Discrepancies Between Osteoporotic Fracture Evaluations in Men Based on German (DVO) Osteoporosis Guidelines or the FRAX Score

Authors
J C Witzel[1, 2*], A Giessel[1], C Heppner[2], A Lamersdorf[2], A Leha[3], C C Glüer[4], H Siggelkow[1, 2]

Affiliations
1 Clinic of Gastroenterology, Gastrointestinal Oncology and Endocrinology, University Medical Center Göttingen, Robert-Koch-Str. Göttingen, Germany
2 MVZ Endokrinologikum Göttingen, Von-Siebold-Str. Göttingen, Germany
3 Institute for Medical Statistics, University Medical Center Göttingen, Humboldtallee Göttingen, Germany
4 Section Biomedical Imaging, Department of Radiology and Neuroradiology, MOINCC, Kiel, Germany

Key words
bone mineral density, therapy threshold, osteoporosis risk factors, secondary osteoporosis, 10-year fracture risk

received 09.03.2022
revised 11.09.2022
accepted 09.11.2022
published online 09.01.2023

Bibliography
Exp Clin Endocrinol Diabetes 2023; 131: 114–122
DOI 10.1055/a-1977-4413
ISSN 0947-7349

Georg Thieme Verlag, Rüdigerstraße 14,
70469 Stuttgart, Germany

Correspondence
Judith Witzel M.D.
Department of Diagnostic and Interventional Radiology
University Medical Center Göttingen
Robert-Koch-Straße 40
37075 Göttingen
Germany
witzel.judith@gmail.com

ABSTRACT

Introduction Established scores estimate 10-year fracture risk in osteoporosis to assist with treatment recommendations. This study compared the risk probabilities of major osteoporotic and hip fractures calculated by the FRAX tool with those of the DVO score, established in German-speaking countries.

Material and Methods This seven-year retrospective study analyzed data of 125 male patients (mean age: 59.2 ± 10.7 years) evaluated for osteoporosis. For the DVO score, the therapy threshold of > 30 % for vertebral and hip fractures suggested by DVO guidelines was implemented. We calculated fracture risks based on FRAX scores with aBMD and applied a common therapy threshold of ≥ 3 % for hip fracture and subsequently determined the "DVO-equivalent risk level" for FRAX-based assessment that would identify as many male patients as identified by the DVO score.

Results Based on DVO score, 60.0 % of patients had a 10-year risk of hip and vertebral fractures > 30 %. The recommendations for individuals based on FRAX scores for hip fracture with aBMD with risk ≥ 3 % overlapped with those based on DVO score in 36 % of patients. Patients identified for treatment only by DVO score presented a higher percentage of spine fractures (65 vs. 41 %). The thresholds for this "DVO-equivalent risk level" for 'FRAX with aBMD' was estimated to be ≥ 6.7 % for major osteoporotic fracture and ≥ 2.1 % for hip fracture.

This study demonstrates that the DVO score was more sensitive than the FRAX score for patients with prevalent spinal fractures. We suggest considering the appropriate score and therapy threshold carefully in the daily care of male patients.

Introduction

Worldwide, 6 % of men and 21 % of women between 50 and 84 years of age suffer from osteoporosis. Osteoporosis is thus one of the four most common health problems, along with cardiovascular disease, stroke, and cancer [1, 2]. The mortality associated with a hip fracture within the first year following occurrence is 8–36 % [3, 4]. In the European Union, most fractures by 2025 are expected to occur in Germany [5], making osteoporosis one of the major health concerns in this country.

* Present affiliation.

Osteoporosis is characterized by a pathological reduction in areal bone mineral density (aBMD) and bone quality, resulting in greater fracture risk [6–8]. The World Health Organization (WHO) guidelines still suggest a T-score ≤ -2.5 quantified at the proximal femur or the lumbar spine by dual X-ray absorptiometry (DXA) as a diagnostic criterion [8, 9]. However, factors other than bone mineral density also influence fracture risk [10–19]. A number of different fracture-risk tools have been developed and published. The GARVAN score is mainly used in Australia, whereas the QFracture score is common in the UK. The score most often applied internationally, however, is the fracture risk assessment or FRAX score [20], first published in 2008. FRAX as a country-specific score is available for 80 % of the global population and has been validated in 11 independent multicenter prospective studies [21, 22]. It can be adopted with or without inclusion of DXA (femoral neck aBMD or T-score). The FRAX score calculates a 10-year probability of suffering a major osteoporotic fracture (MOF) at a specific location, including the spine, hip, shoulder, and wrist, or a 10-year probability of hip fracture (HF), using an unpublished algorithm. The FRAX score does not propose therapy thresholds.

In the development of all these scores, only limited data on men were available. Fracture risk was similar between men and women when using aBMD values or T-scores adjusted for women. However, by using sex-specific T-scores for men, the risk was half that for women. Due to discrepancies in epidemiological study results on the relevance of T-score values for fracture prediction, it was suggested that decisions about treatment to prevent fracture should be based on estimates of risk rather than T-score. In the FRAX algorithm for men, aBMD or the T-score values for women have to be used.

