Fracture risk assessment by the FRAX model

E. V. McCloskey\textsuperscript{a,b}, N. C. Harvey\textsuperscript{c,d}, H. Johansson\textsuperscript{a,e}, M. Lorentzon\textsuperscript{f,g}, E. Liu\textsuperscript{e}, L. Vandenput\textsuperscript{e,g}, W. D. Leslie\textsuperscript{h} and J. A. Kanis\textsuperscript{a,e}

\textsuperscript{a}Centre for Metabolic Bone Diseases, University of Sheffield Medical School, Sheffield, UK; \textsuperscript{b}Centre for Integrated research in Musculoskeletal Ageing (CIMA), Mellanby Centre for Musculoskeletal Research, University of Sheffield, Sheffield, UK; \textsuperscript{c}MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK; \textsuperscript{d}NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, UK; \textsuperscript{e}Mary McKillop Health Institute, Australian Catholic University, Melbourne, VIC, Australia; \textsuperscript{f}Centre for Bone and Arthritis Research (CBAR), Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; \textsuperscript{g}Geriatric Medicine, Department of Internal Medicine and Clinical Nutrition, Institute of Medicine, University of Gothenburg, Sweden; \textsuperscript{h}Department of Internal Medicine, University of Manitoba, Winnipeg, MB, Canada

\textbf{ABSTRACT}

The introduction of the FRAX algorithms has facilitated the assessment of fracture risk on the basis of fracture probability. FRAX integrates the influence of several well-validated risk factors for fracture with or without the use of bone mineral density. Since age-specific rates of fracture and death differ across the world, FRAX models are calibrated with regard to the epidemiology of hip fracture (preferably from national sources) and mortality (usually United Nations sources). Models are currently available for 73 nations or territories covering more than 80% of the world population. FRAX has been incorporated into more than 80 guidelines worldwide, although the nature of this application has been heterogeneous. The limitations of FRAX have been extensively reviewed. Arithmetic procedures have been proposed in order to address some of these limitations, which can be applied to conventional FRAX estimates to accommodate knowledge of dose exposure to glucocorticoids, concurrent data on lumbar spine bone mineral density, information on trabecular bone score, hip axis length, falls history, type 2 diabetes, immigration status and recency of prior fracture.

\textbf{Introduction}

Osteoporosis is operationally defined on the basis of bone mineral density (BMD) assessment by dual-energy X-ray absorptiometry (DXA), with recent refinements of the description focusing on measurements at the femoral neck as a reference standard \cite{1}. The World Health Organization (WHO)-defined T-score of -2.5 standard deviations (SDs) or lower, originally designed for classification in epidemiological studies, has since been widely adopted as both a diagnostic and an intervention threshold. The principal difficulty for fracture risk assessment is that whereas this threshold has high specificity, it has low sensitivity, such that the majority of fragility fractures occur in individuals with BMD values above the osteoporosis threshold \cite{2}. Many risk factors have been identified over the last two decades that contribute to fracture risk, at least partly if not wholly independently of DXA BMD. These include age, sex, a prior fracture, a family history of fracture and lifestyle risk factors such as physical inactivity and smoking \cite{3}. These and other factors have been combined in analyses of individual cohort studies to develop algorithms and scores to characterize future risk at the level of an individual. Such independent risk factors used with BMD can enhance fracture risk assessment; additionally, the incorporation of risk factors that correlate with BMD (e.g. age, fracture, body mass index [BMI]) can also facilitate fracture risk assessment in situations in which DXA is not available. These were the considerations underlying the development of the FRAX tool, which was devised by the former WHO Collaborating Centre at the University of Sheffield \cite{4}.

\textbf{Components of FRAX}

The principal aim of treatments for osteoporosis is to decrease the risk of fragility fractures. Thus, the ability to assess fracture risk is critical in identifying patients who are eligible for therapeutic intervention. FRAX is a fracture risk assessment tool for estimating the individualized 10-year probability of hip and major osteoporotic fracture (hip, clinical spine, distal forearm or proximal humerus) \cite{3,5} and integrates eight clinical risk factors (CRFs): prior fragility fracture, parental hip fracture, smoking, systemic glucocorticoid use, excess alcohol intake, BMI, rheumatoid arthritis and other causes of secondary osteoporosis. These, in addition to age and sex, contribute to a 10-year fracture risk estimate independently of BMD. The BMD at the femoral neck is an optional input variable.
Unlike other fracture risk calculators, FRAX computes fracture probability, accounting for both the risk of fracture and the risk of death. This is important because some of the risk factors affect both of these outcomes. Examples include increasing age, low BMI, low BMD, glucocorticoids and smoking. Other risk engines calculate the risk of a clinical event without taking into account the possibility of death [6–8].

