Journal of Clinical Densitometry, vol. 7, no. 4, 376–381, 2004

1094-6950/04/7:376–381/$25.00

Original Article

Use of the Lowest Vertebral Body T-Score to Diagnose Lumbar Osteoporosis in Men

Is "Cherry Picking" Appropriate?

Karen E. Hansen, MD,*[*,1,2,3] Nellie Vallarta-Ast, RT(R), CDT,[2,6] Diane Krueger, BS, CCRC,[2] Ron Gangnon, PhD,[4] Marc K. Drezner, MD,[2,3,7] and Neil Binkley, MD[2,3,5]

[1]Rheumatology Section, Department of Medicine, University of Wisconsin, Madison, WI and Veterans Hospital, Madison, WI; [2]Osteoporosis Clinical Center and Research Program, [3]Endocrinology, Diabetes and Metabolism Section, Department of Medicine, [4]Biostatistics and Medical Informatics Department, and [5]Geriatrics Section, Department of Medicine, University of Wisconsin, Madison, WI; and [6]Radiology Department and [7]GRECC at the William S. Middleton Veterans Hospital, Madison, WI

Abstract

In this study, we hypothesized that use of the lowest T-score among four lumbar vertebral bodies would lessen the impact of degenerative arthritis and other artifacts on diagnostic categorization at this site and increase study sensitivity, classifying more men with prior fracture as osteoporotic than the other two methods of lumbar spine analysis. Bone density studies of 533 male veterans measured between January and October 2002 were reviewed to determine diagnostic classification using the L1–L4 average, International Society for Clinical Densitometry (ISCD)-determined, and lowest lumbar vertebral body T-score. We calculated sensitivity and specificity of the three methods of spine analysis, using spine osteoporosis to indicate a positive test and prior fracture as the true indicator of osteoporosis. The lowest lumbar T-score performed with similar sensitivity and specificity to that of the lowest hip or wrist T-score in the ability to classify men with prior fracture as osteoporotic, whereas the average L1–L4 and ISCD-determined T-scores performed with lower sensitivity, but better specificity. In conclusion, this retrospective study suggests that use of the lowest vertebral body T-score among men increases diagnostic sensitivity of lumbar spine bone mass measurement. Prospective studies are needed to determine which of these three methods of lumbar spine analysis best predicts future fragility fracture in men and women.

Key Words: Osteoporosis; bone density; lumbar spine; T-score; analysis; sensitivity.

Introduction

Dual-energy X-ray absorptiometry (DXA) is routinely used to diagnose osteoporosis, and guidelines for interpreting bone densitometry in men and women are published by the International Society for Clinical Densitometry (ISCD) *(1).* These guidelines recommend use of the lowest T-score in the hip (excluding Ward's area) to diagnose osteopenia or osteoporosis. Assessment of lumbar spine bone mass is made using the average T-score of L1 through L4 unless a focal structural abnormality, lack of increase in bone mineral content or area from L1 to L4, or an unusual discrepancy in T-score between adjacent vertebrae are noted, allowing exclusion of a given vertebra *(2).*

The ISCD criteria for vertebral body exclusion were created to minimize the impact of artifacts on spine bone density measurement *(2).* However, consensus has not been reached regarding whether a single vertebral body is sufficient for diagnosis of

Received 01/12/04; Revised 04/12/04; Accepted 4/12/04.
* Address correspondence to Karen E. Hansen, MD, VA Hospital Room B5055, 2500 Overlook Terrace, Madison, WI 53750. E-mail: keh@medicine.wisc.edu

spine osteoporosis when three vertebral bodies meet criteria for exclusion. Furthermore, no prospective study has evaluated the ability of the lumbar spine T-score, determined by use of less than three vertebral bodies, to predict fragility fracture *(3–5).*

Importantly, disease processes such as degenerative arthritis, scoliosis, compression fractures, aortic calcification, or hardware could spuriously elevate the DXA-measured lumbar spine bone mass *(6).* Notably, degenerative arthritis is extremely common in the very population that develops osteoporosis, with 60% of women and 80% of men demonstrating spinal osteophytes by age 50 yr *(7).* Furthermore, spine osteoarthritis (OA) artificially elevates lumbar spine bone mass, with one study documenting that among older men, only 4% with, but 10% without, spine OA had lumbar spine osteoporosis *(8).* Thus, bone density measurement at the lumbar spine might miss a large number of people affected by osteoporosis. Consistent with this, a small, cross-sectional study showed that lumbar spine bone mass was not associated with fragility fracture in men *(9).*

