Bone 147 (2021) 115934

Contents lists available at ScienceDirect

Bone

journal homepage: www.elsevier.com/locate/bone

Full Length Article

Validation of the Fracture Risk Evaluation Model (FREM) in predicting major osteoporotic fractures and hip fractures using administrative health data

Michael K. Skjødt [a,b], Sören Möller [a], Nana Hyldig [a], Anne Clausen [a], Mette Bliddal [a], Jens Søndergaard [c], Bo Abrahamsen [a,b], Katrine Hass Rubin [a,*]

[a] OPEN - Open Patient data Explorative Network, Department of Clinical Research, University of Southern Denmark, Odense University Hospital, Heden 16, DK-5000 Odense C, Denmark
[b] Department of Medicine, Holbæk Hospital, Smedelundsgade 60, DK-4300 Holbæk, Denmark
[c] The Research Unit of General Practice, Department of Public Health, University of Southern Denmark, J.B. Winsløws Vej 9, DK-5000 Odense, Denmark

ARTICLE INFO

Keywords:
Effectiveness
Validation
FREM fracture risk
Automated risk calculation

ABSTRACT

Background: Prevention of osteoporotic fractures remains largely insufficient, and effective means to identify patients at high, short-term fracture risk are needed. The FREM tool is available for automated case finding of men and women aged 45 years or older at high imminent (1-year) risk of osteoporotic fractures, based on administrative health data with a 15-year look-back. The aim of this study was to validate the performance of FREM, and the effect of applying a shorter look-back period. We also evaluated FREM for 5-year fracture risk prediction.

Methods: Using Danish national health registers we generated consecutive general population cohorts for the years 2014 through 2018. Within each year and across the full time period we estimated the individual fracture risk scores and determined the actual occurrence of major osteoporotic fractures (MOF) and hip fractures. Risk scores were calculated with 15- and 5-year look-back periods. The discriminative ability was evaluated by area under the receiver operating curve (AUC), and negative predictive value (NPV) and positive predictive value (PPV) were estimated applying a calculated risk cut-off of 2% for MOF and 0.3% for hip fractures.

Results: Applying a 15-year look-back, AUC was around 0.75–0.76 for MOF and 0.84–0.87 for hip fractures in 2014, with minor decreases in the subsequent fracture cohorts (2015 to 2018). Applying a 5-year look-back generated similar results, with only marginally lower AUC. In the 5-year risk prediction setting, AUC-values were 0.70–0.72 for MOF and 0.81–0.84 for hip fractures. Generally, PPVs were low, while NPVs were very high.

Conclusion: FREM predicts the 1- and 5-year risk of MOF and hip fractures with acceptable vs excellent discriminative power, respectively, when applying both a 15- and a 5-year look-back. Hence, the FREM tool may be applied to improve identification of individuals at high imminent risk of fractures using administrative health data.

1. Introduction

Osteoporosis is a prevalent disease and fragility fractures have been estimated to occur in one in three women and one in five men aged 50 years or older [1]. Although safe and efficient anti-osteoporosis medications are available, a considerable treatment gap remains in both the primary and secondary fracture prevention setting [2–5]. Hence, there is a need for efficient detection of individuals at risk of fractures to allow clinical assessment and intervention as appropriate.

Identification of individuals with a high fracture risk can be driven by formalised risk assessment tools. One review identified 48 different tools [6], of which 3 were developed for fracture risk prediction and validated more than once in population-based settings with sufficient methodological quality [6]. These tools - Garvan, FRAX, and QFracture – were developed to predict the 5- or 10-year risk of fractures (though QFracture also allows for shorter term risk prediction) [7–12]. However,

* Corresponding author at: OPEN - Open Patient Data Explorative Network, Department of Clinical Research, University of Southern Denmark, Odense University Hospital, Heden 16, DK-5000 Odense C, Denmark.
E-mail address: krubin@health.sdu.dk (K.H. Rubin).

https://doi.org/10.1016/j.bone.2021.115934
Received 3 July 2020; Received in revised form 12 January 2021; Accepted 17 March 2021
Available online 20 March 2021

for high-risk patients a shorter term perspective on fracture risk may be more clinically relevant [13], and this would also cater to a dynamic fracture risk given the changing individual health [14]. Few studies have addressed this issue [13–15], and only QFracture and the *Fracture Risk Evaluation Model* (FREM) are driven by administrative healthcare data and can currently be employed in the assessment of imminent fracture risk [11,12,16]. Yet use of such readily available data may facilitate identification of individuals at risk on a population level with less manual imputation into the model and less recall bias.

The FREM tool was developed to facilitate automated case finding of individuals aged 45 years or older with a high imminent fracture risk. The model predicts the 1-year risk of hip fractures and MOF, respectively, based on individual patient data from the Danish health registries applying a 15-year look-back period. Evaluation of FREM in terms of discrimination has demonstrated acceptable performance for MOF (AUC 0.750 and 0.752 for women and men, respectively), and excellent performance for hip fractures (0.869 and 0.857 for women and men, respectively) [16]. While further evaluation of the model is needed for clinical application, it may also be argued that a 15-year look-back period is infeasible in some healthcare systems. Hence, the aim of this study was to validate the 1-year predictive value of FREM for hip fractures and MOF using consecutive annualised fracture cohorts from 2014 to 2018 based on administrative health data from the Danish health registries. We further examined the 5-year predictive performance of the model and the effect of applying a shorter look-back period.

2. Materials and methods

This was an observational register-based cohort study using administrative health data from the Danish health registries covering the years 1999 through 2018 to validate the FREM model in five consecutive annualised cohorts (years 2014–2018) with 15- and 5-year look-back periods, respectively, stratified by gender. Furthermore, the study examined the model's ability to predict the 5-year fracture risk (2014–2018) with both 15- and 5-year look-back periods.

2.1. Data sources

Data were sourced from Danish national registries. In Denmark all citizens are provided with a unique personal identification number, which is used as the key identifier in all Danish health and social care registers. Data from the following two national registers were used to identify the study population and to extract data on baseline characteristics, risk factors for the FREM model, and outcomes:

- The Danish Civil Registration System (CRS), which includes all persons living in Denmark [17].
- The Danish National Patient Register (NPR), which comprises all somatic inpatient admissions and outpatient visits including diagnosis codes [18,19]. We extracted information on all ICD-10 codes from 1999 to 2018 for the total population in order to determine risk factors and outcomes.

2.2. Cohort selection

From the CRS we extracted data on all citizens in Denmark aged 45 years or above on 1st January in the year of interest (years 2014 through 2018) in order to define five consecutive annualised cohorts. As these cohorts had a large overlap in individuals and exposure periods (Fig. 1), we investigated each cohort separately to evaluate a possible deterioration of the FREM performance over time, which could be possible due to increasing time difference to the original development period (2013).

