Decision making for osteoporotic treatment using FRAX or DVO risk algorithms in a clinical setting

S. Rendl, C. Lapa, C. Blümel, R.A. Bundschuh, P. Schneider

University Hospital, Department of Nuclear Medicine, Oberdärrbacherstrasse 6, 97080 Würzburg, Germany

Abstract

Objectives: To evaluate the therapy decisive clinical risk factors (CRFs) in tools provided by WHO (WHO-FRAX) and the Head Osteology Organization of Germany (DVO) in a clinical setting, and, the degree of agreement between them. Methods: Three hundred subjects, 40 to 88 years of age, were consecutively referred for an evaluation of osteoporosis-related fracture risk, and therapy was possibly recommended. The evaluation used the 12 CRFs in the FRAX tool and the 21 CRFs in the DVO tool. We analyzed the degree of agreement and the strength of the CRFs in determining the therapy decision. Results: Before evaluation, 52 (17.3%) of the patients took anti-osteoporotic medication. The FRAX tool indicated 36 (12.0%) patients suggested for treatment when hip density was included as a CRF, whereas the DVO tool indicated 80 (26.7%) and 91 (30.3%), depending on bone density site. The pre- and post-test results agreed poorly to fair, whereas agreement was poor to good within both models and using the plain T-score to define the therapy intervention threshold. Conclusions: CRFs with debatable evidence reached significant influence on therapy decision. A considerably divergent number of patients were identified as treatment candidates, deserving further investigation to confirm the usefulness of some CRFs.

Keywords: FRAX, Osteoporosis, Therapy, Fracture Risk, DVO

Introduction

A large 10-year follow-up cohort study has shown that the FRAX tool predicts a 10-year probability of the incidence of hip and major osteoporotic fractures. It was shown that the FRAX tool did not predict prior fractures or “osteoporosis” as defined by the WHO T-score stratification based on bone mineral density (BMD)\(^1\). We used the FRAX tool because it provides a paradigm shift, encouraging physicians and patients to think in terms of fracture probability (=absolute fracture risk). It encourages the consideration of treatment efficacy independent of the BMD categories\(^2\). The Head Osteology Organization of Germany (Dachverband der Osteologischen Gesellschaften, DVO) provided a better risk stratification model based on its S3 guidelines\(^3\); the model potentially facilitates treatment decisions for patients more effectively than the BMD alone. The WHO’s FRAX tool is used worldwide, whereas the DVO tool covers a German-speaking population area of over 100 million people. The partially congruent clinical risk factors (CRFs) implemented in both tools take into account that patients with low bone mass but not yet osteoporotic, particularly women, relate to the concept of fracture probability and primary prevention. Whereas the FRAX tool may be used without entering the femur neck density data, the DVO tool was designed to deliver results only if at least one bone density value is entered.

This study investigated the hypothesis that in a clinical setting, the 2 tools should show agreement on the patients that should be recommended for medication therapy. An investigation of the degree of agreement of the pre-assessment treatment status with the outcomes of the 2 risk assessment tools was needed. Subsequently, to identify the strengths of the CRFs that influence the recommendations, there was a need to address the agreement between the 2 assessment tools and to analyze the deviations between the therapy recommendations of the 2 tools.

Materials and methods

Sample

During a recruitment period of 3 months, 69 men and 261 women, 40 to 88 years of age, were consecutively referred for
evaluation of osteoporosis-related fracture risks and the potential recommendation for medication therapy. All the outpatients had been thoroughly diagnosed by specialists or were referred for diagnoses as inpatients (40% of referrals) in one of the clinics in our hospital. They were subsequently referred to our department for densitometry, with a prior justification for undergoing radiological evaluation. Each patient signed an informed consent form. All the patients referred who did not meet the following exclusion criteria were entered into the study; the CRFs are listed in Table 1. The exclusion criteria were patients with artifacts (implants) in the measurement sites of the radius, lumbar spine or hips; pregnancy; and mental impairment. The study complied with the Declaration of Helsinki standards, and the institutional data safety board of the university hospital agreed with the use of anonymized data of the subjects.

**Bone densitometry**

Dual energy X-ray absorptiometry (DXA) (GE Lunar Prodigy, Lunar Inc., Madison, WI, USA) assessed the areal bone mineral density (aBMD) at the lumbar spine from scans covering thoracic vertebra 12 to lumbar vertebra 5 and the left and right femoral neck. Patients with degenerative vertebrae were excluded from further consideration. Having at least 2 vertebrae without degenerative disease was required for eligibility. DXA provides aBMD information of the bones at relevant fracture sites. The method and the normative database of the T-scores were described in detail elsewhere 4-6. The lower aBMD value of the scans of both femoral necks was used. Peripheral quantitative computed tomography (pQCT) (XCT2000, Stratec GmbH Pforzheim, Germany) was used to measure the total volumetric bone density (vBMD) at the metaphysis (4% of bone length) of the non-dominant radius. We developed the method in the late 1980s to deliver a number of vBMD and strength parameters of the trabecular and cortical bone compartments at the distal radius. We used the total vBMD of a cross sectional slice, which we hypothesize represents a surrogate for the bone strength at exactly this relevant fracture site 7,8. We derived the T-scores of the total vBMD from the reference data of a German multicenter study provided in the pQCT software of the manufacturer 9. The T-score of the vBMD information is useful, but not directly comparable to the T-scores derived from the aBMD. The technologist made the assessments, and the radiologist checked the quality of the

