

Assessing fracture risk in people with MS: a service development study comparing three fracture risk scoring systems

Ruth Dobson, Sara Geraldine Leddy, Sunay Gangadharan, Gavin Giovannoni

To cite: Dobson R, Leddy SG, Gangadharan S, *et al.* Assessing fracture risk in people with MS: a service development study comparing three fracture risk scoring systems. *BMJ Open* 2013;**3**:e002508. doi:10.1136/bmjopen-2012-002508

► Prepublication history for this paper are available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2012-002508>).

Received 18 December 2012
Revised 31 January 2013
Accepted 14 February 2013

This final article is available for use under the terms of the Creative Commons Attribution Non-Commercial 2.0 Licence; see <http://bmjopen.bmj.com>

Centre for Neuroscience and Trauma, Blizard Institute, Queen Mary University of London, Barts and the London School of Medicine and Dentistry, London, UK

Correspondence to

Dr Ruth Dobson;
ruth.dobson@qmul.ac.uk

ABSTRACT

Objectives: Suboptimal bone health is increasingly recognised as an important cause of morbidity. Multiple sclerosis (MS) has been consistently associated with an increased risk of osteoporosis and fracture. Various fracture risk screening tools have been developed, two of which are in routine use and a further one is MS-specific. We set out to compare the results obtained by these in the MS clinic population.

Design: This was a service development study. The 10-year risk estimates of any fracture and hip fracture generated by each of the algorithms were compared.

Setting: The MS clinic at the Royal London Hospital.

Participants: 88 patients with a confirmed diagnosis of MS.

Outcome measures: Mean 10-year overall fracture risk and hip fracture risk were calculated using each of the three fracture risk calculators. The number of interventions that would be required as a result of using each of these tools was also compared.

Results: Mean 10-year fracture risk was 4.7%, 2.3% and 7.6% using FRAX, QFracture and the MS-specific calculator, respectively ($p < 0.0001$ for difference). The agreement between risk scoring tools was poor at all levels of fracture risk.

Conclusions: The agreement between these three fracture risk scoring tools is poor in the MS population. Further work is required to develop and validate an accurate fracture risk scoring system for use in MS.

Trial registration: This service development study was approved by the Clinical Effectiveness Department at Barts Health NHS Trust (project registration number 156/12).

INTRODUCTION

Suboptimal bone health is increasingly recognised as an important cause of morbidity. Recent National Institute of Clinical Excellence (NICE) guidelines in the UK¹ recommend assessing bone health using one of two validated scoring systems, FRAX and QFracture, in those at risk of fragility

ARTICLE SUMMARY

Article focus

- Recent National Institute of Clinical Excellence (NICE) guidelines in the UK recommend assessing bone health using one of two validated scoring systems, FRAX and QFracture, in those at risk of fragility fracture. However, these fracture risk scoring systems do not take into account all risks associated with fragility fractures.
- Multiple sclerosis (MS) has been associated with an increased hazard ratio of hip fracture, and was one of the two neurological conditions significantly associated with an increased risk of fracture in the Global Longitudinal study of Osteoporosis in Women (GLOW) study.
- This study therefore set out to compare existing fracture risk scoring algorithms in a multiple sclerosis clinic, in order to assess both the effect of using each of the algorithms on further investigations and treatment, and to assess whether the algorithms provide similar results in this clinic population.

Key messages

- The agreement between fracture risk calculators is poor, with QFracture consistently giving lower risk estimates than FRAX.
- While reducing fracture risk should be a priority to all clinicians dealing with chronic conditions associated with an increased risk of fracture, there must be consistency in the way in which fracture risk is calculated. A prospective study is urgently required in order that we can determine the best way to predict and prevent fractures.

fracture.^{2 3} These scores, both of which have been generated from and validated against large databases, allow the calculation of an individual's 10-year fracture risk, both in terms of any fracture and hip fracture. NICE guidelines currently recommend the calculation of fracture risk before proceeding to dual x-ray absorptiometry (DXA) imaging.¹

ARTICLE SUMMARY

Strengths and limitations of this study

- To the best of our knowledge, this is the first study to directly compare fracture risk scoring tools in the MS clinic population, assessing the rate of interventions that would be indicated by using each tool.
- This study is of relatively small sample size; however, it provides important pilot data to support further work in this area.
- The lack of longitudinal follow-up does not allow us to fully assess the relative accuracy of each tool. Further longitudinal prospective studies are required.

