Clinical Utility of Using Lumbar Spine Trabecular Bone Score to Adjust Fracture Probability: The Manitoba BMD Cohort

Journal:	*Journal of Bone and Mineral Research*
Manuscript ID	Draft
Wiley - Manuscript type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Martineau, Patrick; University of Ottawa Faculty of Medicine, Department of Radiology Leslie, William; University of Manitoba, Dept of Medicine Johansson, Helena; WHO Collaborating Centre for Metabolic Bone Diseases, , University of Sheffield, Oden, Anders McCloskey, Eugene; The University of Sheffield, Academic Unit of Bone Metabolism, The Mellanby Centre for Bone Research Hans, Didier; Center of Bone Diseases, Department of Bone and Joint, Lausanne University Hospital Kanis, John; University of Sheffield,
Keywords:	DXA < ANALYSIS/QUANTITATION OF BONE, General population studies < EPIDEMIOLOGY, Fracture risk assessment < PRACTICE/POLICY-RELATED ISSUES, Osteoporosis < DISEASES AND DISORDERS OF/RELATED TO BONE
Abstract:	Decreased lumbar spine trabecular bone score (TBS), a DXA-derived image texture measurement, is a risk factor for major osteoporotic fracture (MOF) and hip fracture (HF) independent of ten-year fracture probability estimated using FRAXR. We determined how often applying the TBS adjustment to fracture probability altered treatment qualification. Using a population-based registry containing all clinical DXA results for Manitoba, Canada, we identified 34,316 women with baseline spine and hip DXA, FRAX-based fracture probability measurements (computed with femoral neck bone mineral density), lumbar spine TBS, and minimum 5 years of observation (mean 8.7 years). Population-based health services data were used to identify incident non-traumatic MOF and HF in 3,503 and 945 women, respectively. Baseline MOF and HF probabilities were estimated using FRAX before and after applying the TBS adjustment. Risk re-categorization was assessed using net reclassification improvement (NRI) for individual FRAX-based intervention criteria and three national clinical practice guidelines (CPGs) (US National Osteoporosis Foundation, Osteoporosis Canada, UK National Osteoporosis Guideline Group). Overall proportions of women reclassified with the TBS adjustment to FRAX were small (less than 5%) with over 90% of the reclassification occurring close to the intervention threshold. For women close to an intervention cut-off reclassification rates ranged from 9.0% to 17.9%, and were <1%

	otherwise. There was a small but significant improvement in overall NRI for all individual FRAX-based intervention criteria (range 0.007 to 0.018) and all three national CPGs (range 0.008 to 0.011). NRI was larger in women below age 65 years (up to 0.056 for hip fracture). In summary, a small but significant improvement in MOF and HF risk assessment was seen by using lumbar spine TBS to adjust FRAX probability. An improvement in risk reclassification was observed for CPGs from three different countries, with almost all of the benefit seen in individuals close to an intervention threshold.

Clinical Utility of Using Lumbar Spine Trabecular Bone Score to Adjust Fracture Probability: The Manitoba BMD Cohort

P. Martineau[1], W.D. Leslie[2], H. Johansson[3,5], A. Oden[3], E.V. McCloskey[3], D. Hans[4], J.A. Kanis[3,5]

[1] Department of Radiology, University of Ottawa, Ottawa, Ontario, Canada

[2] Department of Internal Medicine, University of Manitoba, Winnipeg, Manitoba, Canada

[3] Center for Metabolic Bone Diseases, University of Sheffield Medical School, Sheffield, UK.

[4] Bone and Joint Department, Lausanne University Hospital, Lausanne, Switzerland

[5] Institute for Health and Aging, Catholic University of Australia, Melbourne, Australia

Short Title: Clinical Utility of Lumbar Spine TBS

Tables: 2; Figures: 2; Supplemental Tables: 2

Word count: 2841

Address for Correspondence

Dr. William D. Leslie

C5121-409 Tache Ave,

Winnipeg, Manitoba

Canada R2H 2A6

Phone: 204-237-2311, Fax: 204-233-7154

Email: bleslie@sbgh.mb.ca

Abstract:

Decreased lumbar spine trabecular bone score (TBS), a DXA-derived image texture measurement, is a risk factor for major osteoporotic fracture (MOF) and hip fracture (HF) independent of ten-year fracture probability estimated using FRAX^{R}. We determined how often applying the TBS adjustment to fracture probability altered treatment qualification. Using a population-based registry containing all clinical DXA results for Manitoba, Canada, we identified 34,316 women with baseline spine and hip DXA, FRAX-based fracture probability measurements (computed with femoral neck bone mineral density), lumbar spine TBS, and minimum 5 years of observation (mean 8.7 years). Population-based health services data were used to identify incident non-traumatic MOF and HF in 3,503 and 945 women, respectively. Baseline MOF and HF probabilities were estimated using FRAX before and after applying the TBS adjustment. Risk re-categorization was assessed using net reclassification improvement (NRI) for individual FRAX-based intervention criteria and three national clinical practice guidelines (CPGs) (US National Osteoporosis Foundation, Osteoporosis Canada, UK National Osteoporosis Guideline Group). Overall proportions of women reclassified with the TBS adjustment to FRAX were small (less than 5%) with over 90% of the reclassification occurring close to the intervention threshold. For women close to an intervention cut-off reclassification rates ranged from 9.0% to 17.9%, and were <1% otherwise. There was a small but significant improvement in overall NRI for all individual FRAX-based intervention criteria (range 0.007 to 0.018) and all three national CPGs (range 0.008 to 0.011). NRI was larger in women below age 65 years (up to 0.056 for hip fracture). In summary, a small but significant improvement in MOF and HF risk assessment was seen by using lumbar spine TBS to adjust FRAX probability. An improvement in risk reclassification was observed for CPGs from three different countries, with almost all of the benefit seen in individuals close to an intervention threshold.

Abstract word count: 300

Introduction

Trabecular bone score (TBS) was introduced as a method to assess the texture of trabecular bone using a conventional DXA image (1;2) Cross-sectional studies, longitudinal studies and a recent meta-analysis of multiple prospective cohorts have demonstrated that reduced TBS correlates with increased fracture risk (3;4). More recently, lumbar spine TBS was shown to be a risk factor for osteoporotic fracture and also for risk of death independent of FRAX clinical risk factors and femoral neck bone mineral density (BMD) (5). An algorithm was developed to adjust fracture probability to account for the effect of TBS (6) and this approach has subsequently been independently validated in a meta-analysis of international cohorts (7).

