Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Hospital Frailty Risk Score in hip fracture measures documented illness burden, not frailty: an external validation study

Yuanjie Li
First People's Hospital of Yunnan Province
Chunyu Wang
First People's Hospital of Yunnan Province
Liangzhao Ma
First People's Hospital of Yunnan Province
Junhua Fang
First People's Hospital of Yunnan Province
Weibing Liu
First People's Hospital of Yunnan Province
Jianhong Hou
hjh@kust.edu.cn

First People's Hospital of Yunnan Province

Research Article

Keywords: frailty, Hospital Frailty Risk Score, hip fracture, external validation, Charlson Comorbidity Index, coding intensity, administrative data

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11446727/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

The Hospital Frailty Risk Score in hip fracture measures documented illness burden, not frailty: an external validation study

Authors: Yuanjie Li[1]‡, Chunyu Wang[1]‡, Liangzhao Ma[1], Junhua Fang[1], Weibing Liu[1]*, Jianhong Hou[1]†

Author notes: ‡ These authors contributed equally to this work. * Corresponding author. † Co-corresponding author.

Affiliations: [1] Department of Orthopedics, The First People's Hospital of Yunnan Province (Affiliated Hospital of Kunming University of Science and Technology), Kunming, Yunnan, China

Corresponding author: Weibing Liu, Department of Orthopedics, The First People's Hospital of Yunnan Province (Affiliated Hospital of Kunming University of Science and Technology), Kunming, Yunnan, China. E-mail: ynliuweibing@163.com; Tel: +86 13648800214

Co-corresponding author: Jianhong Hou, Department of Orthopedics, The First People's Hospital of Yunnan Province (Affiliated Hospital of Kunming University of Science and Technology), Kunming, Yunnan, China. E-mail: hjh@kust.edu.cn; Tel: +86 18808713303

Author e-mails: Yuanjie Li, 18877590464@163.com; Chunyu Wang, 18487236334@163.com; Liangzhao Ma, 18288233219@163.com; Junhua Fang, fangjunhua0719@163.com; Weibing Liu, ynliuweibing@163.com

Word count (abstract): 333 · **Manuscript type:** Original research (TRIPOD Type 4 external validation)

Abstract

Background. In hip fracture, geriatric co-management and discharge planning are directed at patients judged frail, and the Hospital Frailty Risk Score (HFRS) is increasingly used to make that judgement. Whether it measures frailty has never been tested: every validation has scored the same admission whose outcome it predicts. We therefore scored it twice, at the index and earlier admissions.

Methods. Retrospective cohort from MIMIC-IV v3.1 (2015–2019): 1,594 hospitalizations (1,445 patients) aged ≥65 years with hip fracture (ICD-10 S72.0–S72.2). We compared the HFRS with the Charlson Comorbidity Index (CCI) and Elixhauser indices, quantified incremental value over comorbidity and routine admission laboratory values, classified its 109 codes as acute, chronic or geriatric, and compared the pre-admission score (n=558) with a Charlson index built from the same prior codes.

Results. The HFRS discriminated 30-day mortality similarly to the CCI (0.726 vs. 0.693, p=0.19) and added to it (ΔAUC +0.026 to +0.041) and to laboratory values (+0.078). It correlated with the number of coded diagnoses (ρ=0.76), added nothing beyond it at 30 days (ΔAUC −0.000) and was worse for prolonged stay (0.837 vs. 0.881, p=0.001). Two-thirds of its points (65.6%) came from acute events and complications; the acute component alone reproduced its discrimination (30-day AUC 0.726; chronic 0.606), and removing acute codes left none of the 322 admissions above 15. It flagged 20.2% of admissions yet captured 44.5% of 30-day deaths and 65.9% of prolonged stays. Net benefit for prolonged stay was higher with the HFRS at every threshold (20%: 0.028 vs. 0.007). Before the fracture, the score was at chance at 30 days (0.551), inferior to a Charlson index from the same prior codes (0.649, p=0.018), and correlated only weakly with the index score (ρ=0.28).

Conclusions. At discharge the HFRS can flag the patients whose admission went badly—one in five, yet nearly half of 30-day deaths and two-thirds of prolonged stays—for geriatric follow-up. It summarises what was documented during the admission, not the vulnerability that preceded it, and should not guide surgery or baseline risk adjustment.

Keywords: frailty; Hospital Frailty Risk Score; hip fracture; external validation; Charlson Comorbidity Index; coding intensity; administrative data

Background

Hip fracture is a sentinel event in the life of an older adult, carrying high early mortality [1], prolonged institutionalization, and loss of independence [2], with incidence rising as populations age [3]. Prognosis after hip fracture is driven less by the fracture itself than by the patient's underlying vulnerability—frailty and comorbidity burden—making risk stratification at admission central to geriatric co-management [4,5].

Traditionally, risk stratification has relied on comorbidity burden, exemplified by the Charlson Comorbidity Index (CCI) [6]. However, prognosis in older adults depends not only on how many diseases are present but also on frailty—a distinct syndrome reflecting diminished physiological reserve [7]. Comorbidity and frailty are related but not identical: comorbidity asks "which diseases," whereas frailty captures the organism's overall reduced capacity to withstand a stressor, and their contributions to prognosis may differ [8,9].

