporosis Risk Estimation (SCORE)†	
Points are given for:	
Race/ethnicity	
Other than African American	5
H/O of rheumatoid arthritis	4
H/O of minimal trauma fracture after age 45	
Number of minimal trauma fractures of wrist, hip, or rib multiplied by (maximum of 12)	4
Age	
First digit of age in years multiplied by	3
Estrogen use	
Never used estrogen therapy	1
Weight	
Weight in pounds divided by 10 and truncated to integer multiplied by	-1

*ORAI threshold to refer for dual energy x-ray absorptiometry (DXA) scan is total points ≥ 9.
†Score threshold to refer for DXA scan is total points ≥ 6.

are referred for DXA. The scoring algorithms are summarized in Table 1.

Statistical Analysis

We used one-way ANOVA to compare interval-scaled continuous variables (age and weight) and Pearson chi square statistic for categorical variables e.g., use of estrogen, history of fracture, and history of rheumatoid arthritis, across the racial/ethnic groups.

Based on cutoff values of 9 for the ORAI and 6 for the SCORE instrument, we divided women into *screen* and *do not screen* categories and constructed classification tables to calculate sensitivities, specificities, predictive values, and accuracies for both instruments against DXA results. Blinded classification was assured using a computer algorithm to calculate and classify participants from results of the ORAI and SCORE instrument and DXA results. We compared values from our sample with values published by Lydick et al.[17] and Cadarette et al.[20,24] using the exact binomial test for a single proportion. We report the area under the receiver operator characteristic curve (AUC) for both instruments.

To directly compare the overall accuracy of both instruments, we constructed 2×2 tables comparing *correct* (true positives plus true negatives) versus *incorrect* (false positives plus false negatives) screening decisions for both instruments and obtained values for McNemar's test of equality of paired proportions.[25] We directly compared sensitivities and specificities for both instruments overall and for each racial/ethnic group using sample estimates and 95% confidence intervals.
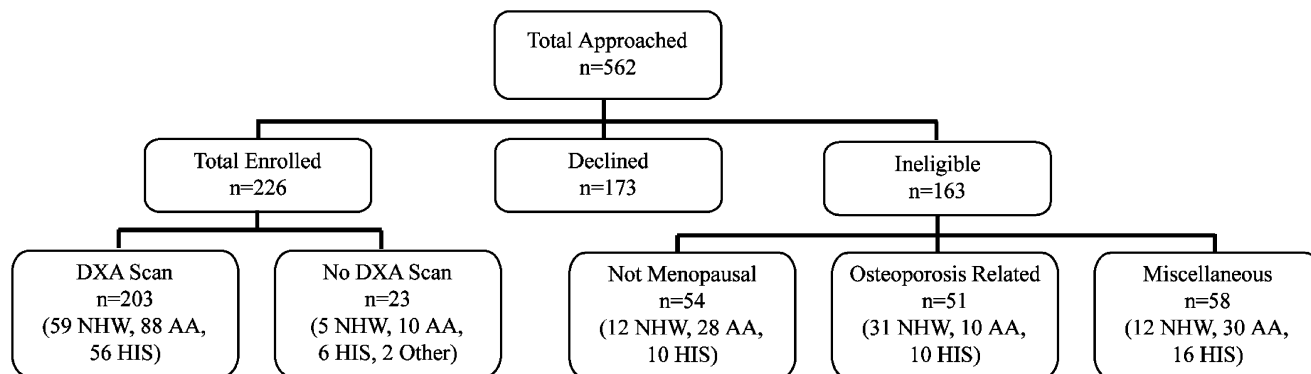
FIGURE 1. Enrollment schema. This figure illustrates the final disposition of all women approached for this study. Osteoporosis-related reasons included prior diagnosis of osteoporosis or current use of bone active medications other than estrogens. Miscellaneous reasons for exclusion included history of hip fracture, weight exceeded limit of DXA scanner, or other bone diseases. DXA, dual-energy x-ray absorptiometry; NHW, non-Hispanic white; AA, African American; HIS, Hispanic.

