Evaluating the performance of the Charlson Comorbidity Index (CCI) in fracture risk prediction and developing a new Charlson Fracture Index (CFI): A register-based cohort study

Word count: 4,764

Author:
Anne Clausen\textsuperscript{a}
Sören Møller\textsuperscript{a, b}
Michael Kriegbaum Skjødt\textsuperscript{b, c}
Bodil Hammer Bech\textsuperscript{d}
Katrine Hass Rubin\textsuperscript{a, b}

Corresponding author:
Katrine Hass Rubin
E-mail address: krubin@health.sdu.dk
Telephone: +45 21261966

\textsuperscript{a}OPEN - Open Patient data Explorative Network, Odense University Hospital, Odense, Denmark.
\textsuperscript{b}Research Unit OPEN, Department of Clinical Research, University of Southern Denmark
\textsuperscript{c}Department of Medicine, Holbæk Hospital, Smedelundsgade 60, DK-4300 Holbæk, Denmark
\textsuperscript{d}Department of Public Health – Department of Epidemiology, Aarhus University, Bartholins Allé 2 building 1260, 232, DK-8000 Aarhus C, Denmark
Mini abstract (50 words)
The Charlson Comorbidity Index (CCI) may be applicable for predicting fracture risk since several diagnoses from the index are predictors of fracture. Main results were that the CCI was updated to predict risk of hip fracture with fair precision and that the index could be useful in detecting high-risk individuals.

Abstract (250 words)
Purpose
Several of the Charlson Comorbidity Index (CCI) diagnoses are validated predictors of fracture. The purpose of this study was to evaluate the performance of the CCI 1987 by Charlson et al. and of the CCI 2011 by Quan et al. in predicting major osteoporotic fracture (MOF) and hip fracture (HF). Furthermore, it was examined if the index could be modified to improve fracture risk prediction.

Methods
The study population included the entire Danish population aged 45+ years as per 1st January 2018. The cohort was split randomly 50/50 into a development and a validation cohort. CCI diagnoses and fracture outcomes were identified from hospital diagnoses. The weighting of diagnoses was updated in a new Charlson Fracture Index (CFI) using multivariable logistic regression. Predictive capabilities of the CCI 1987, the updated CCI 2011, and the new Charlson Fracture index were evaluated in the validation cohort by receiver operating characteristics (ROC) curves and area under the curve (AUC).

Results
In the validation cohort the 1987 and 2011 CCIs resulted in AUCs below or around 0.7 in prediction of MOF and HF in both sexes. The CFI resulted in AUCs < 0.7 in prediction of MOF in both sexes. In prediction of HF the CFI resulted in AUC of 0.755 (95% CI 0.749; 0.761) in women and 0.782 (95% CI 0.772; 0.793) in men.

Conclusion
The 1987 and 2011 CCIs showed overall poor accuracy in fracture risk prediction. The CFI showed fair accuracy in prediction of HF in women and in men.

Keywords
Charlson Comorbidity Index, CCI, Osteoporotic Fractures, Osteoporosis, Automated Risk Calculation, Register data
**Introduction**

Osteoporosis is a disease characterized by low bone mass and deterioration of bone tissue which results in compromised bone strength and increased risk of fragility fractures [1]. Osteoporosis is under-diagnosed, under-treated and causes around 8.9 million fractures worldwide every year [2, 3]. Osteoporotic fractures are considered to be significant causes of mortality and morbidity for patients as well as an economic burden on society [4]. Fractures caused by osteoporosis represent a considerable public health problem which calls for the detection of high-risk individuals to improve screening and treatment decisions in clinical practice.

Numerous risk factors for osteoporosis and fractures have been identified and several tools have been developed to integrate risk factors into a single estimate of fracture risk for individuals [5, 6]. Among those the World Health Organization (WHO) fracture risk assessment tool (FRAX) is one of the most known worldwide [7]. FRAX includes data on clinical and lifestyle risk factors, thus requiring clinical evaluation [8]. Other fracture risk assessment tools also emphasize clinical risk factors while yet other tools are based on administrative data from registers. An advantage of register-based tools is that the use of administrative data eliminates extensive data collections and recall bias. Such tools lend themselves towards automatic risk calculations which may be useful for the systematic detection of high-risk individuals. An example of a register-based fracture risk assessment tool is the Fracture Risk Evaluation Model (FREM) that was developed using hospital diagnoses from Danish registries [9]. Other hospital diagnoses that were not detected in FREM may also be relevant predictors to include in fracture risk assessments performed on administrative data.

