Published in final edited form as:

Osteoporos Int. 2014 September ; 25(9): 2181–2188. doi:10.1007/s00198-014-2744-z.

What Analytic Method Should Clinicians Use to Derive Spine T-scores and Predict Incident Fractures in Men? Results from the MrOS study

Karen E Hansen, MD, MS,

Rheumatology Division, Department of Medicine, University of Wisconsin School of Medicine and Public Health, keh@medicine.wisc.edu

Robert D Blank, MD, PhD,

Endocrinology Division, Department of Medicine, Medical College of Wisconsin; Clement J Zablocki Veterans Affairs Medical Center, Milwaukee, WI, roblank@mcw.edu

Lisa Palermo, MA,

University of California, San Francisco LPalermo@psg.ucsf.org

Howard A Fink, MD, MPH, and

GRECC, Minneapolis Veterans Affairs Medical Center howard.fink@va.gov

Eric S Orwoll, MD

Oregon Health & Science University orwoll@ohsu.edu

for the Osteoporotic Fractures in men (MrOS) Study Research Group

Abstract

Summary—In this study, the area under the curve was highest when using the lowest vertebral body T-score to diagnose osteoporosis. In men for whom hip imaging is not possible, the lowest vertebral body T-score improves ability to diagnose osteoporosis in men who are likely to have an incident fragility fracture.

Purpose—Spine T-scores have limited ability to predict fragility fracture. We hypothesized that using lowest vertebral body T-score to diagnose osteoporosis would better predict fracture.

Corresponding Author Contact Information: Karen E. Hansen, M.D., M.S., Associate Professor of Medicine University of Wisconsin School of Medicine and Public Health Room 4124 MFCB, 1685 Highland Avenue, Madison, WI 53705 608-263-3457, keh@medicine.wisc.edu.

Authors' roles:

Study design: KEH, RDB, HAF, EO.

Study Conduct: KEH, RDB, LP, EO.

Data Collection: KEH, RDB, LP, EO.

Data Analysis: KEH, LP.

Data Interpretation: KEH, LP.

Drafting Manuscript: KEH.

Revising and Approval of Manuscript Content: KEH, RDB, LP, HAF, EO.

Disclosures: Karen E Hansen is a consultant to Takeda Pharmaceuticals and Deltanoid Pharmaceuticals. Eric Orwoll is a consultant to Amgen, Merck and Lilly and receives research support from Amgen, Merck and Lilly. Robert D Blank, Lisa Palermo and Howard Fink report no conflicts of interest.

Methods—Among men enrolled in the Osteoporotic Fractures in Men Study, we identified cases with incident clinical fracture (n=484) and controls without fracture (n=1,516). We analyzed the lumbar spine BMD in cases and controls (n=2,000) to record the L1-L4 (referent), the lowest vertebral body and ISCD-determined T-scores using a male normative database, and the L1-L4 T-score using a female normative database. We compared the ability of method to diagnose osteoporosis and therefore predict incident clinical fragility fracture, using area under the receiver operator curves (AUC) and the net reclassification index (NCI) as measures of diagnostic accuracy. ISCD-determined T-scores were determined in only 60% of participants (n=1205).

Results—Among 1,205 men, the AUC to predict incident clinical fracture was 0.546 for L1-L4 male, 0.542 for the L1-L4 female, 0.585 for lowest vertebral body and 0.559 for ISCD-determined T-score. The lowest vertebral body AUC was the only method significantly different from the referent method (p=0.002). Likewise, a diagnosis of osteoporosis based on the lowest vertebral body T-score demonstrated a significantly better NRI than the referent method (net NRI +0.077, p=0.005). By contrast, the net NRI for other methods of analysis did not differ from the referent method.

Conclusion—Our study suggests that in men, the lowest vertebral body T-score is an acceptable method by which to estimate fracture risk.

Keywords

bone densitometry; fracture; lumbar spine; men; net reclassification index; osteoporosis

Introduction

Prior studies [1-3] report that lumbar spine bone mineral density (BMD) does not predict incident clinical fractures as well as hip BMD, presumably due to spurious elevation of spine BMD by osteoarthritis, aortic calcification or ironically, compression fractures [4,5]. In an attempt to minimize the impact of artifacts on the DXA-measured lumbar spine BMD, the International Society for Clinical Densitometry (ISCD) recommends that interpreters exclude vertebrae with visible focal structural anomalies or T-scores that differ by more than one standard deviation from those of adjacent vertebrae [6]. However, the guidelines state that clinicians must use at least two vertebrae to determine the lumbar spine T-score. A single vertebral body T-score should not be used to diagnose osteoporosis [6] due to concerns about a high precision error when measuring a single vertebral body, and lack of data on the ability of a single vertebral body to assess fracture risk.