2.3. Outcomes (fractures)

The primary outcome was MOF, defined as a hip, clinical vertebral, wrist, or humerus fracture (given as a primary or secondary diagnosis code in the NPR, using the following ICD-10 codes: S120, S121, S122, S220, S221, S320, T08, S422, S423, S720, S721, S722, S525, S526) during the year of analysis (e.g. 1st January 2014 to 31st December 2014, and similarly for the years 2015–2018). The secondary outcome was hip fracture during the year of analysis (ICD-10 codes: S720, S721, S722). Both outcomes were defined as the proportion of individuals with incident MOFs or hip fractures, respectively, dichotomized as having at least one of the ICD-10 codes or none in the year of analysis.

2.4. Risk factors

The original FREM study identified 38 and 43 risk factors for MOF for women and men, respectively, and 32 risk factors for hip fracture for both women and men (Supplementary tables S1-S4). We used ICD-10 codes on primary or secondary diagnosis from the NPR to identify

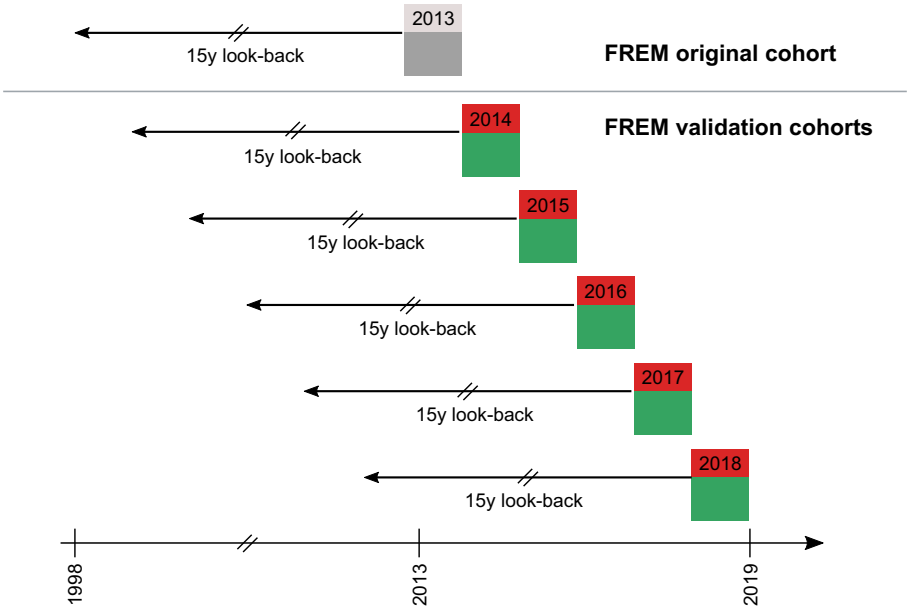

Fig. 1. Study design and time frame for validating the Fracture Risk Evaluation Model in five one-year cohorts (2014–2018).

these risk factors. Each was defined as present or absent within a time period of 15 years or 5 years, respectively, before the index date (i.e. 1st January for each of the years 2014–2018). Age on index date, separately calculated for the different outcome periods, split into 5-year age categories (<50, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, ≥80 years) was included as a separate risk factor.

The Charlson comorbidity index (CCI) was used to classify comorbid conditions in the study population. To calculate CCI, we used ICD-10 codes from the NPR with 15-year look-back using the updated CCI by Quan and colleagues [20].

2.5. Statistical analysis

We investigated the performance of the FREM algorithm separately on five pairs of exposure and outcome periods, starting with exposure period 1999–2013 and outcome year 2014 and repeating the analysis for each following year up to exposure period 2003–2017 and outcome year 2018. In each of these analyses all individuals alive and residing in Denmark, and aged 45 years or older on 1st January in the outcome year, were included. Moreover, we investigated the performance of FREM for 5-year risk, using 1999–2013 as exposure period and 2014–2018 as outcome period including all individuals in the cohort at 1st of January 2014. The cohorts were stratified by sex to take into account the sex-specific risk score in the FREM algorithm. For each outcome period we reported descriptive statistics on age as the median with interquartile range (IQR), and on age groups and Charlson comorbidity index groups (0, 1–2, 3+) as counts with proportions.

For each period, we determined both a FREM-MOF and FREM-Hip risk score based on 5- and 15-year look-back, respectively, for each individual, and determined the predicted and actual occurrence (i.e. cumulative incidence) of MOF and hip fractures during the outcome period. Details on how to apply the FREM algorithm to determine these risk scores are presented supplementary materials. We estimated Receiver Operating Curves (ROC) with Area Under the Curve (AUC, c-statistics) as well as positive and negative predictive values with 95% confidence intervals for predetermined risk score cut-offs (1-year risk of 2% for MOF and 0.3% for hip fractures, and 5-years risk of 10% for MOF and 1.5% for hip fractures). The risk cut-offs are similar to the 1-year cut-offs employed in the original FREM paper and their 5-year accumulation [16]. Moreover, in predicted risk strata (0–1%, 1–2%, 2–3%, 3–4%, 4–5%, ≥5% for 1-year risk and 0–5%, 5–10%, 10–15%, 15–20%, 20–25%, ≥25% for 5-year risk) we compared the predicted risk with the observed occurrence of fractures. Furthermore, we determined the proportion of individuals with a predicted risk above the cut-offs (Supplementary table S5).

We performed the following four sensitivity analyses: Firstly, we investigated positive and negative predictive values for different choices of risk cut-off (5% for MOF and 0.75% for hip fractures for 1-year risk). Secondly, we tested a model only including age as risk factor for comparison (Supplementary table S6). Thirdly, we compared FREM to a model including only age and MOF history as predictors. Finally, we repeated the above analyses stratified by age group to investigate differing validity of the FREM algorithm in different age groups, hence removing the effect of age from the model.

2.6. Ethics

An ethics committee approval is not required for registry-based studies according to Danish law. The study was approved by the Danish Data Protection Agency (jf.nr. 2008-58-0035).

3. Results

As of 1st January 2014, the total population of individuals aged 45 years or above living in Denmark was 2,548,159 (1,317,327 women and 1,230,832 men) increasing to 2,643,369 individuals on 1st January 2018 (1,365,537 women and 1,277,832 men). The cohorts had an overlap of 95% between two consecutive cohorts (e.g. 2014 and 2015) with approximately 55,000 persons (53,302 to 56,671) leaving and 79,000 persons (77,248 to 80,843) entering a consecutive cohort. Overall, the median age was slightly higher for women than for men in accordance with the national demographics. In total, women had a statistically significant higher risk of both MOF and hip fractures ($p < 0.001$) compared to men. The risk increased in both genders with age. No differences were found in number of comorbidities between genders or outcome years. Demographics and fracture incidence for the cohorts are shown in Table 1 (outcome year 2014 and 2014–2018, conjointly stratified by gender) and Supplementary tables S7 and S8 (outcome year 2015, 2016, 2017 and 2018). Prevalence of the risk factors for both 15- and 5-year look-back periods are reported in Supplementary tables S1–S4.