The Charlson Comorbidity Index (CCI) was developed in 1987 as a method of predicting mortality and measuring burden of disease [10]. The 1987 index included 19 comorbidities weighted from 1 to 6 according to the impact of each comorbidity on 1-year mortality [10]. Quan et al. validated and updated the CCI in 2011 on more recent health data where it was still found to be a valid predictor of mortality [11]. The CCI has been investigated as a predictor of mortality, yet has not been evaluated in fracture prediction. A number of studies included CCI diagnoses as predictors for mortality in hip fracture (HF) cohorts, where the outcome was mortality and not subsequent fracture in itself [12-15]. However, several of the diagnoses from the CCI are associated with fracture in the literature [16-30]. These include heart failure [16], peripheral vascular disease [17], cerebrovascular disease [18, 19], Alzheimers disease [20], chronic pulmonary disease [18, 19], systemic sclerosis.
Furthermore, CCI diagnoses are identifiable in administrative data through ICD-10 coding [31]. Hence it should be explored if the CCI is an applicable tool for fracture risk assessments using administrative data and the hypothesis of this study is that the CCI may be applied to fracture risk prediction.

The aim of this study was to evaluate the performance of the CCI 1987 and of the CCI 2011, respectively, in predicting risk of major osteoporotic fracture (MOF) and HF. The second objective was to explore if the weighting of CCI diagnoses could be modified to improve the fracture risk prediction in a new Charlson Fracture Index (CFI). The third objective was to validate the performance of the new CFI.

Methods

Study design and study population

This nationwide register-based cohort study was based on a data extraction from Danish national registers covering all citizens living in Denmark on January 1, 2018 aged 45 years or above. The data extraction included fifteen years of retrospective data (2003 to 2017) and one year of follow-up data (2018).

Danish national registers

Danish national registers offer long-term quality data on the entire population which makes it possible to conduct large population-based studies. All individuals living in Denmark are assigned a unique personal identification number, which enables the linkage of information on diagnoses from registries on the individual level [32]. Since 1968 the Danish Civil Registration System (DCRS) has registered all persons alive and living in Denmark for administrative use [33]. The DCRS was used to identify persons for inclusion in the study population. Data were extracted on all citizens in Denmark aged 45+ years and alive on January 1, 2018 from the DCRS.

The Danish National Patient Registry (DNPR) provides nationwide registration of all inpatient, outpatient and emergency department contacts from Danish hospitals, since 1977 (outpatient since 1995) for both publicly and privately owned hospitals (private hospitals since 2003) [34]. For every patient contact one primary and optional secondary diagnoses are registered according to the International Classification of Diseases (ICD) [34]. The DNPR was used to retrieve information on
ICD-10 codes from 2003 to 2018. ICD-10 codes from the DNPR were used for defining both predictors and outcomes.

Outcomes (fractures)
The primary outcome was MOF, and the secondary outcome was HF. The definition of MOF is identical with the definition used in FRAX and FREM \[9\]. MOF includes hip, clinical vertebral, wrist or humerus fracture (ICD-10 codes: S120, S121, S122, S220, S221, S320, T08, S422, S423, S720, S721, S722, S525, S526) during 2018 (January 1, 2018 to December 31, 2018) derived from DNPR data as primary or secondary diagnosis. HF during 2018 was also derived from DNPR data (ICD-10 codes: S720, S721, S722). Both outcomes were coded as dichotomous variables (0=no, 1=yes).

Predictors (CCI diagnoses)
Predictors were the 17 diagnoses included in the CCI which were coded according to the ICD-10 coding suggested by Quan et al. \[31\]. ICD-10 coding of CCI diagnoses has proven high positive predictive value (PPV) in DNPR data which qualifies the coding of the diagnoses as predictors in the dataset \[35\]. Table S1 shows the specification of ICD-10 coding as well as index weights in the 1987 and 2011 CCI, respectively. ICD-10 codes were derived from the DNPR as primary or secondary diagnoses and the study population’s exposure to CCI diagnoses were followed for a 15-year look-back period from 2003 to 2017 (January 1, 2003 to December 31, 2017). All 17 predictors were coded as dichotomous variables (0=no, 1=yes).