Prior to the international development of the FRAX score in 2008, the Association of German-language Societies in Osteology (DVO) developed and applied a risk-adapted evaluation score in 2003 [23]. This so-called DVO score was revised in 2006, 2009, 2014, and 2017 in accordance with the relevant literature and has been serving as a diagnostic and therapeutic guideline since 2003. The fracture risk estimates are mostly based on European (EVOS study) and Dutch (Rotterdam study) data on fracture incidence [24, 25]. Data on fracture incidence in men was very limited at the time of generating these risk estimates. This score was validated using women of the "FREEDOM" trial into the effects of therapy on fracture risk [26]. The DVO score is a multistep approach, including risk assessment and therapy threshold, and calculates the 10-year fracture risk probability for hip or vertebral fractures. Based on a wide set of risk factors, patients are primarily classified into one of two groups of 10-year fracture risk: < 20 % or > 20 %. All patients with a 10-year fracture probability > 20 % are recommended further diagnostic tests, including DXA measurement. Depending on the DXA T-score, they are then classified as having a 10-year fracture probability of 20–30 % or > 30 %. All patients with a 10-year fracture probability of suffering hip or vertebral fractures > 30 % are advised to undergo specific anti-osteoporotic treatment unless differential diagnosis points to other disorders with different treatment recommendations. Given the number of patients predicted in Germany, it is of interest to compare the German-speaking countries with international studies with respect to fracture risk and therapy threshold. We, therefore, compared the DVO and FRAX scores in a cohort of men evaluated for osteoporosis in an endocrinological center in Germany. Additionally, we compared therapy recommendations based on DVO score with those based on FRAX. With this aim, we selected internationally common therapy thresholds for FRAX (≥ 20 % for MOF and ≥ 3 % for HF) and adapted these for our study cohort [27].

Material and Methods

Patients

This seven-year retrospective study was based on data recorded between July 2007 and June 2014 at the MVZ endokrinologikum in Göttingen. This endocrinological center specializes in treating osteoporosis patients from all over Germany. The patients were mainly referred from other clinics or practices by doctors requesting an osteoporosis-specific evaluation. Only patients providing documented informed consent to use their data for research purposes were included. Permission from the ethics board dates from 18 February 2007 (18/2/07). Out of 395 male osteoporosis patients, 125 met the study requirements, which included a complete data set on the DVO score, FRAX score, clinical risk factors, and a DXA measurement at least at the femoral neck. Of these 125 patients, 117 also underwent DXA measurement of the lumbar spine.

Data

All data were available as electronic medical records (Medistar). Available data included information on history, clinical and laboratory parameters, as well as DXA measurements. The following items were documented: patient age, height and loss of height, body mass, family history of femoral neck fractures, illnesses associated with secondary osteoporosis, lifestyle factors (forms of nutrition, mobility, physical fitness), smoking, alcohol intake, previously experienced fractures (pathological and/or traumatic), and medication.

Areal BMD was measured using DXA. Referring to prior studies, DXA has an accuracy error that ranges from 5 % to 8 % [28]. Just over half (51.2 %) of the DXA procedures were performed at the MVZ endokrinologikum Göttingen using a GE Lunar Prodigy device manufactured by GE Healthcare. GE Healthcare uses a male reference population to calculate T-scores. A number of laboratory values were analyzed to evaluate for secondary osteoporosis. These included serum calcium, serum phosphate, serum sodium (optional), serum alkaline phosphatase, gamma-GT, creatinine clearance, C-reactive protein, blood panel, thyroid-stimulating hormone, total testosterone, vitamin D, parathyroid hormone, and bone turnover markers. Other measurements were taken for analysis depending on the individual patient.

Analysis

We used the data of the patients data to calculate the 10-year fracture probability for hip or vertebral fractures, referring to the DVO guidelines version available during the year of the first presentation to the MVZ endokrinologikum. The DVO guidelines allow a step-by-step 10-year-risk classification into three sub-groups: < 20 %, 20–30 % and > 30 %. All patients with a 10-year fracture probability for hip or vertebral fractures above 20 % were recommended that

they undergo DXA. For the next step, that is, determining whether anti-osteoporosis treatment should be recommended, the DVO guidelines provide a therapy-indication table based on gender, age, and DXA T-score that can be used to determine whether the 10-year risk exceeds 30 %. In Germany, DXA devices employ reference data by default for men to calculate T-scores of male patients. Therefore, all T-scores in this document were calculated relative to reference data for men, unless noted otherwise. The lowest of the gender-specific T-scores at the spine, femoral neck, and total hip is taken when using the therapy-indication table. Patients with a 10-year fracture probability > 30 % for hip or vertebral fractures met the criterion for specific anti-osteoporotic treatment.

We also calculated the FRAX 10-year fracture probabilities for a MOF and HF with and without DXA-based aBMD measurement using the Internet-based German version of the FRAX tool (https://www.sheffield.ac.uk/FRAX/tool.aspx?lang = de). All variables for the FRAX score were defined exactly as prescribed by FRAX, and the aBMD was entered in g/cm^2. Referring to meta-analyses, the internationally common therapy thresholds for the FRAX score were set: ≥ 20 % for MOF and ≥ 3 % for HF [27]. ▶ **Fig. 1** illustrates the most important similarities and differences between the calculation of the DVO and the FRAX scores.

For FRAX, there is a risk that mistakes may occur when entering the DXA-based bone density as T-score and not in g/cm^2. Entering bone density results in FRAX either as a T-score or as aBMD in g/cm^2 should theoretically lead to the same risk probabilities. Importantly, and as stated on the FRAX website, FRAX expects T-scores for both men and women relative to *female* reference data. In Germany, however, DXA devices, by default, use *male* reference data to calculate T-scores *for men*. A T-score calculated relative to a reference population of males will be lower than that calculated relative to a reference population of females because the values of female reference data are lower than the male reference data. Thus, if a T-score is entered with respect to a reference population of males in FRAX (using the data derived from the DXA device), these T-scores are going to be lower than if entered as required by FRAX. It follows that the calculated FRAX-based fracture risk is erroneously higher than the risk calculated from aBMD in g/cm^2. Assuming that many centers may not be aware of these caveats and in order to estimate the magnitude of this effect, we also calculated the FRAX scores using T-scores relative to male reference data (data not shown).