Sample Size and Power

McNemar's test of equality of paired proportions was used to make a direct comparison of the 2 instruments to determine if the instruments perform differently. McNemar's statistic tests the null hypothesis that the proportion of 1 of the discordant paired results is equal to 0.50. A sample size of 193 pairs has 90% power to detect a difference in proportions of 0.10 when the proportion of discordant paired results is expected to be 0.20 based on McNemar's test of equality of paired proportions with a 0.050 2-sided significance level.[26] The difference in proportions of 0.10 was estimated from findings published by Cadarette et al.[24] and represents the difference in the rates of correct screening decisions between the 2 instruments. Discordant paired results in excess of 20% should be sufficient to conclude that the tests are not equivalent.

RESULTS

From February 2002 to April 2003, we invited 562 women to participate in the study and enrolled 226 women (57% of the 399 found eligible). Eligibility and inclusion/exclusion results are found in Figure 1. We found no significant difference between participants and nonparticipants with regard to age, race/ethnicity, or weight. Comparisons of sociodemographic information, osteoporosis risk predictors, and bone mineral densities for the entire sample and 3 racial/ethnic groups are summarized in Table 2.

In our sample, the sensitivity and specificity for the ORAI were 0.68 and 0.66, respectively, and both differed significantly (binomial test for single proportion, *P* value <.001 for each comparison) from the sensitivity and specificity originally reported by Cadarette et al.[20] of 0.94 and 0.41, respectively, and from values reported in a subsequent validation study of 0.97 and 0.28, respectively.[24] The performance of the ORAI revealed minor variations across the 3 racial/ethnic groups. The sensitivity and specificity for the SCORE instrument were 0.54 and 0.72, respectively. Both values differed significantly (binomial test for single proportion, *P* value <.001 for each comparison) from the sensitivity and specificity originally reported by Lydick et al.[17] of 0.94 and 0.43, respectively, and subsequently by Cadarette et al. of 0.99 and 0.18, respectively.[24] The performance of the SCORE instrument was comparable for non-Hispanic white and Hispanic women. The sensitivity for the SCORE instrument was only 0.30 in

Table 2. Demographic Characteristics, Predictor Variables, and Bone Mineral Density Data

Variable	Total *n*=203	NHW *n*=59	HIS *n*=56	AA *n*=88	Significance
Age (y)	60.2 ± 9.6	60.1 ± 9.2	60.8 ± 9.4	59.9 ± 10.0	F, $df_{2,\ 200}$=0.14, P=.87
Weight (lbs)	187.1 ± 44.1	188.5 ± 45.2	171.9 ± 39.4	195.8 ± 44.0	F, $df_{2,\ 200}$=5.32, P=.006*
Never used ERT[†]	31%	20%	46%	28%	χ^2, df_2=9.64, P=.008
Current ERT[‡]	36%	52%	25%	31%	χ^2, df_2=11.08, P=.004
Rheumatoid arthritis[§]	2.5%	0.0%	7.1%	1.1%	χ^2, df_2=7.24, P=.03
H/O of minimal fracture	2.5%	3.4%	3.6%	1.1%	χ^2, df_2=1.14, P=.56
DXA results					
Normal	55%	68%	52%	49%	
Osteopenia	34%	24%	36%	40%	χ^2, df_4=5.67, P=.22
Osteoporosis	11%	8.5%	12%	11%	
Osteoporosis by site					
Hip only	1.0%	1.7%	1.8%	0.0%	
Spine only	7.9%	3.4%	7.1%	11%	χ^2, df_6=7.56, P=.27
Both	2.0%	3.4%	3.6%	0.0%	

*Mean difference between Hispanic and African American was significant (P =.004).

[†]Based on SCORE instrument.

[‡]Based on ORAI definition.

[§]This value represents the women with RA confirmed by review of the medical record and was the data used to calculate the SCORE value.

AA, African American; DXA, dual-energy x-ray absorptiometry; HIS, Hispanic; NHW, non-Hispanic white, ORAI, osteoporosis risk assessment instrument; SCORE, simple calculated osteoporosis risk estimation.