However, these fracture risk scoring systems do not take into account all risks associated with fragility fractures. Multiple sclerosis (MS) has been associated with a hazard ratio (HR) of hip fracture of 1.9–4.08,⁴ and was one of the two neurological conditions significantly associated with an increased risk of fracture in the Global Longitudinal study of Osteoporosis in Women (GLOW) study (HR of any fracture in MS 1.7; 95% CI 1.2 to 2.6).⁵ There have been recent efforts to develop a fracture risk calculator that takes into account the increased risk of fracture and osteoporosis associated with MS.⁶ However, this calculator has been developed from a single database, the UK General Practice Database, and has not been validated to date.

There are common factors associated with an increased risk of developing MS and an increased fracture risk, such as vitamin D deficiency and smoking. It therefore seems likely that the increased fracture risk associated with MS develops early in the disease.⁴ Indeed, it has been shown that the lowered bone mineral density (BMD) associated with MS develops while patients remain fully mobile.⁴ This leads to problems using fracture risk assessment tools, as the FRAX algorithm has the lower age limit set at 40 years, while in the QFracture algorithm the lowest age is 30. The mean age of MS diagnosis is approximately 29,⁷ implying that many patients are first seen at a relatively young age.

It has been argued that the 10-year fracture risk at which intervention becomes cost-effective varies according to the country in which the societal cost is modelled.⁸ For a 50-year-old individual, the 10-year fracture risk at which it becomes cost-effective to intervene may be as low as 0.84% in the UK (a relative risk of osteoporotic fracture of 1.83 compared to the general population, similar to that associated with MS; in the USA, treatment at a relative risk of 1.31 is thought to be cost-effective).⁸ This highlights the importance of fracture risk screening in the MS clinic population.

Given the importance of fracture risk screening in the population with MS and the uncertainty regarding which risk calculator to use, we set out to compare the three fracture risk calculator systems in the MS outpatient clinic population. This study enables a direct comparison of the fracture risk estimates generated by these three studies in addition to examining the number of interventions that the use of each of these calculators would result in.

METHODS

Patient selection and data collection

This service development study was approved by the Clinical Effectiveness Department at Barts Health NHS Trust (project registration number 156/12). One hundred patients with clinically definite MS attending either the MS outpatient clinic or the Neurology Daycase Unit were assessed. Sufficient data to enable full fracture risk scoring were available on 88 patients (see [table 1](#) for details of data required for each fracture risk calculator). The use of an assistive device for walking together with details regarding MS duration and treatment, previous DXA imaging, previous fragility or other fracture, and medications used for the treatment of reduced BMD were also recorded.

Fracture risk scoring

Ten-year risk of both ‘any fracture’ and ‘hip fracture’ were assessed using the FRAX scoring algorithm,² the QFracture algorithm³ and the recently proposed MS-specific fracture risk score algorithm.⁶ As the FRAX score algorithm only allows a minimum age of 40 years, patients aged <40 were assigned an age of 40 for the purposes of this calculation. The QFracture algorithm allows a minimum age of 30, and so patients aged <30 were assigned an age of 30. The MS-specific fracture risk calculator does not have a lower age cut-off. A result of this was that patients aged <40 were assigned different ages in at least two of the risk calculations. A subgroup analysis was performed including only those patients aged 40 or over, in order to assess whether the inclusion of patients younger than the cut-off age had affected the results.

In order to assess the number of patients who would require DXA imaging and/or treatment, an imaging threshold of a 10-year fracture risk for any fracture of >5% was assigned. The treatment threshold was taken to be a 10-year fracture risk of >7% for any fracture, and >4% for hip fracture. The UK National Osteoporosis Guideline Group (NOGG) has estimated that, in the UK, pharmacological treatment is cost-effective at all ages when the 10-year probability of major osteoporotic fracture exceeds 7%.⁹ The UK NOGG recommends an age-dependent intervention threshold, which ranges from 1-year to 10-year fracture risk of 7.5–30% for ages 50–80 years.¹⁰ However, these figures are, if anything, somewhat conservative as discussed above.