Bone density testing plays an important role in the care of patients with low trauma fracture. Because other diseases might cause low-trauma fractures, it is common practice to obtain a bone mineral density (BMD) test when assessing a patient with such a fracture to confirm the diagnosis of osteoporosis. In addition, a bone density test is often useful as a baseline when starting osteoporosis therapy. Thus, the proper use and interpretation of BMD plays an important role in the care of patients with fracture.

In this study, we hypothesized that use of the lowest T-score among four lumbar vertebral bodies would lessen the impact of degenerative arthritis and other artifacts on diagnostic categorization at this site and increase study sensitivity, classifying more men with prior fracture as osteoporotic than the other two methods of lumbar spine analysis. In statistical analyses, men with prior fracture and low-trauma fracture were considered the "true positives" in order to objectively compare the three methods of lumbar spine analysis.

Materials and Methods

Subjects

Five hundred thirty-three male veterans referred for bone densitometry at the William S. Middleton Memorial Veterans Hospital between January 2 and October 5, 2002 were evaluated. The mean and median age of these men was 66 and 68 yr, respectively (range: 25–93 yr). Racial background was predominantly Caucasian (98%), with 2 Hispanic and 10 African-American men in the group. The average body mass index was 27.3 (range: 15.4–45.5 kg/m^2). Current and prior use of alcohol was reported in 107 and 66 men, respectively, and current and prior use of tobacco was noted in 106 and 129 men, respectively. Of the group, 208 men reported a prior clinical fracture, including 184 men with low-trauma fractures, defined as those occurring without trauma or upon falling from standing height.

Sites of fracture included vertebra (n = 63), hip (n = 10), wrist (n = 11), other sites (n = 57), and multiple fractures, including combinations of vertebral, wrist, hip, and other fractures (n = 67). The University of Wisconsin Human Subjects Committee reviewed this study and, as all identifiers were removed from the data, it was concluded that informed consent was not required.

BMD Measurement

One technologist using a GE Medical Systems Lunar Expert-XL densitometer (Madison, WI) acquired bone density measurements of the AP lumbar spine, proximal femur, and nondominant radius. Data were analyzed using software version 1.92 and the GE Lunar male normative database. The technologist recorded patient-reported clinical fractures at the time of DXA scanning, including site and circumstance.

Diagnostic Criteria

The World Health Organization (WHO) criteria were applied to diagnose low bone mass, with a BMD T-score of –1.1 to –2.4 considered osteopenia and a T-score of –2.5 or lower considered osteoporosis *(10).* The hip T-score was determined using the lowest T-score at the femoral neck, trochanter, or total femur, whereas the wrist T-score was the lowest T-score at the 33% or ultradistal radial site.

Bone density images and data were reviewed to determine the lumbar spine T-score by application of the ISCD criteria for vertebral body exclusion *(2).* One author (KEH) applied these criteria, noting which vertebrae were excluded and the indication(s) for exclusion. Specifically, vertebrae were excluded from analysis for a focal structural defect, unusual discrepancy in T-score (greater than one standard deviation), or a lack of increase in bone mineral content or area between adjacent vertebrae *(2).* In the case in which only one vertebra was excluded by applying these criteria, the average T-score of the remaining adjacent vertebrae was noted. For example, if the L4 vertebral body was excluded, then the L1–L3 T-score was recorded. Conversely, if the L2 vertebra was excluded, the L3–L4 T-score was reported. In the case in which only two nonadjacent vertebrae were reportable, the lower T-score was arbitrarily chosen to represent the ISCD-determined T-score. A second author (NB) applied the ISCD exclusion criteria to a random 10% of the bone density tests, and resulting concordance in the ISCD T-score was high (r = 0.98, p-value < 0.0001 by Pearson's correlation coefficient).