Table 3. Comparison of Sensitivities, Specificities, and AUC for ORAI and SCORE Instrument to Predict BMD, *T* score ≤ −2.5

	Sensitivity	Specificity	PPV	NPV	AUC
ORAI					
Overall	0.68 (0.49 to 0.88)	0.66 (0.59 to 0.73)	0.20 (0.11 to 0.29)	0.94 (0.90 to 0.98)	0.74 (0.63 to 0.84)
Non-Hispanic white	0.60 (0.04 to 0.94)	0.72 (0.57 to 0.81)	0.17 (0.0 to 0.33)	0.95 (0.86 to 1.0)	0.76 (0.60 to 0.91)
Hispanic	0.86 (0.47 to 0.99)	0.59 (0.44 to 0.72)	0.23 (0.07 to 0.39)	0.97 (0.87 to 1.0)	0.75 (0.59 to 0.91)
African American	0.60 (0.34 to 0.91)	0.67 (0.55 to 0.76)	0.19 (0.05 to 0.32)	0.93 (0.84 to 0.98)	0.69 (0.52 to 0.87)
SCORE					
Overall	0.54 (0.34 to 0.75)	0.72 (0.65 to 0.78)	0.19 (0.09 to 0.29)	0.93 (0.89 to 0.97)	0.67 (0.54 to 0.79)
Non-Hispanic white	0.80 (0.27 to 0.98)	0.63 (0.50 to 0.76)	0.17 (0.01 to 0.34)	0.97 (0.89 to 1.0)	0.68 (0.44 to 0.93)
Hispanic	0.71 (0.29 to 0.96)	0.49 (0.35 to 0.63)	0.17 (0.03 to 0.29)	0.92 (0.78 to 0.99)	0.69 (0.48 to 0.90)
African American	0.30 (0.00 to 0.56)	0.92 (0.86 to 0.98)	0.33 (0.00 to 0.61)	0.91 (0.85 to 0.97)	0.70 (0.51 to 0.89)

AUC, area under the receiver operator characteristic curve; NPV, negative predictive value; ORAI, osteoporosis risk assessment instrument; PPV, positive predictive value; SCORE, simple calculated osteoporosis risk estimation.

African-American women but the 95% confidence intervals overlapped with the other racial/ethnic groups. The specificity of the SCORE instrument was significantly higher in the African-American women (0.92) than in the other racial/ethnic groups. Table 3 details the overall operating characteristics of both instruments as well as the operating characteristics across racial/ethnic groups.

The AUC for the receiver operator characteristic curves for both instruments were similar in the examined population and across racial/ethnic groups (Table 3). Therefore, over a spectrum of potential cut points, the performance of the ORAI and SCORE instrument appears similar in our sample. However, based on the recommended cut points of 9 for the ORAI and 6 for the SCORE instrument, we found notable differences in performance of the instruments.

In our multiethnic sample, the accuracy (percentage of true positives plus true negatives) of ORAI was 66% compared with 70% for the SCORE instrument. In non-Hispanic white women the ORAI achieved an accuracy of 71% versus 64% for the SCORE instrument. The accuracy of the ORAI in Hispanic women was 62% compared with 52% for the SCORE instrument. In African-American women the ORAI yielded an accuracy of 66%, which was lower than the accuracy of the SCORE instrument, 85%. The accuracy of the 2 instruments when the Hispanic and non-Hispanic white women were combined were 67% for the ORAI and 58% for the SCORE instrument. Table 4 summarizes the results of the statistical analysis of the comparison of paired results for accuracies of the ORAI and SCORE instrument. We observed a statistically significant difference between the instruments in African-American women and a marginally significant difference when the non-Hispanic

and Hispanic women were combined. Hence, based on published cutoff values, the ORAI and SCORE instrument are not equivalent in African-American women.

DISCUSSION

In our study, the operating characteristics of the ORAI and SCORE instrument differed significantly from previous reports. The differences observed in sensitivity and specificity are not necessarily unexpected findings. The initial estimates of operating characteristics for many diagnostic tests tend to change when reevaluated in different settings; however, the magnitude of differences reported here are notable. Although the AUC for both instruments were similar in this study, both were significantly lower than previously reported.[17,20–24] The lower AUC also suggests that both instruments may have less discriminating power than previously assumed.