The Hospital Frailty Risk Score (HFRS) was introduced to capture the second of these axes. Developed by Gilbert and colleagues from the English Hospital Episode Statistics, it sums 109 ICD-10 codes into a 0–99 score and predicts 30-day mortality, prolonged length of stay, and readmission in older inpatients [10]. Its appeal in hip fracture is obvious: no examination, no laboratory, no interviewer—just administrative codes. The score has been applied to hip fracture cohorts and predicts adverse outcomes in them [11,12]. What such studies establish, however, is that the score predicts. They do not establish that a score labelled "frailty" measures frailty. A score built from discharge codes could predict just as well by summarising how much illness was documented in the record—and, in particular, how much acute illness the patient experienced during the very admission whose outcome is being predicted. The volume of documentation is known to vary widely between institutions and countries for reasons unrelated to case-mix

[13,14], and individual codes are an imperfect record of the conditions actually present [15]. The two interpretations carry opposite clinical consequences. If the HFRS measures pre-existing vulnerability, it can be used to triage patients at admission and to adjust for baseline risk in research. If it measures what happened during the admission, neither use is valid, and the score belongs at discharge.

To our knowledge, published validations of the HFRS, from the development study itself [10] to the two in hip fracture [11,12], have all scored the admission whose outcome they predict. That design cannot separate a measure of pre-existing vulnerability from a record of what happened during the stay, because both produce the same number at the same moment.

We therefore did not set out simply to validate the HFRS in hip fracture, but to determine what it measures, through three questions. First, does it carry prognostic information beyond comorbidity and beyond routinely measured admission laboratory values? Second, where does any such information come from—from the sheer number of codes recorded, and from which categories of code? Third, does a score computed from the patient's hospitalizations before the fracture—the version a clinician could actually compute on the morning of surgery—carry the same signal? We reasoned that a measure of pre-existing vulnerability should survive the third test, whereas a measure of documented acute illness should not.

Methods

Study design and data source

This is a retrospective external validation of a previously developed prediction model (TRIPOD type 4) [16], conducted using MIMIC-IV version 3.1. MIMIC-IV is a publicly available, de-identified database of all hospitalizations at the Beth Israel Deaconess Medical Center (BIDMC), a tertiary academic center in Boston, USA, between 2008 and 2019 [17,18]. The study used the hospital module (`admissions`, `patients`, `diagnoses_icd`, `labevents`,

`microbiologyevents`). Because ICD-10 coding begins in October 2015 in MIMIC-IV, the analytic cohort spans 2015–2019 although the database covers 2008–2019. Because the data are de-identified, this analysis was exempt from institutional review board review.

Participants

We included all hospitalizations of patients aged ≥65 years with a discharge diagnosis of hip fracture, defined by ICD-10 codes S72.0 (fracture of neck of femur), S72.1 (pertrochanteric fracture), or S72.2 (subtrochanteric fracture) in any diagnostic position. Each hospitalization was retained as a unit of analysis for in-hospital outcomes; patient identifiers were used for clustering and mortality linkage. In a sensitivity analysis we retained only the first admission per patient to assess the influence of repeated hospitalizations.

Hospital Frailty Risk Score

The HFRS was calculated from ICD-10 diagnostic codes using the original 109 codes and weights of Gilbert et al. [10]. Each 3-character code present during the hospitalization was assigned its weight (0.1–7.1), and weights were summed to produce a continuous score (range 0–99). Patients were categorized using the original cut-offs: low risk (<5), intermediate risk (5–15), and high risk (>15). ICD-10 codes were matched after removing the decimal point.

Comparator

The Charlson Comorbidity Index was calculated from ICD-10 codes using the coding algorithm of Quan et al. [6,19], summed across 17 comorbidity categories as a continuous score.

Outcomes

Mortality was assessed at 30, 90, and 365 days from admission using the date of death recorded in MIMIC-IV, which is compiled from hospital records and the Massachusetts State Registry of Vital Records and Statistics [18]. Prolonged hospital stay was defined as a length of

stay >10 days (the top decile and a threshold used in prior HFRS validation [10]). Thirty-day readmission was defined as a subsequent hospitalization within 30 days of discharge. As an exploratory outcome, venous thromboembolism (VTE) was defined by the presence of ICD-10 codes I26 (pulmonary embolism), I80 (phlebitis and thrombophlebitis), or I82 (other venous embolism) during the hospitalization. Because only 24 events occurred, this outcome was not analysed for discrimination and is reported descriptively only.

Statistical analysis

Continuous variables are reported as median (IQR) and categorical variables as count (%). Discrimination was assessed by the area under the ROC curve (AUC), with 95% confidence intervals (CIs) by the Hanley–McNeil method [20] and comparisons by the DeLong test [21]. Incremental value was quantified with logistic models containing age, sex and CCI, and with the HFRS added, reporting the odds ratio per point, likelihood-ratio test, ΔAUC, IDI and category-based NRI [22], with CIs from 500 patient-level bootstrap resamples; the reverse model added the CCI to age, sex and the HFRS. Survival was assessed with Kaplan–Meier curves, log-rank tests and Cox models adjusted for age, sex and CCI. Because the HFRS has no published risk equation, absolute calibration could not be assessed; we report observed event rates by category and decile calibration of a model refitted in these data [23]. Decision-curve analysis compared the net benefit of the HFRS, the CCI and treat-all/treat-none strategies [24]. Sensitivity analyses repeated the main analyses by diagnosis, age, admission sequence, sex and admission type, and all diagnoses per admission were counted to test whether the HFRS added anything to coding intensity. Analyses used Python (NumPy, SciPy, pandas, matplotlib); two-sided $p<0.05$ was significant.