Models and uptake of FRAX

Fracture probability differs markedly within and across regions of the world [9,10], and thus FRAX models are calibrated to the epidemiology of fracture and mortality in individual countries. Models are currently available for 73 nations or territories, covering more than 80% of the world population [11]. The FRAX website (http://www.shef.ac.uk/FRAX) receives approximately 3 million visits annually, and the tool is available in 35 languages. Website usage markedly underestimates the uptake of FRAX since this is not the sole portal for the calculation of fracture probabilities using the FRAX tool. For example, FRAX is available in BMD equipment, on smartphones and, in some countries, through hand-held calculators. However, access to the website provides a good overview of global usage of the tool [12] and its role in routine daily practice, as evidenced by the impact of the COVID-19 pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [13].

Performance characteristics

The characteristic of major importance, for the purpose of risk assessment, is the ability of a tool to correctly predict the occurrence of new fractures, traditionally expressed as the increase in relative risk per SD unit increase in risk score. This is termed the gradient of risk. The gradient of risk with the use of FRAX is presented in Table 1 for the use of the CRFs alone, femoral neck BMD alone and the combination [14]. Overall, the predictive value compares very favorably with other risk engines such as the Gail score for breast cancer [15].

Table 1. Gradients of risk with the use of bone mineral density (BMD) at the femoral neck, clinical risk factors (CRFs) or the combination [14].

<table>
<thead>
<tr>
<th>Gradient of risk</th>
<th>Age (years)</th>
<th>BMD only</th>
<th>CRFs alone</th>
<th>CRFs + BMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hip fracture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>3.68 (2.61–5.19)</td>
<td>2.05 (1.58–2.65)</td>
<td>4.23 (3.12–5.73)</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>3.07 (2.42–3.89)</td>
<td>1.95 (1.63–2.33)</td>
<td>3.51 (2.85–4.33)</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>2.78 (2.39–3.23)</td>
<td>1.84 (1.65–2.05)</td>
<td>2.91 (2.56–3.31)</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>2.28 (2.09–2.50)</td>
<td>1.75 (1.62–1.90)</td>
<td>2.42 (2.18–2.69)</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>1.70 (1.50–1.93)</td>
<td>1.66 (1.47–1.87)</td>
<td>2.02 (1.71–2.38)</td>
<td></td>
</tr>
<tr>
<td>Other osteoporotic fractures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>1.19 (1.05–1.34)</td>
<td>1.41 (1.28–1.56)</td>
<td>1.44 (1.30–1.59)</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>1.28 (1.18–1.39)</td>
<td>1.48 (1.39–1.58)</td>
<td>1.52 (1.42–1.62)</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>1.39 (1.30–1.48)</td>
<td>1.55 (1.48–1.62)</td>
<td>1.61 (1.54–1.68)</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>1.54 (1.44–1.65)</td>
<td>1.63 (1.54–1.72)</td>
<td>1.71 (1.62–1.80)</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>1.56 (1.40–1.75)</td>
<td>1.72 (1.58–1.88)</td>
<td>1.81 (1.67–1.97)</td>
<td></td>
</tr>
</tbody>
</table>

Data presented as relative risk per standard deviation change (95% confidence interval). Table reproduced with kind permission from Springer Science and Business Media BV.

Whereas both BMD and the CRFs alone provide significant gradients of risk, the best performance (highest gradients of risk) is observed when BMD is also entered into the FRAX model. Importantly, the impact of the CRFs and BMD is not purely multiplicative as there is some interdependence ($r = -0.25$). The importance of this observation is that the selection of patients with high FRAX probability, but without knowing their BMD, will preferentially select patients with low BMD, and that the higher the fracture probability, the lower will be the BMD [16,17]. These findings consistently indicate that the categorization of patients at high risk on the basis of FRAX without use of DXA selects patients with low BMD, and the higher the probability, the lower the BMD. This has obvious significance for case finding in the absence of access to DXA scanning.

Validation

The performance characteristics of FRAX have been evaluated in 11 independent cohorts that did not participate in the model synthesis. In all of the validation cohorts, the use of CRFs alone or in combination with BMD gave gradients of fracture risk that differed significantly from unity and which were comparable to those in the original cohorts used for model building (see Table 1) [14].