Several authors have examined the value of combining TBS with BMD (with or without additional risk factors) for fracture risk prediction, with most of these studies reporting the increase in area under the receiver operating characteristic curve (AUC)(8-11). On the whole, inclusion of TBS led to small increases in AUC; however, as has been pointed out (12;13), the addition of novel factors in risk prediction models can result in deceptively small changes in the AUROC despite an overall improved risk classification. For these purposes, it has been suggested that reclassification analyses may serve as a more appropriate metric for assessing the clinical impact of an additional factor in risk classification models (14;15).

Net reclassification improvement (NRI) is a technique applicable when dealing with risk prediction models (16). The NRI measures the impact of including an additional variable on the classification of predicted risk. In the current population-based study, we used NRI to examine the clinical impact of applying the TBS adjustment to FRAX in terms of risk reclassification and treatment qualification for individual FRAX-based intervention criteria and three national clinical practice guidelines (CPGs) (US National Osteoporosis Foundation [NOF] (15) , Osteoporosis Canada (17), UK National Osteoporosis Guideline Group [NOGG]) (18).

Materials and Methods

Patient population

In the Province of Manitoba, Canada, health services are provided to virtually all residents through a single public health care system. Manitoba Health maintains computerized databases of physician billing claims and hospital separations for all residents of the province eligible to receive health services. The Manitoba Bone Density Program is a targeted case-finding clinical program with the associated database validated and described elsewhere (19;20). This database has been shown to exceed 99% in terms of completeness and accuracy. We performed a historical cohort study in women, age 50 years or older who had undergone baseline BMD measurement of the spine and hip by DXA using a single narrow fan-beam scanner configuration (Prodigy, GE Healthcare, Madison, WI, USA). All participants had medical coverage during the observation period ending March 31, 2013. In cases of multiple eligible visits, only the first record was included in the analysis. The study and was approved by the Health Research Ethics Board for the University of Manitoba.

Measurement of BMD and TBS

All DXA scans were performed and analyzed in accordance with the manufacturer's recommendations. BMD measurements were recorded for the lumbar spine BMD for L_1 through L_4 (L_1–L_4,) and the femoral neck. Hip T- and Z-scores were calculated using the National Health and Nutrition Examination Survey (NHANES) III white female reference values (21). For the lumbar spine, manufacturer reference data for white US women were used. The resulting data approximated a normal distribution. Instruments were cross-calibrated using anthropomorphic phantoms. No clinically significant differences were identified; therefore, all analyses are based on unadjusted numerical results generated by the instrument.

All TBS measurements were performed in the Bone Disease Unit at the University of Lausanne, Lausanne, Switzerland (TBS iNsight Software, Version 2.1, Med-Imaps, Pessac, France), using anonymized spine DXA files from the Manitoba database to ensure blinding of the Swiss investigators to all clinical parameters and outcomes. No significant differences in mean TBS measurements were seen for the three DXA scanners used.

All three instruments used for this study exhibited stable long-term performance (coefficient of variation [CV] < 0.5%) and satisfactory in vivo precision. Short-term reproducibility (CV) for TBS was 2.1% and for spine BMD was 1.7% in 92 individuals with repeat spine DXA scans performed within 28 days.

Fracture Outcomes

Hip fracture (HF) and major osteoporotic fracture (MOF; hip, clinical spine, forearm, and humerus fractures) were studied as these are the basis for the 10-year absolute fracture risk estimates generated by FRAX. Each health system contact includes information on a patient's demographics, date and type of service, and diagnoses from (1) physician billing claims (inpatient, outpatient, and private office) coded using the *International Classification of Disease*, 9th edition, *Clinical Modification* (ICD-9-CM) system and (2) hospital discharge abstracts, for which the diagnoses and procedures have been coded using the ICD-9-CM system prior to 2005 and the *International Classification of Disease*, 10th edition, *Canada* (ICD-10-CA) system thereafter. Anonymous linkage of these databases to the BMD database was possible via a unique scrambled health identification number, thereby allowing for the creation of a longitudinal record of health services and outcomes. Longitudinal health service records were examined for the presence of fracture codes before and after BMD testing that were not associated with trauma codes using previously validated algorithms (22). We required that hip and forearm fractures be accompanied by a site-specific fracture reduction, fixation, or casting code, which enhances the diagnostic and temporal specificity of an acute fracture.

FRAX and TBS-adjusted FRAX Calculations

The ten-year probabilities for MOF and for HF were calculated using the fracture risk assessment tool developed by the World Health Organization Collaborating Centre at Sheffield, Canadian version (FRAX® Desktop Multi-Patient Entry, version 3.7). Data required for calculating fracture probability with FRAX were assessed through a combination of data from the BMD registry, self-reported information at the time of BMD testing, hospital discharge abstracts, physician claims and a province-wide retail pharmacy database as previously described (23). Anthropomorphic data (height and weight) were measured at the time of DXA, and BMI was

calculated. In addition to prior osteoporotic fractures, we identified prior diagnoses of rheumatoid arthritis, chronic obstructive pulmonary disease (COPD, a proxy for smoking), alcohol/substance abuse (a proxy for high alcohol intake), and prolonged (>3 months) systemic corticosteroid use in the last year. The Canadian FRAX tool was calibrated using nationwide hip fracture data (24), and its predictions agree closely with observed fracture rates in Manitoba and the general Canadian population (25;26).

We then derived TBS-adjusted FRAX fracture probability for MOF and HF using the method previously described by McCloskey *et al* (6). This procedure incorporates both competing mortality and an age-TBS interaction in the calculation. This resulted in both FRAX and TBS-adjusted FRAX fracture probabilities for MOF and HF in all subjects.