Statistical analysis

Statistical analysis was performed using PASW V.18.0 (SPSS). Risk score distributions were assessed for normality using a Shapiro-Wilk test, and attempts made to normalise the data using a natural log transformation. As it proved impossible to normalise the data, non-parametric statistical tests were used. The absolute risk scores generated by each fracture risk score were directly compared using the Friedman test. Scores were then

Table 1 Data used in the calculation of fracture risk for each of the three risk scoring algorithms

FRAX	QFracture	MS-specific calculator
Age	Age	Age
Sex	Sex	Sex
Weight; height; BMI	Weight; height; BMI	BMI
Previous fracture	Previous fragility fracture	Previous fracture (any fracture only)
Parental hip fracture	Parental osteoporosis or hip fracture	
Current smoking	Current or previous smoking, number of cigarettes smoked	Current smoking
Glucocorticoid exposure	Regular glucocorticoid exposure	Use of oral/intravenous glucocorticoids in the prior 6 months
Rheumatoid arthritis	Rheumatoid arthritis or SLE	
Secondary osteoporosis		
Alcohol >3 units/day	Alcohol number of units/day	
Femoral neck DXA (if available)		
	Ethnicity	
	Diabetes	
	Nursing/care home residence	
	Falls	History of a fall 3 months–1 year prior
	Dementia	
	Cancer	
	Asthma/COPD	
	Heart attack, angina, stroke, TIA	
	Chronic liver disease	
	Chronic kidney disease	
	Parkinson's disease	
	Malabsorption including Crohn's disease	
	Endocrine problems including thyroid dysfunction	
	Epilepsy/anticonvulsant exposure	Use of anticonvulsants 6 months prior (any fracture risk only)
	Antidepressants	Use of antidepressants in the prior 6 months
	Oestrogen-only HRT	History of fatigue in the prior 6 months (hip fracture only)

BMI, body mass index; COPD, chronic obstructive pulmonary disease; DXA, dual x-ray absorptiometry; HRT, hormone replacement therapy; SLE, systemic lupus erythematosus; TIA, transient ischaemic attack.

Glucocorticoid exposure: defined currently as exposure to oral glucocorticoids or previous exposure to oral glucocorticoids for more than 3 months at a dose of prednisolone of 5 mg daily or more (or equivalent doses of other glucocorticoids).

Secondary osteoporosis: defined as a disorder strongly associated with osteoporosis. These include type I (insulin dependent) diabetes, osteogenesis imperfecta in adults, untreated long-standing hyperthyroidism, hypogonadism or premature menopause (<45 years), chronic malnutrition or malabsorption and chronic liver disease.

compared between pairs of risk scoring systems using the Wilcoxon signed-rank test.

The agreement between individual scores was assessed using a Bland-Altman plot.¹¹ This method allows a visual description of the agreement between scores, in addition to demonstrating any systematic or significant proportional errors between the two sets of results.¹¹ A further analysis was performed by putting the scores obtained into rank order, and separating them into rank order quintiles, an accepted technique used in the MS literature.^{12 13} The agreement between quintiles was then compared using the κ coefficient. Finally, the proportion of individuals meeting the pre-set criteria for DXA imaging and potential treatment intervention was compared using Fisher's exact test.

RESULTS

Subjects

Of the 100 patients recruited, 88 gave sufficient information to allow their 10-year fracture risk to be accurately calculated using the three algorithms. Demographic details of these patients are given in table 2. Mean disease duration was 7.96 years (range 0–30); 76/88 patients were receiving disease-modifying treatment (see table 2 for more information). Of the 100 patients, 42 used a walking aid; of whom 8 required bilateral assistance and 4 used a wheelchair to mobilise. Of the four patients using wheelchairs to mobilise, two were ambulatory with bilateral assistance for short distances, and two were essentially confined to the wheelchair, requiring assistance to transfer. Of the 88 patients who had their fracture risk calculated, 37

Table 2 Patient details

Characteristic	Patients (n=88)
Age (mean, range)	37.5 y (22–56)
Female (n, %)	55 (62.5%)
Disease duration (mean, range)	7.96 y (0–30)
Disease-modifying therapy	76/88 (86.4%) Receiving disease-modifying therapy 5/88 (5.7%) Glatiramer acetate (Copaxone) 15/88 (17.0%) Interferon β preparations 56/88 (63.6%) Natalizumab (Tysabri)
Ambulatory assistance required	37 (42%) Used a walking aid 28 (32%) Unilateral assistance, ie, single stick 6 (7%) Bilateral assistance 3 (3%) Wheelchair
BMI (mean, range)	24.4 (15.5–46.1)
BMI <20 (n, %)	15 (17%)
Current smoking (n, %)	28 (31.8%)
History of falls (n, %)	48 (54.5%)
Previous fragility fracture	0 (0%)
Previous DXA imaging	6 (6.8%)

BMI, body mass index; DXA, dual x-ray absorptiometry.

