



ations of their calibration with observed fractures in community-based studies, current Canadian guidelines suggest the use of either of these tools<sup>2,3,11</sup>.

FRAX is a multiattribute aggregate score designed to calculate the 10-year probability of hip and major osteoporotic fracture for individuals between the ages of 40 to 90<sup>1</sup>. The following indicators are used in its calculation: age, sex, height, weight, previous fragility fracture, parental hip fracture history, current smoking status, current use of glucocorticoids, diagnosis of rheumatoid arthritis (RA), alcohol intake (3 or more units daily), diagnosis of secondary osteoporosis (OP), and femoral neck bone mineral density (BMD) T score. The Canadian FRAX tool has been shown to predict observed fracture rates in large-scale Canadian population-based and cohort studies<sup>12,13</sup>.

The CAROC tool uses data on sex, age, and femoral neck BMD T scores to estimate the semiquantitative risk for a major osteoporotic fracture within the next 10 years in 3 categories: low (< 10%), moderate (10%-20%), and high (> 20%). Two special circumstances also lead individuals to be “bumped” up to the next categorization of risk in the CAROC system — the presence of fragility fractures or prolonged corticosteroid use. Persons with one of these risk factors would move from either low risk to moderate risk or from moderate risk to high risk. In CAROC, when both factors are present, the patient is considered to be at high fracture risk<sup>2,3,11</sup>.

In addition, the 2010 Clinical Practice Guidelines (CPG) for the diagnosis and management of OP<sup>2</sup> and the report by Lentle, *et al* on fracture risk assessment<sup>3</sup> suggest that 2 additional clinical qualifiers should be considered when determining fracture risk. First, regardless of the calculated fracture risk or the tool used, persons who present with hip or vertebral fractures, or those with more than 1 fragility fracture episode, should be considered at high risk<sup>2,14</sup>. Second, individuals with a T score for the lumbar spine or total hip  $\leq -2.5$  are considered to have at least moderate risk; i.e., they are moved from low to moderate risk<sup>2</sup>. Applying these clinical qualifiers to the risk calculator (FRAX or CAROC) is closest to the recommended approach for fracture risk assessment in Canada.

Fracture risk assessment is important because the guidelines link treatment recommendations to the risk of (re)fracture. Misclassification could lead to missed opportunities to reduce fracture risk with proven therapies, or exposure to additional costs and possible side effects and inconveniences of care. Current guidelines suggest that where the risk of major fragility fracture is high, a pharmacologic agent is recommended to reduce fracture risk, in addition to optimization of vitamin D and calcium intake, lifestyle recommendations, and interventions to reduce the risk of falls<sup>2</sup>. These are considered a Grade A recommendation in this high-risk group<sup>2</sup>.

Prior to the release of the 2010 Canadian guidelines,

Leslie, *et al*<sup>11</sup> demonstrated high agreement between the CAROC (2005) and FRAX BMD tools in 2 different community samples (CaMos, 89%; Manitoba BMD cohort, 88%). About 13% and 18% of these cohorts, respectively, had a prior fragility fracture and/or reported glucocorticoid use. For persons with a clinical risk factor (either prior fracture and/or steroid use), agreement was lower (CaMos, 61%; Manitoba BMD cohort, 74%), raising a potential concern about the comparability of these tools when assessing 10-year fracture risk in fragility fracture patients where the prevalence of a CAROC clinical risk factor (the current fragility fracture) is 100%.

The purpose of this study was to examine the level of agreement between 2 fracture risk assessment tools when applied within the context of the Canadian Guidelines and in a population of fragility fracture patients. We aimed to compare our findings to the cohort work of Leslie, *et al*<sup>11</sup>, and to determine the factors that are associated with any disagreement between the 2 fracture risk tools.