Using published cut points, the SCORE instrument tended to be slightly more accurate than the ORAI, but the ORAI was more consistent across racial/ethnic groups. The racial/ethnic differences were most apparent in African-American women. The SCORE instrument achieved superior accuracy by avoiding unnecessary DXA scans (predominately in African-American women), but failed to identify the majority (70%) of African-American women with osteoporosis. Both instruments maintained relatively high negative predictive values, overall and across ethnic groups, but in this setting where the prevalence of osteoporosis was 10.8%, even an indiscriminant test would have a high negative predictive value.

The intent of clinical risk stratification or clinically based prescreening is to strike a balance between minimizing testing

Table 4. Comparison of Accuracy of Paired Results for Screening Recommendations from the ORAI and SCORE Instruments

		SCORE		
	ORAI	Correct	Incorrect	McNemar's Test
Overall sample *N*=203	Correct	57% (116)	9.4% (19)	
	Incorrect	13% (26)	21% (42)	*P*=.37
Non-Hispanic white *n*=59	Correct	61% (36)	10% (6)	
	Incorrect	3.4% (2)	25% (15)	*P*=.29
Hispanic *n*=56	Correct	45% (25)	18% (10)	
	Incorrect	7.1% (4)	30% (17)	*P*=.18
African American *n*=88	Correct	62% (55)	3.4% (3)	
	Incorrect	23% (20)	11% (10)	*P*<.001
Combined Hispanic and non-Hispanic white *n*=115	Correct	53% (61)	14% (16)	
	Incorrect	5.2% (6)	28% (32)	*P*=.05

ORAI, osteoporosis risk assessment instrument; SCORE, simple calculated osteoporosis risk estimation.

to save health care dollars and not missing anyone who might benefit from diagnosis and treatment. The prevalence of disease, burden of illness, sensitivity and specificity, and accuracy of the prescreening algorithm determine this balance. The magnitude of the differences in operating characteristics reported in this study brings into question the usefulness of the instruments studied.[17,20–24] For example, consider a hypothetical cohort of postmenopausal women, 45 years of age and older, which reflects the prevalence of osteoporosis reported in NHANES III[27–29] and represented by 60% non-Hispanic white, 25% African-American, and 15% Hispanic women. Compared with universal screening, the ORAI, based on our findings, would reduce screening by 55%; however, it would miss 32% of women with osteoporosis. Similarly, the SCORE instrument would reduce screening by 64%, but would miss 46% of women with osteoporosis. The false negative rates may not be clinically acceptable, despite the considerable reduction in rates of screening.

Several reasons may explain the differences we observed. First, the SCORE instrument may underestimate the risk of osteoporosis in African-American women. Second, the operating characteristics of the instruments may vary according to anatomic site of osteoporosis. In our study, the anatomic site of osteoporosis differed across racial/ethnic groups. In particular, osteoporosis was limited to the lumbar spine in the African-American women. This observation is consistent with other reports[30–32] that show that in African-American women, the hip is less likely to be involved with osteoporosis than the lumbar spine. The SCORE instrument was actually developed in reference to the hip, but in Cadarette's[24] study, it was applied to the hip and lumbar spine and was associated with similar sensitivity but lower specificity. Finally, another potential source of inaccuracy for both instruments is the way in which weight is modeled. Both instruments attribute an increasing "protective" effect as weight increases. The women in our sample were 30 pounds heavier on an average than the women in the ORAI and SCORE development studies. African-American women were the heaviest group in our study. Therefore, in our study population, both instruments would yield lower scores, especially in African-American women, which could have resulted in lower sensitivities.

The prevalence of osteoporosis also differed from the expected. The observed prevalence in non-Hispanic white women was surprisingly low. This finding may be explained partially by the fact that non-Hispanic white women were more likely to have used estrogen/progesterone therapies, which could also have lowered the prevalence of osteoporosis in this group. They also tended to weigh more, which could have enhanced BMD, especially in the hip. Finally, a large proportion of the non-Hispanic white women were excluded based on a previous diagnosis of osteoporosis or current treatment with bone active medications (Figure 1). This leads us to suspect that a clinical bias in screening, operating before the initiation of the study, favored DXA screening for non-Hispanic white women. Taking the excluded cases into account and the average age of the sample, the number of women with osteoporosis more closely approximates the expected racial/ethnic distribution for osteoporosis.