Theoretically, the ISCD-determined T-score would increase the diagnostic sensitivity of DXA for detecting osteoporosis, compared to the T-score derived using four lumbar vertebrae. However, no prospective studies have compared the diagnostic sensitivity and specificity of the ISCD-determined T-score to that of the L1-L4 T-score. Additionally, interpreters often disagree on which vertebrae to exclude when applying the ISCD criteria [7-9] potentially leading to different diagnostic categorization for a given patient. Finally, when three or more vertebrae have focal structural anomalies and/or T-score discrepancies, an ISCD-derived lumbar spine T-score cannot be reported for that patient [7]. All of these issues limit the utility of the ISCD guidelines in clinical practice.

Our prior research [10] suggested that use of lowest vertebral body BMD might maximize sensitivity of lumbar spine BMD for fracture prediction, but the study was limited by a small sample size and patient recall of fractures without adjudication of fracture events. A Canadian study [8] reported that in 20,478 women, the lowest vertebral body T-score improved clinical fracture prediction compared to the L1-L4 T-score, however this was not true in men (n=1,534). The Osteoporotic Fractures in Men (MrOS) Study is a prospective cohort study designed to determine risk factors for osteoporosis and fracture in older men [11,12]. The study provides a heretofore-unavailable opportunity to clarify the optimal method by which to analyze lumbar spine BMD in men. We designed a case-cohort study, using data from a subset of men enrolled in MrOS. We hypothesized that in men, the lowest vertebral body T-score would predict incident clinical fractures better than the ISCD-determined and mean L1-L4 T-score. We further hypothesized that a male normative database would be superior to a female normative database, in diagnosing osteoporosis and assessing fracture risk in men.

Materials and Methods

MrOS is a multi-center prospective cohort study designed to identify risk fractures for osteoporotic fracture [11,12]. Men were eligible for enrollment in MrOS if they were 65 years old, able to walk without assistance from another person, and had no history of bilateral total hip arthroplasty. 5,994 men from six centers (Birmingham, AL; Minneapolis, MN; Palo Alto, CA; Pittsburgh, PA; Portland, OR and San Diego, CA) were enrolled in the study between 2000 and 2002. Researchers measured participants' BMD at the spine and proximal femur during the baseline study visit, using Hologic QDR 4500 densitometers.

Study personnel contacted participants every four months for ~4.5 years to inquire about incident clinical fractures. Study personnel verified self-reported fractures by reviewing community radiology reports and clinical notes. If a radiology report was unavailable or unclear, a study radiologist reviewed copies of the X-rays to confirm or refute the fracture. A study physician reviewed and adjudicated all fracture events.

For the current study, we identified cases as men with any incident clinical fracture during the study. We defined a fragility fracture as one occurring by any of three low-trauma events: a fall from standing height or less, a fall from a minor change in positional height (e.g. stepping off a curb), or a minor traumatic event other than a fall, such as a cough-induced compression fracture. We randomly selected controls from the remaining men without incident clinical fracture.

The MrOS Steering Committee and the University of Wisconsin (UW) Human Subjects Committee reviewed and approved our study protocol. To preserve confidentiality among the men selected for the current study, the Data Coordinating Center (San Francisco, California, USA) de-identified the lumbar DXA scan images and ancillary results and replaced identifiers with a study code. Personnel shipped the baseline lumbar DXA scan images and ancillary results for cases and controls to the UW for scan examination and data extraction. The UW interpreters were blinded to the case-control status of each scan.

Using a standardized worksheet, one ISCD-certified physician (KEH) analyzed 2,000 lumbar spine BMD studies using a male normative database, and recorded the L1-L4, lowest vertebral body and the ISCD-determined T-score. The ISCD-determined T-score was derived by applying published criteria to define when vertebral bodies should be excluded from analysis [6,13]. To explore whether a female or male normative database should be used to determine the L1-L4 T-score in men, all lumbar spine DXA images were re-analyzed using the female NHANES normative database to derive and record the L1-L4 female T-score. We entered the four T-scores into a database in duplicate, and checked the data to verify accuracy.