Statistical Analysis

Sensitivity and specificity are two important features of a diagnostic test and are often utilized when determining whether a new test is acceptable in clinical practice *(11).* Thus, sensitivity and specificity of the three methods in the ability to classify lumbar osteoporosis among those with prior fracture was calculated using the formulas Sensitivity = Probability (Test positive/ Fracture present) and Specificity = Probability (Test negative/ Fracture absent). In these analyses, men with prior fracture were considered "true positives" with osteoporosis, in order to objectively compare the three methods of spine analysis. For each method of spine analysis, the presence of lumbar spine osteoporosis was considered a positive test

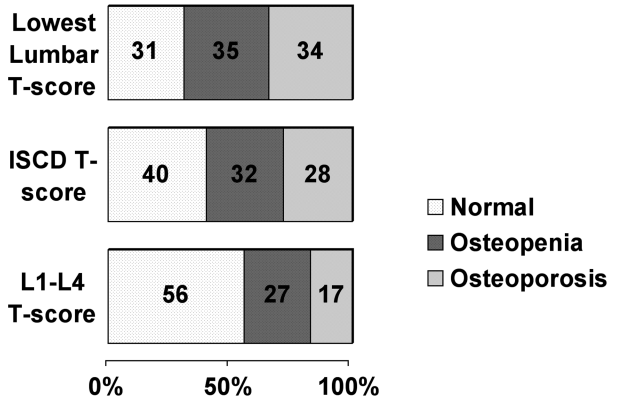

Fig. 1. Diagnostic categorization of lumbar spine bone mass using the average, ISCD, and lowest lumbar T-score.

result. The proportion of men with prior fracture found to be osteoporotic using the three methods of spine analysis was compared, with corresponding 95% confidence intervals (CIs). Calculations of sensitivity and specificity were also made using only men with low-trauma fractures. Comparisons among these three methods were analyzed by the χ^2 test, using Analyse-it Software (Leeds, UK).

Results

Numbers of Men With Osteoporosis Using Three Methods of Spine Analysis

Figure 1 summarizes the proportion of these 533 men with normal or low bone mass, using the three methods of lumbar spine analysis. Seventeen percent of men were classified as osteoporotic by the L1–L4 average lumbar T-score, 28% with application of the ISCD criteria, and 34% by lowest vertebral body T-score ($p < 0.0001$ for difference between the L1–L4 average and both of the two other methods of lumbar spine analysis, $p = 0.0551$ for difference between the ISCD-determined and lowest vertebral body T-score). Thus, there was a trend, albeit not significant, toward increased classification of osteoporosis with use of the lowest vertebral body T-score.

The Venn diagrams in Fig. 2 illustrate the proportion of men classified as osteoporotic, utilizing femur, radius, and each of the three methods of lumbar spine analysis. Notably, with the L1–L4 average T-score, only 14 additional men were classified as osteoporotic by lumbar spine measurement when femur and radius bone mass did not demonstrate osteoporosis. These data are very similar to a prior study showing that in men, bone densitometry of the L1–L4 spine adds little increased diagnostic ability to classify osteoporosis when radial and hip BMD are assessed *(12)*.

Using the ISCD-determined lumbar spine T-score, 36 additional men were classified as osteoporotic at the spine when T-scores at the femur and radius did not show osteoporosis. Thus, the ISCD-determined T-score classified an additional 23 men as osteoporotic (including 7 with fracture) compared to