Additional comparators and analyses

To place the HFRS within the wider family of administrative indices, we additionally derived the Elixhauser comorbidity index from ICD-10 codes using the Quan algorithm, both as an unweighted count of the 30 categories [25] and with the van Walraven weights [26]. Discrimination of the HFRS, the CCI and both Elixhauser representations was compared on identical outcomes with the DeLong test.

To determine whether the HFRS adds prognostic information beyond routinely available admission data, we fitted a laboratory-based model containing age, sex and the first 24-hour values of haemoglobin (minimum), serum creatinine (maximum), platelet count (minimum) and serum sodium (maximum). We then compared this model with the same model plus the HFRS using ΔAUC, the likelihood-ratio test, IDI and category-based NRI. The analysis was restricted to complete cases (n=1,113).

We also characterised patients with low comorbidity but high frailty (CCI 0–1 and HFRS >15) across demographics, admission characteristics, laboratory values, microbiology, resource use and outcomes, and compared them with low-comorbidity, low-frailty patients (CCI 0–1 and HFRS <5).

Pre-admission analysis. To test whether the score carries information that is available before the fracture, we computed a second HFRS for each index admission using only diagnoses recorded during that patient's earlier hospitalizations. The primary prior window was the 730 days preceding the index admission, mirroring the two-year look-back used in the original derivation [10]; a patient entered the analysis if at least one prior admission in the window carried an ICD-10 diagnosis. Codes were de-duplicated across prior admissions and weighted exactly as in the original score, and a Charlson index was computed from the same set of codes; the number of unique codes served as the coding-intensity comparator. We compared discrimination of the pre-admission HFRS, the pre-admission CCI and the index HFRS, tested whether the pre-admission HFRS added to age, sex and the pre-admission CCI (ΔAUC, IDI and NRI with 500 clustered bootstrap resamples), and quantified how much of the index score had already been coded before

admission. Sensitivity analyses used all available prior records, excluded prior admissions within 30 days, and retained only the first index admission per patient.

Code-level composition analysis

To determine which components of the HFRS carried its prognostic information, we classified all 109 codes into three mutually exclusive groups from the code definitions alone and before any outcome was examined: chronic or pre-existing conditions and established deficits (for example dementia, hemiplegia, parkinsonism, blindness or deafness, osteoporosis, arthrosis and chronic organ failure); acute events, injuries or complications of the index admission (for example falls, fractures other than the index fracture, delirium, acute kidney injury, fluid or electrolyte disorders, infection, respiratory failure and procedure complications); and geriatric syndromes, function or care status (gait and cognitive symptoms, incontinence, care and social Z codes). Points were summed within class for each admission using the same one-code-one-count rule as the original score. We described the share of total points contributed by each class and the codes responsible, correlated each class score with the total number of coded diagnoses, compared the discrimination of each class score with that of the total score and of the diagnosis count, fitted a model containing all three class scores simultaneously, and tested whether each class score added discrimination and model fit (likelihood-ratio test) to a base model of age, sex, the CCI and the diagnosis count. We then recomputed the score without acute codes to determine how many patients in each HFRS category remained above the original thresholds. Because 19 codes could reasonably be assigned to more than one class (for example urinary tract infection versus incontinence, duodenal ulcer, pathological fracture, resistant organism and social admission codes), we repeated the key analyses with all 19 counted as chronic and with all 19 excluded. The full per-code classification and the number of admissions carrying each code are provided in Additional file 2: Table S2.

Use of AI tools

A large language model (Claude, Anthropic) was used to assist with drafting and language editing of the text and with writing and debugging of the analysis code. All AI-generated text and code were reviewed, verified and edited by the authors, who take full responsibility for the content of the manuscript.

Results

Cohort characteristics

We identified 1,594 hospitalizations among 1,445 patients with hip fracture (Fig. 1, Table 1). Median age was 82 years (IQR 75–88), 1,109 (69.6%) were women, and 696 (43.7%) were admitted emergently or urgently; hip fracture was the primary diagnosis in 1,328 (83.3%). The median HFRS was 8.8 (IQR 5.3–13.7); 346 (21.7%) were low risk, 926 (58.1%) intermediate, and 322 (20.2%) high risk. The median CCI was 2 (IQR 1–4, range 0–14). Median length of stay was 4 days (IQR 3–7). Cumulative mortality was 9.2% at 30 days, 17.4% at 90 days, and 30.6% at 365 days; prolonged stay occurred in 164 (10.3%), 30-day readmission in 79 (5.0%), and venous thromboembolism in 24 (1.5%).