## MATERIALS AND METHODS

A convenience sample of participants was drawn from the Osteoporosis Exemplary Care Program (OEC) at our tertiary level teaching hospital. The OEC is a coordinator-based program that identifies, investigates, and initiates care for patients with fragility fracture (gold-level fracture liaison service, according to the International Osteoporosis Foundation). OEC women over 40 years and men aged 50 years or older who were not receiving OP care at the time of screening and who had BMD testing completed at the hospital between December 2008 and November 2011 were considered for inclusion in our study. Only treatment-naive patients were considered because the FRAX tool is calibrated only for those not receiving treatment<sup>15</sup>. Further details regarding OEC screening, baseline-measurement, and followup procedures have been reported<sup>16</sup>.

*Final sample for analysis.* As of November 2011, 181 individuals who were not receiving treatment had BMD test results and height and weight data available on the electronic medical records database at the hospital, making it possible to compute a 10-year osteoporotic fracture probability using both FRAX and CAROC. Of these, 135 had completed a questionnaire with self-report data on the risk factors required to compute an accurate FRAX fracture probability estimate (response rate of 74.6%). Generally, patients with self-report data did not differ significantly from those without survey data. However, the latter group had a higher proportion of patients with hip fractures, who would therefore be at higher risk of refracture.

*Fracture risk calculators.* Ten-year probabilities of any major osteoporotic fracture were calculated using the World Health Organization's online Canadian FRAX (with BMD) tool ([www.shef.ac.uk/FRAX/tool.jsp?country=19](http://www.shef.ac.uk/FRAX/tool.jsp?country=19)). Femoral neck hip T scores from the participant's bone density dual-energy x-ray absorptiometry (DEXA) image were extracted from the image reports and used in the FRAX calculations<sup>2,12,17</sup>. Information on the clinical risk factors was extracted from individual self-report on the OEC baseline measures. We only collected information on maternal hip fracture history, which was used as the indicator of parental hip fracture history in FRAX BMD. In the questionnaire, we asked patients if they currently smoke cigarettes, and whether on average they drink more than 2 alcoholic beverages a day. One alcoholic beverage is equivalent to 1 glass of wine (150 ml), 1 beer (341 ml), or 1 spirit (30-40 ml). More than 2 alcoholic beverages would therefore mean 3 or more units/day. A diagnosis of RA was determined by asking patients to indicate whether they had been diagnosed by a doctor with any of several diseases, including RA. Patients were asked whether they were taking oral glucocorticoids. An individual's height and

weight for FRAX were extracted from their BMD DEXA image. Secondary OP was not measured in the questionnaire and we did not have the blood work necessary to determine whether patients had a disorder strongly associated with OP. This risk factor was therefore considered missing in the FRAX calculation. A typology of categorical risk classification that included low, moderate, and high risk categories was created based on Leslie, *et al*<sup>11</sup>. Specifically, each individual's FRAX BMD probability of a major osteoporotic fracture in the next 10 years was categorized as either low (< 10%), moderate (10%-20%), or high (> 20%).

For CAROC 2010, each individual was similarly assigned a fracture risk category (low < 10%, moderate 10%-20%, or high > 20%) using relevant information available from the BMD DEXA image and OCEP questionnaire. No individual was classified at low risk using CAROC because they had all experienced at least their current fragility fracture. Recent prolonged use of glucocorticoids also increased a patient's risk to the next category. Patients were considered to have a high risk of refracture over the next 10 years when both risk factors were present<sup>11</sup>.

**Additional clinical qualifiers (Canadian CPG).** Consideration was also given to the additional clinical qualifiers as suggested by the 2010 CPG<sup>2</sup> described above. In our analysis, these were applied as a first step, as is recommended by the guidelines (Figure 1). We therefore assigned patients to high, moderate, or low risk using a combination of the clinical qualifiers and the risk calculators (FRAX or CAROC). Persons with hip, vertebral, or multiple fractures were considered high risk of refracture regardless of the fracture risk tool used (CAROC or FRAX). Further, those with T scores  $\leq -2.5$  at either the lumbar spine or total hip were moved up to moderate risk if they were previously low risk<sup>2</sup>. In our comparison of fracture risk tools, agreement was evaluated after the application of the additional clinical qualifiers for each risk assessment calculator, to be closest to the recommendations of the Canadian guidelines.