Treatment Qualification

Participants were categorized according to treatment qualification based on both the FRAX and TBS-adjusted FRAX probabilities using individual FRAX-based intervention criteria as well as clinical practice guidelines (CPGs) from three different countries – the US National Osteoporosis Foundation (NOF) (15), Canadian Osteoporosis guidelines (17), and the UK National Osteoporosis Guideline Group (NOGG) (18). Each of these guidelines suggests pharmacological intervention thresholds based, in part, on FRAX-predicted fracture risk, with specific intervention thresholds unique to each set of guidelines. Recommendations proposed by the US NOF suggest that therapy be initiated in those patients with BMD in the osteopenic range with 10-year risk of MOF ≥20% or risk of hip fracture ≥3%; additional criteria for treatment are prior non-trauma HF, vertebral fracture or osteoporotic T-score (15). The Canadian Osteoporosis Guidelines recommend that pharmacological therapy be initiated in those patients with a ten-year risk of MOF in excess of 20%; additional criteria for treatment are prior non-trauma HF, vertebral fracture or multiple fracture episodes (17). The NOGG (UK) also suggests initiating treatment with pharmacological agents based on the 10-year fracture risk, but with an intervention threshold that is dependent on age (equivalent to the risk for a woman with a prior fracture and no other risk factors); additional criteria for treatment are prior non-trauma fracture (18). In routine practice, hip fracture probability is the primary determinant of FRAX-initiated treatment under the US NOF guideline, whereas MOF is the primary determinant of FRAX-initiated treatment under the

Canadian and UK NOGG guidelines. Therefore, the primary analysis for the US NOF guideline was based upon HF, and was based upon MOF for the other guidelines.

Statistical Analysis

The NRI for moving from treatment qualification based upon FRAX-based probabilities without TBS to TBS-adjusted FRAX probabilities was determined for the individual FRAX criteria as well as the overall CPGs which include the additional criteria for treatment. NRI was computed separately for individuals with and without incident fractures, the sum reflecting the overall reclassification improvement. Given the known age-TBS interaction, we also performed analyses stratified by age (<65 years versus ≥65 years) (6). The NRI was calculated as per the method detailed by Pencina *et al* (16) and reported as recommended by Leening *et al* (27). For individuals who sustain a fracture in follow up, NRI fracture is the probability of moving to a higher FRAX risk category minus the probability of moving to a lower FRAX risk category. Conversely, for individuals who remain fracture-free in follow up, NRI non-fracture is the probability of moving into a lower FRAX risk category minus the probability of moving into a higher FRAX risk category. Values of NRI fracture and NRI non-fracture greater than zero indicate an improvement in risk classification, whereas negative values indicate worse risk classification. An asymptotic test of significance for the null hypothesis of NRI=0 based upon the multinomial distribution was performed (16). In addition to NRI, the integrated discrimination improvement (IDI) was also determined where IDI can be viewed as a difference between improvement in average sensitivity corrected for any potential decrease in average (16). The subgroup consisting of those patients whose FRAX results were close to an intervention threshold were tabulated separately, where this was operationally defined as: HF 3% ± 1%; MOF fixed 20% ± 5%; MOF age-dependent ± 5%. Risk reclassification was also determined for subgroups based on femoral neck T-score (WHO category) and age (5 year groupings). Statistical analyses were performed using Statistica (Version 12.7, StatSoft, Inc., Tulsa, OK, USA).

Results

A total of 34,316 women, with a mean age of 64 years, satisfied the eligibility criteria. Over a mean follow-up period of 8.7 (±2.7) years, 3503 (10.2%) of these women suffered incident MOFs, with a total of 945 (2.8%) incident HFs recorded. **Table 1** demonstrates the baseline characteristics of these subjects. On average, TBS was significantly lower in subjects with incident fractures than in those without. The majority of subjects with incident fractures had T-scores in the osteoporotic range. Inclusion of TBS in the fracture risk assessment did not significantly change the average MOF or HF risk calibration.

The effect of including TBS in the fracture risk assessment on treatment reclassification is shown in **Table 2.** Overall proportions of women reclassified with the TBS adjustment to FRAX were relatively small (less than 5%). However, for women close to an intervention cut-off treatment reclassification rates were much higher (range 9.0% to 17.9%), with reclassification rates <1% otherwise. Indeed, the vast majority of the treatment reclassification (range 90.0% to 98.4%) occurred in the narrow bands around the intervention cut-offs. A small but statistically significant improvement in NRI for all individual FRAX-based intervention criteria (range 0.007 to 0.018) and all three national CPGs (range 0.008 to 0.011) was seen. The improvement in risk assessment was attributable to a positive effect in identifying fracture cases. When stratified by age, a larger overall NRI was seen in women less than age 65 years (NRI up to 0.056 for hip fracture, statistically significant improvement in 5 of 6 analyses) versus older women. The TBS adjustment significantly increased the IDI for MOF risk prediction (0.002, P<0.001) but not for HF risk prediction (P=0.357).

Figure 1 demonstrates the proportion of reclassified subjects according to individual FRAX-based intervention criteria and CPGs for increasing distance from the intervention cutoff. In all scenarios the majority of treatment reclassifications occurred in subjects with initial FRAX probability just below or just above the intervention threshold. For example, based on a fixed 3% hip fracture intervention threshold, treatment reclassification occurred in 21.1% and 13.0% of those just below and just above (within 1%) of the intervention cutoff. A smaller percent reclassification was seen for the US NOF guideline in those with initial hip fracture probability closer to the intervention threshold (9.4% and 4.6%, respectively), reflecting the multiple pathways to treatment

qualification. Large treatment reclassification was also seen close to the fixed and age-dependent MOF cutoffs (range from 13.0% to 20.1%). Slightly lower percent reclassification was seen when the FRAX-based intervention criteria were integrated into the Canadian and NOGG CPGs. Percentage treatment reclassification by age and femoral neck T-score category are shown in **Supplemental Table 1** and **Supplemental Table 2**, respectively.

Figure 2 illustrates the potential clinical impact of the TBS adjustment on two hypothetical patients of different ages (50 years and 75 years) with the same baseline FRAX predicted fracture risk (MOF 15% and HF 2.5%). The TBS adjustment is seen to have a larger effect on HF than MOF, and also a larger effect for younger than older patients. The MOF intervention cutoff of 20% would be exceeded for a TBS value of 1.140 at age 50 but would need to be below 1.000 at age 75 years. The HF intervention cutoff of 3% would be exceeded for a TBS value of 1.210 at age 50 but would need to be 1.140 at age 75 years.

Discussion

Our results show that the use of TBS-adjusted FRAX leads to a small but significant improvement in HF and MOF risk assessment, as measured by NRI. The impact on risk reclassification was found to be greatest in those women close to an intervention threshold; however, the apparent impact on reclassification from including TBS in fracture risk assessment varied between the various guidelines examined in this study. Additionally, it was found that the utility of using TBS-adjusted FRAX was greatest in younger women (<65 years of age).