Interpreters can have differing thresholds for vertebral body exclusion, primarily due to disagreement regarding the presence of focal structural anomalies.[7,9] For the current study, we randomly selected 10% of the spine images for interpretation by a second ISCD-certified physician (RDB) to assess the inter-rater reproducibility of the ISCD-determined T-score.

Statistical Analysis

Our primary study outcome was the ability of four methods of lumbar spine analysis to diagnose osteoporosis and predict clinical fragility fracture among men enrolled in the MrOS Study. The four methods were the L1-L4, lowest vertebral body and ISCD-determined T-scores derived using a male normative database, and the L1-L4 T-scores derived using a female normative database. We used the receiver operating characteristic area under the curve (AUC) values as our principal measure of diagnostic accuracy, with statistical comparisons to the L1-L4 male T-score as the referent. Based upon prior work [10] we estimated that the AUC would be 0.60 for the L1-L4 T-score derived using a male normative database, and 0.65 for the lowest vertebral body T-score and ISCD-determined T-score [10,7]. We estimated that correlation between the different methods of lumbar spine analysis on the same patient was 0.66 by Kendall's τ test [10,7]. Thus, a sample size of 2,000 men including 400 with incident fracture provided ~98% power to detect a significant difference in diagnostic accuracy between the analytic methods, using a two-sided test with 5% level. We next analyzed the ability of the different analytic methods to predict incident clinical fractures based on odds ratios. We also performed an exploratory analysis to compare the ability of the four methods of spine analysis to predict incident clinical vertebral fractures. For this analysis, we compared men with incident clinical vertebral fractures (n=63) to men without clinical fractures (n=1,509).

The net reclassification index (NRI) is a statistical tool used to determine whether a new method improves classification of individuals with an event as having the disease, and/or improves classification of individuals **without** an event as **not** having the disease.[14] We determined the NRI[14] when using the three methods of spine analysis to diagnose osteoporosis, compared to the referent (L1-L4 T-score derived using a male normative database). The NRI was used to evaluate whether each new method of spine analysis correctly reclassified men with a fracture as having osteoporosis (NRIevent), and reclassified men without an incident fracture as not having osteoporosis (NRI no event), compared to the referent method.

Results

We received DXA scans for 2,000 men (484 cases with incident clinical fracture and 1,516 controls without clinical fracture) from the MrOS Data Coordinating Center. We excluded seven scans due to foreign bodies in or near the lumbar spine (n=6) or incorrect labeling of vertebrae (n=1). An ISCD-determined T-score could not be determined for 788 men (40%), as three or more vertebrae had focal structural anomalies and/or T-score discrepancies. Thus, we recorded L1-L4 male, L1-L4 female and lowest vertebral body T-scores for 1,993 men and an ISCD-determined T-score for 1,205 men. Not surprisingly, the T-scores by various methods of analysis were highly correlated ($r=0.99$, L1-L4 male vs. L1-L4 female; $r=0.97$, L1-L4 male vs. lowest vertebral body and $r=0.96$, L1-L4 male vs. ISCD-determined T-score).

The 1,993 men with DXA scans permitting inclusion in this analysis had a mean age and body mass index of 74 ± 6 years and 27 ± 4 kg/m², respectively (Table 1). The majority of men (90%) were Caucasian. Compared to MrOS participants who were not included in the analysis (n=4,001), men selected for analysis were older, more likely to report fractures at or after age 50 years, had lower spine, total hip and femoral neck T-scores, lower measures of neuromuscular function, and were more likely to use bisphosphonate or tricyclic antidepressant therapy (Table 1). Among the 484 men with incident clinical fracture, there were 79 hip, 22 pelvis, 48 wrist and 63 vertebral fractures; the remaining fractures affected miscellaneous sites.

Not surprisingly, spine osteoporosis was diagnosed most often (Table 2) using the lowest vertebral body T-score, and least often using the L1-L4 T-score derived from a female normative database. Table 2 summarizes the percent of men with osteoporosis based on each method of spine analysis. A diagnosis of osteoporosis was more likely in men with fracture ($p<0.001$) by all four methods of spine analysis.