3.1. Validation of FREM (15-year look-back period)

The performance of FREM with 15-year look-back period was stable over time for both MOF and hip fractures in both women and men (Table 2 and Supplementary table S9). The AUC decreased slightly from

Table 1
Demographics and fracture incidence for the study population by outcome year 2014 and years 2014–2018, stratified by gender.

	Year 2014		Year 2014–2018	
	Women *N* = 1,317,327	Men N = 1,230,832	Women N = 1,317,327	Men N = 1,230,832
Age				
Median (Q1-Q3)	62.2 (53.0–71.7)	60.6 (52.2–69.5)	62.2 (53.0–71.7)	60.6 (52.2–69.5)
Age categories N (%)				
45–49	212,203 (16)	219,211 (18)	212,203 (16)	219,211 (18)
50–54	192,683 (15)	196,293 (16)	192,683 (15)	196,293 (16)
55–59	178,703 (14)	178,991 (14)	178,703 (14)	178,991 (14)
60–64	171,860 (13)	168,395 (14)	171,860 (13)	168,395 (14)
65–69	181,652 (14)	176,070 (14)	181,652 (14)	176,070 (14)
70–74	133,809 (10)	122,550 (10)	133,809 (10)	122,550 (10)
75–79	98,537 (7)	82,158 (7)	98,537 (7)	82,158 (7)
80+	147,880 (11)	87,164 (7)	147,880 (11)	87,164 (7)
Incident major osteoporotic fractures (MOF) N (%)				
MOF total	19,371 (1.47)	7657 (0.62)	88,427 (6.71)	34,849 (2.83)
45–49	559 (0.26)	657 (0.30)	3387 (1.60)	3185 (1.45)
50–54	967 (0.50)	682 (0.35)	5878 (3.05)	3254 (1.66)
55–59	1548 (0.87)	769 (0.43)	7963 (4.46)	3610 (2.02)
60–64	1911 (1.11)	863 (0.51)	9678 (5.63)	3958 (2.35)
65–69	2520 (1.39)	1007 (0.57)	12,745 (7.02)	4903 (2.78)
70–74	2454 (1.83)	905 (0.74)	11,742 (8.78)	4227 (3.45)
75–79	2541 (2.58)	899 (1.09)	11,806 (11.98)	4154 (5.06)
80+	6871 (4.65)	1875 (2.15)	25,228 (17.06)	7558 (8.67)
Incident hip fractures N (%)				
Hip fractures total	6101 (0.46)	2876 (0.23)	27,389 (2.08)	13,742 (1.12)
45–49	36 (0.02)	77 (0.04)	206 (0.10)	384 (0.18)
50–54	93 (0.05)	119 (0.06)	513 (0.27)	581 (0.30)
55–59	140 (0.08)	178 (0.10)	858 (0.48)	861 (0.48)
60–64	286 (0.17)	202 (0.12)	1465 (0.85)	1148 (0.68)
65–69	429 (0.24)	311 (0.18)	2480 (1.37)	1759 (1.00)
70–74	588 (0.44)	355 (0.29)	3175 (2.37)	1816 (1.48)
75–79	857 (0.87)	428 (0.52)	4586 (4.65)	2223 (2.71)
80+	3672 (2.48)	1206 (1.38)	14,106 (9.54)	4970 (5.70)
Comorbidity N (%)				
CCI = 0	1,051,964 (80)	1,004,630 (82)	1,051,964 (80)	1,004,630 (82)
CCI = 1–2	216,041 (16)	179,006 (14)	216,041 (16)	179,006 (14)
CCI ≥ 3	49,322 (4)	47,196 (4)	49,322 (4)	47,196 (4)

CCI: Charlson comorbidity index, MOF: major osteoporotic fractures.

Table 2
Gender stratification: Performance and predictive capabilities of FREM in women and men by outcome year with 15-year look-back period.

Performance metric	Men		Women	
	MOF	Hip	MOF	Hip
One year risk, outcome year 2013[a]				
C-statistic (95% CI)	0.752 (0.743–0.761)	0.857 (0.848–0.867)	0.750 (0.741–0.795)	0.869 (0.864–0.874)
2014				
C-statistic (95% CI)	0.746 (0.740–0.752)	0.844 (0.837–0.851)	0.757 (0.754–0.760)	0.867 (0.863–0.871)
PPV (95% CI)	3.8 (3.7–4.0)	1.0 (0.9–1.0)	4.3 (4.2–4.4)	1.4 (1.3–1.4)
NPV (95%CI)	99.5 (99.5–99.5)	99.9 (99.9–99.9)	99.2 (99.1–99.2)	99.9 (99.9–99.9)
2015				
C-statistic (95% CI)	0.744 (0.738–0.750)	0.847 (0.840–0.854)	0.750 (0.746–0.753)	0.866 (0.862–0.870)
2016				
C-statistic (95% CI)	0.742 (0.736–0.748)	0.848 (0.841–0.855)	0.748 (0.745–0.751)	0.862 (0.858–0.867)
2017				
C-statistic (95% CI)	0.742 (0.736–0.748)	0.851 (0.844–0.857)	0.745 (0.741–0.748)	0.859 (0.854–0.863)
2018				
C-statistic (95% CI)	0.740 (0.733–0.746)	0.844 (0.837–0.852)	0.742 (0.738–0.745)	0.860 (0.856–0.865)
PPV (95% CI)	3.5 (3.4–3.7)	0.9 (0.8–0.9)	3.9 (3.8–4.0)	0.9 (0.8–0.9)
NPV (95% CI)	99.5 (99.5–99.5)	99.9 (99.9–99.9)	99.1 (99.1–99.1)	99.9 (99.9–99.9)
Five year risk, outcome years 2014–18				
C-statistic (95% CI)	0.713 (0.711–0.716)	0.819 (0.816–0.823)	0.722 (0.720–0.723)	0.842 (0.840–0.844)
PPV (95% CI)	17.3 (15.6–19.2)	7.4 (7.1–7.8)	22.0 (21.0–23.1)	9.9 (9.8–10.1)
NPV (95% CI)	97.2 (97.2–97.2)	99.0 (99.0–99.0)	93.4 (93.3–93.4)	98.9 (98.9–98.9)

[a] Data from the validation sample of the previously published FREM development study [16]. CI, confidence interval; MOF, major osteoporotic fractures; NPV, negative predictive value; PPV, positive predictive value.

the original 2013 cohort to the 2018 cohort. The PPV and NPV remained stable over time. With a cut-off of 2% for MOF and 0.3% for hip fractures the PPV was low (MOF < 4.4; Hip < 1.5) and the NPV high (MOF ≥ 99.0; Hip $= 99.9$) for both genders (Table 2 and Supplementary table S9). An alternative cut-off of 5% for MOF and 0.75% for hip fractures resulted in PPVs ranging from 5.1 to 7.1 and NPVs ranging from 82.8 to 96.7 for both genders for MOF, and lower with PPVs around 2 and NPVs around 80 for hip fractures (Supplementary table S10). FREM generally outperformed models using age as the only risk factor (Supplementary table S6), as these provided lower AUCs and worse PPV. FREM generally performed better than models only including age and history of MOF as risk factors (Table 3).