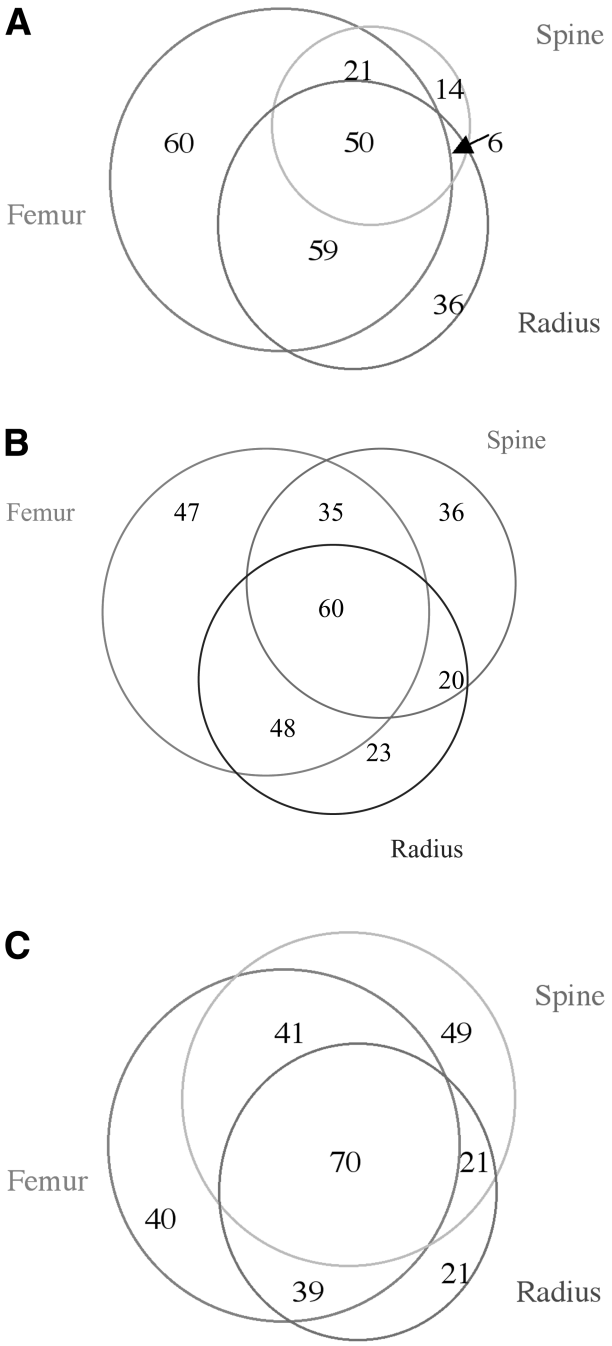

Fig. 2. Diagnostic categorization with three methods of lumbar spine analysis. **(A)** This Venn diagram illustrates the number of men with osteoporosis at various sites ($n = 246$, including 121 with fracture) when using the average lumbar T-score. **(B)** This Venn diagram illustrates the number of men with osteoporosis at various sites ($n = 269$, including 128 with fracture) using the ISCD-determined lumbar spine T-score. **(C)** This Venn diagram illustrates the number of men with osteoporosis ($n = 281$, including 133 with prior fracture) when using the lowest lumbar body T-score.