Previous analyses have shown that lumbar spine TBS is a predictor of fracture risk, independent of BMD, but its inclusion in a multivariate risk model leads to only a small improvement in performance as measured by the AUC (8;9;28-32). Prior studies have also examined the impact of TBS on risk assessment using NRI (10;31;33;34); however, the large NRI obtained in some of these studies raises concerns over model miscalibration as other performance metrics were non-significant (35). The results of the current study are in line with those of Schousboe *et al* (36) who reported NRI in fracture cases of +0.033 (P<0.005) for TBS-adjusted fracture risk prediction in older men (MrOS), with only a very small net decrease for NRI in men without fracture.

Minimal net reclassification of hip fracture cases or non-cases was seen with the addition of TBS in the MrOS cohort.

Our study was able to identify those patient populations for which incorporation of TBS in the risk assessment would be most useful in terms of impact on management - specifically, in those patients where conventional risk assessment placed them close to an intervention threshold. That this would be the case can be easily understood for a number of reasons. TBS adjustments are unlikely to move someone with a very low fracture risk to above a treatment threshold. Likewise, in those patients at very high risk of fracture and for whom treatment is already indicated, TBS will rarely be contributory.

An important result of our study is that the findings were not specific to any particular set of CPGs but were applicable to guidelines from three different countries. Indeed, the uniformity of this finding confirms the broad potential utility of TBS as a clinical tool for risk assessment and patient decision making. We found that the effect of applying TBS within the NOF guidelines was relatively less than with the Canadian and UK guidelines, presumably reflecting the multiple pathways to treatment eligibility under the former and the more limited role for FRAX probability assessment. In clinical practice, more than 90% of reclassifications arose when the assessed fracture probability (without TBS) lay within a narrow band around the intervention threshold ($\pm$1% for hip fracture probability and $\pm$5% for MOF probability). There is a useful analogy with the utility of BMD, which in the NOGG guidelines, limit the use of BMD in individuals close to the intervention threshold (37).

Certain limitations of this study should be acknowledged. In particular, the population studied was from a clinical registry, and was composed largely of Caucasians, which may limit applicability to other populations. In addition, our subjects consisted entirely of women and whether these findings can be generalized to men is unknown; however, a recent meta-analysis has shown that TBS-adjusted FRAX performs similarly in men and women (4).

In summary, we have shown that incorporating TBS in the FRAX risk assessment paradigm led to an improvement in fracture risk classification, and was most effective in women close to an intervention threshold from the traditional assessment and in women below age 65 years.

Acknowledgments:

The authors acknowledge the Manitoba Centre for Health Policy for use of data contained in the Population Health Research Data Repository (HIPC 2012/2013 -18). The results and conclusions are those of the authors and no official endorsement by the Manitoba Centre for Health Policy, Manitoba Health, Seniors and Active Living, or other data providers is intended or should be inferred. This article has been reviewed and approved by the members of the Manitoba Bone Density Program Committee.

Funding:

This study had no external funding body.

Disclosures:

John A. Kanis: Grants from Amgen, grants from Lilly, non-financial support from Medimaps, grants from Unigene, non-financial support from Asahi, grants from Radius Health, outside the submitted work. Dr Kanis is the architect of FRAX but has no financial interest. Governmental and NGOs: National Institute for health and clinical Excellence (NICE), UK; International Osteoporosis Foundation; INSERM, France; Ministry of Public Health, China; Ministry of Health, Australia; Ministry of Health, Abu Dhabi; National Osteoporosis Guideline Group, UK; WHO.

Didier Hans: Co-ownership in the TBS patent. Stock options or royalties: Med-Imaps. Research grants: Amgen, Eli Lilly, Servier, Nycomed-Takeda.

Eugene McCloskey: Nothing to declare for FRAX and the context of this paper, but numerous ad hoc consultancies / speaking honoraria and/or research funding from Amgen, Bayer, General Electric, GSK, Hologic, Lilly, Merck Research Labs, Novartis, Novo Nordisk, Nycomed, Ono, Pfizer, ProStrakan, Roche, Sanofi-Aventis, Servier, Tethys, UBS and Warner-Chilcott.

P Martineau, W. Leslie, H. Johansson, A. Oden declare that they have no conflict of interest.

Roles:

Authors' roles: conception, design and analysis (WDL), interpretation of data (All Authors); drafting the article (PM); critically revising the article for important intellectual content (All Authors); final approval of the version to be published (All Authors); and agreement to be accountable for all aspects of the work (All Authors). WDL had full access to all the data in the study and takes the responsibility for the integrity of the data and the accuracy of the data analysis.

TABLE 1: Baseline characteristics of the study population

	Overall	No Fracture	Incident MOF	Incident Hip Fracture
N	34,316	30,813	3503	945
Age (years)	63.5 ± 10.8	62.9 ± 10.6	68.5 ± 10.9	74.0 ± 9.2
Femoral neck T-score	-1.4 ± 1.0	-1.4 ± 0.9	-2.0 ± 0.9	-2.3 ± 0.8
Femoral neck T-score osteoporotic	4211 (12.3)	3227 (10.5)	984 (28.1)	407 (43.1)
Prior fracture	4656 (13.6%)	3766 (12.2%)	890 (25.4%)	262 (27.7%)
Prior hip or spine fracture	1378 (4.0%)	1065 (3.5%)	313 (8.9%)	95 (10.1%)
Lumbar spine TBS	1.317 ± 0.121	1.323 ± 0.120	1.269 ± 0.120	1.245 ± 0.118
Percent FRAX MOF risk without TBS	10.1 ± 7.4	9.6 ± 6.9	14.9 ± 9.5	19.3 ± 9.8
Percent FRAX MOF risk with TBS	10.2 ± 7.4	9.7 ± 6.9	15.2 ± 9.3	19.4 ± 9.4
Percent FRAX hip fracture risk without TBS	2.3 ± 3.9	2.1 ± 3.6	4.6 ± 5.4	6.8 ± 6.0
Percent FRAX hip fracture risk with TBS	2.4 ± 3.8	2.1 ± 3.5	4.7 ± 5.3	6.9 ± 5.8

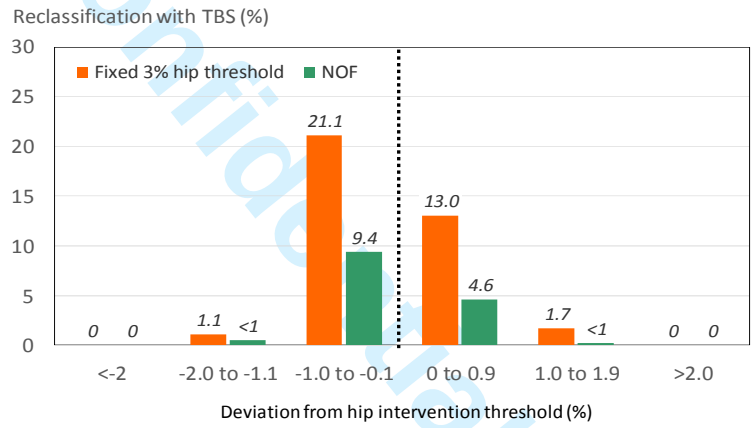
Data are mean ± SD for continuous variables, N (percent) for categorical variables. MOF = major osteoporotic fracture.