Two physicians derived the ISCD-determined T-score for 200 randomly selected scans. With application of the ISCD criteria for vertebral body exclusion, both interpreters excluded all vertebrae in 35% of scans (n=71). In 44% of scans (n=87), both interpreters derived T-scores. There was disagreement in 21% of scans (n=42), with one interpreter recording an ISCD-determined T-score and the other excluding all vertebrae. For the 87 scans in which both interpreters derived an ISCD-determined T-score, the T-score was highly correlated ($R=1.0$, $P<0.001$) with 100% agreement on the diagnosis of osteoporosis (spine T-score -2.5).

Compared to men without fracture (n=1,509), those who had experienced an incident fracture (n=484) were older, had lower spine and hip T-scores, were more likely to be Caucasian, report a fracture at or after age 50 and/or use bisphosphonates or serotonin uptake inhibitor therapy, and had lower measures of neuromuscular function (Table 1).

Table 3 summarizes the AUC values for four methods of spine analysis as a means of predicting incident clinical fractures, including clinical vertebral fractures. When predicting all clinical fractures, the AUC for the lowest vertebral body method was the only value significantly higher than the referent, whether comparing values in 1,993 men (AUC 0.558

vs 0.533, $p=0.004$) or values in the subset of 1,205 men for whom an ISCD-determined T-score was derived (AUC 0.585 vs. 0.546, $p=0.002$). When predicting clinical vertebral fracture, the AUC for the lowest vertebral body method was higher than the referent, in the whole group (AUC 0.666 vs. 0.567, $p<0.001$) and in the subset of 1,205 men (AUC 0.655 vs. 0.566, $p=0.007$). Likewise, the AUC for ISCD-determined method was higher than the referent in the subset of 1,205 men for prediction of vertebral fracture (AUC 0.643 vs. 0.566, $p=0.007$).

NRI results are summarized in Table 4. The lowest vertebral body T-score approach to diagnosing osteoporosis was the only method significantly different than the referent, in the larger group of 1993 men and in the subset of 1205 with an ISCD-determined T-score. Among 1993 men including 484 with incident clinical fracture, the lowest vertebral body T-score reclassified 14% of men with fracture as having osteoporosis, and inappropriately reclassified 9% of men without fracture as having osteoporosis (net NRI +0.050, $p=0.007$). In the subset of 1205 men, appropriate reclassification of men with fracture as having osteoporosis occurred in 7% using the ISCD-determined T-score, 19% using the lowest vertebral body T-score and none when using the T-score derived from a female normative database. In 894 of 1205 men without fracture, inappropriate reclassification occurred in 4% using the ISCD-derived T-score, 12% using the lowest vertebral body T-score and none using the female normative database to determine the T-score. The lowest vertebral body T-score was significantly different from the referent method for the overall NRI (+0.077, $p=0.005$).

Table 5 (on-line supplemental material) summarizes odds ratios for incident clinical fracture, and incident clinical vertebral fracture, based on the four methods of spine analysis. In general, a T-score -2.5 based on the L1-L4 T-score and male normative database demonstrated the lowest odds ratios for prediction of fractures. However, all methods demonstrated similar odds ratios and overlapping confidence intervals for prediction of fracture. Not surprisingly, [15] odds ratios for fracture were higher when each method was used to predict incident clinical vertebral fracture, as opposed to all clinical fractures.

Table 6 (on-line supplemental material) summarizes sensitivity, specificity, positive and negative predictive values for clinical fracture by four methods of spine analysis. The lowest vertebral body T-score method demonstrated highest sensitivity, but lowest specificity, compared to the other methods of spine analysis. In general, the positive predictive value was low with all four methods, whether predicting all clinical fractures or clinical vertebral fractures. By contrast, negative predictive value was 95% to 98%, when each method was used to judge the risk of incident clinical vertebral fracture.

Discussion

Using data from the MrOS study, we evaluated whether three methods of analyzing bone bone mineral density would predict incident clinical fracture better than the referent method, the mean L1-L4 T-score. We categorized men as osteoporotic for each method of spine analysis. We next compared the ability of each method to predict incident fracture, using area under the receiver operator curves (AUC) and the net reclassification index (NCI) as

measures of diagnostic accuracy. Among 1,205 men with a T-score by all four methods of analysis, the AUC for the lowest vertebral body was the only method significantly different from the referent method ($p=0.002$). Likewise, a diagnosis of osteoporosis based on the lowest vertebral body T-score demonstrated a significantly better NRI than the referent method (net NRI +0.077, $p=0.005$), while other analytic methods did not differ from the referent method. As reported by others [15,8], all methods worked better when predicting vertebral fracture [15] compared to performance characteristics when predicting all clinical fractures.