Comparing the predicted and observed fracture risk showed overall good agreement (Supplementary tables S11 and S12). Approximately 18.6% of women and 3.5% of men had a predicted risk above the cut-off of 2% for MOF. For hip fractures the numbers were 31.2% of women and 18.4% of men with a cut-off of 0.3% (Supplementary table S5).

We observed an increasing risk of MOF and hip fractures by age with the mean predicted risk closely following the observed risk (Fig. 2 for the outcome year 2014, with similar results seen for the other outcome

Table 3
Comparison of the performance of FREM with a model only including age and MOF history as predictors for 1- and 5-year fracture risk, using a 15-year look-back.

Performance metric	Men		Women	
	Age + MOF history	FREM	Age + MOF history	FREM
One year risk, outcome year 2014				
Major Osteoporotic Fracture (MOF)				
C-statistic (95% CI)	0.702 (0.696–0.709)	0.746 (0.740–0.752)	0.743 (0.740–0.747)	0.757 (0.754–0.760)
PPV (95% CI)	3.6 (3.33–3.77)	3.8 (3.7–4.0)	3.7 (3.63–3.77)	4.3 (4.2–4.4)
NPV (95% CI)	99.4 (99.43–99.46)	99.5 (99.5–99.5)	99.2 (99.18–99.21)	99.2 (99.1–99.2)
Hip fracture				
C-statistic (95% CI)	0.814 (0.806–0.822)	0.844 (0.837–0.851)	0.856 (0.851–0.860)	0.867 (0.863–0.871)
PPV (95% CI)	1.0 (0.91–0.99)	1.0 (0.9–1.0)	1.3 (1.25–1.31)	1.4 (1.3–1.4)
NPV (95% CI)	99.9 (99.90–99.91)	99.9 (99.9–99.9)	99.9 (99.91–99.92)	99.9 (99.9–99.9)
Five year risk, outcomes years 2014–18				
Major Osteoporotic Fracture (MOF)				
C-statistic (95% CI)	0.677 (0.674–0.680)	0.713 (0.711–0.716)	0.710 (0.708–0.712)	0.722 (0.720–0.723)
PPV (95% CI)	13.5 (12.88–14.05)	17.3 (15.6–19.2)	14.8 (14.70–14.96)	22.0 (21.0–23.1)
NPV (95% CI)	97.3 (97.26–97.31)	97.2 (97.2–97.2)	95.6 (95.51–95.59)	93.4 (93.3–93.4)
Hip fracture				
C-statistic (95% CI)	0.793 (0.789–0.796)	0.819 (0.816–0.823)	0.833 (0.830–0.835)	0.842 (0.840–0.844)
PPV (95% CI)	4.2 (4.09–4.27)	7.4 (7.1–7.8)	5.6 (5.54–5.68)	9.9 (9.8–10.1)
NPV (95% CI)	99.5 (99.43–99.46)	99.0 (99.0–99.0)	99.5 (99.45–99.48)	98.9 (98.9–98.9)

CI, confidence interval; MOF, major osteoporotic fractures; NPV, negative predictive value; PPV, positive predictive value.