Calibration

Since age-specific rates of fracture and death differ across the world, all FRAX models are calibrated with regard to the epidemiology of hip fracture (preferably from national sources) and mortality (usually United Nations sources). Thus, if the population of each country was to be ‘FRAXed’, the numbers of hip fractures and deaths estimated would match those indicated by the source data. It follows that the calibration of the FRAX algorithms is only as good as the epidemiology with which the tools are populated. Additionally, any validation exercise will be critically dependent on the representativeness of the population tested for the index country. There are several studies that have examined populations that are nationally representative. The first was based on a UK prospective open cohort study of more than 2 million men and women aged 30–85 years using routinely collected data from 357 general practices [8]. The area under the receiver operating characteristic curve for the FRAX algorithm in hip fracture prediction was 0.85 for women and 0.82 for men. Given the small differences in the incidence of hip fracture assumed by FRAX and that observed in the cohort, FRAX appears well calibrated for the UK. Similar findings were reported from Norway: the area under the receiver operating characteristic curve for hip fracture prediction was 0.81 (95% confidence interval (CI) 0.78–0.83) for women and 0.79 (95% CI 0.76–0.83) for men [18]. In two separate studies from Israel, the area under the receiver operating characteristic curve for hip fracture was 0.82 (95% CI 0.81–0.82) in both [19,20].

Fracture probabilities based on the Canadian FRAX tool (both without and with BMD) were compared with observed 10-year fracture incidence from men and women in the
Leslie et al. [22]. BMD (open circles). Dashed line depicts the line of identity. Redrawn from observed 10-year fracture rates (95% confidence interval) by risk category (low, Canadian FRAX tool with and without bone mineral density (BMD) versus Figure 1. Ten-year fracture probability for a major fracture derived from the Canadian FRAX tool with and without BMD at the femoral neck in Caucasian women aged 20–29 years (the international T-score refer-ence used in FRAX). Adapted from Orimo et al. [25].

The use of FRAX in assessment guidelines
FRAX has been incorporated into more than 80 guidelines worldwide [23], although the nature of this application has been heterogeneous. Several guidelines have adopted FRAX into pre-existing guidelines. In the USA, for example, the gateway to treatment includes either a prior fracture (hip or spine fracture) or a BMD T-score of less than −2.5 SDs irrespective of FRAX probability [24]. FRAX is reserved for individuals in whom the T-score is in the osteopenic range and treatment recommended if the probability of a major fracture or hip fracture lies at 20% or more or 3% or more, respectively. Similarly in Japan, the use of FRAX is reserved for individuals without a prior fracture and a BMD that lies between a T-score of −1.8 and −2.7 SDs, with treatment recommended if the probability of a major fracture is 15% or more [25] (Figure 2).

The setting of intervention thresholds is complex and has been approached in a variety of ways (for a detailed review, see Kanis et al. [23]). In the USA, for example, the thresholds (20% for a major osteoporotic fracture and 3% for hip fracture probability) were based on an economic analysis [26], which is both time and health-care system dependent (i.e. costs of treatment change with time, sometimes rapidly, and health-care systems will have differing fracture risks, costs of fracture, willingness to pay and a myriad of other factors to be considered). Other countries (e.g. Finland, Switzerland, Sweden) have determined intervention thresholds more appropriate to the local health-care setting [23]. Other approaches have decided inappropriately to use fracture risk assessment tools, such as FRAX, as the means of identifying patients with BMD-defined osteoporosis, an aim for which such tools were not primarily designed [27–29].

The use of BMD alone or BMD with prior fracture as a gateway to assessment is not without problems, as recently reviewed [30]. First, although reduced bone mass is easily quantifiable and strongly related to fracture risk, most fragility fractures occur in individuals with a BMD T-score above the operational threshold for osteoporosis [2]. Second, the significance of any given T-score threshold differs by age. For example, at age 65 years, a T-score of −2.5 SDs confers a modest increase in the probability of fracture compared with women with no CRFs and in whom BMD is not measured. With advancing age, the difference in the probability of fracture between the general population and those with a T-score of −2.5 SDs reduces; indeed, from the age of 78 years in the USA, fracture probability becomes progressively lower than that of the age and sex-matched general population (Figure 3) [30]. Thus, a T-score of −2.5 SDs becomes a protective factor from the age of 78 years in the USA, relative to the general population. Third, fracture rates differ widely between countries, much more so than can be explained by variations in BMD [9]. Thus, the T-score at a given probability will vary from country to country; for example, for an
intervention threshold set at a 10-year probability of a major fracture of 20% in women aged 65 years, the femoral neck T-score ranges from –4.6 SDs in Venezuela to –2.0 SDs in Iceland [30].