Our study had several limitations. First, the limited number of participants in our sample yielded wide confidence intervals for the sensitivities and specificities of the instruments for each racial/ethnic group. However, the overall sensitivities and specificities were clearly different from other published reports. The small sample size may have contributed to the differences in prevalence of osteoporosis. Second, we did not extensively confirm the self-reported data contained in the 2 instruments. This was particularly problematic for subjects with rheumatoid arthritis, where we did confirm the diagnosis in the medical record. On further review we determined that the problem in assessing rheumatoid arthritis occurred primarily in the Hispanic population and was likely due to translation in the Spanish version. Finally, a preexisting clinical bias that favored previous DXA screening for non-Hispanic white women may have influenced the prevalence of osteoporosis in that group. This bias suggests that African-American and Hispanic women may not have been referred for DXA screening as frequently as non-Hispanic white women. This observation is consistent with other reports.[15,33]

The ORAI requires only a simple checklist. The SCORE instrument is much more cumbersome and requires mathematical manipulations and truncations. Moreover, the inclusion of rheumatoid arthritis and history of non-traumatic fractures adds another dimension to the SCORE instrument that goes beyond primary screening and risk stratification. Rheumatoid arthritis or a history of non-traumatic fractures probably justifies DXA scanning as a diagnostic test rather than a primary screening test.

Considering the overall performance of both instruments, the ease of use in the clinical setting, and the more consistent performance of the ORAI across racial/ethnic groups, we believe that the ORAI, which nearly replicates the 2002 recommendations of the USPSTF, is the better instrument for identifying women, from an ethnically diverse population, who should be referred for DXA scans. However, the poorer performance of the ORAI in our sample, compared with previous reports, precludes recommending widespread use of the instrument until more research is conducted in other diverse populations. The USPSTF recommendations offer an alternative prescreening strategy. However, the USPSTF recommendations were not developed in clinical studies, and to our knowledge, have not been validated in clinical studies.

Additional studies, conducted in larger populations that reflect the racial/ethnic distribution of other primary care populations and particularly include Asian women, are needed to compare and evaluate the utility of clinical risk assessment instruments and guidelines for osteoporosis screening. Similarly, the concept of clinical risk stratification for osteoporosis should be expanded to include men. Finally, clinical trials are needed to determine if screening algorithms, including the recommendations of the USPSTF, affect health-related quality of life, and the morbidity and mortality associated with osteoporosis and related fractures.

REFERENCES

1. **Ray N, Chan J, Thamer M, Melton M.** Medical expenditures for the treatment of osteoporotic fractures in the United States in 1995: report

from the National Osteoporosis Foundation. J Bone Miner Res. 1997;12:24–35.

2. **Black D, Cummings S, Karpf D, et al.** Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. Lancet. 1996;347:1535–41.

3. **Black D, Thompson D, Bauer D, et al.** Fracture risk reduction with alendronate in women with osteoporosis: the fracture intervention trial. J Clin Endocrinol Metab. 2000;85:4118–24.

4. **Bone H, Greenspan S, McKeever C, et al.** Alendronate and estrogen effects in postmenopausal women with low bone mineral density. J Clin Endocrinol Metab. 2000;85:720–6.

5. **Downs RJ, Bell N, Ettinger M, et al.** Comparison of alendronate and intranasal calcitonin for treatment of osteoporosis in postmenopausal women. J Clin Endocrinol Metab. 2000;85:1783–8.

6. **Ensrud K, Black D, Palermo L, et al.** Treatment with alendronate prevents fractures in women at highest risk: results from the fracture intervention trial. Arch Intern Med. 1997;157:2617–24.

7. **Delmas P.** Treatment of postmenopausal osteoporosis. Lancet. 2002; 359:2018–26.

8. **Hodsman A, Hanley D, Robert J.** Do bisphosphonates reduce the risk of osteoporotic fractures? An evaluation of the evidence to date. Can Med Assoc J. 2002;166:1426–30.

9. **Hosking D, Chilvers C, Christiansen C, et al.** Prevention of bone loss with alendronate in postmenopausal women under 60 years of age. N Engl J Med. 1998;338:485–92.