---

**Table 1.** Clinical risk factor (CRF) prevalence as assessed for the FRAX and the DVO score input. The values in parentheses are a % of the total sample where indicated or standard deviation, range in brackets. *FRAX CRF; +DVO CRF.
scan results, the absence of external or internal artifacts, the exclusion of fractured vertebrae or osteophytes, and the accuracy of the measurement site. The long-term precision of both devices was within 0.1% standard deviation using the calibration phantoms.

Risk input

The FRAX tool comprises 12 CRFs: age, gender, weight, height, prevalence of low trauma fractures, parental hip fractures, smoking, >5 mg prednisolon or equivalent glucocorticosteroids medication regime, therapy for an established diagnosis of rheumatism defined by the American College of Rheumatology guidelines, secondary osteoporoses, more than 3 units of alcohol (1 unit=285 ml of beer daily), and femoral neck aBMD http://www.sheffield.ac.uk/FRAX/). The DVO tool features 21 CRFs: age, gender, hip and vertebral aBMD, vertebral fracture status, >5 mg prednisolon or equivalent glucocorticosteroids, parental hip fracture, smoking, falls, immobility, epilepsy, primary hyperparathyroidism, hypercortisolism, gastrostomy, diabetes, antiandrogen therapy, hypogonadism, thyroid-stimulating hormone (TSH), aromatase inhibitors (AIs), and an established diagnosis of rheumatism (http://www.dv-osteologie.org/). These CRFs were identified during the development of the S3 guidelines, which are based on a systematic evidence research, study outcome analysis, evaluation of clinical relevance and frequent review1. There was a discrepancy between the definitions of fracture in the FRAX and DVO tools. The DVO tool discriminated between the grade 1 and grades 2-3 single vertebral fractures. This distinction resulted in an additional 9 cases in the DVO CRFs and features “vertebral fractures” compared to those featured by the FRAX tool as “previous fractures”. The FRAX tool provides a fracture probability estimator but no treatment recommendation10, and the DVO tool results in a “yes” or “no” decision regarding treatment. It appeared that the DVO tool uses a -2.0 aBMD T-score threshold to switch from “no” to “yes” independently from the CRFs in the treatment decision process.

To compare the dichotomous treatment recommendation, we developed a simple decision visualization tool for the FRAX probability estimators. The percentages for major osteoporotic or hip fractures were plotted into the graph provided by Kanis et al.11, allowing for visualization of whether the two probability values exceeded the recommendation threshold (Figure 1). When 1 of the 2 values exceeded the threshold, treatment was suggested.

Statistical analyses

Cohen’s Kappa statistics were applicable to correctly determine the significant differences between the FRAX and the DVO treatment recommendations, thereby correcting the proportion of agreement because of chance by using a 2x2 table of dependent variables12. We calculated the agreement between the pre- and post-risk evaluation of the therapy decision and the proportion of agreement between the assessments. The strength of agreement is defined by the categories poor, fair, moderate, and good, according to the Kappa value.

Logistic regression was used to determine the influence of one or more variables on the bivariate target variable determining the “yes” or “no” therapy decision. The independent variables were the risks entered into the FRAX or DVO scoring systems. The mixed binomial and continuous target variables were estimated using the logistic regression approach, resulting in the logarithm of the chances (log odds ratio) varying between – and + . The statistical analyses were performed using STATISTICA® 10 (StatSoft Inc., Tulsa, OK, USA) and SPSS V.20 (IBM Corp., New York, US). The significance threshold was set at p<0.01.