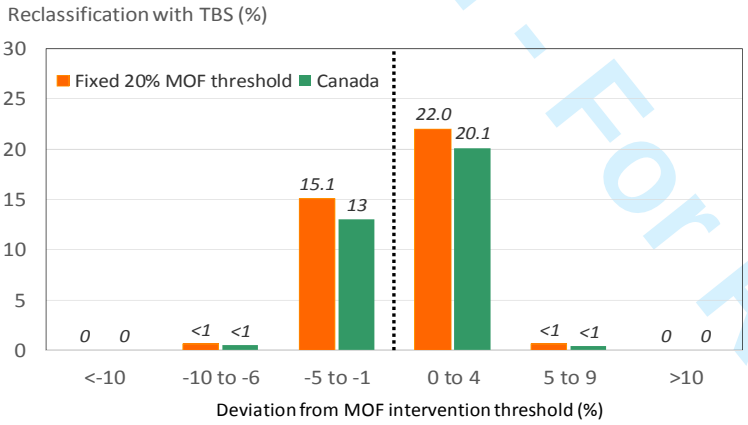
TABLE 2: Percentage reclassified and net reclassification improvement (NRI) for individual FRAX intervention criteria and three national clinical practice guidelines (CPG).

	Fixed Hip 3%[a]	Fixed MOF 20%	Age-dependent MOF	US NOF CPG [a]	Canadian CPG	UK NOGG CPG
Percentage with reclassification, all subjects	2.8%	2.6%	4.5%	1.2%	2.3%	2.8%
Percentage with reclassification, close to cutoff[b]	17.9%	17.5%	9.0%	7.5%	15.4%	5.6%
Percentage of all reclassification, close to cutoff[b]	90.7%	95.3%	98.4%	90.0%	95.4%	97.8%
NRI fracture, all subjects	+3.0%**	+1.4%***	+1.7%***	+1.5%**	+1.1%**	+2.1%***
NRI non-fracture, all subjects	-1.1%***	-0.4%***	-1.0%***	-0.5%***	-0.3%***	-0.9%***
NRI overall, all subjects	+0.018***	+0.011**	+0.007	+0.009*	+0.008*	+0.011**
NRI overall, age <65	+0.056*	+0.016***	+0.019*	+0.012	+0.012**	+0.022**
NRI overall, age ≥65	+0.002	+0.007	+0.002	+0.004	+0.006	+0.007

[a]For hip fracture. [b]Hip ±1%, MOF ±5%. NOF = US National Osteoporosis Foundation. NOGG = UK National Osteoporosis Guidelines Group.
* P<0.05, ** P<0.01, *** P<0.001

FIGURE 1: Treatment reclassification (%) after TBS adjustment for increasing deviation from the intervention cutoff. Top: Treatment based upon fixed 3% 10-year hip fracture probability and NOF CPG. Middle: Treatment based upon fixed 20% major osteoporotic fracture (MOF) fracture probability and Canadian CPG. Bottom: Treatment based upon age-dependent 10-year MOF probability and NOGG CPG. The dotted vertical line indicates the treatment threshold.





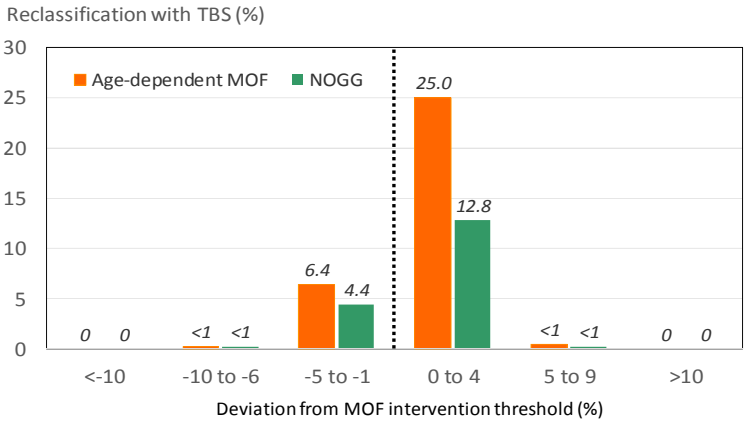
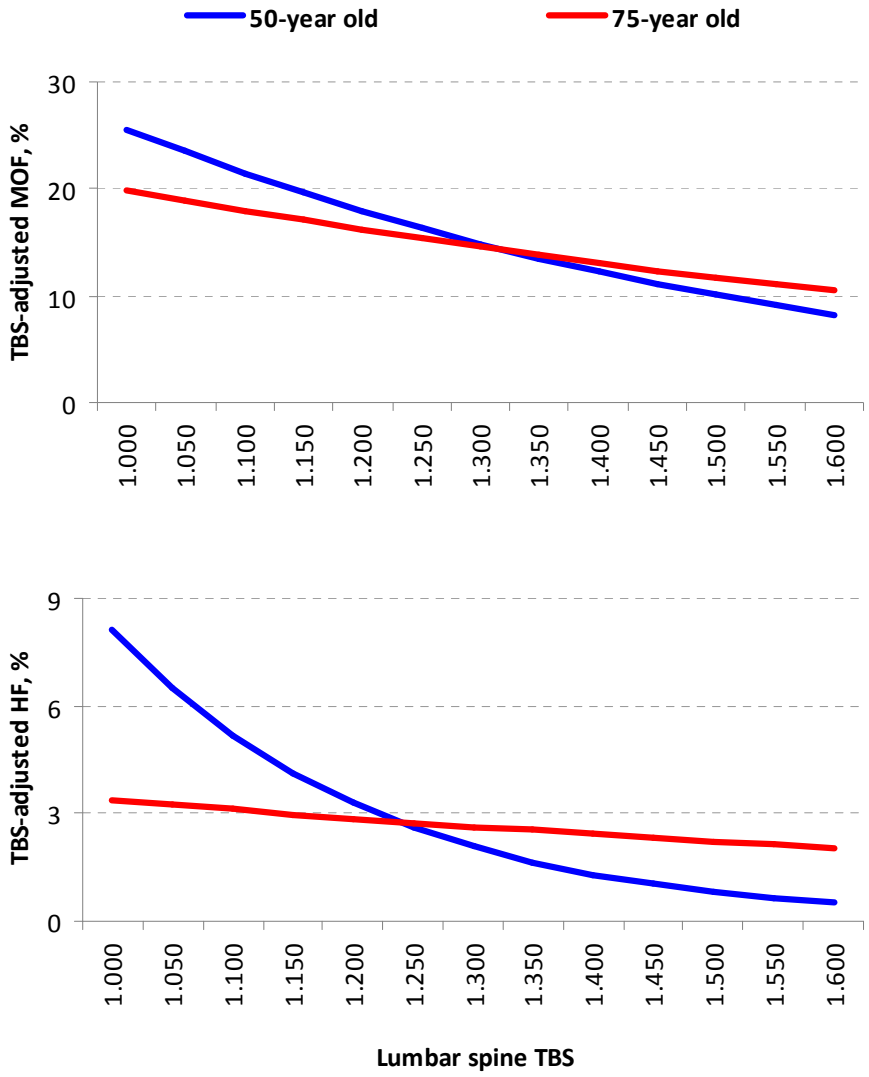