Few groups have investigated the ability of the lowest vertebral body T-score to predict incident fracture. In a small retrospective study [10], the lowest lumbar vertebral body T-score showed higher sensitivity (but lower specificity) for diagnosing men with prior clinical fracture as having osteoporosis, compared to the L1-L4 or ISCD-determined T-score. In a study [9] of 187 women, the ISCD-determined T-score was better than the L1-L4 or the lowest vertebral body T-score at predicting prevalent grade 2 or 3 fractures detected on vertebral fracture assessment images. However, careful scrutiny of the study reveals that as many as 51 of 187 (27%) women's ISCD-determined T-score was actually the lowest vertebral body T-score. Among adults undergoing routine DXA in Manitoba, Canada, the lowest vertebral body T-score demonstrated a higher ability to predict all clinical fractures, and clinical spine fractures, compared to the L1-L4 T-score in women ($n=20,478$) but not in men ($n=1,534$). The authors were not able to determine which clinical fractures occurred from minimal or low trauma. Additionally, it was unclear whether fractures were prevalent or incident, in temporal relation to the bone density test. Researchers [16] used baseline BMD measurements from placebo-treated women with osteoporosis participating in the MORE study, to compare the ability of the lowest vertebral body, L1-L4, femoral neck, total hip, and the lowest of the aforementioned T-scores to predict incident radiographic vertebral fracture. All five T-score approaches demonstrated similar ROC curves and odds ratios for incident vertebral fracture, but only women with osteoporosis were included, raising concern about the applicability of results in a clinical practice. Our study avoided a number of the weaknesses of prior studies: eligibility for the MrOS study did not require a diagnosis of osteoporosis, low-trauma fractures were carefully tracked and adjudicated, and we compared all four methods of spine analysis in a robust sample size of men, providing adequate power to detect differences between the methods.

In the current study, the ISCD-determined T-score predicted incident clinical fractures similar to the referent using the AUC, but was superior in its ability to predict vertebral fractures. By contrast, NRI data indicated that the ISCD-determined T-score was not significantly different than the referent. Of note, three issues limit the applicability of the ISCD approach in clinical practice. First, interpreters demonstrate only fair to moderate agreement on which vertebrae should be excluded in a DXA report [7-9], potentially leading to disagreement between interpreters regarding final diagnostic categorization for the patient. Second, based on the current and prior studies [8], it is not possible to derive an ISCD-determined T-score in 30- 40% of scans, as 3 vertebrae have T-score discrepancies and/or focal structural anomalies [8]. Finally, the ISCD-determined T-score requires clinicians to subjectively judge whether focal structural anomalies are present; we have previously documented suboptimal agreement between interpreters on which vertebrae to

exclude [7]. By contrast, the lowest vertebral body T-score is an objective data point, making the method highly reproducible and consistent within and between interpreters.

Our study has several strengths, including a prospective study design and measurement of baseline bone mineral density in a standardized fashion at all MrOS study centers. Regular contact with subjects was used to identify incident fragility fractures, with adjudication of fracture events by study personnel. Though we excluded ~40% of spine scans when applying the ISCD exclusion criteria, our sample size of 1,205 men still provided ~94% power to detect a significant difference between the methods of spine analysis. We have a number of study weaknesses as well. We had small numbers of men with clinical vertebral fractures. Our study results might not apply to other groups of patients, including women and children. We did not determine the ISCD T-score or lowest vertebral body T-score using a female normative database. Finally, the study was not designed to test the ability of a single vertebral body to monitor response to osteoporosis therapy.