(42%) used a walking aid; of these, 6 (7%) required bilateral assistance, 2 (2%) used a wheelchair for longer distances only and 1 (1%) was essentially wheelchair bound. Of the 100 patients, 49 reported falling in the preceding 6 months; of the 88 patients with full fracture risk data, 48 (54.5%) reported a history of falls. Twenty two (52%) patients using a walking aid reported falls in the preceding 6 months compared to 28% of those who did not require a walking aid. No patients had a history of a prior fracture meeting the definition of a fragility fracture.¹⁴

Fracture risk

Ten-year risk of any fracture

Mean 10-year fracture risk was 4.7% assessed by FRAX (SD 3.20, range 2.3–19.0), 2.3% assessed by QFracture (SD 2.14, range 0.4–13.0) and 7.6% using the MS-specific calculator (SD 5.05, range 2.0–25.0) (table 3). Despite efforts, it was not possible to normalise the distribution for any of the fracture risk scores. There was an overall significant difference between the scores generated by the three algorithms ($p < 0.001$; Friedman test), which was preserved on pairwise testing ($p < 0.001$ for all comparisons, Wilcoxon signed rank test) (table 3 and figure 1A).

Bland-Altman plots revealed a reasonable agreement between FRAX and QFracture at lower fracture risk scores, but for those patients with a higher fracture risk, a systematic error was apparent with QFracture consistently giving lower risk estimates than FRAX (mean difference 2.68) (figure 2A). When FRAX and the MS-specific risk score were compared, the agreement was poor, with FRAX consistently lower than the MS-specific score (mean difference 2.97) (figure 2B). The same could be seen when QFracture and the MS-specific risk score were compared (mean difference 5.60; data not shown).

Given that 50 of the patients were younger than 40, the minimum age used in the FRAX calculation, the results

obtained for the 38 patients aged 40 or over were compared in a sub-group analysis. This revealed similar results to those obtained when all patients were included (figure 1B). The highly significant difference in the results obtained by all three fracture risk scores remained ($p < 0.001$ for all comparisons, Wilcoxon signed rank test).

The κ coefficient was calculated for the agreement between rank quintiles for pairs of fracture risk scores. All comparisons generated a low κ value, indicative of poor agreement between rank quintile assignments (FRAX vs QFracture: κ 0.065, 95% CI –0.05 to 0.181, weighted κ 0.133; FRAX vs MS-specific score: κ 0.084, 95% CI –0.029 to 0.197, weighted κ 0.225, QFracture vs MS-specific score: κ 0.114 95%, CI –0.006 to 0.235, weighted κ 0.057).

The number of patients who met the predetermined criteria for DXA imaging and treatment is given in table 3. There was a significant difference between all three groups for both DXA imaging ($p < 0.0001$ for all comparisons) and treatment ($p = 0.03$ when comparing FRAX and QFracture, otherwise $p < 0.0001$). Of the six patients who had previously undergone DXA imaging, three met the criteria for imaging using either the FRAX or MS-specific risk score, but none met the fracture risk cut-off for imaging using the QFracture algorithm. Of the six patients who had undergone DXA imaging, four were on no treatment, one was taking calcium supplementation and one hormone replacement therapy. None of the patients had been diagnosed with osteoporosis.

Ten-year risk of hip fracture

Mean 10-year hip fracture risk was 0.7% as assessed by FRAX (SD 0.95, range 0.1–5.6), 0.2% assessed by QFracture (SD 0.55, range 0.0–4.8) and 3.4% using the MS-specific calculator (SD 7.78, range 0.0–55.0) (table 3). Again, it was not possible to normalise the distribution for any of the fracture risk scores. There was an

Table 3 10-year fracture risks generated

	FRAX any fracture (10-year % risk)	QFracture any fracture (10-year % risk)	MS-specific any osteoporotic fracture (10-year % risk)	FRAX hip fracture (10-year % risk)	QFracture hip fracture (10-year % risk)	MS-specific hip fracture (10-year % risk)
SD	4.69 (3.20)	2.04 (2.14)	7.64 (5.05)	0.66 (0.95)	0.23 (0.55)	3.39 (7.78)
Median	3.45	1.20	6.00	0.30	1.20	0
Range	2.3–19.0	0.4–13.0	2.0–25.0	0.10–5.60	0–4.80	0–55.00
Patients meeting criteria for DXA (n, %)	27 (30.7%)	6 (6.8%)	65 (73.9%)			
Patients meeting criteria for treatment (n, %)	12 (13.6%)	3 (3.4%)	38 (43.2%)	2 (2.3%)	1 (1.1%)	22 (25.0%)

DXA, dual x-ray absorptiometry; MS, multiple sclerosis.