Other study parameters (age)
Age on January 1, 2018 was categorized into the following categories: 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, ≥80 years.

Statistical analyses
The cohort was split randomly 50/50 into a development and a validation cohort, respectively. Descriptive statistics were reported stratified by sex and in development/validation cohorts. Categorical variables (CCI diagnoses, age groups, MOF and HF) were reported as counts and proportions. Median and quartiles were reported for numerical age. Death during 2018 was reported as counts and proportions.

CCI 1987 and CCI 2011
In the 1987 and 2011 CCIIs generated in this study, CCI diagnoses were assigned weights using a similar methodology as reported in publications by Charlson et al. and by Quan et al., respectively \[10, 11\] (Table S1). Age and sex were not assigned weights. The performance of the CCI 1987 and
the CCI 2011 in prediction of MOF and HF were evaluated in the validation cohort by producing receiver operating characteristics (ROC) curves and estimating area under the curve (AUC) stratified by sex.

**Development of a new Charlson Fracture Index: CFI**

Higher age is known to be associated with increased fracture risk and categorical age was for that reason included in the CFI as a predictor [1]. Further, fracture risk differs between sexes and the index was therefore split by sex. The CFI resulted in four indexes:

- **CFI MOF women**
- **CFI MOF men**
- **CFI HF women**
- **CFI HF men**

Weights for inclusion in the CFI indexes were derived from odds ratios from multivariable logistic regression analyses performed in the development cohort. For the CFI MOF women a multivariable logistic regression analysis was performed with women only from the development cohort, which included the dichotomous categorical outcome MOF, the 17 dichotomous categorical CCI diagnoses and the 8 age groups. The same analysis was performed with men only from the development cohort as well as with HF as outcome instead of MOF with separate analyses for both sexes.

Using the odds ratios from the multivariable logistic regression analysis, weights were updated using a methodology similar to those applied in the 1987 and 2011 articles [10, 11].

Weights were assigned in the following manner: variables with a p-value of $\geq 0.05$ or an odds ratio of $< 1.2$ were assigned a weight of 0; variables with an odds ratio $\geq 1.2$ and $< 1.5$ were assigned a weight of 1; variables with an odds ratio $\geq 1.5$ and $< 2.5$ were assigned a weight of 2; variables with an odds ratio of $\geq 2.5$ and $< 3.5$ were assigned a weight of 3, variables with an odds ratio of $\geq 3.5$ and $< 4.5$ were assigned a weight of 4, variables with an odds ratio of $\geq 4.5$ and $< 6$ were assigned a weight of 5, and variables with an odds ratio of $\geq 6$ were assigned a weight of 6. Only variables with p-values $< 0.05$ were assigned weights 1-6.

The following comorbid conditions were mutually exclusive following the methodology reported by Quan et al. [11]: Diabetes with organ damage and diabetes without organ damage; mild liver disease and moderate or severe liver disease; and any malignancy and metastatic solid tumor.
To evaluate the four resulting indexes (CFI MOF women, CFI MOF men, CFI HF women and CFI HF men) on the development cohort ROC curves and AUCs with 95% confidence intervals were produced. Univariable logistic regression analyses were performed for descriptive purposes.

Validation of the CFI
The four final CFIs were validated internally in the validation cohort. The two indexes for women with MOF and HF, respectively, were validated in women only from the validation cohort. The two indexes for men with MOF and HF, respectively, were validated in men only from the validation cohort. The validation was performed by producing ROC curves and calculating AUCs with 95% confidence intervals.

Sensitivity analyses
Firstly, an age-only index was developed by repeating the logistic regression analysis with categorical age only and weights were assigned to each age category.
Secondly, a CCI-only index was developed by repeating the regression analysis with the 17 CCI diagnoses adjusted for age ≥ 65 years but only assigning weights to the 17 CCI diagnoses and not age in the index.
Thirdly, a CFI-age-stratified index was developed by repeating the regression analysis with interactions between CCI diagnoses and age categories. The regression analysis thereby resulted in odds ratios for the 17 CCI diagnoses within each of the eight age categories as well as in odds ratios for categorical age. In the CFI-age-stratified the 17 CCI diagnoses were weighted within each age group and the age group itself were also weighted in the index.
Lastly, the internal validity of the CFI for HF prediction was evaluated by excluding individuals who sustained a HF before baseline January 1st, 2018 and also when excluding HFs that where sustained within a 6 months grace period of a former HF, respectively. These sensitivity analyses were performed with HF rather than MOF as the outcome, because diagnosis codes for HFs have recently been validated against surgical codes in the DNPR [36].
All sensitivity analyses were performed by re-running the logistic regression analyses on the development cohort for each sensitivity index by including the predictors for each index and assigning new weights according to odds ratio results.