In summary, we found that use of a single vertebral body T-score to classify older men with osteoporosis and predict incident clinical fragility fracture demonstrated similar performance characteristics to other methods of spine analysis, and was superior to other methods, when data were analyzed using the net reclassification index. Our study suggests that clinicians can interpret spine BMD in men by using the lowest vertebral body T-score. We recommend additional studies to assess this method of spine analysis in postmenopausal women and young adults, and to assess whether the method can be used to monitor response to osteoporosis medication.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the National Osteoporosis Foundation for grant support for the sub-study. The Osteoporotic Fractures in Men (MrOS) Study is supported by National Institutes of Health funding. The following institutes provide support: the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), the National Institute on Aging (NIA), the National Center for Research Resources (NCRR), and NIH Roadmap for Medical Research under the following grant numbers: U01 AR45580, U01 AR45614, U01 AR45632, U01 AR45647, U01 AR45654, U01 AR45583, U01 AG18197, U01-AG027810, and UL1 TR000128.

Grant Support: National Osteoporosis Foundation, National Institutes of Health (details, p. 14)

References

1. Cummings SR, Black DM, Nevitt MC, Browner W, Cauley J, Ensrud K, Genant HK, Palermo L, Scott J, Vogt TM. Bone density at various sites for prediction of hip fractures. The Study of Osteoporotic Fractures Research Group. *Lancet*. 1993; 341(8837):72–75. [PubMed: 8093403]
2. Melton LJ 3rd, Atkinson EJ, O'Connor MK, O'Fallon WM, Riggs BL. Bone density and fracture risk in men. *J Bone Miner Res*. 1998; 13(12):1915–1923. [PubMed: 9844110]
3. Cummings SR, Cawthon PM, Ensrud KE, Cauley JA, Fink HA, Orwoll ES. BMD and risk of hip and nonvertebral fractures in older men: a prospective study and comparison with older women. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research*. 2006; 21(10):1550–1556. doi:10.1359/jbmr.060708.

4. Liu G, Peacock M, Eilam O, Dorulla G, Braunstein E, Johnston CC. Effect of osteoarthritis in the lumbar spine and hip on bone mineral density and diagnosis of osteoporosis in elderly men and women. *Osteoporos Int*. 1997; 7(6):564–569. [PubMed: 9604053]
5. Genant HK, Engelke K, Fuerst T, Gluer CC, Grampp S, Harris ST, Jergas M, Lang T, Lu Y, Majumdar S, Mathur A, Takada M. Noninvasive assessment of bone mineral and structure: state of the art. *J Bone Miner Res*. 1996; 11(6):707–730. [PubMed: 8725168]
6. Hamdy RC, Petak SM, Lenchik L. Which central dual X-ray absorptiometry skeletal sites and regions of interest should be used to determine the diagnosis of osteoporosis? *J Clin Densitom*. 2002; 5(Suppl):S11–18. [PubMed: 12464707]
7. Hansen KE, Binkley N, Christian R, Vallarta-Ast N, Krueger D, Drezner MK, Blank RD. Interobserver Reproducibility of Criteria for Vertebral Body Exclusion. *J Bone Miner Res*. 2005; 20(3):501–508. [PubMed: 15746996]
8. Tsang JF, Leslie WD. Exclusion of focal vertebral artifacts from spine bone densitometry and fracture prediction: a comparison of expert physicians, three computer algorithms, and the minimum vertebra. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research*. 2007; 22(6):789–798. doi:10.1359/jbmr.070319.
9. Vokes TJ, Gillen DL, Lovett J, Favus MJ. Comparison of T-scores from different skeletal sites in differentiating postmenopausal women with and without prevalent vertebral fractures. *Journal of clinical densitometry : the official journal of the International Society for Clinical Densitometry*. 2005; 8(2):206–215. [PubMed: 15908709]
10. Hansen KE, Vallarta-Ast N, Krueger D, Gangnon R, Drezner MK, Binkley N. Use of the lowest lumbar vertebral body T-score to diagnose lumbar osteoporosis in men: is “cherry picking” appropriate? *J Clin Densitom*. 2005; 7(4):376–381. [PubMed: 15618597]
11. Orwoll E, Blank JB, Barrett-Connor E, Cauley J, Cummings S, Ensrud K, Lewis C, Cawthon PM, Marcus R, Marshall LM, McGowan J, Phipps K, Sherman S, Stefanick ML, Stone K. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study--a large observational study of the determinants of fracture in older men. *Contemp Clin Trials*. 2005; 26(5):569–585. [PubMed: 16084776]
12. Blank JB, Cawthon PM, Carrion-Petersen ML, Harper L, Johnson JP, Mitson E, Delay RR. Overview of recruitment for the osteoporotic fractures in men study (MrOS). *Contemporary clinical trials*. 2005; 26(5):557–568. doi:10.1016/j.cct.2005.05.005. [PubMed: 16085466]
13. ISCD Official Positions of the International Society for Clinical Densitometry. ISCD Position Development Conference. Vancouver, BC, Canada: 2005.
14. Kerr KF, Wang Z, Janes H, McClelland RL, Psaty BM, Pepe MS. Net reclassification indices for evaluating risk prediction instruments: a critical review. *Epidemiology*. 2014; 25(1):114–121. doi: 10.1097/EDE.000000000000018. [PubMed: 24240655]
15. Fink HA, Harrison SL, Taylor BC, Cummings SR, Schousboe JT, Kuskowski MA, Stone KL, Ensrud KE. Differences in site-specific fracture risk among older women with discordant results for osteoporosis at hip and spine: study of osteoporotic fractures. *Journal of clinical densitometry : the official journal of the International Society for Clinical Densitometry*. 2008; 11(2):250–259. doi:10.1016/j.jocd.2007.12.018. [PubMed: 18296090]
16. Chen P, Miller PD, Binkley NC, Kendler DL, Wong M, Krohn K. Use of lowest single lumbar spine vertebra bone mineral density T-score and other T-score approaches for diagnosing osteoporosis and relationships with vertebral fracture status. *Journal of clinical densitometry : the official journal of the International Society for Clinical Densitometry*. 2008; 11(4):525–531. doi: 10.1016/j.jocd.2008.04.009. [PubMed: 18599331]