Statistics

Data were analyzed with IBM SPSS Software Version 20 and Microsoft Excel. We used mean values ± standard error of the mean (SEM) to analyze the basic characteristics of our study group and to compare the DVO and FRAX scores. Additionally, we employed the Wilcoxon signed-rank and McNemar´s tests. McNemar´s test was specifically used to compare the therapy recommendations. We set the alpha level at p = 0.01 for both tests.

We used a confusion matrix to visualize and summarize the performance of these scores. We analyzed the respective sensitivities and specificities of the DVO and FRAX scores with respect to the predictive value of prevalent spinal fractures in patients recommended therapy. The sensitivity based on the confusion matrix is an effect size itself. We used the binomial test to prove the statistical significance of sensitivity by demonstrating that the distribution of results deviates from the expected statistical distribution of 50 %.

Results

The study sample comprised 125 male patients with a complete set of data. The baseline characteristics are summarized in ▶ **Table 1**. The FRAX Score lists eight secondary osteoporosis causes, whereas the DVO score lists a variety of additional secondary osteoporosis causes, diseases, and medications that foster secondary osteoporosis. Some risk factors were categorized differently in each score. Rheumatoid arthritis is treated as a secondary osteoporosis factor in the DVO score, whereas it is not included in the secondary osteoporosis causes in the FRAX score and is listed separately. Alcohol consumption and smoking are treated as general risk factors in both scores. Glucocorticoid (GC) therapy is considered an independent treatment fostering fractures in DVO and FRAX scores.

DVO score	FRAX score
Many (version 2017= 40) risk factors listed in the DVO guidelines	12 risk factors plus optionally DXA
Different fracture types	Prior fractures
T-score at lumbar spine OR total femur OR femoral neck	DXA at femoral neck only optional
Step-by-step evaluation of the 10-year fracture probability	One step to the 10-year fracture probability
Consistent therapy threshold > 30 %	No therapy threshold
Dependent variable: 10-year fracture incidence of hip and vertebral fracture	Dependent variable: 10-year fracture probability of major osteoporotic fracture or hip fracture
Published 10-year fracture risk calculation method	Unknown proprietary 10-year fracture risk calculation algorithm
Additional risk factors with direct therapy recommendation	

▶ **Fig. 1** Comparison of DVO and FRAX scores.

▶ **Table 1** Baseline characteristics of patients (male, N = 125).

Baseline characteristics of patients	n	„MVZ endokrinologikum“ Göttingen
Age in years ± SD	125	59.18 ± 10.67
BMI in kg/m² ± SD	125	26.56 kg/m² ± 4.59
Prior fracture	74	56.92 %
Parent hip fractures	14	10.77 %
Current smoking	29	22.30 %
Oral glucocorticoids (current intake < 5 mg)	16	12.31 %
Rheumatoid arthritis	3	2.31 %
Alcohol three or more units a day	21	16.15 %
Femoral neck T-score	125	−2.056 ± 1.017
Secondary osteoporosis (8 conditions as described in FRAX)	17	13.08 %
Secondary osteoporosis (as described in the DVO guidelines, 40 risk factors in total)	58	44.0 %

BMI: bone mass index; SD: standard deviation.

In FRAX, these risk factors are entered separately into the input screen form of the FRAX tool. Focusing on the secondary osteoporosis causes according to the DVO guidelines, 17 of the 55 patients with secondary causes had evidence of steroid-induced osteoporosis (30.9 %), the most common secondary cause we documented. There were fewer patients with secondary osteoporosis causes according to the FRAX criteria (p = 0.01) (13.08 %).

All 125 patients had at least one documented aBMD value and T-score at the femoral neck. In comparison, the average T-score at the lumbar spine (−2.4 ± 1.1) was significantly lower than the average T-score at the femoral neck (−2.0 ± 1.0; p ≤ 0.05, Wilcoxon test). Taking into account the lowest T-scores of all three measurement sites, the average T-score was even lower (−2.8 ± 0.9).

Taking all patient data into account and referring to the DVO score, 75 patients (60.0 %) had a 10-year fracture probability for hip or vertebral fractures > 30 % and were recommended specific anti-osteoporotic treatment (▶ **Fig. 2**). Forty-three (57.3 %) patients with a risk > 30 % had primary osteoporosis, and 32 (42.7 %) had secondary osteoporosis, according to the DVO criteria. Sixteen (12.8 %) patients had a 10-year fracture probability for hip or vertebral fractures below 20 % according to their DVO-score, all of them having a T-score above −2.5. Clinical risk factors for secondary osteoporosis were detected in nine of these men with a 10-year fracture probability < 20 %.

The average 10-year fracture probabilities for the FRAX scores are summarized in ▶ **Table 2**. As expected, there was a difference not only with or without including the DXA values. There was a significant (p < 0.001) but weak (r = 0.378) correlation for FRAX MOF with and without aBMD input and a significantly (p < 0.001) strong (r = 0.517) correlation for FRAX HF with and without aBMD input.

Additionally, we calculated the mean individual 10-year fracture probabilities for the FRAX scores for all male patients with a DVO score ≤ 20 %, 20–30 %, or ≥ 30 %. The median FRAX 10-year probabilities when including DXA were slightly but not significantly greater (p < 0.0001) compared to not including aBMD.