Clinical risk stratification: low comorbidity does not exclude high frailty

Cross-tabulating the cohort by HFRS and CCI categories (Table 2) showed that the two scores carry different risk information. Among the 749 patients with low comorbidity burden (CCI 0–1), 74 (9.9%) were high-risk by frailty (HFRS >15); these "low-comorbidity, high-frailty" patients had 30-day mortality of 13.5% and prolonged stay of 21.6%. By contrast, patients with high comorbidity but low frailty (CCI ≥4 and HFRS <5) had 4.2% (1 of 24) 30-day mortality and 0% prolonged stay. Prolonged stay tracked frailty rather than comorbidity: across the three CCI tiers, it was 0–0.4% in the low-frailty row versus 22–44% in the high-frailty row. Mortality, in contrast, increased along both axes, rising from 1.2% (low frailty, low comorbidity) to 27.0% (high frailty, high comorbidity) at 30 days, and to 69.2% at 365 days in the latter group. These

results indicate that comorbidity-only screening would miss approximately one in ten high-risk frail patients.

Patients missed by comorbidity-only screening

An HFRS >15 flagged 20.2% of admissions yet captured 44.5% of 30-day deaths and 65.9% of prolonged stays. Patients with CCI 0–1 accounted for 47.0% of the cohort and 21.9% of its 30-day deaths, and 74 of them (9.9%) were flagged. Those 74 patients differed from the 260 low-comorbidity, low-frailty patients mainly in age and acute illness rather than in organ-specific laboratory abnormalities (Table 3). They were older (median 84.5 vs. 78.0 years; $p<0.001$), stayed longer (6 vs. 3 days; $p<0.001$), more often had a positive culture (27.0% vs. 3.1%; $p<0.001$), were more often admitted to intensive care (20.3% vs. 3.5%) and had higher blood urea nitrogen (20 vs. 18 mg/dL; $p=0.014$), whereas haemoglobin, creatinine, platelet count and sodium did not differ ($p=0.13$–0.78). Thirty-day mortality was 13.5% vs. 1.2% ($p<0.001$), 365-day mortality 32.4% vs. 6.5%, and prolonged stay 21.6% vs. 0.4%. Non-home discharge (92.8% vs. 85.3%; $p=0.11$) and 30-day readmission (5.4% vs. 2.3%; $p=0.24$) did not differ significantly. The mirror-image group with high comorbidity but low frailty (CCI ≥4, HFRS <5; n=24) had 30-day mortality of 4.2%.

Discrimination against the Charlson Comorbidity Index

For mortality, the HFRS and CCI discriminated similarly, with the HFRS numerically but not significantly superior at every horizon and the gap narrowing with longer follow-up: 30-day AUC 0.726 (95% CI 0.678–0.775) vs. 0.693 (0.644–0.742), $\Delta+0.034$, $p=0.19$; 90-day 0.729 (0.694–0.765) vs. 0.709 (0.672–0.745), $\Delta+0.021$, $p=0.25$; 365-day 0.725 (0.696–0.753) vs. 0.713 (0.684–0.742), $\Delta+0.011$, $p=0.45$ (Fig. 2, Table 4).

For prolonged stay the HFRS was significantly superior: 0.837 (0.799–0.876) vs. 0.716 (0.670–0.762), $\Delta+0.122$, $p<0.001$. Neither score discriminated 30-day readmission (0.615 vs.

0.618, p=0.94). Venous thromboembolism was recorded in only 24 hospitalizations, so discrimination was not analysed for this outcome.

Comparison with other administrative indices

The HFRS was not consistently superior to the Elixhauser indices (Table 5). For 30-day mortality, AUROCs were 0.726 (95% CI 0.678–0.775) for the HFRS, 0.693 (0.644–0.742) for the CCI, 0.698 (0.648–0.747) for the unweighted Elixhauser count and 0.752 (0.705–0.799) for the van Walraven-weighted Elixhauser index. No comparison reached significance: HFRS vs. CCI Δ+0.034, p=0.19; HFRS vs. unweighted Elixhauser Δ+0.029, p=0.16; CCI vs. unweighted Elixhauser Δ−0.005, p=0.81. At 90 days the HFRS was superior to the unweighted Elixhauser count (0.729 vs. 0.691; Δ+0.039, p=0.019), and at 365 days all pairwise differences were non-significant (p>0.08). The four indices were moderately correlated (Spearman ρ 0.49–0.76). For prolonged stay, the HFRS (AUC 0.837) outperformed the CCI (0.716; Δ+0.122, p<0.001) and the van Walraven-weighted Elixhauser index (0.791; Δ+0.046, p=0.018), but did not differ significantly from the unweighted Elixhauser count (0.809; Δ+0.029, p=0.12) (Table 5). These results indicate that the HFRS is one reasonable representation of accumulated vulnerability rather than a uniquely informative instrument.

Calibration

Observed mortality increased monotonically across HFRS categories: 30-day mortality 1.7% (low), 8.1% (intermediate), 20.2% (high); 90-day 4.0%, 16.3%, 35.1%; and 365-day 8.4%, 30.7%, 54.3% (Fig. 3). In logistic models fitted in this cohort, each HFRS point corresponded to an odds ratio of 1.13 (95% CI 1.10–1.16) for 30-day mortality, 1.14 (1.11–1.16) at 90 days and 1.14 (1.12–1.17) at 365 days. Apparent decile calibration of these refitted models was acceptable at 30 days (Hosmer–Lemeshow p=0.088) but deviated at longer horizons (p=0.007 and p<0.001; Fig. S3). Because no risk equation was published with the original score, these estimates describe

the behaviour of the HFRS in this cohort and do not constitute an external calibration of absolute risk.