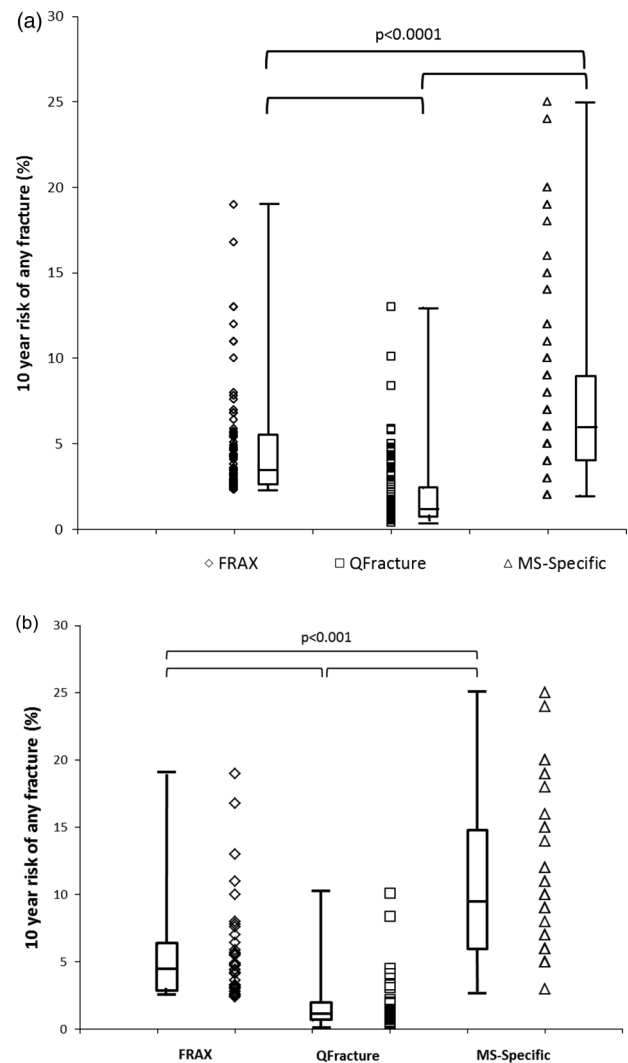


Figure 1 (A) Combined scatter and box-and-whisker plot demonstrating a 10-year any fracture risk generated by each of the three risk scoring algorithms. The box represents the 25th–75th centile bisected by the median, with the whiskers the range. (B) 10-year any fracture risks generated by each of the three risk scoring algorithms for those patients aged 40 or over only. The box represents the 25th–75th centile bisected by the median, with the whiskers the range.

overall significant difference between the scores generated by the three algorithms ($p < 0.001$; Friedman test), which was preserved on pairwise testing ($p = 0.004$ for comparison of FRAX and MS-specific risk calculator, $p < 0.001$ for other comparisons, Wilcoxon signed rank test) (table 3 and figure 3A). Again, when only those aged 40 or over were analysed separately, the significant difference in the 10-year fracture risks generated by the fracture risk calculators differed significantly (figure 3B). Agreement between rank quintiles was poor (FRAX vs QFracture: κ 0.022, 95% CI -0.108 to 0.152 , weighted κ 0.033; FRAX vs MS-specific score: κ 0.016, 95% CI -0.096 to 0.129 , weighted κ 0.107, QFracture vs MS-specific score: κ 0.165 95% CI 0.035 to 0.295 , weighted κ 0.235).