Figure 1.

The figure demonstrates the receiver operator characteristic area under the curves (AUC) for diagnosing osteoporosis and predicting clinical fracture in 1,205 men using the L1-L4 (blue, referent), lowest vertebral body (red) and ISCD-determined T-scores (brown), and the L1-L4 T-score based on a female normative database (green). The lowest vertebral body T-score was the only method superior to the referent ($p=0.002$).

Figure 2.

The figure demonstrates the receiver operator characteristic curves for diagnosing osteoporosis and predicting clinical vertebral fracture in 1,205 men using the L1-L4 (blue, referent), the lowest vertebral body (red) and ISCD-determined T-scores (brown), and the L1-L4 T-score based on a female normative database (green). The lowest vertebral body and ISCD T-score methods were both superior to the referent ($p=0.007$, both comparisons).

Table 1

Baseline Characteristics of Study Participants ^a

	MrOS not in sub-study N=4001	MrOS in Sub-Study				p-value, Cases vs Controls
		All N=1993	p-value, MrOS vs Sub-study	Controls N=1509	Cases N=484	
Clinical Features	Age, years	74.1 ± 6.1	0.000	73.6 ± 5.9	75.7 ± 6.5	<0.001
	Race/Ethnicity, White	89%	0.052	89%	94%	0.016
	Black	5%		4%	1%	
	Asian	3%		3%	2%	
	Hispanic	2%		3%	2%	
	Other	1%		1%	<1%	
Neuromuscular Function	Body Mass Index, kg/m ²	27.4 ± 3.9	0.324	27.4 ± 3.7	27.1 ± 4.0	0.164
	Fracture at/after Age 50	21%	0.001	21%	40%	<0.001
	T-score, L1-L4	-0.1 ± 1.7	<0.001	-0.2 ± 1.7	-0.8 ± 1.6	<0.001
	T-score, Total Hip	-0.5 ± 0.9	<0.001	-0.5 ± 0.9	-1.0 ± 1.0	<0.001
	T-score, Femoral Neck	-1.1 ± 0.9	<0.001	-1.1 ± 0.9	-1.5 ± 0.9	<0.001
	Leg Power, watts	210 ± 63	0.002	209 ± 63	190 ± 59	<0.001
Drugs	Maximum Grip Strength, Kg	42 ± 9	<0.001	42 ± 8	40 ± 9	<0.001
	Walking Pace, m/second	1.2 ± 0.2	0.005	1.2 ± 0.2	1.1 ± 0.2	<0.001
	Narrow Walk Test, Unable	7%	0.001	8%	14%	<0.001
	Use of Arms during Five Chair Stands	3%	0.023	3%	6%	0.002
	Bisphosphonate	1%	0.001	2%	5%	<0.001
	Serotonin Uptake Inhibitor	3%	0.767	3%	4%	0.041
	Tricyclic Antidepressant	2%	0.007	3%	3%	0.959

^a Characteristics that were not different between groups (e.g. tobacco use) were excluded from the Table.