Incremental value beyond comorbidity

Adding the HFRS to a base model of age, sex, and CCI yielded independent and significant incremental value for mortality at every horizon. The OR was 1.10 per point at all three horizons; ΔAUC ranged from +0.026 to +0.041 (all $p \leq 0.006$), IDI from +0.029 to +0.042 (all $p \leq 0.011$), and NRI from +0.124 to +0.218 (all $p \leq 0.006$). Incremental value was larger for prolonged stay: OR 1.22 per point (95% CI 1.18–1.26), ΔAUC +0.118, IDI +0.148 (Fig. 4, Table 6).

The reverse incremental analysis clarified the relationship between the two scores. Adding the CCI to a base model of age, sex, and HFRS retained significant incremental value for mortality (OR 1.17–1.31 per point, ΔAUC +0.028 to +0.036, all $p<0.003$), but added essentially nothing for prolonged stay (ΔAUC +0.004, $p=0.30$) (Table 7). Thus, prolonged hospitalization was captured by the HFRS rather than by the CCI, whereas for mortality the two scores were complementary; as shown below, the advantage of the HFRS for prolonged stay is itself shared with overall coding burden.

Incremental value beyond routine laboratory data

In the complete-case subset (n=1,113), a model containing age, sex, haemoglobin, creatinine, platelet count and sodium achieved an AUROC of 0.679 for 30-day mortality, compared with 0.596 for age and sex alone. Adding the HFRS raised the AUROC to 0.756 (Δ+0.078; DeLong $p=0.0006$), with an OR of 1.11 per point (95% CI 1.07–1.14), a likelihood-ratio χ^2 of 37.8 ($p<0.001$), an IDI of +0.034 (95% CI 0.013–0.065) and an NRI of +0.328 (0.178–0.584). The corresponding increments were +0.064 (IDI +0.055, NRI +0.357) at 90 days and +0.064 (IDI +0.076, NRI +0.360) at 365 days (Table 8). The incremental value of the HFRS is

therefore not confined to comparisons with other comorbidity indices; it persists after adjustment for routinely measured admission laboratory values.

Decision-curve analysis

Across threshold probabilities of 5–40%, acting on the HFRS gave higher net benefit than acting on the CCI for prolonged stay, and the HFRS stayed above both the treat-all and the treat-none strategies throughout, whereas the CCI crossed zero at a threshold of about 30%. At a threshold of 20% the net benefit was 0.028 for the HFRS and 0.007 for the CCI, a difference equivalent to 2.1 additional net true positives per 100 patients assessed (Fig. S4). For 30-day mortality the two scores were similar, and net benefit was positive below a threshold of about 20% for the CCI and about 25% for the HFRS.

Survival

Kaplan–Meier curves separated significantly by HFRS category (log-rank $p<0.001$; Fig. 5): 365-day survival was 91.6% (low), 69.3% (intermediate), and 45.7% (high). In Cox models, the fully adjusted (age, sex, CCI) HR was 1.07 per HFRS point (95% CI 1.06–1.09, $p<0.001$); the CCI in the same model had an HR of 1.17 (1.13–1.21), with both independently associated. Categorically, intermediate and high risk (vs. low) had fully adjusted HRs of 2.98 (2.03–4.39) and 4.99 (3.33–7.50) (Table 7).

Sensitivity and subgroup analyses

All sensitivity analyses were consistent (Fig. S2, Fig. S5; Table 7). Restricting to primary-diagnosis hip fracture (n=1328), age ≥85 years (n=616), or first admission only (n=1445), the HFRS advantage persisted for prolonged stay (AUC 0.798–0.862 vs. 0.634–0.722); the same held after stratification by sex (women: 0.836 vs. 0.722; men: 0.840 vs. 0.684) and admission type (emergency: 0.857 vs. 0.718; elective: 0.819 vs. 0.710). For mortality, the HFRS was slightly

superior in men (365-day 0.770 vs. 0.705, p=0.011) and emergency admissions (0.744 vs. 0.694, p=0.022), and equivalent in women, elective admissions, and the age ≥85 years subgroup.

Discrimination was higher in patients younger than 75 years (30-day AUROC 0.805, 95% CI 0.702–0.909) than in older patients (0.703, 0.649–0.757). Discrimination was lower in patients admitted to intensive care (0.625, 0.538–0.711), in those with a positive culture (0.651, 0.554–0.749) and within HFRS strata (intermediate 0.625; high 0.577). Attenuation within strata is expected from range restriction and indicates that the HFRS stratifies risk across categories more reliably than within them.