10. **Epstein S.** Postmenopausal osteoporosis: fracture consequences and treatment efficacy vary by skeletal site. Aging Clin Exp Res. 2000; 12:330–41.

11. **Harris S, Watts N, Genant H, et al.** Effects of risedronate treatment on vertebral and nonvertebral fractures in women with postmenopausal osteoporosis: a randomized controlled trial. JAMA. 1999;282:1344–52.

12. **McClung M, Geusens P, Miller P, et al.** Effect of risedronate on the risk of hip fracture in elderly women. N Engl J Med. 2001;344:333–40.

13. **U.S. Preventive Services Task Force.** Screening for osteoporosis in postmenopausal women: recommendations and rationale. Ann Intern Med. 2002;137:526–8.

14. **National Osteoporosis Foundation.** NOF applauds USPSTF recommendations on BMD tests. NOF Press Release, September 16, 2002

15. **Wanko S, Inouye L.** Gender and ethnic disparities in osteoporosis risk factors and screening. J Gen Intern Med. 2003;18(Suppl 1):266.

16. **Michaëlsson K, Bergstrom R, Mallmin H, Holmberg L, Wolk A, Ljunghall S.** Screening for osteopenia and osteoporosis: selection by body composition. Osteoporos Int. 1996;6:120–6.

17. **Lydick E, Cook K, Turpin J, Melton M, Stine R, Byrnes C.** 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.

18. **Black D, Palermo L, Pearson J, Abbott T, Johnell O.** SOFSURF: a simple, useful risk factor system can identify the large majority of women with osteoporosis. Bone. 1998;23(Suppl 5):S605.

19. **Weinstein L, Ullery B, Bourguignon C.** A simple system to determine who needs osteoporosis screening. Obstet Gynecol. 1999;93(Part 1): 757–60.

20. **Cadarette S, Jaglal S, Kreiger N, McIsaac W, Darlington G, Tu J.** Development and validation of the osteoporosis risk assessment instrument to facilitate selection of women for bone densitometry. Can Med Assoc J. 2000;162:1289–94.

21. **Koh L, Sedrine W, Torralba T, et al.** A simple tool to identify Asian women at increased risk of osteoporosis. Osteoporos Int. 2001;12:699–705.

22. **National Osteoporosis Foundation.** Physician's Guide to Prevention and Treatment of Osteoporosis. Washington, DC: National Osteoporosis Foundation; 2003.

23. **World Health Organization.** Assessment of Fracture Risk and its Application to Screening for Postmenopausal Osteoporosis. Geneva: World Health Organization; 1994.

24. **Cadarette S, Jaglal S, Murray T, et al.** Evaluation of decision rules for referring women for bone densitometry by dual-energy x-ray absorptiometry. JAMA. 2001;286:57–63.

25. **McNemar Q.** Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika. 1947;12: 153–7.

26. **Elashoff JD.** nQuery Advisor Version 4.0 User's Guide. Los Angeles, CA: Statistical Solutions Ltd; 2000.

27. **Looker A, Johnston C, Wahner H, et al.** Prevalence of low femoral bone density in older US women from NHANES III. J Bone Miner Res. 1995;10:796–802.

28. **Anonymous.** Osteoporosis among estrogen-deficient women—United States, 1988–1994. MMWR. 1998;47:969–73.

29. **Looker A, Wahner H, Dunn W, et al.** Updated data on proximal femur bone mineral levels of US adults. Osteoporos Int. 1998;8: 468–89.

30. **Looker A, Orwoll E, Johnston C, et al.** Prevalence of low femoral bone density in older US adults from NHANES III. J Bone Miner Res. 1997;12:1761–8.

31. **Silverman S, Madison R.** Decreased incidence of hip fracture in Hispanics, Asians, and Blacks: California hospital discharge data. Am J Public Health. 1988;78:1482–3.

32. **Finkelstein J, Lee M, Sowers M, et al.** Ethnic variation in bone density in premenopausal and early perimenopausal women: effects of anthropppometric and lifestyle factors. J Clin Endocrinol Metab. 2002; 87:3057–67.

33. **Binstock M.** Osteoporosis: risk factor prevalence and drug and densitometry utilization. Obstet Gynecol. 2000;95(Suppl 4):50S.