For the reasons presented, FRAX rather than BMD is increasingly used as the principal gateway for assessment. The approach is summarized in Figure 4 [5]. The management process begins with the assessment of fracture probability and the categorization of fracture risk on the basis of age, sex, BMI and the CRFs. Using this information alone, some patients at high fracture risk may be offered treatment without use of BMD testing (e.g. prior fracture). There will be other instances where the probability will be so low that a decision not to treat can be made without BMD. An example might be the well woman at menopause with no CRFs. Thus, not all individuals require a DXA scan and are thus excluded from the intermediate category in Figure 4.

The size of the latter category will vary in different countries. In the USA, this would be a large category, whereas in a large number of countries with limited or no access to DXA, the size of the intermediate group will necessarily be small. In other countries (e.g. the UK), where provision for BMD testing is suboptimal, the intermediate category will lie between the two extremes. It has been conservatively estimated that a minimum of 10 DXA units are required per million of the population and such provision is available for fewer than 20 countries worldwide [31].

The first step in defining the intermediate group is to establish an intervention threshold and target DXA scans to those lying at or around this threshold, in order to maximize the impact of the scan on decision-making. Nearly all guidelines internationally recommend that women with a prior fragility fracture should be considered for intervention without the necessity for a DXA scan (other than to monitor treatment). Since a prior fracture is associated with sufficient risk that treatment can be recommended, the intervention threshold in women without a prior fracture can be set at the age-specific fracture probability equivalent to women with a prior fragility fracture and therefore rises with age, for example from a 10-year probability of 8% to 33% in the UK [5]. This may be termed the ‘fracture threshold’. This approach to intervention thresholds, first used by the UK National Osteoporosis Guideline Group (NOGG) [32], has since been adopted into European guidelines and elsewhere [23,33,34]. The same intervention threshold is applied to men, since the effectiveness and cost-effectiveness of interventions in men are broadly similar to that in women for equivalent risk. In the UK, a subsequent amendment to flatten the threshold from the age of 70 years and upward addressed a possible inequality in those selected for treatment with and without prior fracture at older ages [35]. More recently, new evidence from head-to-head trials of anabolic versus antiresorptive treatment [36–39] has led to the concept of stratifying high risk to delineate a very high risk category where considerations might include first-line anabolic treatment [40,41].

**FRAX and efficacy of intervention**

European guidelines on the evaluation of medicinal products in the treatment of primary osteoporosis place an emphasis on the study of patients at high fracture risk [42]. As a consequence, FRAX has been applied in predominantly post hoc analyses of several phase III studies to determine the enrolment characteristics of patients. This information has also been used to determine whether treatment efficacy varies according to baseline fracture risk. Interventions studied include abaloparatide, raloxifene, bazedoxifene, clodronate,
daily and weekly teriparatide, denosumab, alendronate, strontium ranelate and, most recently, romosozumab [43–53].

Greater efficacy against fracture in individuals at higher risk treated with clodronate, denosumab, bazedoxifene or romosozumab has been demonstrated. This FRAX-dependency has marked economic consequences, illustrated in Table 2 when comparing two hypothetical treatments with similar overall effectiveness on fracture risk but the efficacy of one increases in women with higher baseline fracture probability (treatment B). In contrast, the relative risk reduction with the other treatment is constant over the range of fracture probabilities studied (treatment A) (Table 2). As a consequence, treatment A has better cost-effectiveness in terms of fractures saved at low fracture probabilities whereas treatment B has the better cost-effectiveness at high baseline fracture probabilities [54].

These results have a number of important implications. First, they remove any concern that patients identified on the basis of CRFs with FRAX would not respond to pharmacologic interventions. Indeed, these studies showed that high FRAX probabilities are associated with efficacy, even when BMD is not used to characterize risk. Second, they support the views of the regulatory agencies that treatments should be targeted preferentially to men and women at high fracture risk. Third, the finding of greater efficacy at higher fracture probabilities with some interventions has important implications for health technology assessments and challenges the current meta-analytic approach; greater efficacy in the higher risk groups will improve still further the budget impact and the cost-effectiveness of intervention.