Coding burden

The HFRS was closely related to how much was coded. The median number of diagnoses per hospitalization was 16 (IQR 11–22), and the HFRS correlated strongly with this count (Spearman ρ=0.76; ρ=0.67 after removing the rank of length of stay). Median diagnosis counts were 9, 16 and 26.5 across the low, intermediate and high HFRS categories. As a discriminator, the diagnosis count alone achieved AUCs of 0.748 (95% CI 0.701–0.795) for 30-day mortality and 0.881 (0.847–0.916) for prolonged stay. Formal comparison showed no difference between the HFRS and the count for 30-day mortality (0.726 vs. 0.748; p=0.17) or 365-day mortality (0.725 vs. 0.727; p=0.80), whereas for prolonged stay the count was significantly better (0.881 vs. 0.837; p=0.001). Within quintiles 2–5 of the diagnosis count, the HFRS discriminated 30-day mortality with AUCs of 0.49–0.65 and prolonged stay with 0.62–0.69 (quintile 1, which contained only two cases of prolonged stay, could not be evaluated); range restriction attenuates discrimination within strata, but these values indicate that little prognostic information remained once the number of coded diagnoses was fixed. After adjustment for age, sex, the CCI and the diagnosis count, the HFRS added no discrimination for 30-day mortality (ΔAUC −0.000; DeLong p=0.38) or prolonged stay (ΔAUC −0.004; p=0.12); a small association persisted for 365-day mortality (likelihood-ratio p=0.001) without a gain in AUC (Additional file 1: Table S1).

What the score is made of

Of the 100 HFRS codes that occurred in this cohort, 51 described acute events, injuries or complications, 32 chronic or pre-existing conditions, and 17 geriatric syndromes, function or care status; the remaining nine codes did not occur. Every one of the 1,594 hospitalizations carried at least one acute code (Table 9). Acute codes contributed 6.56 points per admission and 65.6% of all points, chronic codes 2.69 points and 26.9%, and geriatric codes 0.75 points and 7.5%. The largest single contributions were the index femur fracture (S72, present in every admission and worth 1.4 points by definition of the cohort), falls (W19 and W18, 3.2 and 2.1 points in 21.1% and 28.8% of admissions), delirium (F05, 3.2 points in 13.8%), electrolyte disorders (E87, 2.3 points in 18.9%) and acute kidney injury (N17, 1.8 points in 18.9%). The largest chronic contributions were personal history codes (Z87, 1.5 points in 30.7%), dementia (F03, 2.1 points in 19.1%), chronic kidney disease (N18, 1.4 points in 25.3%) and osteoporosis (M81, 1.4 points in 20.9%).

Discrimination was carried almost entirely by the acute component. For 30-day mortality the acute score achieved an AUC of 0.726, matching the total score, whereas the chronic score achieved 0.606 and the combination of chronic and geriatric codes 0.622 (total vs. non-acute, $p<0.001$). The pattern held at 365 days (acute 0.699, chronic 0.653, total 0.725) and for prolonged stay (0.848, 0.611 and 0.837; non-acute 0.668). In a model containing all three components, the adjusted odds ratio per point for 30-day mortality was 1.16 (95% CI 1.12–1.20) for acute codes, 1.06 (0.98–1.15; $p=0.13$) for chronic codes and 1.05 (0.93–1.18; $p=0.44$) for geriatric codes. For prolonged stay, acute codes (1.30, 1.25–1.35) and geriatric codes (1.25, 1.11–1.40) were associated, whereas chronic codes were not (1.02, 0.94–1.11).

Adding any component to a model containing age, sex, the CCI and the diagnosis count produced no gain in discrimination (ΔAUC −0.010 to +0.000), although the acute component retained a small likelihood-ratio contribution for 365-day mortality ($p=0.006$) and prolonged stay

(p<0.001). Consistent with these results, removing acute codes moved most patients below the conventional thresholds: of the 1,248 admissions with HFRS ≥5, 824 (66.0%) had fewer than 5 non-acute points, and none of the 322 admissions with HFRS >15 remained above 15 (maximum non-acute score 14.3). The result was robust to the classification: when all 19 ambiguous codes were counted as chronic, the reading most favourable to a frailty interpretation, the chronic share rose only to 40.1%, 52.4% of admissions with HFRS ≥5 still fell below 5 without acute codes, and only 13 admissions retained a score >15 (Table 9).

A score computed before the fracture does not carry the same information

To determine whether the score reflects information available before the fracture, we recomputed it from each patient's earlier hospitalizations. Of the 1,594 index admissions, 652 (40.9%) were preceded by at least one hospitalization of the same patient that contained ICD-10 diagnoses. Restricting this to the 730 days before the fracture left 558 admissions in 495 patients (median one prior admission; median 108 days before the index admission). These patients were sicker than the full cohort (median index HFRS 10.0 vs. 8.8; 365-day mortality 36.6% vs. 30.6%).

In this subgroup the score computed from prior diagnoses differed substantially from the index score: it was higher in 57.9% of admissions (median 11.0 vs. 10.0), the two scores correlated only weakly (Spearman ρ=0.28), and a median of 34.5% (IQR 17.8–51.1) of the points in the index score came from codes that had already been recorded before admission. Of the 131 admissions with an index HFRS >15, 60 (45.8%) also had a pre-admission score >15; the remaining 71 crossed the threshold only during the index admission.