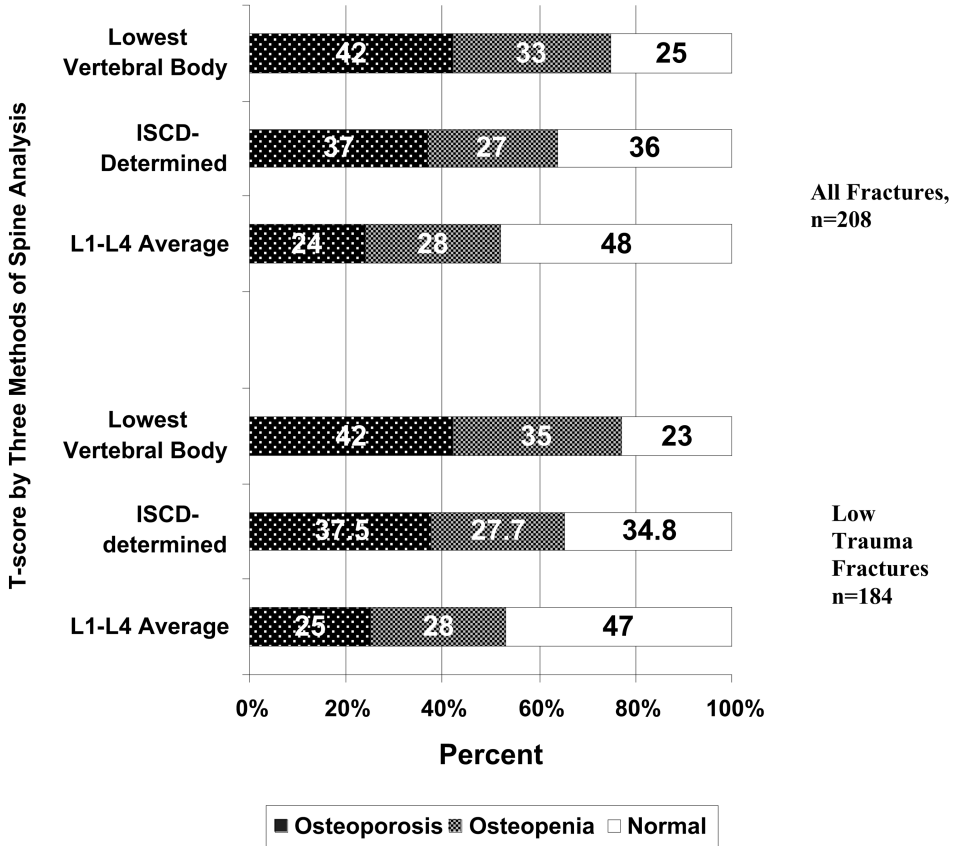

Fig. 3. Diagnostic categorization of men with prior fracture, using three methods of spine analysis. The lowest lumbar T-score was significantly more likely to classify men with fracture (n = 208) as osteoporotic than the L1–L4 T-score (18%; 95% CI = 13–24%) or the ISCD-determined T-score (5%; 95% CI = 2–8%). Likewise, the lowest lumbar T-score was significantly more likely to classify men with prior low-trauma fracture (n = 184) as osteoporotic than the L1–L4 T-score (17%; 95% CI = 12–23%) or the ISCD-determined T-score (4%; 95% CI = 2–8%). By χ^2 testing, the proportion of men with prior fracture (n = 208) found to have normal bone mass was significantly higher with the L1–L4 average T-score (p = 0.0171 compared to ISCD-determined T-score and $p < 0.0001$ compared to lowest vertebral body T-score) and significantly higher with the ISCD-determined T-score compared to the lowest lumbar vertebral body T-score (p = 0.0141). Likewise, among the subset of men with low-trauma fracture, the L1–L4 average T-score found a higher proportion with normal bone mass (p = 0.0259 compared to ISCD-determined T-score and $p < 0.0001$ compared to the lowest vertebral body T-score). Similarly, the ISCD-determined T-score classified a greater proportion of men with prior fragility fracture as having normal spine bone mass than the lowest vertebral body T-score (p = 0.0156).

the L1–L4 average T-score in combination with femur and radius data. Using the lowest vertebral body T-score, 49 additional men were osteoporotic when femur and radius testing did not show osteoporosis. Thus, use of the lowest vertebral body T-score classified an additional 12 men as osteoporotic compared to the ISCD-determined T-score, and an additional 35 men (including 13 with fracture) compared to the L1–L4 average T-score in combination with femur and radius data.

Diagnosis of Lumbar Osteoporosis in Men With Prior Fracture

In 208 men with prior fracture, lumbar osteoporosis was diagnosed in 24% (n = 49) using the L1–L4 average T-score, 37% (n = 76) using the ISCD-determined lumbar *T*-score, and 42% (n = 87) using the lowest lumbar vertebral body T-score. In men with fragility fracture (n = 184), lumbar osteoporosis was diagnosed in 25% (n = 46) using the L1–L4 average T-score, 38% (n = 69) using the ISCD-determined T-score, and 42% (n = 78) using the lowest lumbar vertebral body T-score. These data are summarized in Fig. 3.

Importantly, among all men with fracture (n = 208), normal bone mass was noted in 48% (n = 100) using the L1–L4 average T-score, 36% (n = 75) using the ISCD-determined T-score, and 25% (n = 51) using the lowest vertebral body T-score ($p <$ 0.05 by the χ^2 test for differences in proportions among the three groups; see caption of Fig. 3). Likewise, for the subset of men with low-trauma fracture (n = 184), normal bone mass was noted in 47% (n = 86) using the L1–L4 average T-score, 35%

Table 1
Sensitivity and Specificity of Bone Mass Measurement in the Diagnosis of Osteoporosis Among Men With Prior Fracture

Method	Sensitivity	Specificity
L1–L4 average T-score	24% (CI = 18–29%)	87% (CI = 83–91%)
ISCD-determined T-score	37% (CI = 30–44%)	77% (CI = 72–82%)
Lowest lumbar T-score	42% (CI = 35–49%)	71% (CI = 66–76%)
Lowest femur T-score	45% (CI = 38–52%)	70% (CI = 66–75%)
Lowest radius T-score	39% (CI = 36–43%)	79% (CI = 75–83%)

Note: The presence of lumbar spine osteoporosis (T-score ≤–2.5 was considered a positive test result.)