All analyses were conducted using Stata 16 [37]. AUC results of all analyses were interpreted as suggested by Carter et al. [38].
Results

Descriptive statistics
On January 1st, 2018, Denmark had a total population of 2,643,369 individuals aged 45 years or above (1,365,537 women and 1,277,832 men). The development cohort consisted of 682,770 women and 638,919 men. The validation cohort consisted of 682,767 women and 638,913 men (Table 1). The study population held more women than men which was expected as the life expectancy of Danish women is higher than the life expectancy of Danish men. For the same reason, the median age of women in the study population was higher than that of the men (62.7 versus 61.2 years in the development cohort and 62.7 versus 61.3 years in the validation cohort). By the end of 2018, 2.0% of women and 2.2% of men had died. During 2018 1.5% of women and 0.6% of men experienced a MOF, while 0.4% of women and 0.2% of men experienced a HF (Table 1). The cumulative incidence of CCI diagnoses from 2003-2017 ranged from 0.0% for AIDS/HIV to 8.7% for any malignancy in women in the development cohort. For men in the development cohort the cumulative incidence of CCI diagnoses from 2003-2017 ranged from 0.1% for AIDS/HIV to 7.7% for any malignancy (Tables 2-3).

Performance of the CCI 1987 and of the CCI 2011 in the validation cohort
In the validation cohort the CCI 1987 resulted in AUC=0.591 (95% CI 0.586; 0.596) for women and AUC=0.638 (95% CI 0.629; 0.646) for men in prediction of MOF (Table 4). The CCI 1987 resulted in AUC=0.658 (95% CI 0.649; 0.668) for women and AUC=0.711 (95% CI 0.697; 0.724) for men in prediction of HF. The CCI 2011 resulted AUC=0.569 (95% CI 0.564; 0.574) for women and AUC=0.609 (95% CI 0.601; 0.617) for men in prediction of MOF. The CCI 2011 resulted in AUC=0.619 (95% CI 0.609; 0.628) for women and AUC=0.665 (95% CI 0.652; 0.679) for men in prediction of HF.

Development of a new Charlson Fracture Index: CFI
On the basis of results from the multivariable logistic regression analyses each CCI diagnosis and age group were assigned weights according to the method described in the statistics section. Results of unadjusted analyses are shown in supplemental tables 2-3. The CFI MOF women weighted 8 CCI diagnoses with weights 1-2 and age categories with weights 2-6 (Table 2). The maximum possible index score for risk of MOF in women was 16. The CFI MOF women resulted in an AUC of 0.704 (95% CI 0.700; 0.709) in the development cohort (Table 4). The CFI MOF men weighted 10 CCI diagnoses with weights 1-3 and age categories with weights 1-5. The maximum possible index score for risk of MOF in men was 20. The CFI MOF men resulted in an AUC of 0.684 (95%
CI 0.676; 0.693) in the development cohort. The *CFI HF women* weighted 8 CCI diagnoses with weights 1-2 and age categories with weights 3-6 (Table 3). The maximum possible index score for risk of HF in women was 16. The *CFI HF women* resulted in an AUC of 0.758 (95% CI 0.752; 0.764) in the development cohort (Table 4). The *CFI HF men* weighted 9 CCI diagnoses with weights 1-5 and age categories with weights 2-6. The maximum possible index score for risk of HF in men was 24. The *CFI HF men* resulted in an AUC of 0.776 (95% CI 0.765; 0.787) in the development cohort.

**Validation of the CFI**

In the validation cohort the CCI 1987, the CCI 2011 and the new CFI produced ROC curves (Figure 1: A-D) and AUCs (Table 4). For both sexes and outcomes, the CFI performed better than the 1987 and 2011 CCIs in the validation cohort in predicting fracture risk. In addition, the CCI 1987 performed better than the CCI 2011 in predicting risk of both MOF and HF (Figure 1). The *CFI* resulted in AUC=0.699 (95% CI 0.695; 0.704) for women and AUC=0.693 (95% CI 0.684; 0.702) for men in prediction of MOF in the validation cohort. The *CFI* performed better in HF prediction with AUC=0.755 (95% CI 0.749; 0.761) for women and AUC=0.782 (95% CI 0.772; 0.793) for men in the validation cohort.