FIGURE 2: Potential clinical effect of the TBS adjustment on two hypothetical patients of different ages (50 years and 75 years) with the same baseline FRAX predicted fracture risk. Upper panel assumes baseline 10-year major osteoporotic fracture [MOF] probability 15%; lower panel assumes baseline 10-year hip fracture (HF) probability 2.5%.



SUPPLEMENTAL TABLE 1: Percentage treatment reclassification by age

Age range (years)	N	Hip 3%	MOF 20%	Age-dependent	NOF CPG	Canada CPG	NOGG CPG
40-44	1145	0.3%	0.1%	2.4%	0.1%	0.1%	1.3%
45-49	2496	0.4%	0.2%	3.4%	0.0%	0.1%	2.2%
50-54	4727	1.0%	0.3%	4.8%	0.1%	0.1%	3.4%
55-59	5573	1.9%	0.6%	5.9%	0.5%	0.5%	4.1%
60-64	5043	3.2%	1.9%	5.4%	1.0%	1.5%	3.7%
65-69	5223	4.7%	2.9%	5.2%	1.9%	2.5%	3.2%
70-74	4386	4.8%	4.1%	3.6%	2.4%	3.7%	1.8%
75-79	3331	3.6%	6.9%	3.4%	2.3%	6.1%	1.3%
80-85	1680	2.1%	7.6%	2.4%	1.6%	7.3%	0.7%
>85	712	1.1%	6.5%	2.8%	0.6%	6.5%	1.1%

SUPPLEMENTAL TABLE 2: Percentage treatment reclassification by femoral neck T-score

Femoral neck T-score	N	Hip 3%	MOF 20%	Age-dependent	NOF CPG	Canada CPG	NOGG CPG
Normal	10672	0.1%	0.1%	1.4%	0.1%	0.1%	0.8%
Osteopenic	19433	3.8%	2.4%	4.8%	2.0%	2.0%	2.7%
Osteoporotic	4211	4.6%	9.6%	11.0%	0.0%	8.9%	8.4%

Reference List

1. Pothuaud L, Carceller P, Hans D 2008 Correlations between grey-level variations in 2D projection images (TBS) and 3D microarchitecture: applications in the study of human trabecular bone microarchitecture. Bone 42:775-787.

2. Silva BC, Leslie WD, Resch H, Lamy O, Lesnyak O, Binkley N, McCloskey EV, Kanis JA, Bilezikian JP 3/2014 Trabecular bone score: a noninvasive analytical method based upon the DXA image. J Bone Miner Res 29:518-530.

3. Silva BC, Broy SB, Boutroy S, Schousboe JT, Shepherd JA, Leslie WD 7/2015 Fracture Risk Prediction by Non-BMD DXA Measures: the 2015 ISCD Official Positions Part 2: Trabecular Bone Score. J Clin Densitom 18:309-330.

4. McCloskey EV, Oden A, Harvey NC, Leslie WD, Hans D, Johansson H, Barkmann R, Boutroy S, Brown J, Chapurlat R, Elders PJ, Fujita Y, Gluer CC, Goltzman D, Iki M, Karlsson M, Kindmark A, Kotowicz M, Kurumatani N, Kwok T, Lamy O, Leung J, Lippuner K, Ljunggren O, Lorentzon M, Mellstrom D, Merlijn T, Oei L, Ohlsson C, Pasco JA, Rivadeneira F, Rosengren B, Sornay-Rendu E, Szulc P, Tamaki J, Kanis JA 5/2016 A Meta-Analysis of Trabecular Bone Score in Fracture Risk Prediction and Its Relationship to FRAX. J Bone Miner Res 31:940-948.

5. Leslie WD, Johansson H, Kanis JA, Lamy O, Oden A, McCloskey EV, Hans D 9/2014 Lumbar spine texture enhances 10-year fracture probability assessment. Osteoporos Int 25:2271-2277.

6. McCloskey EV, Oden A, Harvey NC, Leslie WD, Hans D, Johansson H, Kanis JA 6/2015 Adjusting fracture probability by trabecular bone score. Calcif Tissue Int 96:500-509.

7. McCloskey EV, Oden A, Harvey NC, Leslie WD, Hans D, Johansson H, Barkmann R, Boutroy S, Brown J, Chapurlat R, Elders PJ, Fujita Y, Gluer CC, Goltzman D, Iki M, Karlsson M, Kindmark A, Kotowicz M, Kurumatani N, Kwok T, Lamy O, Leung J, Lippuner K, Ljunggren O, Lorentzon M, Mellstrom D, Merlijn T, Oei L, Ohlsson C, Pasco JA, Rivadeneira F, Rosengren B, Sornay-Rendu E, Szulc P, Tamaki J, Kanis JA 5/2016 A Meta-Analysis of Trabecular Bone Score in Fracture Risk Prediction and Its Relationship to FRAX. J Bone Miner Res 31:940-948.

8. Touvier J, Winzenrieth R, Johansson H, Roux JP, Chaintreuil J, Toumi H, Jennane R, Hans D, Lespessailles E 4/2015 Fracture discrimination by combined bone mineral density (BMD) and microarchitectural texture analysis. Calcif Tissue Int 96:274-283.