(n = 64) using the ISCD-determined T-score, and 23% (n = 42) using the lowest vertebral body T-score ($p < 0.05$ for differences among proportions in all three groups, *see* caption of Fig. 3).

Sensitivity and Specificity of Three Methods of Lumbar Spine Analysis

In the classification of osteoporosis among men with prior fracture, the L1–L4 average T-score showed a sensitivity of 24% and specificity of 87%, the ISCD-determined T-score showed a sensitivity of 37% and specificity of 77%, and the lowest lumbar vertebral body T-score showed a sensitivity of 42% and specificity of 71% (Table 1). By comparison, the lowest femur T-score showed a sensitivity of 45% and a specificity of 70%, whereas the lowest radial T-score showed a sensitivity of 39% and specificity of 79% in the ability to classify men with prior fracture as osteoporotic. In summary, the lowest lumbar vertebral body T-score demonstrated a diagnostic sensitivity similar to that of femur and radius analysis, whereas the L1–L4 average and ISCD-determined T-scores showed a lower sensitivity, but better specificity, in the ability to diagnose osteoporosis among those with prior fracture. The sensitivity and specificity of the three methods were not significantly altered when using the subset of men with fragility fracture (data not shown).

Discussion

In the current study, use of the lowest vertebral body T-score classified an additional 18% (95% CI = 13–24%) of men with prior fracture as osteoporotic than the T-score as determined by the L1–L4 average value and an additional 5% (95% CI = 2–8%) when compared to the ISCD-determined lumbar spine T-score. Thus, the lowest vertebral body T-score method had a higher sensitivity, but lower specificity, than the other two methods of lumbar spine analysis.

Measurement of bone mass at the spine and hip is recommended to increase the sensitivity of DXA in the diagnosis of osteoporosis *(13).* However, DXA-measured bone mass in the lumbar spine might be affected by degenerative arthritis, scoliosis, compression fractures, aortic calcification, or hardware *(6).* Likely as a direct result of this observation, a study among men living in Rochester, MN found that lumbar spine bone mass was not associated with fragility fracture *(9).* This lack of clinical utility demands evaluation of alternative strategies to detect low lumbar spine bone mass in men.

In this regard, the ISCD-determined or lowest lumbar vertebral body T-score could theoretically minimize the impact of other diseases affecting spine BMD, thereby making imaging at this site more sensitive in low bone mass detection. However, neither the ISCD-determined T-score nor the T-score by "cherry picking" the lowest lumbar vertebral body, has been prospectively studied to determine its ability to predict future fracture. Such prospective studies are necessary and, furthermore, will evaluate the possibility that these two approaches might excessively classify men as osteoporotic.

Indeed, excessive diagnosis of osteoporosis could potentially result from alternative methods of spine analysis. However, in this study, the ISCD-determined or lowest vertebral body T-score in combination with femur and radius BMD measurement did not dramatically increase the number of men classified as osteoporotic. Specifically, the ISCD-determined and lowest vertebral body T-score approaches classified an additional 4.3% and 6.5% of these men, respectively, as having osteoporosis. These results suggest that the potential for excessive diagnosis is low.

Appropriately, concerns have been raised regarding use of a single vertebral body to guide diagnosis and therapy of osteoporosis because of lower accuracy and precision when measuring smaller areas of bone *(2).* Although the DXA accuracy of three vertebral bodies is approx 9% based on a study of 11 cadavers, to our knowledge single vertebral body DXA accuracy has not been studied *(14).* Thus, it is possible that an isolated process could cause decreased bone mass in a single vertebral body without causing systemic osteoporosis.