**Index score distributions and cut off points**

The highest numbers of MOFs in women were detected in those with a baseline *CFI* score of 6+, the highest numbers of MOFs in men were detected in those with a baseline *CFI* score of 2. The highest numbers of HFs were detected in those with a baseline *CFI* score of 6+ in both women and men (Table S4). *CFI* score distributions for the entire study population irrespective of fracture outcome showed similar results as the largest proportions of the study population received a *CFI* score of 6+ due to the weighting of age in the index (Table S5). Oppositely, in the 1987 and 2011 CCIs the highest numbers of MOF and HF were seen in patients with a score of 0 in both sexes and cohorts. Age is not weighted in the 1987 and 2011 CCIs which explains this distribution (Table S4). In the validation cohort a *CFI* score of 6+ detected 65.4% of MOFs in women, a *CFI* score of 2+ detected 79.4% of MOFs in men and a *CFI* score of 6+ detected 97% of HFs in women and 85.4% of HFs in men (Table 5). Further increases in *CFI* scores as cut offs resulted in significant drops in sensitivity. Additionally, a *CFI* score of 6+ resulted in a Positive Predictive Value (PPV) of 2.8% for MOFs in women, a *CFI* score of 2+ resulted in a PPV of 0.9% for MOFs in men and a *CFI* score of 6+ resulted in a PPVs of 0.7% and 0.5% for HFs in women and men, respectively (Table 5).
Sensitivity analyses

The *age-only* index weighted age categories with weights 0-6 (Table S6) and resulted in AUC=0.688 (95% CI 0.683; 0.692) for women and AUC=0.669 (0.660; 0.678) for men in MOF prediction in the validation cohort (Table 4). Further, the *age-only* index resulted in AUC=0.700 (95% CI 0.696; 0.703) for women and AUC=0.727 (95% CI 0.718; 0.736) for men in prediction of HF in the validation cohort. The CCI-*only* index with MOF weighted 11 diagnoses for women with weights 1-2 and 12 diagnoses for men with weights 1-3 (Table S7). The CCI-*only* index with HF weighted 10 diagnoses for women with weights 1-4 and 11 diagnoses for men with weights 1-4. It resulted in AUC=0.581 (95% CI 0.576; 0.586) for women and AUC=0.630 (95% CI 0.622; 0.639) for men in MOF prediction in the validation cohort. The CCI-*only* index resulted in AUC=0.660 (95% CI 0.650; 0.670) for women and in AUC=0.707 (95% CI 0.694; 0.721) for men in prediction of HF in the validation cohort (Table 4).

Weight assignment of the CCI diagnoses in the CFI-*age-stratified* are reported in supplemental tables S8-S11. The CFI-*age-stratified* with MOF for women resulted in AUC=0.689 (0.684; 0.693) in the validation cohort (Table 4). The CFI-*age-stratified* with MOF for men resulted in AUC=0.687 (0.678; 0.696) in the validation cohort. The CFI-*age-stratified* with HF for women resulted in AUC=0.709 (0.704; 0.715) in the validation cohort. The CFI-*age-stratified* with HF for men resulted in AUC=0.760 (0.750; 0.771) in the validation cohort.

Lastly, the CFI for HF prediction was tested for both sexes in analyses excluding individuals who sustained a HF before baseline January 1\textsuperscript{st}, 2018 and also in analyses excluding HFs sustained within a 6 months grace period of a former HF. These analyses did not change the results. Results were reported in supplemental table S12.