Results

To compare our results with those obtained from the 2 CRF tools, we counted the number of patients who would have received a treatment recommendation according to the basic WHO T-score threshold definition (-2.5) for osteoporosis (Figure 2, left 3 bars). The mean (±standard deviation) of the 10-year fracture probability for major osteoporotic low trauma fractures obtained from the FRAX tool was 11.3% (8.4) in women and 7.2% (4.2) in men in our sample. For hip fractures, it was 4% (5.6) in women and 3% (3.0) in men. These probabilities were the results of the FRAX CRF input as shown in Table 1. Before the evaluation, 52 (17.3%) of the patients referred took the following medication: 37 took bisphosphonates, 3 took teripartide, 3 took strontium ranelate, 2 took alphacalcidol, 4 took estrogen, 2 took raloxifen, and 1 took calcitonin. These agents were prescribed in combination with calcium in 23 patients, with 1,25 dihydroxyvitamin D3 in 22, or with both in 73.
The percentages of patients recommended for treatment are shown in Figure 2. The FRAX tool produced treatment recommendations for 36 patients (12.0%), and the DVO tool produced treatment recommendations for 80 (26.7%) patients with lumbar spine aBMD and 91 (30.3%) patients with femur neck aBMD. Cohen’s Kappa test revealed the strength of agreement between previous treatments and the treatment recommendations after employing the FRAX and the DVO tools with lumbar spine aBMD to be fair and with the femur aBMD to be fair (Table 2). Table 3 shows the strength of agreement...
between the different risk assessment tools or the WHO definition of osteoporosis based on the T-scores.

The log odds ratios of the risk factors provided a measure of how effectively these factors had influenced the therapy decision. The results are shown in Tables 4-6 for the FRAX tool and for the two DVO runs, using the aBMD of the spine or hip. The significant CRFs are listed.

Table 4. Logistic regression of CRF impacting on therapy recommendation yes/no within the FRAX tool. CRF= clinical risk factor. Only significant CRFs are shown.

<table>
<thead>
<tr>
<th>Regression coefficient B</th>
<th>Standard error</th>
<th>p</th>
<th>Log OR [95% CL]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-.149</td>
<td>.037</td>
<td>.000</td>
</tr>
<tr>
<td>Height</td>
<td>-.124</td>
<td>.053</td>
<td>.020</td>
</tr>
<tr>
<td>Previous fracture</td>
<td>3.871</td>
<td>.762</td>
<td>.000</td>
</tr>
<tr>
<td>Parental hip fracture</td>
<td>3.600</td>
<td>.822</td>
<td>.000</td>
</tr>
<tr>
<td>Smoking</td>
<td>1.954</td>
<td>.876</td>
<td>.026</td>
</tr>
<tr>
<td>Glucocorticosteroids &gt;5 mg prednisol</td>
<td>3.539</td>
<td>.752</td>
<td>.000</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>2.281</td>
<td>.725</td>
<td>.002</td>
</tr>
<tr>
<td>Femoral neck density</td>
<td>.800</td>
<td>.282</td>
<td>.005</td>
</tr>
</tbody>
</table>

Table 5. Logistic regression of CRF impacting on therapy recommendation yes/no within the DVO tool using lumbar spine density. CRF=clinical risk factor. Only significant CRFs are shown.

<table>
<thead>
<tr>
<th>Regression coefficient B</th>
<th>Standard error</th>
<th>p</th>
<th>Log OR [95% CL]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.350</td>
<td>.079</td>
<td>.000</td>
</tr>
<tr>
<td>Gender</td>
<td>1.932</td>
<td>.954</td>
<td>.043</td>
</tr>
<tr>
<td>T-score lumbar spine</td>
<td>-5.826</td>
<td>1.155</td>
<td>.000</td>
</tr>
<tr>
<td>Glucocorticosteroids &gt;5 mg prednisol</td>
<td>3.980</td>
<td>1.415</td>
<td>.005</td>
</tr>
<tr>
<td>Falls</td>
<td>5.525</td>
<td>1.571</td>
<td>.000</td>
</tr>
<tr>
<td>TSH &lt;0.3 mU/l</td>
<td>6.493</td>
<td>2.418</td>
<td>.007</td>
</tr>
<tr>
<td>Aromatase inhibitors</td>
<td>-5.832</td>
<td>2.416</td>
<td>.016</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>6.087</td>
<td>1.561</td>
<td>.000</td>
</tr>
</tbody>
</table>

Table 6. Logistic regression of CRF affecting the therapy recommendation yes/no within the DVO tool using femur density. CRF=clinical risk factor. Only significant CRFs are shown.

<table>
<thead>
<tr>
<th>Regression coefficient B</th>
<th>Standard error</th>
<th>p</th>
<th>Log OR [95% CL]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.263</td>
<td>.050</td>
<td>.000</td>
</tr>
<tr>
<td>Gender</td>
<td>1.885</td>
<td>.752</td>
<td>.012</td>
</tr>
<tr>
<td>T-score Femur</td>
<td>-4.093</td>
<td>.691</td>
<td>.000</td>
</tr>
<tr>
<td>Glucocorticosteroids &gt;5 mg prednisol</td>
<td>4.788</td>
<td>1.131</td>
<td>.000</td>
</tr>
<tr>
<td>Multiple falls</td>
<td>2.772</td>
<td>.940</td>
<td>.003</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>3.377</td>
<td>.785</td>
<td>.000</td>
</tr>
</tbody>
</table>

Discussion

The study sought to determine whether the two CRF tools showed agreement in the patients recommended for treatment. The plain T-scores showed a poor to good agreement in comparison to the CRF-based decisions (Table 3). The pre-assessment therapy showed a poor to fair relationship to the outcome.
suggested by the CRF-driven tools (Table 2). Despite a good agreement between the DVO tool with the spinal aBMD input and the WHO T-score spine definition for osteoporosis, our expectations that the FRAX and the DVO tools should have good agreement regarding patients recommended for treatment were not met, with poor to fair agreement observed between the FRAX and DVO tools.