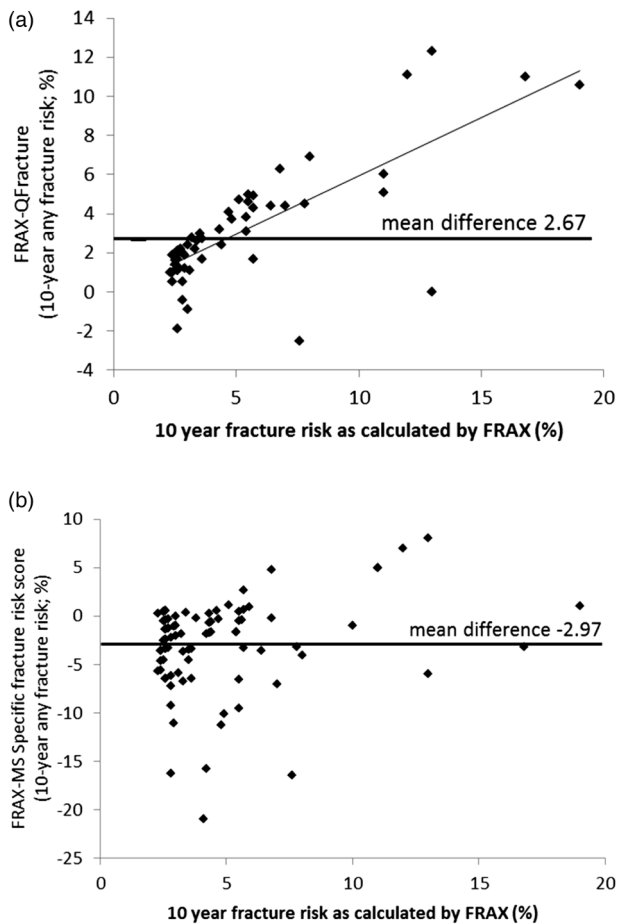


Figure 2 (A) Bland-Altman plot comparing FRAX and QFracture scores for 10-year risk of any fracture. (B) Bland-Altman plot comparing FRAX and MS-specific scores for 10-year risk of any fracture.

The comparison of scores generated similar Bland-Altman plots to those seen when examining the data for overall fracture risk (data not shown). There was no significant difference between FRAX and QFracture in the number of patients who met the predetermined treatment threshold; however, both differed significantly from the MS-specific fracture risk calculator ($p < 0.0001$ for both comparisons).

DISCUSSION

From the results presented above, it can be seen that the agreement between the three fracture risk calculators is poor in this population, both in absolute terms and when examining rank quintiles. There have been previous attempts to compare the results obtained by FRAX and QFracture,^{15 16} and these have highlighted similar issues.¹⁷ When the authors of the MS-specific score compared the scores generated by their model to those generated by the FRAX algorithm, they found that FRAX appeared to significantly underestimate fracture risk for patients with MS, especially with regard to hip fracture.⁶

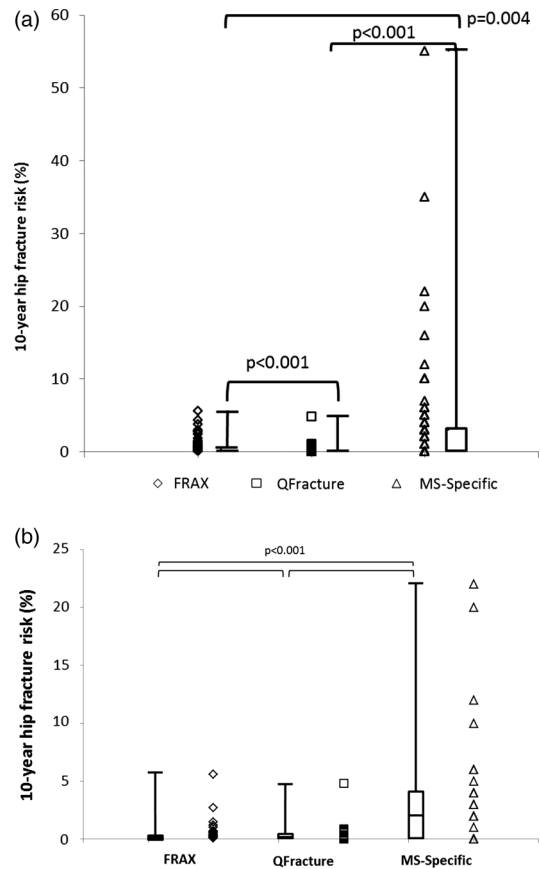


Figure 3 (A) 10-year hip fracture risks generated by each of the three risk scoring algorithms for all patients. The box represents the 25th–75th centile bisected by the median, with the whiskers the range. (B) 10-year hip fracture risks generated by each of the three risk scoring algorithms for those patients aged 40 or over only. The box represents the 25th–75th centile bisected by the median, with the whiskers the range.