Discussion

Main findings

In this study the performances of the CCIs reported by Charlson et al. in 1987 and by Quan et al. in 2011, respectively, were evaluated in terms of fracture risk prediction. The CCI was originally developed to predict mortality and measuring burden of disease and the current study is to our knowledge the first study to evaluate the performance of the CCI in fracture risk prediction. The 1987 and 2011 CCIs showed overall poor accuracy in prediction of MOF and HF in both sexes. It was explored if the weighting of diagnoses in the CCI could be updated to improve fracture risk prediction. On the basis of results from four multivariable logistic regression analyses the Charlson
Fracture Index (CFI) was developed by assigning new weights to the 17 CCI diagnoses and adding age to prediction in four resulting indexes: CFI MOF women, CFI MOF men, CFI HF women and CFI HF men. The CFI showed poor accuracy in prediction of MOF and fair accuracy in prediction of HF in both sexes. Age was the strongest predictor in the CFI and age categories were assigned the highest weights. In sensitivity analyses an age-only index resulted in AUCs similar to the CFI in MOF prediction in both sexes, indicating that the inclusion of CCI diagnoses did not contribute substantially to MOF prediction. The CFI resulted in markedly higher AUCs in prediction of HF in both sexes compared to the age-only index, indicating that the inclusion of CCI diagnoses contributed beyond age in prediction of HF. Due to the poor performance in predicting risk of MOF the CFI did not prove to be an efficient tool in identifying individuals with increased risk of MOF. The CFI proved to be a useful tool in identifying individuals with increased risk of HF. The CFI score of 6+ identified 93% of all HFs and it could be speculated that a CFI score of 6 may be a relevant cut-off for screening decisions.

Comparison of the CFI with other prediction models and studies
The above-mentioned Fracture Risk Evaluation Model (FREM) used Danish administrative register data for fracture risk prediction, similar to the CFI [9]. The same registers, study population and time period definition applied to the current CFI study were recently applied in a validation study of FREM [39]. In the validation study, FREM resulted in AUC=0.757 (95% CI 0.754; 0.760) for women and AUC=0.746 (95% CI 0.740; 0.752) for men in prediction of MOF, as well as AUC=0.867 (95% CI 0.863; 0.871) for women and AUC=0.844 (95% CI 0.837; 0.851) for men in prediction of HF with a 15-year look-back period and 1-year risk prediction period [39]. FREM performed better than CFI in MOF and HF prediction, however a limitation to FREM is that it results in a high number of diagnoses as risk factors (38 in women and 43 in men for MOF and 32 for both sexes for HF). In comparison, the CFI resulted in a much lower number of diagnoses as risk factors (8 in women and 10 in men for MOF, 8 in women and 9 in men for HF). The low number of risk factors is a strength to the CFI as it provides a manageable number of diagnoses to screen for in clinical practice.

Other fracture risk assessment tools include FRAX, Garvan and Qfracture, amongst others [8]. These fracture risk assessment tools use data that are not readily available in Danish health registers because data collection requires laboratory testing, clinical measurements or self-report [8]. FRAX and Garvan include data on bone mineral density (BMD) from DXA (Dual energy X-ray...
Absorptiometry) scans, and FRAX and Qfracture include data on lifestyle risk factors including smoking habits and alcohol consumption [8]. Such data provides information on known risk factors for osteoporosis and fractures which may improve CFI fracture risk prediction. Examples of studies applying prediction models with clinical and lifestyle risk factors to fracture prediction include Akyea et al. and Dennison et al. [40, 41]. Akyea et al. used FRAX and Qfracture to predict risk of MOF and HF in a cohort of UK patients (n=72,559) with chronic obstructive pulmonary disease (COPD) [40]. FRAX resulted in AUC=0.714 (95% CI 0.706; 0.722) and Qfracture in AUC=0.614 (95% CI 0.605; 0.623) for prediction of MOF. Both models resulted in an AUC=0.761 (95% CI 0.749; 0.772) for HF. Akyea et al. used UK General Practice health records to identify patients with COPD, MOF and HF, and further included records on history of falls, smoking status, alcohol consumption, BMI (Body Mass Index) and use of prescription drugs [40]. In addition, Dennison et al. derived a fracture risk prediction model resulting in AUC= 0.636 for prediction of incident fracture by using data from the Global Longitudinal Study of Osteoporosis in Women (GLOW) which included self-reported data on comorbidities, BMI, history of fractures and falls, smoking status and alcohol consumption [41]. Comparisons between these studies and the CFI should nonetheless be made with caution since the results may not be comparable between studies that differ in study populations and methodology. To determine whether inclusion of clinical and lifestyle risk factors would improve predictive capabilities of the CFI, the model should be tested with and without inclusion of such risk factors. For instance, Marques et al. tested the ability of the FRAX tool with and without BMD to predict 10 year fracture risk in a cohort study to determine whether the addition of BMD improved fracture risk prediction [42].