**Analysis.** Univariate analyses were used to provide a description of the patients, including frequency distributions of demographic variables, fracture type, and clinical risk factors. Categorical fracture risk classifications (high, moderate, low) were described for both of the assessment tools with the CPG clinical qualifiers.

Using these guideline-based classifications, a 3 by 3 table was constructed to identify the discordance between CAROC and FRAX. Percent agreement and a description of any discordance were provided. A quadratic

weighted  $\kappa$  ( $K_w$ )<sup>18,19</sup> and 95% CI were calculated to estimate the chance corrected agreement between the risk assessment tools. There are no absolute values used to interpret  $\kappa$  statistics. However, according to Landis and Koch<sup>20</sup>,  $\kappa$  values > 0.80 indicate excellent agreement, while Fleiss, *et al*<sup>21</sup> report that estimates > 0.75 should be considered very good or excellent agreement. More recently, Aaronson, *et al*<sup>22</sup> suggest that a  $\kappa$  of 0.70 would be acceptable for group level analyses, and a higher level of agreement is needed for individual clinical decision making, such as in the classification of individual patient's refracture risk. It is generally agreed that lower levels of agreement can be tolerated when applied to a large cohort or randomized trial, and higher values are needed when the results are linked to clinical decisions at an individual level where there is less tolerance for misclassification error. Given that our practice guidelines offer the strongest message to those at high risk versus other (moderate or low), we also dichotomized risk into high versus non-high (i.e., moderate and low) according to each system and tested the agreement across these intervention thresholds.

We were also interested in the characteristics of persons who were found in discordant cells in the 3 by 3 table and whether this was a predictable discordance between the risk assessment tools. Logistic regression was used to model the concordant versus discordant cells when comparing CAROC and FRAX with the clinical qualifiers. We evaluated the factors associated with overall concordance in the classification of risk. We then assessed factors associated with concordance versus being classified higher on CAROC than FRAX. Only 3 patients were classified lower on CAROC than FRAX, therefore logistic regression was not performed for that outcome. The independent variables entered into the logistic regression equations were participant sex, age, and reported risk factors other than the previous broken bone (yes/no) as this was invariant (100%) in our sample. Both unadjusted and adjusted models, which included all of the independent variables, were analyzed using SAS 9.3 (SAS Institute).

## RESULTS

**Description of sample.** Details regarding the demographics of the sample can be found in Table 1. The average age of our sample was 66.5 years (SD 10.7), ranging from 50 to 90 years. In total, 49.6% of the sample was aged 64 years or younger, and 64.4% of the sample identified as female. The