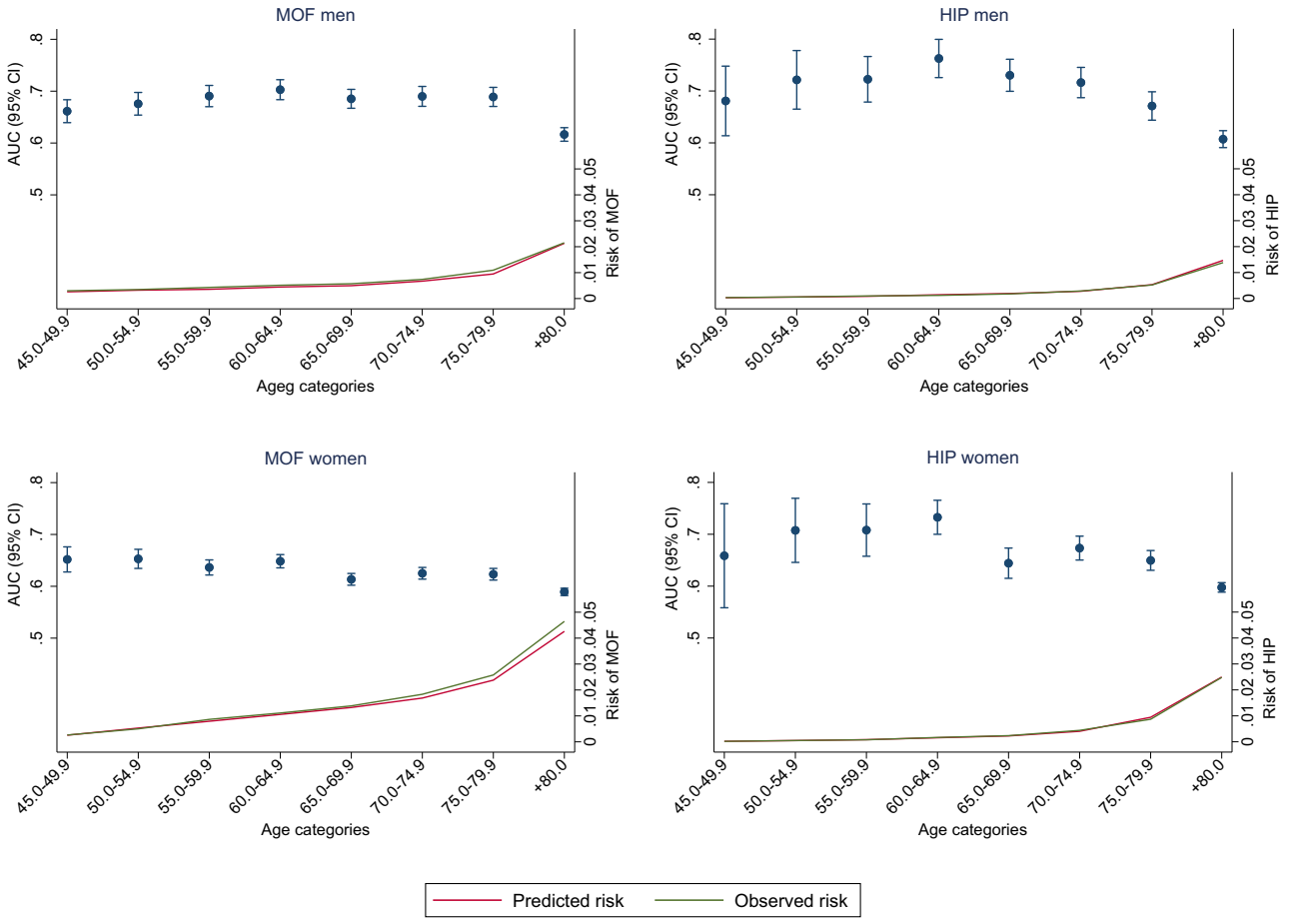

Fig. 2. AUC, the mean predicted and the observed 1-year risk of MOF and hip fractures by age with 15-year look-back in the 2014 outcome year. Analyses stratified by age, implying that age itself does not contribute to the AUC.

Table 4
Gender stratification: Performance and predictive capabilities of FREM in women and men by outcome year with 5-year look-back period.

Performance metric	Men		Women	
	MOF	Hip	MOF	Hip
One year risk, outcome year 2014				
C-statistic (95% CI)	0.731 (0.725–0.738)	0.835 (0.827–0.843)	0.750 (0.746–0.753)	0.863 (0.859–0.867)
PPV (95% CI)	4.4 (4.1–4.7)	1.0 (1.0–1.1)	4.6 (4.5–4.7)	1.4 (1.3–1.4)
NPV (95% CI)	99.5 (99.4–99.5)	99.9 (99.9–99.9)	99.1 (99.1–99.1)	99.9 (99.9–99.9)
2015				
C-statistic (95% CI)	0.729 (0.723–0.735)	0.837 (0.830–0.845)	0.742 (0.739–0.746)	0.860 (0.856–0.864)
2016				
C-statistic (95% CI)	0.728 (0.722–0.734)	0.837 (0.830–0.844)	0.741 (0.737–0.744)	0.858 (0.854–0.862)
2017				
C-statistic (95% CI)	0.728 (0.722–0.734)	0.844 (0.837–0.851)	0.737 (0.734–0.740)	0.853 (0.849–0.858)
2018				
C-statistic (95% CI)	0.728 (0.721–0.734)	0.833 (0.826–0.840)	0.734 (0.730–0.737)	0.855 (0.851–0.859)
PPV (95% CI)	4.1 (3.9–4.4)	0.9 (0.9–0.9)	4.2 (4.1–4.3)	1.1 (1.1–1.1)
NPV (95% CI)	99.5 (99.5–99.5)	99.9 (99.9–99.9)	99.0 (99.0–99.0)	99.9 (99.9–99.9)
Five year risk, outcome years 2014–18				
C-statistic (95% CI)	0.700 (0.697–0.703)	0.810 (0.801–0.814)	0.714 (0.712–0.716)	0.837 (0.835–0.839)
PPV (95% CI)	17.9 (14.4–21.9)	7.9 (7.5–8.3)	23.3 (21.1–25.6)	9.8 (9.6–9.9)
NPV (95% CI)	97.2 (97.2–97.2)	99.0 (99.0–99.0)	93.3 (93.3–93.4)	98.9 (98.9–98.9)

CI, confidence interval; MOF, major osteoporotic fractures; NPV, negative predictive value; PPV, positive predictive value.

years).

3.2. Validation of FREM (5-year look-back period)

Similar results were found for 5-year look-back periods, but with slightly lower AUC than for 15-year look-back periods (Table 4 and Supplementary table S9). The observed risk was generally slightly higher than the predicted, indicating an under-prediction of the true risk (Supplementary tables S13 and S14). Approximately 14.8% (MOF) and 30.7% (hip) of women and 1.9% (MOF) and 16.4% (hip) of men had a predicted risk above the cut-offs (Supplementary table S5).

3.3. Five-year risk (15-year look-back period)

When testing the model's ability to predict the 5-year risk (2014–2018) of MOF with a 15-year look-back period, we derived ROC curves with an AUC of 0.722 (95% CI: 0.720–0.723) for women and 0.713 (95% CI: 0.711–0.716) for men (Table 2). For hip fractures the model produced AUCs of 0.842 (95% CI: 0.840–0.844) for women and 0.819 (95% CI: 0.816–0.823) for men (Table 2). Finally, when assessing the performance of FREM in different age groups, we observed AUCs around 0.6 for most age groups, but lower values for the highest age groups. Further, we also observed an increasing risk of MOF and hip fractures by age with the mean predicted risk closely following the observed risk (Fig. 3). Again, FREM generally performed better than models only including age and MOF history (Table 3).

3.4. Five-year risk (5-year look-back period)

Similarly, the AUC was 0.714 (95% CI: 0.712–0.716) for women and 0.700 (95% CI: 0.697–0.703) for men for MOF with a 5-year look-back and 0.837 (95% CI: 0.835–0.839) for women and 0.810 (95% CI: 0.801–0.814) for men for hip fractures (Table 4).

4. Discussion

We have validated the performance of FREM using the Danish health registries, finding excellent and acceptable discriminative power in the prediction of hip fractures and MOF, respectively, when applying a 15-year look-back. Results were comparable across genders, and we observed only a minimal decline in the c-statistics with increasing temporal distance to the derivation cohort. The FREM tool performed better than models only including age respectively age and history of MOF. Importantly, we have also demonstrated that FREM performs well with a 5-year look-back (though with slightly lower c-statistics). This demonstrates the potential for using this tool also in persons where only a shorter look-back period is available or in the setting of registers that were established more recently.