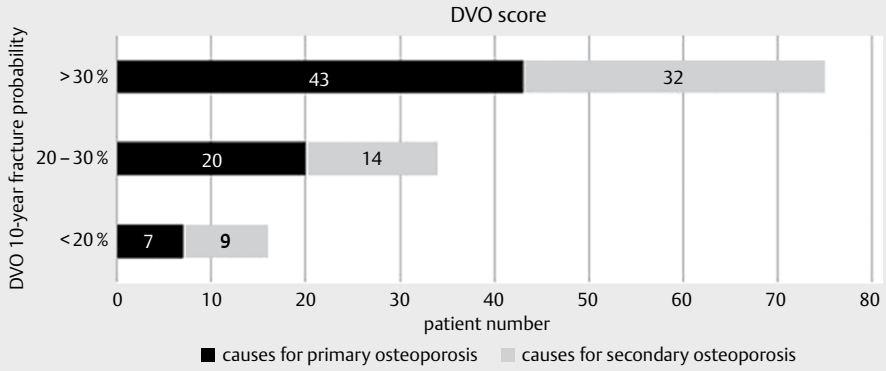
We also applied commonly implemented published therapy thresholds (for MOF ≥ 20 % and for HF ≥ 3 %) to compare the frequency of therapy indications for the DVO and the FRAX risk calculator (▶ **Fig. 3**). The DVO score identified the greatest number of patients requiring treatment (60.0 %). The largest number of patients was 49.6 %, who were recommended therapy by one of the FRAX-based calculations, namely by the FRAX score for HF, including aBMD. There was no significant (p = 0.106) difference between therapy recommendations based on the DVO and FRAX score for HF (therapy threshold ≥ 3 %), including aBMD. All other FRAX scores differed greatly from the DVO score with respect to therapy recommendations. The fewest of all therapy recommendations were made based on the FRAX score without aBMD for a MOF with a therapeutic threshold of ≥ 20 %: 1.6 %. All patients recommended therapy for a MOF also received the corresponding recommendation of therapy for HF, with or without DXA.

HF with aBMD and a threshold of ≥ 3 % was revealed as having the greatest percentage agreement between FRAX and DVO scores in terms of the actual individuals the scores identified as requiring therapy (36 %). In this case, the FRAX score for HF with aBMD classified another 13.6 % of patients as requiring therapy not classified as such by the DVO score. The DVO score, on the other hand, classified another 24.0 % of patients as needing therapy, who were not classified as such by the FRAX score for HF with aBMD. When viewing the recommendations of therapy based on the DVO score and FRAX score for HF with aBMD, a total of 74.4 % were in need of specific anti-osteoporotic therapy. The FRAX score for MOF without DXA with a therapeutic threshold of ≥ 20 % revealed the fewest of all agreements (1.6 %) with a DVO score.

Referring to the 60.0 % of patients requiring therapy according to the DVO score, we calculated the corresponding therapy threshold for the FRAX scores as depicted in ▶ **Table 3**. Using those thresholds, the scores would classify the same percentage of patients in need of therapy as the DVO score. All newly adapted therapy thresholds revealed markedly lower values in our male study cohort than the internationally used therapy thresholds of ≥ 20 % for MOF and ≥ 3 % for HF.

To better understand the patient-specific differences in the two scores with the greatest overlap but also with unidentified patients in each score, we compared patient parameters identified by both DVO and FRAX HF with aBMD with those of each single score (▶ **Table 4**).

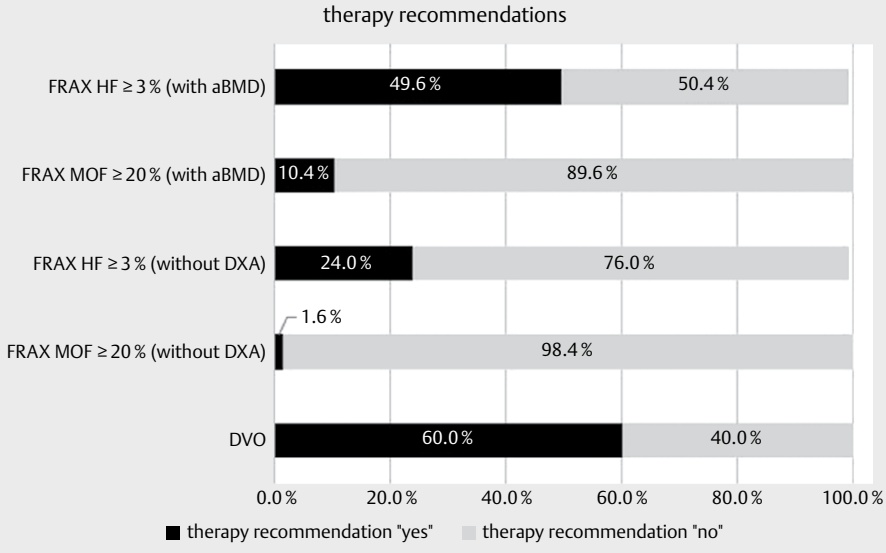
We conclude from our results that a number of risk factors have different influences on the scores, resulting in different treatment recommendations to our patients. Those only identified by FRAX HF with aBMD were shown to have a higher percentage of the risk factors “parent hip fracture” (12 versus 3 %), “current smoking” (47 versus 13 %), and “alcohol three or more units a day” (29 versus 6 %). On the other hand, patients identified by DVO score presented a higher percentage of prior fractures (65 vs.41 %). When considering fracture sub-types, men identified by the DVO score as in need of treatment demonstrated more than double the number of prevalent spinal fractures (65 vs. 29 %). The percentage of peripheral fractures was only slightly higher in those identified by the FRAX score (24 vs. 23 %). We used a confusion matrix (▶ **Tables 5,6**) to analyze the respective sensitivities and specificities of the DVO and FRAX scores with respect to the predictive value of prevalent



▶ **Fig. 2** 10-year fracture probabilities for hip or vertebral fractures according to the DVO score, split into primary (black) and secondary (gray) causes of osteoporosis.