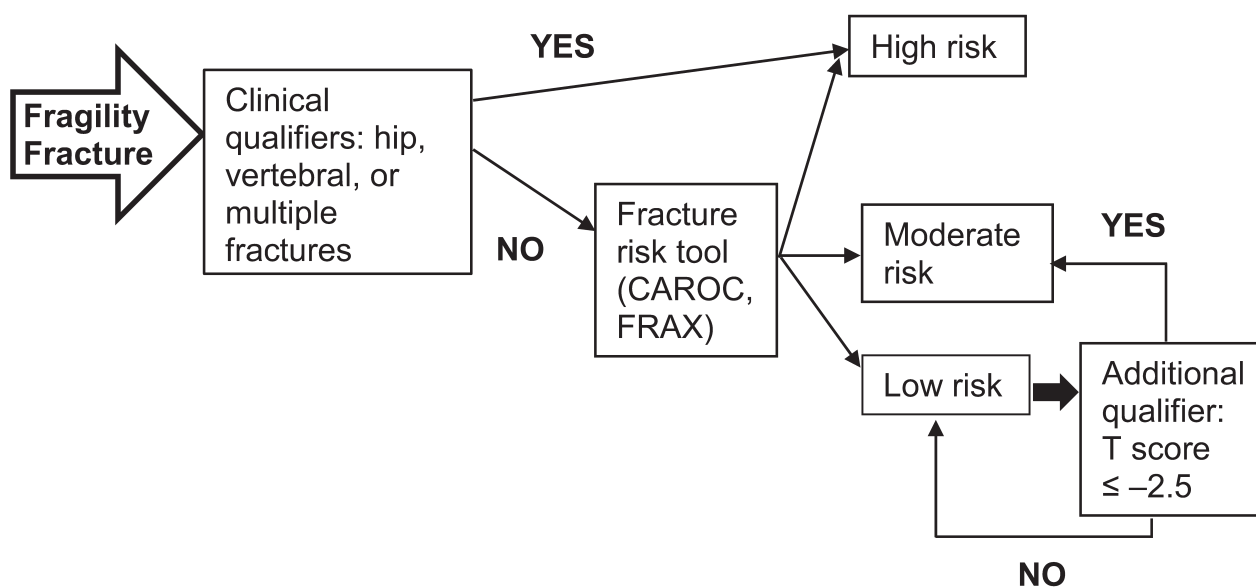


Figure 1. Flow diagram for ascertaining re-fracture risk by applying the Canadian Clinical Practice Guidelines for the diagnosis and management of osteoporosis. CAROC: Canadian Association of Radiologists and Osteoporosis Canada; FRAX: Canadian Fracture Risk Assessment.

Table 1. Study sample characteristics, n = 135. Values are n (%) unless otherwise indicated.

Characteristics	Values
Age, yrs, mean (SD)	66.5 (10.7)
64 and under	67 (49.6)
65 and over	68 (50.4)
Sex	
Male	48 (35.6)
Female	87 (64.4)
Fracture location	
Wrist	73 (54.5)
Shoulder	34 (25.4)
Hip and other	21 (15.7)
Spine	6 (4.5)
Additional risk factors used in calculating FRAX score	
Maternal hip fracture	19 (14.1)
Current smoker	24 (17.8)
Daily alcohol intake (3 or more units/day)	10 (7.4)
Diagnosed with RA	5 (3.7)
Taking oral glucocorticoids	2 (1.5)
Weight, kg, mean (SD)	69.3 (15.6)
Height, cm, mean (SD)	168.8 (10.4)
T score femoral neck BMD, mean (SD)	-1.5 (1.1)

FRAX: Canadian Fracture Risk Assessment; RA: rheumatoid arthritis; BMD: bone mineral density.

most common fracture location was participant's wrist (54.5%), followed by the shoulder (25.4%), hip (15.7%), and spine (4.5%). About 34% of the sample self-reported additional risk factors for inclusion in their FRAX fracture probability assessments. The most common additional risk factor reported by participants was current smoking, followed by maternal history of hip fracture.

*Fracture risk and agreement between CAROC and FRAX.* A 3 by 3 comparison of the categorical risk classifications for CAROC and FRAX (integrating the additional clinical qualifiers) is presented in Table 2. For CAROC, 0%, 53.3%, and 46.7% were found to be at low, moderate, and high risk, respectively. Using FRAX, 23.0%, 36.3% and 40.7% were low, moderate and high risk, respectively. We found 45

discordant cases out of a total of 135 (33.3%). Of these, 42 were classified higher on CAROC than FRAX, and 3 were classified lower on CAROC than FRAX. The chance corrected agreement between the 2 tools was  $Kw = 0.64$  (95% CI 0.58–0.71). The narrow CI obtained for the  $Kw$  estimate indicates that our sample size of 135 patients allowed adequate precision in our estimation of  $Kw$  for our 3 category tools<sup>23</sup>. We also performed a sensitivity analysis by collapsing the cells to compute a pairwise  $\kappa$  estimate across the intervention thresholds, i.e., high risk versus non-high risk (low plus moderate). In this case, the agreement between CAROC and FRAX was  $Kw = 0.79$  (95% CI 0.69–0.89).