One failure of the FRAX algorithm is that it does not allow the calculation of accurate risk for those aged < 40 ; as the majority of our patients were aged between 20 and 40, this could represent a source of error. However, the findings did not differ significantly when only those patients aged 40 or over were included. A significant limitation of this study was the relatively small sample size ($n = 88$), which was further reduced in the subgroup analysis of those patients aged 40 or over ($n = 38$). This must be borne in mind when interpreting the results; however, the magnitude of the differences cannot be ignored.

The FRAX algorithm has been criticised previously for not incorporating factors such as falls¹⁵; while this is a significant omission, it would be expected that this would lead to an underestimation of risk when using this calculator. However, the QFracture algorithm, which does incorporate falls into the calculation, gave consistently lower fracture risk scores than the FRAX calculator. This may be due to a more accurate estimate of age-specific risk, as the QFracture allows the imputation of age from age 30. However, despite including Parkinson’s

disease as a factor, the QFracture calculator does not include MS.

The MS-specific calculator appeared to be consistently overestimating fracture risk. While one might imagine that this would be the most accurate of the risk calculators, the number of patients judged to be over either the investigation or treatment threshold was far higher than expected. This risk calculator was generated from the UK General Practice (GP) database. It incorporates a number of factors into the risk calculator that are not captured by other risk calculators, such as recent steroid use (as a surrogate for relapses) and fatigue (hip fracture risk calculation only). However, it has been shown previously that coding is of variable accuracy in the GP database.¹⁸ As most of the short courses of intravenous steroids are given in secondary care, it is not inconceivable that these would not be captured accurately. Additionally, while more than half of the patients with MS report fatigue when directly questioned,¹⁹ it is quite likely that only those with the very highest levels of fatigue have this recorded by their GPs.

The lack of agreement between the fracture risk calculation tools is likely to be, at least in part, a result of the ways in which they have been developed. FRAX was developed using fracture incidence rates in the UK general population, while the other two calculators have been generated using the UK GP database. Differences in fracture reporting and recording, together with differences in the recording of fracture risk factors between these databases, are likely to contribute to the differences between the results generated in the population studied. This study highlights the fact that the results generated by one fracture risk scoring tool cannot be substituted for those generated by another, and consistent use of a single tool within a population is required to stratify risk in that population. Similarly, the thresholds for further investigation or treatment are likely to vary between the risk scoring tools. The NOGG guidance⁹ has been developed with reference to the FRAX tool, and so should be used in conjunction with this.

To date, there are no papers examining the primary prevention of osteoporosis or osteoporotic fractures in MS. There is also a need to assess the effect of MS disease-modifying treatments on fracture risk. Whether MS disease-modifying treatments have any effect on BMD outside of a general beneficial effect on bone health through the maintenance of weight-bearing mobility remains controversial. Theoretically, interferon-beta preparations should protect against bone mineral loss in MS through induction of the tumour necrosis factor-related apoptosis-inducing ligand (TRAIL).²⁰ There is a single paper demonstrating that people with MS treated with interferon-beta had z-scores significantly greater than zero,²¹ but there was no control group in this study, meaning that it is impossible to draw any firm conclusions regarding this. There remains much work to be done with regard to assessing fracture risk in the MS population, who are at a high risk of fracture and

associated complications. While reducing fracture risk should be a priority to all clinicians dealing with chronic conditions associated with an increased risk of fracture, there must be consistency in the way in which fracture risk is calculated. A prospective study on the MS population, encompassing both fracture risk calculation and bone densitometry estimation using DXA, is urgently needed in order that we can determine the best way to assess the risk, and act to prevent fractures.

Contributors RD and GG conceived the idea of this study. RD, SGL and SG performed fracture risk scoring. RD performed the statistical analysis and initially drafted the paper. All authors provided input into the final manuscript. GG is the guarantor.

Funding RD is funded by an Association of British Neurologists/MS Society of Great Britain Clinical Research Fellowship. GG receives grant support from the MRC, National MS Society, MS Society of Great Britain and Northern Ireland, AIMS2CURE and the Roan Charitable Trust.