▶ **Table 2** FRAX-based 10-year risk of fracture (median), grouped according to DVO-based 10-year fracture incidence rates for hip or vertebral fractures.

FRAX		10-year risk of fracture (mean ± SD)				Minimum/Maximum
		Total	DVO score < 20 % (n = 16)	DVO score 20–30 % (n = 34)	DVO Score > 30 % (= 75)	DVO Score > 30 % (= 79)
Without aBMD n = 125	**MOF**	5.3	3.65	4.2	6.6	2/22
	HF	0.9	0.65	0.5	1.5	0.1/16
With aBMD n = 125	**MOF**	7.85	4.6	7.45	9.3	3.5/68
	HF	2.8	1.3	2	3.8	0.3/62



▶ **Fig. 3** Therapy recommendations “yes” (black) and “no” (gray) based on DVO guidelines and FRAX score (grouped by therapeutic thresholds for major osteoporotic fracture [MOF] and hip fracture [HF]).

spinal fractures. Whereas the DVO score was significantly (p = 0.011) sensitive (0.65) for patients with prevalent spinal fractures, the FRAX score for HF with aBMD was not significantly (p = 0.704) sensitive (0.53) for this patient group. In contrast, alcohol intake, smoking, and the parental hip fracture had a greater impact on FRAX HF with aBMD.

Discussion

In this study of a pre-selected cohort comprising 125 male patients, we compared the DVO score, established mainly in German-speaking countries, with the internationally validated and implemented FRAX score with and without bone mineral density determination. These two scores differ in their sets of risk factors and their 10-year

▶ **Table 3** Adapted therapy: FRAX risk thresholds resulting in the identification of 60.0 % of patients as in need of therapy, to match the sensitivity of the DVO score criterion.

Score	Cut off 60.0 % as in need of therapy
FRAX MOF without aBMD	4.8 %
FRAX HF without aBMD	0.7 %
FRAX MOF with aBMD	6.9 %
FRAX HF with aBMD	2.1 %

fracture-risk-calculation algorithms. Whereas the DVO score presents a step-by-step, replicable 10-year fracture assessment of the risk of suffering a hip or vertebral fracture, the FRAX score uses an unpublished algorithm to calculate the 10-year fracture probability of a MOF or HF. The DVO score is based on the T-score (using a male reference population), whereas FRAX can be used without DXA or with aBMD or T-score values (using a female reference population, in men and women alike).

The DVO therapy threshold is set at a fixed fracture-risk level, and the guidelines recommend specific anti-osteoporotic therapy for patients when their 10-year fracture probability of suffering a hip or vertebral fracture is > 30 %. The FRAX therapy thresholds vary across countries. A systematic review revealed that the most commonly used therapy threshold for MOF is ≥ 20 % and for HF ≥ 3 % [27].

In comparison with the male participants of other studies in the literature, the pre-selected patients of the "MVZ endokrinologikum" Göttingen were younger (average age 59.22 years). The "MrOS" study included community-dwelling patients in Sweden aged over 65 years with an average age of 75.4 years [29]. Marques et al. included male patients older than 40 years with an average age of 60.3 years [30]. Focusing on risk profiles, the study population from the "MVZ endokrinologikum" Göttingen presented a greater number of risk factors except for "current smoking" and "alcohol intake" when compared to the study groups from the "MrOS" study and the work of Marques et al. The average T-score measured at the femoral neck (–2.1 ± 1.0) was distinctly lower in the MVZ endokrinologikum Göttingen study group than in the other male study groups. As a result of the greater risk and the lower average T-scores at the femoral neck, the male patients of the "MVZ endokrinologikum" demonstrated a higher mean and median FRAX 10-year fracture probability for all FRAX scores, with and without DXA, in comparison with other published study groups.

Comparing DVO for hip, or vertebral fractures and FRAX scores for the 10-year fracture probability of MOF and HF, all the FRAX scores differed clearly from the DVO score in the calculation of the percentage of 10-year fracture risk, with much lower values resulting in the FRAX scores. This may be due to several reasons. First, the fracture outcomes are not the same. In the DVO score, hip and vertebral fracture risks are estimated (including radiologically identified vertebral fractures), whereas, in FRAX scores, only clinical vertebral fractures are included in the MOF category, in which shoulder and forearm fractures are also included. In addition, our study cohort represented a male osteoporosis population with severe forms of osteoporosis, 56.9 % with fractures and a large number of secondary reasons for osteoporosis, and not an epidemiological study group. Consequently, the mean T-score of –2.1 at the femoral neck and a total minimal T-score of –2.8 (including the lumbar spine, femoral neck, and total femoral DXA) was very low compared to the T-score of, e. g., –0.9 in the Swedish population of the "MrOS" study [29]. From these data, we may also speculate that, for example, younger men with severe osteoporosis at high risk of fracture may not be properly identified by FRAX scores owing to the fact that a number of types of secondary osteoporosis have been proven to be associated with enhanced fracture risk are not listed among the types of secondary osteoporosis for the FRAX calculator, as also suggested by others [31].

With reference to internationally commonly used therapy thresholds for FRAX, the thresholds for DVO and FRAX scores differed similarly to the discrepancy in the 10-year fracture risk probability. The DVO score identified more patients in need of therapy than all of the different FRAX scores. Only the FRAX score for HF with aBMD and DVO score were highly similar but still did not identify the same individual patients.