The commonly used fracture risk assessment tools apply a 5- or 10-year perspective on fracture prediction [7–12]. This study indicates that FREM performs well in a 5-year risk prediction setting, as well as in 1-year prediction, and may offer an efficient administrative data driven alternative to these tools with demonstrated performance even in the imminent risk scenario. Comparing tools that have been evaluated in populations with different risk demographics is clearly a challenge, but attempts have been made to collate and compare performance data in a

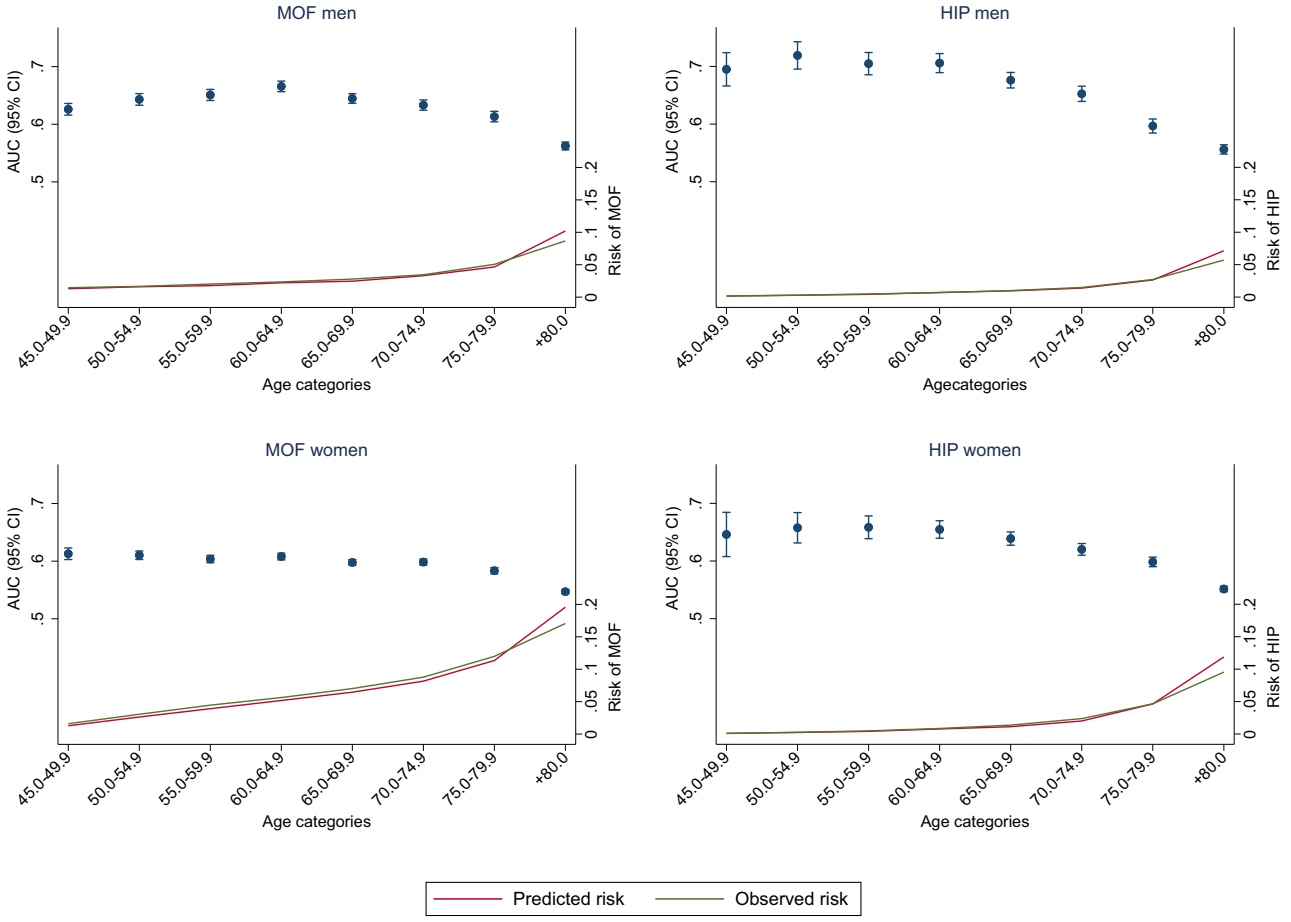

Fig. 3. AUC, the mean predicted and the observed 5-year risk of MOF and hip fractures by age with 15-year look-back. Analyses stratified by age, implying that age itself does not contribute to the AUC.

systematic way. Recently, Beaudoin et al. [21] have reported meta-regression comparisons for the prediction of hip fractures and MOF with some of the commonly applied risk assessment tools [21]. For hip fractures, Garvan with and without BMD, FRAX with and without BMD, and QFracture (2009-version) were compared, yielding unadjusted AUC-values between 0.72 (Garvan without BMD) and 0.88 (QFracture), and adjusted AUC-values between 0.70 (Garvan without BMD) and 0.81 (FRAX with BMD). For major osteoporotic fractures (defined according to FRAX or QFracture 2009), QFracture (2009), and FRAX with and without BMD and adjusted for TBS were compared, yielding unadjusted and adjusted AUC-values of 0.64–0.79 and 0.67–0.77, respectively [21]. Informal comparison with these results suggests that FREM performs on par with these tools. Another study applied administrative data in the FRAX tool which, if successful, would potentially allow for the use in population-based screening in line with the aim of FREM. This study demonstrated c-statistics for hip and major osteoporotic fractures of 0.76 and 0.66, respectively [22], which is numerically lower than the results of FREM, although this could of course be due to different study characteristics rather than the actual performance of the prediction models. Comparison with QFracture is highly relevant, given this also being an administrative data driven tool allowing 1-year fracture risk prediction (though, to our knowledge, QFracture has not been externally validated for 1-year risk assessments). However, several of the clinical risk factors driving QFracture are not readily available in the Danish health registers, including height and weight, ethnicity, parental history of osteoporosis or hip fracture, smoking, alcohol habits, and falls. This limits the potential for using QFracture in Denmark, and potentially in other countries as well.

Several studies have evaluated predictors of imminent fracture risk using either clinical or administrative data [13–15,23]. While some between-study variation exists, all the studies except Hannan et al. [23] identify falls or falls-associated risk factors as associated with an increased risk of fractures, and some include lifestyle factors (e.g. smoking, wheelchair use) and clinically derived measures (e.g. the mini-mental status examination, MMSE) in the models [13–15,23]. In comparison, FREM is based on ICD-10 diagnosis codes obtained in daily clinical practice, which have been linked statistically to an increased imminent fracture risk, with no requirement for the underlying diseases to be pathophysiologically associated with osteoporotic fractures. In addition to identifying imminent fracture risk factors, some of the studies mentioned above also developed predictive models, yet with c-statistics ranging from poor to acceptable [13,14,23]. We are aware of at least one more imminent fracture risk prediction model currently under development, with data on performance and discriminative ability yet to be published in full [24].

A number of studies have recently investigated the use of register data in automated risk assessments [25–31]. A key learning point from some of these studies concerns the need for transparent reporting of the methods applied in the model to enable applicability and clinical understanding [25,28–31]. Further, these studies highlight the variability in application of registry data in predictive models. Hence, beyond information on diagnosis codes from the National Patient Register, Dworzynski et al. [25] included drug prescription data in their machine learning model to predict the 5-year risk of type 2 diabetes comorbidities, Ahlström et al. [26] included data on employment and educational level in their algorithm to predict HIV-status, and Vistisen et al. [27] included clinical data in their prediction model of the 5-year risk of cardiovascular disease in type 1 diabetes patients [25–27]. All three studies presented AUC results similar to FREM [25–27]. The FREM algorithm holds the potential for inclusion of additional registry data, pending confirmation that such data will add additional predictive value to the model.