The log odds ratios in our study express the strength of the risk factors that directly influenced the therapy decision. To understand the differences in the effects observed in the logistic regression analysis, one must consider that the CRFs of the FRAX tool are based on a series of meta analyses that identified factors associated with an increased fracture risk independently of age and aBMD at the femoral neck and other major fracture sites. The DVO tool uses a system based on a literature review that extracted the factors known to affect the bone density and/or fracture. The logistic regression analysis of the FRAX tool CRFs in our sample suggested that a subset of 8 CRFs within the entire set of CRFs had a highly significant effect on the treatment recommendations (Table 4). There is some lack of transparency to the user concerning the mechanism of the DVO algorithm. Within the DVO CRFs modulated by the lumbar spine T-score, 8 CRFs significantly contributed to the treatment recommendations in our sample (Table 5). Using the femur T-score as a modulating CRF, 6 CRFs showed a significant influence on the recommendations for treatment (Table 6). One should consider that this finding is a relative effect of the CRFs in our sample, which contributes in the given combination to the recommendation outcome. Generally, the significant CRFs should not be confused with their expected effect on fracture risk. They express the effect on decision making within our sample and reflect the unequal numbers of subjects with those specific risks.

There is an intense controversy concerning whether TSH suppression is a risk factor for fracture. Recently, it was shown that it does not significantly influence bone density. There is no evidence that AIs increase the fracture risk. Measurement of spinal aBMD may be obsolete because of a number of systematic errors, some of which, such as the soft tissue error, cannot be identified by physicians. Our analyses suggested that inclusion of these risk factors significantly influenced the treatment recommendations without having overestimation in the model. Their small contribution in our regression analysis may be explained by the T-score range of the aBMDs measured in our patients, with means of -0.7 to -1.6, thus exceeding the threshold defined as “osteoporosis” in a fraction of our patients (Table 1).

There are reasons for the need for tools that produce consistent outcomes. Assuming that 3-5 years of treatment with an osteotropic agent would typically be of benefit to a patient by a 5% increase in hip density, it would result only in small effects within the FRAX or the DVO tools. This CRF and others should be considered with caution, and their importance should not be overestimated without sufficient evidence. We considered testing the slight input variations of the FRAX and the DVO tools. Whereas the FRAX tool transparently results in slight changes of the probabilities, the DVO switched to a “yes” suggestion after an invisible score threshold was exceeded. An 80-year-old woman with no additional risk factors and a T-score of -1.99 would not be a treatment candidate, but with a score of -2.0 there would be a “yes” indication for treatment as well as by adding only one of the CRFs, such as TSH suppression, to the T-score of -1.99. We suggest reconsidering implementation of the guideline evaluation into the DVO tool with more care. It does not reveal an insight into the foundation of the scores nor into the mechanism that generates the “yes/no” decision. In contrast, the FRAX tool provided a more transparent system using the CRFs, which appears to be epidemiologically better substantiated, and its primary result, the fracture probability, avoids a dichotomous decision that has a tendency to strongly influence physicians.

In this study, we did not assess the reasons for the pre-test prescriptions for medications. Compared to another study, the pre-FRAX frequency of therapy was higher in our study, and the post-FRAX frequency of therapy was lower. In our study, the post-DVO frequency of medical treatment increased. We did not follow up on whether our risk assessment and recommendations had affected the prescribing behavior of the physicians. We agree that the FRAX tool may have value for guiding the decision regarding the need to continue or withdraw treatment.

The results of the decision-making tools in our patient sample proved to be unequal. This finding may be different in a sample that does not primarily originate from a clinical environment. The inconsistent results of our study would have a great effect on cost-effectiveness issues.

We could demonstrate that two tools that use CRFs to make decisions regarding treatment recommendations produced poor to moderate agreement in their outcomes and identified a considerably divergent number of patients as treatment candidates. Further investigations are needed to confirm the usefulness of the significant CRFs to justify their application for the identification of patients for whom osteoporotic treatment should be considered.

Acknowledgements

The was no external funding for the study. The authors’ roles were as follows: RS did the data acquisition, analysis and interpretation; LC, BRA and BC did the data acquisition; and PS drafted and revised the manuscript.

References