The demonstrated difference in DVO and FRAX scores in our study group might be explained by the calculation method of the FRAX score, the algorithm of which has never been published. Another possible explanation could be its implementation of the lowest DXA value, including the spine value and the inclusion of fracture risk not only for hip but also for spinal fractures. However, an epidemiological study in Canada revealed that simply applying the minimal T-score at the lumbar spine as a parameter for FRAX instead of the T-score at the femoral neck overestimated the actual fracture risk for a MOF [32]. Other epidemiological studies also failed to demonstrate any considerable improvement in fracture risk stratification using the lowest value at several measurement sites versus the hip alone [33, 34]. Therefore, by using the lowest aBMD or T-score value, including the spine, the DVO score may overestimate the 10-year fracture risk probability. However, not all studies report the number of spinal fractures, because not all patients were admitted to a hospital or underwent an X-ray of the spine.

To clarify this assumption, we analyzed differences between patients identified by different scores comparing patient-specific risk factors. In contrast to our assumption, the number of secondary osteoporosis risk factors made no difference between the scores. However, patients identified only by the DVO score as in need of therapy had more than double the number of prevalent spinal fractures than patients identified solely by the FRAX score. This would imply that the DVO algorithm identified patients with the clinically important risk factor "spinal fracture" more sensitively compared to FRAX scores. Our data further suggest that more focus is placed on alcohol intake, smoking, and parent hip fracture in the calculation of fracture risk in FRAX. We conclude that DVO, in comparison to the FRAX Score, is significantly sensitive in identifying male patients with spinal fractures.

For those treating osteoporosis patients according to the DVO score in German-speaking countries, it is of interest to compare the therapy threshold for FRAX score values for these 60.0 % of patients with a DVO 10-year fracture probability > 30 % for hip or vertebral fractures. In our study, this subgroup of men had a mean age of 61.4 years with a FRAX 10-year fracture probability between 2.48 % ± 2.84 (HF without aBMD) and 12.24 % ± 9.58 (MOF with

▶ **Table 4** Comparison of patients identified by DVO and FRAX HF with aBMD, only by DVO score, and only by FRAX HF aBMD as in need of therapy.

Baseline characteristics of patients	DVO AND FRAX HF aBMD "YES" (n = 45)	DVO "YES" FRAX HF aBDM "NO" (n = 31)	FRAX HF aBMD "YES" DVO "NO" (n = 17)
Age in years ± SD	63.3 ± 10.5	58.43 ± 10.75	56.1 ± 11.0
BMI in kg/m² ± SD	25.2 ± 4.9	27.7 ± 4.7	27.3 ± 3.6
Prior fracture	80 %	65 %	41 %
– Spine fracture	64 %	65 %	29 %
– Peripheral fractures at the age of 50 years	42 %	23 %	24 %
Parent hip fractures	13 %	3 %	12 %
Current smoking	24 %	13 %	47 %
Oral glucocorticoids (current intake > 5 mg)	18 %	6 %	0 %
Rheumatoid arthritis	2 %	3 %	6 %
Alcohol three or more units a day	27 %	6 %	29 %
Secondary osteoporosis (8 conditions as described in FRAX)	11 %	13 %	12 %
Secondary osteoporosis (as described in the DVO guidelines, 40 risk factors in total)	40 %	45 %	47 %
minimal aBMD femoral neck	0.69 ± 0.01	0.86 ± 0.07	0.72 ± 0.065
minimal aBMD lumbar spine	0.83 ± 0.12	0.92 ± 1.45	0.87 ± 0.14

▶ **Table 5** Confusion matrix DVO Score regarding the predictive value of spinal fractures.

Confusion matrix DVO Score		**Predicted values (p-value)**		
Actual values		Positive	Negative	
	Positive	49	11	Sensitivity: 0.65 (p = 0.011)
	Negative	26	39	Specificity: 0.22 (p = 0.00)
		Precision: 0.62	Negative predictive value: 0.78	Accuracy: 0.81

▶ **Table 6** Confusion matrix FRAX Score regarding the predictive value of spinal fractures.

Confusion matrix FRAX Score		**Predicted values (p-value)**		
Actual values		Positive	Negative	
	Positive	33	27	Sensitivity: 0.53 (p = 0.707)
	Negative	29	36	Specificity: 0.43 (p = 0.314)
		Precision: 0.53	Negative predictive value: 0.57	Accuracy: 0.55

aBMD). If these 60.0 % of patients identified by the DVO score as in need of therapy were to be treated according to FRAX scores, therapy thresholds would be ≥ 4.8 % (FRAX MOF without DXA), ≥ 0.7 % (FRAX HF without DXA), ≥ 6.9 % (FRAX MOF with aBMD), and ≥ 2.1 % (FRAX HF with aBMD). Therefore, the German DVO score already recommends treatment of male patients at a FRAX 10-year fracture risk of ≥ 2.1 %, and not ≥ 3 % as suggested in most other countries.

The average FRAX 10-year fracture probabilities, therefore, differed from the internationally adapted therapy thresholds of ≥ 3 % for HF and ≥ 20 % for MOF. As demonstrated in previous studies, the FRAX score appears to have greater discrimination ability in HF than in MOF [35, 36]. Therefore, in relation to the hip fracture risk, patients in Germany are classified more generously as in need of treatment at a lower risk of ≥ 2.1 % already using the DVO score, compared to the FRAX score with international therapy thresholds of ≥ 3 %. The mentioned differences in therapy recommendation between DVO and FRAX scores imply that recommendations resulting from FRAX thresholds are perhaps inadequate for our study group [35, 36]. In general, FRAX has been seen to perform better in women than in men [37], supporting our conclusion, as we included only male patients in this analysis.