Discrimination of the pre-admission score was close to chance for 30-day mortality (AUC 0.551, 95% CI 0.471–0.631) and poor at 90 and 365 days (0.598 and 0.594). A CCI calculated from exactly the same prior codes discriminated better at every horizon (0.649, 0.674 and 0.683; ΔAUC −0.098, p=0.018; −0.075, p=0.015; −0.090, p<0.001), while the index score remained the strongest of the three (0.674, 0.710 and 0.686) and was markedly better for prolonged stay (0.789

vs. 0.622, p<0.001). Adding the pre-admission HFRS to a model containing age, sex and the pre-admission CCI produced no gain in discrimination (ΔAUC −0.003 to +0.011, all p≥0.16), IDI (all 95% CI included zero) or NRI (all 95% CI included zero); the per-point OR for 90-day mortality reached p=0.007 (1.03 per point), but this was not accompanied by any change in discrimination. Event rates rose across pre-admission categories only weakly and non-monotonically for 30-day mortality (10.5%, 8.6% and 12.9% for low, intermediate and high; p=0.32), whereas the index score showed a clear gradient (2.6%, 8.9% and 19.1%; p<0.001) (Table 10, Fig. 6). Findings were unchanged when all available prior records were used (n=652), when prior admissions within 30 days were excluded (n=479), and when only the first index admission per patient was analysed (n=495) (Additional file 8: Tables S3–S6).

Discussion

Main findings

The clinical question behind this study was whether the HFRS can identify, while decisions are still open, the hip fracture patient who will do badly. In 1,594 older adults hospitalized with hip fracture, the HFRS discriminated 30-day mortality with an AUC of 0.726 and prolonged stay with 0.837, and it added modestly to the Charlson Comorbidity Index. Three findings qualify that performance. Almost all of its prognostic information was shared with the number of diagnoses recorded: it added nothing over that count for 30-day mortality and discriminated less well than the count for prolonged stay. Two-thirds of its points were contributed by acute events of the index admission, and the acute component alone reproduced the discrimination of the whole score. And when it was computed from diagnoses recorded before the fracture it discriminated close to chance (AUC 0.551) and worse than a Charlson index built from the same prior codes. In hip fracture, therefore, the HFRS is better read as a weighted summary of the diagnoses documented during the admission than as a measurement of the reserve a patient brought to it.

One clinical observation follows directly. Patients with low comorbidity but an HFRS >15—9.9% of those with a CCI of 0–1—had a 30-day mortality of 13.5%, above that of patients with high comorbidity and a low HFRS (4.2%, 1 of 24), and they are missed by comorbidity-only screening. What marks them out, however, is the acute burden coded during their admission, not a frail state that preceded it. Its clinical value and its limitation are the same fact: it summarises what happened during the admission.

Comparison with previous studies

The HFRS has been applied across acute-care and surgical populations since 2018 [11,12,14], but evidence in hip fracture remains limited. Shimizu and Maeda [11] found that it predicted in-hospital mortality and adverse events in a nationwide Japanese inpatient database, and the index-admission results reported here point in the same direction. Two observations from outside orthopaedics are also consistent with our reading. Against a bedside assessment, the HFRS identified only 27% of patients judged frail on the Clinical Frailty Scale, with a specificity of 93% [27], and in intensive care unit patients it has been described as of limited value [28]—the pattern expected if the score counts documented illness rather than measuring reserve. And in a general acute-care population, Gilbert et al. found that combining the HFRS with the Charlson and Elixhauser indices identified older patients at risk of poor outcomes better than any one index alone, while the HFRS itself added little to the Charlson index for 30-day in-hospital mortality [29]; the latter is consistent with the modest incremental value we observed for 30-day mortality (ΔAUC +0.026 to +0.041), in contrast to its larger advantage for prolonged stay (+0.122).

One result runs against ours. Wong et al. [12] reported that the HFRS was associated with length of stay, complications and cost but not with mortality, whereas our index score discriminated 30-day mortality with an AUC of 0.726. At least three differences could account for this, and none can be resolved with the published data. Case-mix varies widely between hip fracture cohorts—the proportion of patients in the low HFRS category was 21.7% here and 68.4%

in the Japanese series [11]—and discrimination falls when the outcome is rare or the sample small. Their study related the score to outcomes as an association rather than evaluating discrimination, so the two results are not directly comparable. And neither study separated the risk a patient carried into the admission from the risk generated during it, which is the distinction our coding-burden analyses address.

Two of our findings are, to our knowledge, new. First, no previous study has computed the HFRS from diagnoses recorded before the fracture, which is the only version that could serve as a preoperative frailty measure; the poor agreement between the two versions (ρ=0.28) and the near-chance discrimination of the pre-admission score indicate that the prognostic value of the score is generated by the admission it is used to predict. Second, the advantage of the HFRS over the CCI for prolonged stay (ΔAUC +0.122) has not been quantified in hip fracture; Gilbert et al. observed the same direction [29], and our analysis suggests that the advantage reflects the volume of documented illness rather than a specific frailty effect. In our cohort the van Walraven-weighted Elixhauser index achieved the highest 30-day AUC numerically (0.752 vs. 0.726 for the HFRS), although that comparison was not formally tested. Taken with their proposal, this argues for combining administrative indices rather than replacing one with another [29].

Why the score behaves as it does

The behaviour we observed follows from three properties of the instrument and of the data that feed it.

The first is design. The HFRS was built to flag older inpatients at risk of adverse outcomes from hospital records [10]; its items were selected because they mark documented severity, and a large share of them denote acute events and complications rather than stable conditions. Fifty-one of the 100 codes that occurred in our cohort described acute events, injuries or complications. None of the 109 codes in the instrument measures gait speed, muscle mass, exhaustion or any

other component of the frailty phenotype [8], so whatever it captures of frailty it captures indirectly.