The Danish health registries lack data on certain powerful clinical fracture risk factors, e.g. smoking, paternal hip fracture, and bone mineral density which is the main limitation of FREM. Further, the registries do not reflect the intensity of the exposure and as such the FREM model measures the mere presence or absence of a risk factor. Hence, it could be speculated that more severe and longer lasting diseases increase the risk of fractures more than the average lasting, average severity version of the same diseases. Other limitations pertain to the outcomes of the model which are limited to clinical fractures only, as is most often the case for fracture risk prediction algorithms, yet with two-thirds to three-quarters of vertebral fractures remaining unreported [32], we are likely to miss a substantial number of such fractures. Reassuringly, this misclassification of outcome is independent of exposure and may merely lead the bias towards the null. Using administrative/registry data makes the performance of the algorithm dependent on the quality and consistency of coding practices, which is likely to differ over time and between countries.

Further, our results seem to indicate that the predictive value of FREM decreases in the older parts of the population, as given by the lower AUC in the older age strata observed in Figs. 2 and 3. It should be noted that these analyses are stratified by age, implying that age itself does not contribute to the AUC, which explains the generally lower AUC in these figures as compared to the main results (Tables 2 and 4). In terms of the decreasing AUC by age, it is likely that unmeasured frailty becomes increasingly important as a predictor of fracture risk in the oldest old with a smaller role for the FREM risk factor set. Also, a similar tendency of higher AUCs in younger cohorts has been reported elsewhere, with the authors suggesting that this may be due to a higher competing risk of death with increasing age [21]. In an analysis of 20- to 25-year fracture risk prediction by bone mineral density and fracture history in women aged 67 years or older, based on data from the Study of Osteoporotic Fractures, Black et al. [33] also evaluated if the impact of the risk factors changed over time. Splitting time from baseline into 5-year age intervals, they found an attenuation over time in the predictive value of both risk factors for hip fractures [33]. Though several reasonable explanations for this exists, it could be speculated that the importance of even established clinical fracture risk factors diminishes in the older age groups, supporting our finding of a decreasing predictive capability of FREM in older age.

The strengths of this study include the high number of patients and patient-years covered, the use of validated registries allowing a substantial lookback, and the nationwide coverage of the cohort omitting selection bias. Moreover, the relatively simple and transparent structure of the FREM model ensures that it can be applied in clinical settings relatively easy.

Future research for FREM should include the validation of the model in other administrative datasets, as well as head-to-head studies against other fracture risk prediction algorithms. In addition, randomized clinical trials (RCTs) evaluating the performance of fracture risk prediction models in a standardized research setting would be of value in terms of understanding the effectiveness of such models in reducing fracture outcomes. Currently, only two recent RCTs have evaluated the effect of such screening [34,35], one demonstrating that screening may be efficient in preventing hip fractures [34].

In conclusion, the evaluation of FREM in predicting the 1-year risk of hip fractures and MOF in men and women aged 45 years or older, demonstrated excellent and acceptable discriminative performance, respectively, when applying administrative health data from the Danish health registries using a 15-year look-back. Applying a 5-year look-back provided similar results, and the FREM tool demonstrated capability of predicting the 5-year risk of hip fractures and MOF with only a minor reduction in the discriminative ability as compared to the one-year risk predictions. These results demonstrate that the FREM tool can be applied to identify individuals at high imminent risk of fracture using administrative health data, though validation in non-Danish registries is needed.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bone.2021.115934.

Author contributions

Rubin, Hyldig, Möller had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design. Rubin, Abrahamsen, Bliddal, Möller.

Acquisition, analysis, or interpretation of the data: All authors.

Drafting of the manuscript: Rubin, Skjødt, Hyldig, Clausen.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis, Möller. Hyldig.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

MKS: Institutional research contract, UCB; Educational grant, UCB; Scientific advisory board member, Hedia.

SM: No conflicts of interest.

NH: No conflicts of interest.

AC: No conflicts of interest.

MB: No conflicts of interest.

JS: No conflicts of interest.

BA: Institutional research contracts with Novartis, UCB and Kyowa-Kirin UK. Consulting or speakers fees from UCB, Amgen, Kyowa-Kirin UK and Eli-Lilly.

KHR: No conflicts of interest.