As pointed out in the methods section, there is some room for confusion when using FRAX input based on T-scores. As specified on the FRAX website, the T-scores should be relative to a female reference population. However, in most centers in Germany, the reference T-scores provided by the manufacturers for men are based on male data. The situation may be similar in other countries as well (e. g., in Austria). If such "wrong" T-scores are used to calculate FRAX, the resulting risks are higher. Centers implementing FRAX may not be aware of this caveat. Our data reveal that risk values were higher when including the T-score in the calculation of FRAX for HF (56.0 %) compared to using aBMD (49.6 %), although the difference was not significant. Nor was there a significant difference when using aBMD (16.0 %) or T-score (17.7 %) in the calculation for MOF.

The strength of our study lies in the amount and availability of information on each patient, allowing us to analyze the different scores retrospectively. In addition, each patient was clinically examined for secondary osteoporosis risk factors intensively and was, therefore, part of a well-characterized male study group. These men may better represent the corresponding daily patient group in specialized secondary or tertiary osteoporosis centers than participants in epidemiological studies. However, the study group is rather small as a consequence of the inclusion criteria (n = 125). Furthermore, it is impossible to judge which of the adapted scores is the most accurate to predict the actual 10-year fracture risk. A patient survey concerning the incident fractures may thus help to evaluate the accuracy of the scores.

In summary, our study demonstrates that the DVO score and FRAX score identified different individual male patients at risk of fracture. The DVO score was more sensitive (0.65) than the FRAX score (0.53) for patients with prevalent spinal fractures; in contrast, alcohol intake, smoking, and parent hip fracture had a higher impact on FRAX HF with aBMD. In this severely osteoporotic male study group, when implementing the most common therapy thresholds for FRAX scores, FRAX for HF with aBMD identified fewer patients compared to DVO criteria. Male patients in Germany identified by the DVO score as in need of treatment corresponded to a FRAX with aBMD hip-fracture risk threshold of ≥ 2.1 %, lower than the international therapy threshold of ≥ 3 %; similarly, the corresponding FRAX threshold for major osteoporotic fracture would have to be set to 6.9 % instead of ≥ 20 %, respectively.

Our data reveal interesting aspects of risk identification for the different scores. The DVO score identified patients for treatment with higher sensitivity. It may be argued that this higher sensitivity is combined with a lower specificity. However, we could clearly demonstrate that the main difference was the number of prevalent spinal fractures in those 31 patients not identified by FRAX scores as in need of treatment. Hence, the DVO score may simply be more accurate when classifying patients at high risk of spinal fracture compared to FRAX scores. Fractures of the spine are an important clinical risk factor in the prediction of future fractures, and their identification would be advantageous in the treatment evaluation of patients in daily practice. Our study cohort was too young, and numbers were too small to investigate the difference in the hip or peripheral fractures between the different scores, but this would need to be evaluated in future studies.

In conclusion, the comparison of the German DVO and the international FRAX scores in a male study cohort with severe osteoporosis revealed clear differences in risk assessment and therapy thresholds. We suggest considering the appropriate score, DXA value, and therapy threshold carefully in the daily care of male patients at risk of fracture.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Hernlund E, Svedbom A, Ivergard M et al. Osteoporosis in the European Union: Medical management, epidemiology and economic burden. A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). Arch Osteoporos 2013; 8: 136. doi:10.1007/s11657-013-0136-1

[2] Shariati-Sarabi Z, Rezaie HE, Milani N et al. Evaluation of bone mineral density in perimenopausal period. Archives Bone Jt Surg 2018; 6: 57–62

[3] Lorentzon M, Cummings SR. Osteoporosis: The evolution of a diagnosis. J Intern Med 2015; 277: 650–661. doi:10.1111/joim.12369

[4] Misiorowski W. Osteoporosis in men. Przeglad menopauzalny = Menopause review 2017; 16: 70–73. doi:10.5114/pm.2017.68596

[5] Hernlund E, Svedbom A, Ivergård M et al. Osteoporosis in the European Union: Medical management, epidemiology and economic burden. Arch Osteoporos 2013; 8: 136. doi:10.1007/s11657-013-0136-1

[6] Tu KN, Lie JD, Wan CKV et al. Osteoporosis: A review of treatment options. P T 2018; 43: 92–104

[7] Weaver CM, Gordon CM, Janz KF et al. The National Osteoporosis Foundation's position statement on peak bone mass development and lifestyle factors: A systematic review and implementation recommendations. Osteoporos Int 2016. doi:10.1007/s00198-015-3440-3. doi:10.1007/s00198-015-3440-3

[8] NIH Consensus Development Panel on Osteoporosis Prevention D, and Therapy. Osteoporosis prevention, diagnosis, and therapy. JAMA 2001; 285: 785–795

[9] WHO Consensus development conference prophylaxis and treatment of osteoporosis 1991

[10] Kanis JA, Johnell O, De Laet C et al. A meta-analysis of previous fracture and subsequent fracture risk. Bone 2004; 35: 375–382. doi:10.1016/j.bone.2004.03.024

[11] Kanis JA, Harvey NC, Johansson H et al. Overview of fracture prediction tools. J Clin Densitom 2017; 20: 444–450. doi:10.1016/j.jocd.2017.06.013

[12] 2017 LO. DVO Leitlinien 2017 zur Prophylaxe, Diagnostik und Therapie der Osteoporose bei postmenopausalen Frauen und Männern.2018.