*Factors associated with discordance.* Table 3 presents the (un)adjusted OR and 95% CI for the factors associated with overall concordance in the classification of risk between CAROC and FRAX. In the adjusted model, women, older individuals, and those reporting additional risk factors in FRAX were more likely to be concordant overall. We found similar results when comparing those classified higher on CAROC versus FRAX (data not shown) — sex (being female), age (being older), and having additional risk factors in FRAX were significant predictors in the adjusted model. Thus, younger persons and men were more likely to be found in discordant cells.

## DISCUSSION

Among fragility fracture patients, the agreement between CAROC and FRAX, both integrating the additional clinical qualifiers suggested by the Canadian guidelines<sup>2</sup> was  $Kw = 0.64$  (95% CI 0.58–0.71), with 45 of 135 cases in the cells reflecting disagreement. The  $Kw$  estimate and its confidence limits were lower than the cutoffs used to determine very good or excellent agreement at the group level [0.80 (20) or 0.75 (21)]. These cutoffs are not absolute<sup>19</sup> and others would accept values as low as  $\kappa$  of 0.7<sup>21</sup> as reasonable for diagnostic classification. Our results nonetheless raise a caution that the fracture risk assessments carried out using CAROC and FRAX in conjunction with the Canadian CPG diverge in the

Table 2. Agreement between categorical risk classifications\* according to CAROC and FRAX BMD in conjunction with the clinical qualifiers, n = 135.

CAROC 2010*	FRAX BMD*			Total
	Low, < 10%	Moderate, 10%–20%	High, > 20%	
Low, < 10%	0	0	0	0
Moderate, 10–20%	31	38	3	72
High, > 20%	0	11	52	63
Total	31	49	55	135

\*Risk classifications incorporate additional clinical qualifiers suggested by the Canadian Clinical Practice Guidelines. Hip and vertebral fractures or multiple fragility fractures are automatically assigned to high risk regardless of calculated risk. Anyone with a T score for the lumbar spine or total hip  $\leq -2.5$  is at least moderate risk. Percent agreement = 66.7%; weighted  $\kappa = 0.64$  (95% CI 0.58–0.71). CAROC: Canadian Association of Radiologists and Osteoporosis Canada; FRAX: Canadian Fracture Risk Assessment; BMD: bone mineral density.

Table 3. Unadjusted and adjusted models: factors associated with concordance in the classification of risk between CAROC and FRAX BMD, both incorporating additional clinical qualifiers (n = 135).

Variables	Overall Concordant (n = 90) vs Discordant (n = 45)			
	Unadjusted Estimates		Adjusted <sup>a</sup> Estimates	
	OR	95% CI	OR	95% CI
Sex				
Female	2.35*	1.12–4.94	3.26**	1.40–7.57
Male	1.00		1.00	
Age, yrs	1.05**	1.01–1.09	1.08**	1.03–1.12
Additional risk factors used for FRAX <sup>b</sup>				
Yes	3.38**	1.41–8.08	5.23**	1.99–13.76
No	1.00		1.00	

\*p < 0.05, \*\* p < 0.01. <sup>a</sup> Includes all 3 independent variables. <sup>b</sup> In addition to previous broken bone status. Additional risk factors include current smoking, maternal hip fracture, alcohol consumption, rheumatoid arthritis diagnosis, and glucocorticoid use. CAROC: Canadian Association of Radiologists and Osteoporosis Canada; FRAX: Canadian Fracture Risk Assessment; BMD: bone mineral density.

fragility fracture population, where the prevalence of a prior fracture is 100%.