Competing interests RD, SGL and SG have no conflicts of interest to declare. GG has received research grant support from Bayer-Schering Healthcare, Biogen-Idec, GW Pharma, Merck Serono, Merz, Novartis, Teva and Sanofi-Aventis. GG has received personal compensation for participating on Advisory Boards in relation to clinical trial design, trial steering committees and data and safety monitoring committees from: Bayer-Schering Healthcare, Biogen-Idec, Eisai, Elan, Fiveprime, Genzyme, Genentech, GSK, Ironwood, Merck-Serono, Novartis, Pfizer, Roche, Sanofi-Aventis, Synthon BV, Teva, UCB Pharma and Vertex Pharmaceuticals.

Ethics approval This service development study was approved by the Clinical Effectiveness Department at Barts Health NHS Trust (project registration number 156/12).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Unpublished raw data (anonymised) will be made available to those contacting the corresponding authors.

REFERENCES

1. <http://guidance.nice.org.uk/CG/Wave25/2> (accessed 1 Nov 2012).
2. Kanis JA, Johnell O, Oden A, *et al*. FRAX and the assessment of fracture probability in men and women from the UK. *Osteoporos Int* 2008;19:385–97.
3. Hippisley-Cox J, Coupland C. Derivation and validation of updated QFracture algorithm to predict risk of osteoporotic fracture in primary care in the United Kingdom: prospective open cohort study. *BMJ* 2012;344:e3427.
4. Dobson R, Ramagopalan S, Giovannoni G. Bone health and multiple sclerosis. *Mult Scler* 2012;18:1522–8.
5. Dennison EM, Compston JE, Flahive J, *et al*. Effect of co-morbidities on fracture risk: findings from the Global Longitudinal Study of Osteoporosis in Women (GLOW). *Bone* 2012;50:1288–93.
6. Bazellier MT, van Staa TP, Uitdehaag BM, *et al*. A simple score for estimating the long-term risk of fracture in patients with multiple sclerosis. *Neurology* 2012;79:922–8.
7. Weinshenker BG, Bass B, Rice GP, *et al*. The natural history of multiple sclerosis: a geographically based study. I. Clinical course and disability. *Brain* 1989;112(Pt 1):133–46.
8. Borgstrom F, Johnell O, Kanis JA, *et al*. At what hip fracture risk is it cost-effective to treat? International intervention thresholds for the treatment of osteoporosis. *Osteoporos Int* 2006;17:1459–71.
9. Kanis JA, McCloskey EV, Johansson H, *et al*. Case finding for the management of osteoporosis with FRAX—assessment and intervention thresholds for the UK. *Osteoporos Int* 2008;19:1395–408.
10. <http://www.shef.ac.uk/NOGG/index.html> (accessed 1 Nov 2012).
11. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
12. Levin LI, Munger KL, Rubertone MV, *et al*. Temporal relationship between elevation of Epstein-Barr virus antibody titers and initial onset of neurological symptoms in multiple sclerosis. *JAMA* 2005;293:2496–500.
13. Munger KL, Levin LI, Hollis BW, *et al*. Serum 25-hydroxy vitamin D levels and risk of multiple sclerosis. *JAMA* 2006;296:2832–8.

14. Murad MH, Drake MT, Mullan RJ, *et al.* Clinical review. Comparative effectiveness of drug treatments to prevent fragility fractures: a systematic review and network meta-analysis. *J Clin Endocrinol Metab* 2012;97:1871–80.
15. Johansen A. QFracture is better than FRAX tool in assessing risk of hip fracture. *BMJ* 2012;345:e4988.
16. Cummins NM, Poku EK, Towler MR, *et al.* Clinical risk factors for osteoporosis in Ireland and the UK: a comparison of FRAX and QFractureScores. *Calcif Tissue Int* 2011;89:172–7.
17. Collins GS, Michaelsson K. Fracture risk assessment: state of the art, methodologically unsound, or poorly reported? *Curr Osteoporos Rep* 2012;10:199–207.
18. Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the general practice research database: a systematic review. *Br J Gen Pract* 2010;60:e128–36.
19. Wood B, van der Mei I, Ponsonby AL, *et al.* Prevalence and concurrence of anxiety, depression and fatigue over time in multiple sclerosis. *Mult Scler* 2013;19:217–24.
20. Weinstock-Guttman B, Hong J, Santos R, *et al.* Interferon-beta modulates bone-associated cytokines and osteoclast precursor activity in multiple sclerosis patients. *Mult Scler* 2006;12:541–50.
21. Shuhaibar M, McKenna MJ, Au-Yeong M, *et al.* Favorable effect of immunomodulator therapy on bone mineral density in multiple sclerosis. *Ir J Med Sci* 2009;178:43–5.