9. Boutroy S, Hans D, Sornay-Rendu E, Vilayphiou N, Winzenrieth R, Chapurlat R 1/2013 Trabecular bone score improves fracture risk prediction in non-osteoporotic women: the OFELY study. Osteoporos Int 24:77-85.

10. Briot K, Paternotte S, Kolta S, Eastell R, Reid DM, Felsenberg D, Gluer CC, Roux C 11/2013 Added value of trabecular bone score to bone mineral density for prediction of osteoporotic fractures in postmenopausal women: the OPUS study. Bone 57:232-236.

11. Hans D, Goertzen AL, Krieg MA, Leslie WD 11/2011 Bone microarchitecture assessed by TBS predicts osteoporotic fractures independent of bone density: the Manitoba study. J Bone Miner Res 26:2762-2769.

12. Cook NR 2/20/2007 Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 115:928-935.

13. Kanis JA, Od+®n A, Johansson H, McCloskey E 2012 Pitfalls in the external validation of FRAX. Osteoporosis International 23:423-431.

14. Pencina MJ, D'Agostino RB, Sr., Steyerberg EW 1/15/2011 Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med 30:11-21.

15. Cosman F, de Beur SJ, LeBoff MS, Lewiecki EM, Tanner B, Randall S, Lindsay R 10/2014 Clinician's Guide to Prevention and Treatment of Osteoporosis. Osteoporos Int 25:2359-2381.

16. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS 1/30/2008 Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 27:157-172.

17. Papaioannou A, Morin S, Cheung AM, Atkinson S, Brown JP, Feldman S, Hanley DA, Hodsman A, Jamal SA, Kaiser SM, Kvern B, Siminoski K, Leslie WD 11/23/2010 2010 clinical practice guidelines for the diagnosis and management of osteoporosis in Canada: summary. CMAJ 182:1864-1873.

18. Compston J, Bowring C, Cooper A, Cooper C, Davies C, Francis R, Kanis JA, Marsh D, McCloskey EV, Reid DM 2013 Diagnosis and management of osteoporosis in postmenopausal women and older men in the UK: National Osteoporosis Guideline Group (NOGG) update 2013. Maturitas 75:392-396.

19. Leslie WD, Metge C 2003 Establishing a regional bone density program: lessons from the Manitoba experience. Journal of Clinical Densitometry 6:275-282.

20. Leslie WD, Caetano PA, Macwilliam LR, Finlayson GS 2005 Construction and validation of a population-based bone densitometry database. J Clin Densitom 8:25-30.

21. Looker AC, Wahner HW, Dunn WL, Calvo MS, Harris TB, Heyse SP, Johnston Jr CC, Lindsay R 1998 Updated data on proximal femur bone mineral levels of US adults. Osteoporosis International 8:468-490.

22. Lix LM, Azimaee M, Osman BA, Caetano P, Morin S, Metge C, Goltzman D, Kreiger N, Prior J, Leslie WD 2012 Osteoporosis-related fracture case definitions for population-based administrative data. BMC Public Health 12:301.

23. Leslie WD, Lix LM, Johansson H, Oden A, McCloskey E, Kanis JA 11/2010 Independent clinical validation of a Canadian FRAX tool: fracture prediction and model calibration. J Bone Miner Res 25:2350-2358.

24. Leslie WD, Lix LM, Langsetmo L, Berger C, Goltzman D, Hanley DA, Adachi JD, Johansson H, Oden A, McCloskey E, Kanis JA 3/2011 Construction of a FRAX(R) model for the assessment of fracture probability in Canada and implications for treatment. Osteoporos Int 22:817-827.

25. Leslie WD, O'Donnell S, Lagace C, Walsh P, Bancej C, Jean S, Siminoski K, Kaiser S, Kendler DL, Jaglal S 8/2010 Population-based Canadian hip fracture rates with international comparisons. Osteoporos Int 21:1317-1322.

26. Fraser LA, Langsetmo L, Berger C, Ioannidis G, Goltzman D, Adachi JD, Papaioannou A, Josse R, Kovacs CS, Olszynski WP 2011 Fracture prediction and calibration of a Canadian FRAX-« tool: a population-based report from CaMos. Osteoporosis International 22:829-837.

27. Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW 1/21/2014 Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. Ann Intern Med 160:122-131.

28. Breban S, Briot K, Kolta S, Paternotte S, Ghazi M, Fechtenbaum J, Roux C 7/2012 Identification of rheumatoid arthritis patients with vertebral fractures using bone mineral density and trabecular bone score. J Clin Densitom 15:260-266.

29. Iki M, Tamaki J, Sato Y, Winzenrieth R, Kagamimori S, Kagawa Y, Yoneshima H 1/2015 Age-related normative values of trabecular bone score (TBS) for Japanese women: the Japanese Population-based Osteoporosis (JPOS) study. Osteoporos Int 26:245-252.

30. Nassar K, Paternotte S, Kolta S, Fechtenbaum J, Roux C, Briot K 1/2014 Added value of trabecular bone score over bone mineral density for identification of vertebral fractures in patients with areal bone mineral density in the non-osteoporotic range. Osteoporos Int 25:243-249.

31. Iki M, Fujita Y, Tamaki J, Kouda K, Yura A, Sato Y, Moon JS, Winzenrieth R, Okamoto N, Kurumatani N 6/2015 Trabecular bone score may improve FRAX(R) prediction accuracy for major osteoporotic fractures in elderly Japanese men: the Fujiwara-kyo Osteoporosis Risk in Men (FORMEN) Cohort Study. Osteoporos Int 26:1841-1848.

32. Popp AW, Meer S, Krieg MA, Perrelet R, Hans D, Lippuner K 5/27/2015 Bone mineral density (BMD) and vertebral trabecular bone score (TBS) for the identification of elderly women at high risk for fracture: the SEMOF cohort study. Eur Spine J.

33. Iki M, Tamaki J, Kadowaki E, Sato Y, Dongmei N, Winzenrieth R, Kagamimori S, Kagawa Y, Yoneshima H 2/2014 Trabecular bone score (TBS) predicts vertebral fractures in Japanese women over 10 years independently of bone density and prevalent vertebral deformity: the Japanese Population-Based Osteoporosis (JPOS) cohort study. J Bone Miner Res 29:399-407.