Table 2

Osteoporosis Diagnosis by Four Methods of Spine Analysis

Analytic Method	All Men n=1993	Men with Fracture n=484	Men without Fracture n=1509	P-value, fracture vs. no fracture
L1-L4, male	9%	14%	7%	<0.001
L1-L4, female	5%	10%	4%	<0.001
Lowest Vertebral Body	19%	28%	16%	<0.001
ISCD-determined	12%	18%	10%	<0.001

A male normative database was used to determine the L1-L4 “male”, lowest vertebral body and the ISCD-determined T-scores. “L1-L4, female” indicates the T-score derived using a female normative database.

Receiver Operator Area Under the Curve Values for Four Methods of Spine Analysis as a Means of Diagnosing Osteoporosis and Predicting Incident Clinical Fracture

Table 3

		L1-L4, male (referent)	L1-L4, female	Lowest Vertebral Body	ISCD-Determined
1993 men	All Clinical Fractures	0.533 (0.516, 0.550)	0.529 (0.515, 0.543) p = 0.427	0.558 (0.536, 0.579) p = 0.004	
	Clinical Vertebral Fractures	0.567 (0.516, 0.618)	0.561 (0.515, 0.606) p = 0.649	0.666 (0.603, 0.729) p < 0.001	-
1205 men	All Clinical Fractures	0.546 (0.521, 0.571)	0.542 (0.521, 0.563) p = 0.613	0.585 (0.554, 0.615) p = 0.002	0.559 (0.531, 0.587) p = 0.131
	Clinical Vertebral Fractures	0.566 (0.500, 0.627)	0.566 (0.511, 0.622) p = 0.989	0.655 (0.584, 0.725) p = 0.007	0.643 (0.573, 0.713) p = 0.007

A male normative database was used to determine the L1-L4 “male”, ISCD and lowest vertebral body T-score, whereas the L1-L4 “female” T-score was derived using a female normative database. P-values for AUC values are in comparison to the current standard (L1-L4, “male” T-score indicating osteoporosis).

Table 4

Net Reclassification Index by Three Methods of Spine Analysis, compared to the Male L1-L4 T-score

Compared to Male L1-L4 T-score (referent) in 1993 men	ISCD T-score	Lowest Vertebral Body T-score	Female L1-L4 T-score
Men with Fracture, n=484			
Appropriately reclassified	-	14 % (66)	0% (0)
Inappropriately reclassified	-	0% (0)	4% (21)
No change in classification	-	86% (418)	96% (463)
NRI, fractures	-	+0.136	-0.043
Men without Fracture, n=1509			
Appropriately reclassified	-	0 (0)	4% (53)
Inappropriately reclassified	-	9% (131)	0% (0)
No change in classification	-	91% (1378)	96% (1456)
NRI, no fractures	-	-0.087	+0.035
Net NRI	-	+0.050 p=0.007	-0.008 p=0.437
Compared to Male L1-L4 T-score (referent) in 1205 men	ISCD T-score	Lowest Vertebral Body T-score	Female L1-L4 T-score
Men with Fracture, n=311			
Appropriately reclassified	7% (22)	19% (60)	0% (0)
Inappropriately reclassified	0% (1)	0% (0)	7% (21)
No change in classification	93% (288)	81% (251)	93% (290)
NRI, fractures	+0.068	+0.193	-0.068
Men without Fracture, n=894			
Appropriately reclassified	0% (2)	0% (0)	6% (53)
Inappropriately reclassified	4% (40)	12% (104)	0% (0)
No change in classification	95% (852)	88% (790)	93% (841)
NRI, no fractures	-0.043	-0.116	+0.059
Net NRI	+0.025 p=0.142	+0.077 p=0.005	-0.008 p=0.625

"ISCD" denotes the T-score derived by applying the guidelines for vertebral body exclusion published by the International Society for Clinical Densitometry. The "Female" L1-L4 T-score indicates the T-score derived when using a female normative database. "NR" indicates the net reclassification index, reported for men with events (fractures), men without events (no fractures) and as a net value (net NR).