References

[1] C. Cooper, S. Ferrari, IOF Compendium of Osteoporosis, 2nd edition, 2019.
[2] D.M. Black, C.J. Rosen, Clinical practice. Postmenopausal osteoporosis, N. Engl. J. Med. 374 (2016) 254–262, https://doi.org/10.1056/NEJMcp1513724.
[3] J.A. Kanis, A. Svedbom, N. Harvey, E.V. McCloskey, The osteoporosis treatment gap, J. Bone Miner. Res. 29 (2014) 1926–1928, https://doi.org/10.1002/jbmr.2301.
[4] B. Abrahamsen, M.K. Skjødt, P. Vestergaard, Hip fracture rates and time trends in use of anti-osteoporosis medications in Denmark for the period 2005 to 2015: missed opportunities in fracture prevention, Bone 120 (2019) 476–481, https://doi.org/10.1016/j.bone.2018.12.016.
[5] M.K. Skjødt, S. Khalid, M. Ernst, et al., Secular trends in the initiation of therapy in secondary fracture prevention in Europe: a multi-national cohort study including data from Denmark, Catalonia, and the United Kingdom, Osteoporos. Int. (2020), https://doi.org/10.1007/s00198-020-05358-4.
[6] K.H. Rubin, T. Friis-Holmberg, A.P. Hermann, et al., Risk assessment tools to identify women with increased risk of osteoporotic fracture: complexity or simplicity? A systematic review, J. Bone Miner. Res. 28 (2013) 1701–1717, https://doi.org/10.1002/jbmr.1956.
[7] N.D. Nguyen, S.A. Frost, J.R. Center, et al., Development of a nomogram for individualizing hip fracture risk in men and women, Osteoporos. Int. 18 (2007) 1109–1117, https://doi.org/10.1007/s00198-007-0362-8.
[8] N.D. Nguyen, S.A. Frost, J.R. Center, et al., Development of prognostic nomograms for individualizing 5-year and 10-year fracture risks, Osteoporos. Int. 19 (2008) 1431–1444, https://doi.org/10.1007/s00198-008-0588-0.
[9] J.A. Kanis, O. Johnell, A. Oden, et al., FRAX and the assessment of fracture probability in men and women from the UK, Osteoporos. Int. 19 (2008) 385–397, https://doi.org/10.1007/s00198-007-0543-5.
[10] J.A. Kanis, A. Oden, H. Johansson, et al., FRAX® and its applications to clinical practice, Bone 44 (2009) 734–743, https://doi.org/10.1016/j.bone.2009.01.373.
[11] J. Hippisley-Cox, C. Coupland, Predicting risk of osteoporotic fracture in men and women in England and Wales: prospective derivation and validation of QFractureScores, BMJ 339 (2009) b4229.
[12] J. Hippisley-Cox, C. Coupland, 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 344 (2012), e3427, https://doi.org/10.1136/bmj.e3427.
[13] J.D. Adachi, C. Berger, R. Barron, et al., Predictors of imminent non-vertebral fracture in elderly women with osteoporosis, low bone mass, or a history of fracture, based on data from the population-based Canadian multicentre osteoporosis study (CaMos), Arch. Osteoporos. 14 (2019) 53, https://doi.org/10.1007/s11657-019-0598-x.
[14] D. Weycker, J. Edelsberg, R. Barron, et al., Predictors of near-term fracture in osteoporotic women aged ≥65 years, based on data from the study of osteoporotic fractures, Osteoporos. Int. 28 (2017) 2565–2571, https://doi.org/10.1007/s00198-017-4103-3.
[15] M. Bonafede, N. Shi, R. Barron, et al., Predicting imminent risk for fracture in patients aged 50 or older with osteoporosis using US claims data, Arch. Osteoporos. 11 (2016) 26, https://doi.org/10.1007/s11657-016-0280-5.
[16] K.H. Rubin, S. Möller, T. Holmberg, et al., A new fracture risk assessment tool (FREM) based on public health registries, J. Bone Miner. Res. 33 (2018) 1967–1979, https://doi.org/10.1002/jbmr.3528.
[17] C.B. Pedersen, The Danish civil registration system, Scand J Public Health 39 (2011) 22–25, https://doi.org/10.1177/1403494810387965.
[18] M. Schmidt, S.A.J. Schmidt, J.L. Sandegaard, et al., The Danish national patient registry: a review of content, data quality, and research potential, Clin. Epidemiol. 7 (2015) 449–490.
[19] A. Pottegård, S.A.J. Schmidt, H. Wallach-Kildemoes, et al., Data resource profile: the Danish National Prescription Registry, Int. J. Epidemiol. 46 (2017) 798, https://doi.org/10.1093/ije/dyw213.
[20] H. Quan, B. Li, C.M. Couris, et al., Updating and validating the charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries, Am. J. Epidemiol. 173 (2011) 676–682, https://doi.org/10.1093/aje/kwq433.
[21] C. Beaudoin, L. Moore, M. Gagné, et al., Performance of predictive tools to identify individuals at risk of non-traumatic fracture: a systematic review, meta-analysis, and meta-regression, Osteoporos. Int. 30 (2019) 721–740, https://doi.org/10.1007/s00198-019-04919-6.
[22] S. Yang, W.D. Leslie, S.N. Morin, L.M. Lix, Administrative healthcare data applied to fracture risk assessment, Osteoporos. Int. 30 (2019) 565–571, https://doi.org/10.1007/s00198-018-4780-6.
[23] M.T. Hannan, D. Weycker, R.R. McLean, et al., Predictors of imminent risk of nonvertebral fracture in older, high-risk women: the Framingham osteoporosis study, JBMR plus 3 (2019), e10129, https://doi.org/10.1002/jbm4.10129.
[24] Multinational observational database study on imminent osteoporotic fracture risk: Stage 1. In: Eur. Union Electron. Regist. Post-Authorisation Stud. (EU PAS Regist). http://www.encepp.eu/encepp/viewResource.htm?id=24100, 2018. Accessed 23 Jun 2020.
[25] P. Dworzynski, M. Aasbrenn, K. Rostgaard, et al., Nationwide prediction of type 2 diabetes comorbidities, Sci. Rep. 10 (2020) 1776, https://doi.org/10.1038/s41598-020-58601-7.
[26] M.G. Ahlström, A. Ronit, L.H. Omland, et al., Algorithmic prediction of HIV status using nation-wide electronic registry data, EClinicalMedicine 17 (2019) 100203, https://doi.org/10.1016/j.eclinm.2019.10.016.
[27] D. Vistisen, G.S. Andersen, C.S. Hansen, et al., Prediction of first cardiovascular disease event in type 1 diabetes mellitus: the Steno type 1 risk engine, Circulation 133 (2016) 1058–1066, https://doi.org/10.1161/CIRCULATIONAHA.115.018844.
[28] A. Witteveen, G.F. Nane, I.M.H. Vliegen, et al., Comparison of logistic regression and Bayesian networks for risk prediction of breast Cancer recurrence, Med. Decis. Mak. 38 (2018) 822–833, https://doi.org/10.1177/0272989X18790963.
[29] T. Ahmad, L.H. Lund, P. Rao, et al., Machine learning methods improve prognostication, identify clinically distinct phenotypes, and detect heterogeneity in response to therapy in a large cohort of heart failure patients, J. Am. Heart Assoc. 7 (2018), e008081, https://doi.org/10.1161/JAHA.117.008081.
[30] J. Wallert, M. Tomasoni, G. Madison, C. Held, Predicting two-year survival versus non-survival after first myocardial infarction using machine learning and Swedish National Register Data, BMC Med Inform Decis Mak 17 (2017) 99, https://doi.org/10.1186/s12911-017-0500-y.
[31] J.L. Biccler, S. Eloranta, Brown P. de Nully, et al., Optimizing outcome prediction in diffuse large B-cell lymphoma by use of machine learning and Nationwide lymphoma registries: a Nordic lymphoma group study, JCO Clin Cancer Informatics 2 (2018) 1–13, https://doi.org/10.1200/cci.18.00025.
[32] J.T. Schousboe, Epidemiology of vertebral fractures, J. Clin. Densitom. 19 (2016) 8–22, https://doi.org/10.1016/j.jocd.2015.08.004.
[33] D.M. Black, J.A. Cauley, R. Wagman, et al., The ability of a single BMD and fracture history assessment to predict fracture over 25 years in postmenopausal women: the study of osteoporotic fractures, J. Bone Miner. Res. 33 (2018) 389–395, https://doi.org/10.1002/jbmr.3194.
[34] L. Shepstone, E. Lenaghan, C. Cooper, et al., Screening in the community to reduce fractures in older women (SCOOP): a randomised controlled trial, Lancet 391 (2018) 741–747, https://doi.org/10.1016/S0140-6736(17)32640-5.
[35] K. Rubin, M. Rothmann, T. Holmberg, et al., Effectiveness of a two-step population-based osteoporosis screening program using FRAX: the randomized risk-stratified osteoporosis strategy evaluation (ROSE) study, Osteoporos. Int. 29 (2018) 567–578, https://doi.org/10.1007/s00198-017-4326-3.