We found similar levels of agreement when we looked again at the work of Leslie, *et al*<sup>11</sup> to calculate  $\kappa$  values in the overall sample and the subset of patients with a clinical risk factor (in their case either a prior fracture or prior steroid use). Using data from Leslie, *et al*<sup>11</sup> published tables, we calculated Kw estimates of 0.81 (95% CI 0.80–0.82) in the overall CaMOS cohort and 0.83 (95% CI 0.83–0.84) in the overall Manitoba BMD cohort. The agreement was lower in both cohorts when considering participants with a single clinical risk factor: 0.42 (95% CI 0.37–0.46) for CaMOS and 0.59 (95% CI 0.58–0.61) for the Manitoba BMD cohort. However, Leslie, *et al*<sup>11</sup> did not classify risk by integrating the additional qualifiers from the guidelines as we did, which automatically bump individuals with fragility fracture of the vertebra or hip, or more than 1 fragility fracture, to high risk regardless of calculated risk. K values for the Leslie data would likely be greater had these qualifiers been included.

Our findings correspond to findings of a recent study by Beattie, *et al*, which concluded that FRAX and CAROC provide different fracture risk classifications in their cohort of 65 patients with wrist fracture<sup>24</sup>. They also found more discordance in younger persons. This group did not include the special qualifiers in their assessment, which could have allowed for wrist fracture patients with additional fractures to be moved to high risk. We included these qualifiers to be consistent with the model suggested for clinical practice in Canada.

Regression analyses further revealed that when comparing CAROC and FRAX BMD, concordance was more likely in females, older individuals, and those reporting additional risk factors in FRAX. In terms of sex, it is known that fracture risk is higher in women<sup>25,26,27,28</sup>, and this is accounted for in the calculations of risk in both FRAX and CAROC. FRAX classifications among women will therefore be more similar to CAROC, thereby contributing to better concordance

between the tools. In addition, fracture probabilities are known to increase with age<sup>11</sup> in both men and women. This convergence in refracture risk as individuals age may thus explain the higher likelihood of concordance in older patients when comparing CAROC and FRAX BMD. Finally, including additional risk factors in the calculation of FRAX probabilities allows for classifications of risk to higher categories. This process is similar to the reclassification of risk in CAROC when a prior fragility fracture after age 40 and recent prolonged use of glucocorticoids is taken into consideration, resulting in greater concordance between CAROC and FRAX BMD. Overall, these findings indicate that males, younger people, and those without additional risk factors used in FRAX were more likely to experience disagreement in risk classifications.

There are limitations to our study. In the fragility fracture population, CAROC does not allow for low-risk classifications because the presence of fragility fractures increases the risk to the next level. The zero prevalence low-risk categories in CAROC may therefore contribute to asymmetrical disagreement between the 2 tools. To address this issue, the pairwise  $\kappa$  estimate was computed (intervention threshold), which may also be considered more clinically relevant, resulting in a slightly improved  $\kappa$  value. An additional limitation is the focus on the 4 fracture types (wrist, hip, shoulder, or vertebra) because fragility fractures may also occur in other locations, and agreement may have differed. These are considered the stereotypical fragility fractures. In estimating FRAX scores, we did not have information on paternal hip fracture, and therefore relied on self-reports of maternal hip fracture as a proxy for parental hip fracture. However, we examined another OECF dataset that asked about both parents, and we found that very few had a paternal history of hip fracture (3.5%). Thus we would not expect this to have a major effect on our estimate of parental hip fracture. Further, in our current analysis we classified FRAX risk categories by integrating the additional clinical qualifiers, i.e.,

persons with hip, vertebral, or multiple fractures were considered at high risk of refracture, to reflect the use of risk assessments as recommended by the Canadian 2010 CPG for OP<sup>2</sup>. We recognize that applying these additional clinical qualifiers may be inconsistent with the use of FRAX internationally. However, we wished to apply them in our study to allow us to review FRAX and CAROC scores in a manner consistent with the way Canadian clinicians would apply them in practice.