Limitations of this retrospective study include the use of self-reported clinical fractures; serial radiographs to classify asymptomatic fragility fracture were not performed. Prior fractures and low-trauma fractures were considered the true indicator of osteoporosis, as a means of objectively comparing the three methods of spine analysis; some of these fractures might have been related to diseases other than osteoporosis. In addition, if ISCD criteria required exclusion of nonadjacent lumbar vertebrae, all remaining vertebrae could not be averaged to provide a final T-score, because of the limitations of the software used for this study. However, we arbitrarily chose the lowest of the remaining lumbar vertebrae when determining the ISCD T-score, to avoid biasing the lowest lumbar vertebral body T-score over that of the ISCD-determined T-score, as the better method to classify men with prior fracture as osteoporotic. With respect to decisions on vertebral body exclusion, one notable advantage of "cherry picking" was the

removal of subjective decisions by interpreters on whether to exclude a vertebral body for focal structural defect.

In conclusion, this retrospective study suggests that use of the lowest vertebral body T-score among men increases the diagnostic sensitivity of lumbar spine bone mass measurement. However, a lower specificity was noted when using the lowest vertebral T-score. We hope that our study will spur an interactive discussion among the bone community regarding appropriate lumbar spine analysis. A prospective study is needed to compare all three methods of lumbar spine analysis in the ability to predict future fracture.

References

1. Lenchick L, Leib ES, Hamdy RC, Binkley NC, Miller PD, Watts NB. 2002 Executive summary: International Society for Clinical Densitometry Position Development Conference, Denver, CO, J Clin Densitom 5:S1–S3.
2. Hamdy RC, Petak SM, Lenchik L. 2002 Which central dual X-ray absorptiometry skeletal sites and regions of interest should be used to determine the diagnosis of osteoporosis? J Clin Densitom 5:S11–S17.
3. Jergas M, Breitenseher M, Gluer CC, et al. 1995 Which vertebrae should be assessed using lateral dual-energy X-ray absorptiometry of the lumbar spine. Osteoporos Int 5:196–204.
4. Petley GW, Taylor PA, Murrills AJ, Dennison E, Pearson G, Cooper C. 2000 An investigation of the diagnostic value of bilateral femoral neck bone mineral density measurements. Osteoporos Int 11:675–679.
5. Hodges D, Lenchik L. 2003 Relative precision error of various regions of interest at every skeletal site measured with DXA in clinical practice. J Clin Densitom 6(2):185 (abstract).
6. Genant HK, Engelke K, Fuerst T, et al. 1996 Noninvasive assessment of bone mineral and structure: state of the art. J Bone Miner Res 11(6):707–730.
7. Resnick D. 1996 Degenerative Disease of the Spine. In: Bone and Joint Imaging. 2nd ed. Philadelphia, PA: WB Saunders, 355–377.
8. Liu G, Peacock M, Eilam O, Dorulla G, Braunstein E, Johnston CC. 1997 Effect of osteoarthritis in the lumbar spine and hip on bone mineral density and diagnosis of osteoporosis in elderly men and women. Osteoporos Int 7(6):564–569.
9. Melton LJ, Atkinson EJ, O'Connor MK, O'Fallon WM, Riggs BL. 1998 Bone density and fracture risk in men. J Bone Miner Res 13;1915–1923.
10. The WHO Study Group. 1994 Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Geneva: World Health Organization.
11. Lu Y, Jin H, Genant HK. 2003 On the non-inferiority of a diagnostic test based on paired observations. Statist Med 22; 3029–3044.
12. Vallarta-Ast N, Krueger D, Binkley N. 2003 Densitometric diagnosis of osteoporosis in men. J Clin Densitom 5(4):383–389.
13. Pouilles JM, Tremollieres F, Ribot C. 1993 Spine and femur densitometry at the menopause: are both sites necessary in the assessment of the risk of osteoporosis? Calcif Tissue Int 52:344–347.
14. Ho CP, Kim RW, Schaffler MB, Sartoris DJ. 1990 Accuracy of dual energy radiographic absorptiometry of the lumbar spine: cadaver study. Radiology 176:171–173.