The widespread uptake of FRAX for case-finding and its ease of application have raised the question of whether FRAX might be used in population-based screening. Three large randomized prospective studies have recently been published of potential population screening strategies [55–57]. Despite important differences in study design and approaches to intervention thresholds, two of the studies showed significant reductions in hip fractures [55,56]. While the third study failed to show such an effect, a meta-analysis of all three studies showed a 20% reduction in hip fractures with smaller but significant reductions in major osteoporotic fractures and all osteoporotic fractures, despite treatment being targeted at relatively small proportions of the populations studied [57,58]. The approach utilized in the SCOOP study in the UK has been shown to be highly cost-effective or cost-saving [59,60].

### Addressing the limitations of FRAX

The limitations of FRAX have been extensively reviewed and are only briefly addressed here [23,61]. The risk factors included in FRAX were carefully chosen to limit complexity, for ease of input, and to include only well-established, independent contributors to fracture risk. In addition, it was important that the factors used identified a risk that was amenable to an intervention. While appreciated for its simplicity, FRAX has also been criticized for the same reason because it does not take account of exposure response. For example, the risk of fracture increases with exposure to glucocorticoids, but FRAX only accommodates a yes/no response to the relevant question. Other well-researched examples of ‘dose–response’ include the number of prior fractures and the consumption of alcohol. Other concerns are the lack of provision for lumbar spine BMD (which is commonly recommended in treatment guidelines) and the absence of measurements of the material or structural properties of bone. A concern that treatment might invalidate the interpretation of FRAX appears misplaced [62].

The reason that such factors have not been accommodated within FRAX is that there is a lack of international data that would allow validation of their inclusion, including their interaction with other FRAX risk factors. Nonetheless, arithmetic adjustments have been proposed to address some of these limitations, which can be applied to conventional FRAX estimates of probability. These include exploratory adjustment for knowledge of: high, moderate and low exposure to glucocorticoids [63]; concurrent data on lumbar spine BMD [64,65]; information on the trabecular bone score (TBS) [66,67]; hip axis length [68]; falls history [69]; type 2 diabetes [70]; immigration status [71]; and recency of prior fracture [72].

Such analyses can inform the clinician how to temper clinical judgment on the existing output of the FRAX models.

### Summary

The FRAX fracture risk assessment tool, launched in 2008, provides country-specific algorithms for estimating individualized 10-year probability of hip and major osteoporotic fracture (hip, clinical spine, distal forearm or proximal humerus) [73]. FRAX has been incorporated into more than 80 guidelines worldwide, with heterogeneous approaches to setting intervention thresholds. The relationship between FRAX probability of fracture and efficacy of intervention is now well established and is expected to further influence treatment guidelines in the future.

### Acknowledgements

Some text is reproduced from Kanis et al. [54,73,74].

---

**Table 2. Contrasting effects on the number of fractures saved with an intervention, the efficacy of which via the relative risk reduction (RRR) is independent of or dependent on FRAX.**

<table>
<thead>
<tr>
<th>Fracture probability (%)</th>
<th>FRAX-independent</th>
<th></th>
<th>FRAX-dependent</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fractures saved (n)</td>
<td>RRR (%)</td>
<td>Fractures saved (n)</td>
<td>RRR (%)</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>40</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>40</td>
<td>4</td>
<td>27</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>40</td>
<td>6</td>
<td>40</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>40</td>
<td>8</td>
<td>54</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>40</td>
<td>10</td>
<td>68</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>40</td>
<td>12</td>
<td>80</td>
</tr>
<tr>
<td>Total</td>
<td>42</td>
<td>40</td>
<td>62</td>
<td></td>
</tr>
</tbody>
</table>

Average RRR set at 40%.
Potential conflict of interest  Prof. Kanis led the team that developed FRAX as director of the WHO Collaborating Centre for Metabolic Bone Diseases; he has no financial interest in FRAX. Prof. McCloskey, Prof. Harvey, Dr Johansson, Dr Vandenput and Dr Liu are members of the FRAX team. Prof. Kanis, Prof. Harvey and Prof. McCloskey are members of the advisory body to the NOGG, UK. W. D. Leslie declares no competing interests.

Source of funding  Nil.

ORCID
E. V. McCloskey  http://orcid.org/0000-0003-0177-8140

References