Strengths and limitations
A considerable strength of this study is that it covers the entire Danish population aged 45 years or older (> 2.6 million), thus minimizing the risk of selection bias. The large study population adds statistical strength to findings. Both women and men were included in the study population and the CFI was developed stratified by sex, hence taking account of gender-specific differences in comorbidities.

Another strength of this study is the transparent methodology of the development of the CFI and the use of ICD-10 coding which enables evaluation and application of the model in other study populations with available administrative health records. Lastly, the CFI is a fairly simple index that resulted in a small number of risk factors which is manageable to screen for in clinical practice.
However, calculation of the *CFI* requires information on medical history, which may not be readily available in all settings hence potentially limiting the application of the *CFI* (in line with the CCI). A limitation to this study is that the register data used did not contain information on known clinical risk factors of osteoporotic fracture such as BMI, BMD, history of falls/fractures and drug use; nor does it contain information on lifestyle risk factors including smoking habits, alcohol consumption or physical activity, which are also associated with fracture risk [1, 8]. Inclusion of such data may enhance the *CFI*.

A second limitation to this study is that death is not accounted for as a competing risk to fractures in the follow-up year of 2018 by using logistic regression to develop the new *CFI*. Leslie et al. investigated the effect of not accounting for death as a competing risk in a 10-year fracture risk prediction in Canadian administrative data [43]. Conclusions were that failure to account for competing mortality risk resulted in considerably higher estimates of 10-year fracture risk than if adjustments were made to account for death in analyses with Cox regression or Kaplan Meier methods [43]. In the current study, the effect of not accounting for competing mortality risk is assumably smaller compared to the study by Leslie et al. since the fracture risk prediction was calculated for a 1-year risk period in the *CFI*, whereas Leslie et al. calculated a 10-year fracture risk. However, since it is possible to account for death as a competing risk by using time-to-event methods it could be considered in future work with fracture risk prediction models instead of applying logistic regression methods.

A third limitation is that the definition of fracture outcomes with diagnosis codes from the DNPR does not take into account whether the fracture is new or a check-up for earlier fractures, as these data are not available in the DNPR. Thus, a diagnosis code for a fracture in the follow-up year of 2018 may describe a check-up for a fracture that occurred in for instance 2017. This limitation may have resulted in inflated estimates of fracture incidence in the outcome year 2018. Our sensitivity analyses for HF indicated that the effect on AUCs was small.

**Conclusion**

In conclusion, the 1987 and 2011 Charlson Comorbidity Indexes (CCIs) resulted in overall poor performances in predicting risk of major osteoporotic fractures (MOFs) and hip fractures (HFs) in both women and men. The weighting of CCI diagnoses was updated in a new Charlson Fracture Index (*CFI*) which resulted in a poor performance in prediction of MOF in both sexes in the validation cohort. The CCI diagnoses did not contribute substantially to the prediction of MOF in the *CFI* and the prediction was primarily driven by age. The *CFI* demonstrated fair performance in
prediction of HF in both sexes, showing predictive capabilities beyond age alone. The CFI proved to be an applicable tool for identifying individuals with increased risk of HF. Further, the CFI resulted in a low number of diagnoses as predictors of HF (8 in women and 9 in men), which is a manageable number to screen for. External validation of the CFI would be the next step in evaluating clinical applicability.

Declarations

Funding: This study did not receive any specific grants.

Conflicts of interest:
AC: No conflicts of interest
SM: No conflicts of interest
MKS: Educational grant, UCB Institutional research grant, UCB/Amgen.
BHM: No conflicts of interest
KHR: No conflicts of interest

Availability of data and material: The register data used for this study are not publicly available. Data were available to AC and SM with permission from the Danish Health Data Authority.

Code availability: Not applicable.

Authors’ contributions:
Concept and design: AC, SM, KHR
Statistical analysis: AC, SM
Interpretation of results: All authors
Drafting of manuscript: AC, MKS
Critical revision of manuscript: All authors

Ethics approval: This was a register-based study which do not require an ethics committee approval according to Danish law [44]. The study was approved by the Danish Data Protection Agency (jf.nr. 2008-58-0035).

Consent to participate: Not applicable.

Consent for publication: Not applicable.

References

https://doi.org/10.2147/TCRM.S138000


[37] StataCorp (2019) Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC.