[13] Lems WF. Clinical relevance of vertebral fractures. Ann Rheum Dis 2007; 66: 2–4. doi:10.1136/ard.2006.058313

[14] Kanis JA, Johansson H, Johnell O et al. Alcohol intake as a risk factor for fracture. Osteoporos Int 2005; 16: 737–742. doi:10.1007/s00198-004-1734-y

[15] Kanis JA, Johansson H, Oden A et al. A meta-analysis of prior corticosteroid use and fracture risk. J Bone Miner Res 2004; 19: 893–899. doi:10.1359/jbmr.040134

[16] Kanis JA, Oden A, Johnell O et al. The use of clinical risk factors enhances the performance of BMD in the prediction of hip and osteoporotic fractures in men and women. Osteoporos Int 2007; 18: 1033–1046. doi:10.1007/s00198-007-0343-y

[17] Pluijm SM, Koes B, de Laet C et al. A simple risk score for the assessment of absolute fracture risk in general practice based on two longitudinal studies. J Bone Miner Res 2009; 24: 768–774. doi:10.1359/jbmr.081244

[18] Zhang X, Yu Z, Yu M et al. Alcohol consumption and hip fracture risk. Osteoporos Int 2015; 26: 531–542. doi:10.1007/s00198-014-2879-y

[19] Kanis JA, Johansson H, Oden A et al. A family history of fracture and fracture risk: A meta-analysis. Bone 2004; 35: 1029–1037. doi:10.1016/j.bone.2004.06.017

[20] Oden A, McCloskey EV, Kanis JA et al. Burden of high fracture probability worldwide: Secular increases 2010-2040. Osteoporos Int 2015; 26: 2243–2248. doi:10.1007/s00198-015-3154-6

[21] Leslie WD, Lix LM, Johansson H et al. Independent clinical validation of a Canadian FRAX tool: Fracture prediction and model calibration. J Bone Miner Res 2010; 25: 2350–2358. doi:10.1002/jbmr.123

[22] Kanis JA, Oden A, Johansson H et al. Pitfalls in the external validation of FRAX: Response to Bolland et al. Osteoporos Int 2013; 24: 391–392. doi:10.1007/s00198-012-1985-y

[23] Faßbender WJ, Scheidt-Nave C, Pfeilschifter J. Die neuen Leitlinien zur Osteoporose. Dtsch Med Wochenschr 2003; 128: 1615–1617. doi:10.1055/s-2003-40930

[24] Raspe A, Matthis C, Scheidt-Nave C et al. [European Study of Vertebral Osteoporosis (EVOS): Design and implementation in 8 German study centers]. Medizinische Klinik (Munich, Germany: 1983) 1998; 93: 12–16,18. doi:10.1007/bf03041993

[25] Hofman A, Breteler MMB, van Duijn CM et al. The Rotterdam Study: Objectives and design update. Eur J Epidemiol 2007; 22: 819–829. doi:10.1007/s10654-007-9199-x

[26] Hadji P, Bock O, Resch H et al. Reduktion des frakturrisikos unter denosumab bei postmenopausalen frauen in abhängigkeit der behandlungsbedürftigkeit nach DVO-Leitlinien 2009. Osteologie 2013; 1: 39–45

[27] Kanis JA, Harvey NC, Cooper C et al. A systematic review of intervention thresholds based on FRAX: A report prepared for the National Osteoporosis Guideline Group and the International Osteoporosis Foundation. Arch Osteoporos 2016; 11: 25. doi:10.1007/s11657-016-0278-z

[28] El Maghraoui A, Roux C. DXA scanning in clinical practice. QJM 2008; 101: 605–617. doi:10.1093/qjmed/hcn022

[29] Harvey NC, Oden A, Orwoll E et al. Falls predict fractures independently of FRAX probability: A Meta-Analysis of the Osteoporotic Fractures in Men (MrOS) Study. J Bone Miner Res 2017. doi:10.1002/jbmr.3331. doi:10.1002/jbmr.3331

[30] Marques A, Lucas R, Simoes E et al. Do we need bone mineral density to estimate osteoporotic fracture risk? A 10-year prospective multicentre validation study. RMD open 2017; 3: e000509. doi:10.1136/rmdopen-2017-000509

[31] Valentini A, Cianfarani MA, De Meo L et al. FRAX tool in type 2 diabetic subjects: The use of HbA(1c) in estimating fracture risk. Acta diabetol 2018; 55: 1043–1050. doi:10.1007/s00592-018-1187-y

[32] Leslie WD, Lix LM, Johansson H et al. Spine-hip discordance and fracture risk assessment: A physician-friendly FRAX enhancement. Osteoporos Int 2011; 22: 839–847. doi:10.1007/s00198-010-1461-5

[33] Kanis JA, Johnell O, Oden A et al. The use of multiple sites for the diagnosis of osteoporosis. Osteoporos Int 2006; 17: 527–534. doi:10.1007/s00198-005-0014-9

[34] Leslie WD, Lix LM, Tsang JF et al. Single-site vs multisite bone density measurement for fracture prediction. Arch Intern Med 2007; 167: 1641–1647. doi:10.1001/archinte.167.15.1641

[35] Giangregorio LM, Leslie WD, Lix LM et al. FRAX underestimates fracture risk in patients with diabetes. J Bone Miner Res 2012; 27: 301–308. doi:10.1002/jbmr.556

[36] Reid IR. A broader strategy for osteoporosis interventions. Nat Rev Endocrinol 2020; 16: 333–339. doi:10.1038/s41574-020-0339-7

[37] Sandhu SK, Nguyen ND, Center JR et al. Prognosis of fracture: Evaluation of predictive accuracy of the FRAX algorithm and Garvan nomogram. Osteoporos Int 2010; 21: 863–871. doi:10.1007/s00198-009-1026-7