34. Vasic J, Petranova T, Povoroznyuk V, Barbu CG, Karadzic M, Gojkovic F, Elez J, Winzenrieth R, Hans D, CulaficVojinovic V 2014 Evaluating spine micro-architectural texture (via TBS) discriminates major osteoporotic fractures from controls both as well as and independent of site matched BMD: the Eastern European TBS study. Journal of bone and mineral metabolism 32:556-562.

35. Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW 2014 Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. Annals of internal medicine 160:122-131.

36. Schousboe JT, Vo T, Taylor BC, Cawthon PM, Schwartz AV, Bauer DC, Orwoll ES, Lane NE, Barrett-Connor E, Ensrud KE 3/2016 Prediction of Incident Major Osteoporotic and Hip Fractures by Trabecular Bone Score (TBS) and Prevalent Radiographic Vertebral Fracture in Older Men. J Bone Miner Res 31:690-697.

37. Compston J, Cooper A, Cooper C, Francis R, Kanis JA, Marsh D, McCloskey EV, Reid DM, Selby P, Wilkins M 2009 Guidelines for the diagnosis and management of osteoporosis in postmenopausal women and men from the age of 50 years in the UK. Maturitas 62:105-108.

Journal of Bone and Mineral Research
JBMR®
2025 M Street
Suite 800
Washington, DC 2008
P: 202-367-1161
E: jbmroffice@wiley.com
E: asbmr@asbmr.org

Corresponding Author Name: W. D. Leslie

Manuscript Title: Clinical Utility of Using Lumbar Spine Trabecular Bone Score to Adjust

Strengthening the Reporting of Observational Studies in Epidemiology

When reporting results of studies involving humans, the *JBMR*® adheres to the STROBE statement, and authors must provide the information required by the STROBE checklist as adapted by JBMR. Since statistical analyses vary according to study design and prohibit the establishment of standard rules, the following guidelines are provided to assist authors in performing the most appropriate statistical analyses and reporting those results in accord with *JBMR*® standards. The checklist aids authors in providing detailed information within their submission that meet a standard to achieve reproducibility and transparency of research, and assists reviewers in their effort determine whether the necessary information is present. If you have any questions, please contact us at jbmroffice@wiley.com.

JBMR® has adapted the STROBE Guidelines. To fulfill the guideline requirements for JBMR®, please select the statement that describes your study and fill out only the required STROBE checklist questions for the study.

- [x] STROBE CHECKLIST A- I am reporting the results of a cohort study.
- [ ] STROBE CHECKLIST B- I am reporting the results of a case-control study.
- [ ] STROBE CHECKLIST C- I am reporting the results of a cross-sectional study.

STROBE CHECKLIST A

Cohort Study

Journal of Bone and Mineral Research JBMR

Recommendation- Cohort Study	Page
Title/Abstract/Introduction- Indicate the study's design with a commonly used term in the title or the abstract. State specific objectives, including any prespecified hypotheses in introduction.	1-3
Methods- Present key elements of study design early in the paper. Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection. ❖ Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up. For matched studies, give matching criteria and number of exposed and unexposed	4
Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable. ❖ For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. ❖ Describe any efforts to address potential sources of bias ❖ Explain how the study size was arrived at	4-6
Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why ❖ Describe all statistical methods, including those used to control for confounding. Describe any methods used to examine subgroups and interactions. ❖ Explain how missing data were addressed ❖ If applicable, explain how loss to follow-up was addressed ❖ Describe any sensitivity analyses	6-7
Results- Report numbers of individuals at each stage of study—numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed ❖ Give reasons for non-participation at each stage ❖ Indicate number of participants with missing data for each variable of interest ❖ Summarize follow-up time (eg, average and total amount)	8
Report numbers of outcome events or summary measures over time	8-9
Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included. ❖ Report category boundaries when continuous variables were categorized ❖ If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period ❖ Report other analyses done—e.g. analyses of subgroups and interactions, and	8-9

Recommendation- Cohort Study	Page
sensitivity analyses	
Discussion- Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias. Discuss the generalisability (external validity) of the study results	9-11

STROBE CHECKLIST B

Case-control study

Journal of Bone and Mineral Research JBMR

Recommendation- Case-control study	Page
Title/Abstract/Introduction- Indicate the study's design with a commonly used term in the title or the abstract. State specific objectives, including any prespecified hypotheses in introduction.	
Methods- Present key elements of study design early in the paper. Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection. ❖ Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls. For matched studies, give matching criteria and the number of controls per case.	
Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable. ❖ For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group ❖ Describe any efforts to address potential sources of bias ❖ Explain how the study size was arrived at	
Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why ❖ Describe all statistical methods, including those used to control for confounding. Describe any methods used to examine subgroups and interactions. ❖ Explain how missing data were addressed ❖ If applicable, explain how matching of cases and controls was addressed ❖ Describe any sensitivity analyses	
Results- Report numbers of individuals at each stage of study—numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed ❖ Give reasons for non-participation at each stage ❖ Indicate number of participants with missing data for each variable of interest	
Report numbers in each exposure category, or summary measures of exposure	
Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included. ❖ Report category boundaries when continuous variables were categorized ❖ If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period ❖ Report other analyses done—e.g. analyses of subgroups and interactions, and sensitivity analyses	

Recommendation- Case-control study	Page
Discussion- Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias. Discuss the generalisability (external validity) of the study results	

STROBE CHECKLIST C

Cross-sectional study

Recommendation- Cross-sectional study	Page
Title/Abstract/Introduction- Indicate the study's design with a commonly used term in the title or the abstract. State specific objectives, including any prespecified hypotheses in introduction.	
Methods- Present key elements of study design early in the paper. Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection. ❖ Give the eligibility criteria, and the sources and methods of selection of participants	
Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable. ❖ For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group ❖ Describe any efforts to address potential sources of bias ❖ Explain how the study size was arrived at	
Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why ❖ Describe all statistical methods, including those used to control for confounding. Describe any methods used to examine subgroups and interactions. ❖ Explain how missing data were addressed ❖ If applicable, describe analytical methods taking account of sampling strategy ❖ Describe any sensitivity analyses	
Results- Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed ❖ Give reasons for non-participation at each stage ❖ Indicate number of participants with missing data for each variable of interest	
Report numbers of outcome events or summary measures	
Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included. ❖ Report category boundaries when continuous variables were categorized ❖ If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period ❖ Report other analyses done—e.g. analyses of subgroups and interactions, and sensitivity analyses	

Recommendation- Cross-sectional study	Page
Discussion- Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias. Discuss the generalisability (external validity) of the study results	