The level of agreement between 2 commonly used fracture risk assessment tools was not as high in the fragility fracture patients as it was in general community-based samples. In our study, the agreement for the 3 by 3 comparison was in the range of “substantial” or “good,” and it improved slightly when we assessed the pairwise  $\kappa$ . These estimates may however fall short of the values typically used in large epidemiological studies or in clinical trials, and are lower than necessary when thinking about their comparability in a clinical decision-making setting. Discordance is highest among less-typical patients with OP (younger, male), who may be in greater need of clearer messaging regarding their refracture risk and treatment options. Overall, the agreement between CAROC and FRAX in the fragility fracture population is suboptimal; our results in combination with those of Beattie, *et al*<sup>24</sup> and those extracted from Leslie, *et al*<sup>11</sup> confirm this. It is beyond the scope of this paper to provide evidence of which of the risk assessments provide the more accurate assessment of risk, and we encourage future work to evaluate the predictive validity of these assessments, particularly where they disagree. Users of these fracture risk tools should nonetheless be aware of the potential for discordance in certain subgroups and note differences in risk classifications that may affect treatment decisions. Future work should examine the predictive value of these fracture risk assessment tools, specifically in cases where there is disagreement. This may provide an evidence base for a decision regarding the preferred fracture risk assessment tool in fragility fracture patients, at least in the Canadian context.

## ACKNOWLEDGMENT

The authors thank Osteoporosis Canada for their thoughtful comments on the paper and the rest of the Fracture Clinic Screening Program Evaluation Team for their contributions to the overall evaluation of the program: Ravi Jain, Susan Jaglal, Gillian Hawker, Muhammad Mamdani, Monique Gignac, Suzanne Cadarette, Famida Jiwa, and Cathy Cameron.

## REFERENCES

1. Kanis JA (World Health Organization Collaborating Centre for Metabolic Bone Disease, University of Sheffield Medical School, Sheffield, UK). Assessment of osteoporosis at the primary health-care level. Technical report. Geneva: World Health Organization Scientific Group; 2007.
2. Papaioannou A, Morin S, Cheung AM, Atkinson S, Brown JP, Feldman S, et al. 2010 clinical practice guidelines for the diagnosis and management of osteoporosis in Canada: summary. *CMAJ* 2010;182:1864-73.
3. Lentle B, Cheung AM, Hanley DA, Leslie WD, Lyons D, Papaioannou A, et al. Osteoporosis Canada 2010 guidelines for the assessment of fracture risk. *Can Assoc Radiol J* 2011;62:243-50.
4. Papaioannou A, Wiktorowicz ME, Adachi JD, Goeree R, Papadimitropoulos E, Bedard M, et al. Mortality, independence in living and re-fracture, one year following hip fracture in Canadians. *J Soc Obstet Gynaecol Can* 2010;22:591-7.
5. van Helden S, Cals J, Kessels F, Brink P, Dinant GJ, Geusens P. Risk of new clinical fractures within 2 years following a fracture. *Osteoporos Int* 2006;17:348-54.
6. Papaioannou A, Kennedy CC, Cranney A, Hawker G, Brown JP, Kaiser SM, et al. Risk factors for low BMD in healthy men age 50 years or older: a systematic review. *Osteoporos Int* 2009;20:507-18.
7. Adachi JD, Ioannidis G, Pickard L, Berger C, Prior JC, Joseph L, et al. The association between osteoporotic fractures and health-related quality of life as measured by the Health Utilities Index in the Canadian Multicentre Osteoporosis Study (CaMos). *Osteoporos Int* 2003;14:895-904.
8. Bliuc D, Nguyen ND, Milch VE, Nguyen TV, Eisman JA, Center JR. Mortality risk associated with low-trauma osteoporotic fracture and subsequent fracture in men and women. *JAMA* 2009;301:513-21.
9. Jaglal SB, Hawker G, Cameron C, Canavan J, Beaton D, Bogoch E, et al. The Ontario Osteoporosis Strategy: implementation of a population-based osteoporosis action plan in Canada. *Osteoporos Int* 2010;21:903-8.
10. Wiktorowicz ME, Goeree R, Papaioannou A, Adachi JD, Papadimitropoulos E. Economic implications of hip fracture: health service use, institutional care and cost in Canada. *Osteoporos Int* 2001;12:271-8.
11. Leslie WD, Berger C, Langsetmo L, Lix LM, Adachi JD, Hanley DA, et al. Construction and validation of a simplified fracture risk assessment tool for Canadian women and men: results from the CaMos and Manitoba cohorts. *Osteoporos Int* 2011;22:1873-83.
12. Leslie WD, Lix LM, Johansson H, Oden A, McCloskey E, Kanis JA, et al. Independent clinical validation of a Canadian FRAX tool: fracture prediction and model calibration. *J Bone Miner Res* 2010;25:2350-8.
13. Fraser LA, Langsetmo L, Berger C, Ioannidis G, Goltzman D, Adachi JD, et al. Fracture prediction and calibration of a Canadian FRAX(R) tool: a population-based report from CaMos. *Osteoporos Int* 2011;22:829-37.
14. Baim S, Leslie WD. Assessment of fracture risk. *Curr Osteoporos Rep* 2012;10:28-41.
15. Kanis JA, Hans D, Cooper C, Baim S, Bilezikian JP, Binkley N, et al. Interpretation and use of FRAX in clinical practice. *Osteoporos Int* 2011;22:2395-411.
16. Bogoch ER, Elliot-Gibson V, Beaton DE, Jamal SA, Josse RG, Murray TM. Effective initiation of osteoporosis diagnosis and treatment for patients with a fragility fracture in an orthopaedic environment. *J Bone Joint Surg Am* 2006;88:25-34.
17. Kanis JA, McCloskey EV, Johansson H, Oden A, Melton LJ 3rd, Khaltaev N. A reference standard for the description of osteoporosis. *Bone* 2008;42:467-75.
18. Norman GR, Streiner DL. Biostatistics: the bare essentials. 3rd ed. Shelton, CT: People's Medical Publishing; 2008.
19. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005;85:257-68.
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
21. Fleiss JL, Levin B, Cho Paik M. Statistical methods for rates and proportions. 3rd ed. Hoboken, NJ: Wiley-Interscience; 2003.
22. Aaronson N, Alonso J, Burnam A, Lohr KN, Patrick DL, Perrin E, et al. Assessing health status and quality-of-life instruments:

- attributes and review criteria. *Qual Life Res* 2002;11:193-205.
23. Rotondi MA, Donner A. A confidence interval approach to sample size estimation for interobserver agreement studies with multiple raters and outcomes. *J Clin Epidemiol* 2012;65:778-84.
  24. Beattie K, Adachi J, Ioannidis G, Papaioannou A, Leslie WD, Grewal R, et al. Estimating osteoporotic fracture risk following a wrist fracture: a tale of two systems. *Arch Osteoporos* 2015;10:13.
  25. Werner P. Knowledge about osteoporosis: assessment, correlates and outcomes. *Osteoporos Int* 2005;16:115-27.
  26. Juby AG, Davis P. A prospective evaluation of the awareness, knowledge, risk factors and current treatment of osteoporosis in a cohort of elderly subjects. *Osteoporos Int* 2001;12:617-22.
  27. Giangregorio L, Thabane L, Cranney A, Adili A, deBeer J, Dolovich L, et al. Osteoporosis knowledge among individuals with recent fragility fracture. *Orthop Nurs* 2010;29:99-107.
  28. Gaines JM, Marx KA, Caudill J, Parrish S, Landsman J, Narrett M, et al. Older men's knowledge of osteoporosis and the prevalence of risk factors. *J Clin Densitom* 2010;13:204-9.