The second is the coding process. Diagnoses in administrative data are assigned at or after discharge and describe what happened during the stay; the number recorded rises with length of stay [30]; coding accuracy differs by condition and by hospital [15]; and MIMIC-IV carries no present-on-admission flag [18]. A patient who develops delirium, acute kidney injury or an electrolyte disorder therefore accrues points that a patient with identical reserve and an uncomplicated course does not. In our cohort a median of only 34.5% of the points in the index score came from codes already recorded before admission; the remainder was generated by the stay itself.

The third is the population. The mechanism is close to its maximum in hip fracture, because the fracture and the operation guarantee acute injury and complication codes: every admission in our cohort carried at least one acute code, and the fracture alone contributed 1.4 points by definition of the cohort. In such a population a score assembled from acute codes cannot do otherwise than track the severity of the hospital course, which is also why the raw diagnosis count matched or exceeded it as a discriminator (AUC 0.748 for 30-day mortality and 0.881 for prolonged stay).

These properties do not make the score uninformative; they define what its information is about. Beyond a comorbidity index it does add prognostic information, but beyond the number of coded diagnoses it adds almost none (ΔAUC −0.000 for 30-day mortality and −0.004 for prolonged stay after adjustment for age, sex, the CCI and the count), and a small association with 365-day mortality persisted without any gain in discrimination. Its apparent advantage over the CCI is therefore best understood as the coding burden that a complicated hospital course leaves behind. That information is generated by the admission it is meant to explain, and therefore too late to guide the decisions taken at its start.

Clinical implications

These findings support a discharge-time rather than an admission-time use case, and the practical question is what to do with the number when it appears on a discharge summary. Used at discharge, the score answers a question worth asking: which patients left hospital after a complicated course and will need help afterwards. Flagging one admission in five captured 44.5% of 30-day deaths and two-thirds of prolonged stays; among the 749 patients with a CCI of 0–1—47.0% of the cohort and 21.9% of its deaths—the 74 who were flagged had 30-day mortality of 13.5% (95% CI 7.5–23.1) against 1.2% (0.4–3.3) in the 260 with low comorbidity and an HFRS <5. The cross-tabulation (Table 2) converts the score into absolute risks that can be quoted to patients and families (HFRS >15: 20.2% die within 30 days, 27.0% when comorbidity is also high), and for prolonged stay the net benefit of acting on the score exceeded that of acting on the CCI at every threshold examined. That is the sense in which it has clinical value: a case-finding and risk-communication tool for geriatric follow-up, discharge planning and quality monitoring [5], not a physiological measure.

Used at admission, it cannot do the job it is most often asked to do. Two-thirds of its points arise during the stay, so the score available on the morning of surgery is a partial one, and the version computable from earlier records does not work (AUC 0.551). It should therefore not be used as a preoperative frailty screen or to inform the surgical or anaesthetic plan, and it should not serve as a baseline covariate in studies of hip fracture outcomes, because its points are generated by the admission those studies mean to predict. Where a baseline measure is required, the diagnosis count should be reported alongside the score, or a comorbidity index should be used, which in our data extracted more prognostic information from the same prior records.

Two cautions apply to any use. Because the score overlaps with coding intensity and has no published risk equation, it supports relative ranking rather than absolute prediction; and an elevated score describes what happened to the patient, not what the patient brought to surgery—it

marks the group for whom geriatric co-management, discharge planning and follow-up may be targeted [5], but not those who were frail before they fell. If the score is to inform decisions taken before surgery, what is needed is a measure built from information available on admission: a bedside instrument such as the Clinical Frailty Scale, whose disagreement with the HFRS is documented [27], or an administrative score restricted to diagnoses flagged present on admission [18]. Whether either adds to the diagnosis count and the CCI, and whether acting on the addition changes management or outcome, are questions our design cannot answer.

Limitations

Several limitations should be noted. First, this was a single-centre retrospective study in a case mix skewed towards severe patients (21.7% low-risk, against 68.4% in the Japanese cohort [11]), so discrimination is easier to demonstrate here than in healthier populations and the findings may not transfer to other hospitals or coding systems; we did not adjust for operative timing, anaesthetic technique or surgical approach. Second, the data are administrative: the timing of each diagnosis is not recorded [18], our acute/chronic/geriatric classification of codes is our own, and frailty itself was not measured, so the conclusion that the HFRS is not a frailty measure rests on indirect evidence; and, as the original score had no risk equation, calibration was assessed in a model refitted here—a recalibration-based validation rather than a test of the original coefficients. Third, both scores derive from the index admission, so predictor and outcome share one hospital course; the pre-admission analysis breaks this circle only approximately (558 admissions, 58 deaths within 30 days), and non-significant comparisons are inconclusive rather than evidence of equivalence.

Conclusions

In older adults with hip fracture, comorbidity-only screening misses about one in ten patients at high risk, and the HFRS does add prognostic information beyond the Charlson Comorbidity

Index and beyond routine admission laboratory values. That information is not, however, information about pre-existing vulnerability: it is almost entirely shared with the number of coded diagnoses, two-thirds of the score's points arise from acute events and complications of the admission, the acute component alone reproduces its discrimination, and a score computed from earlier hospitalizations discriminated poorly (AUC 0.55–0.62). For the clinician, the HFRS answers a discharge-time question—not how frail was this patient before the fracture, but how badly did this admission go, and who will need help next. It is a tool for case-finding and for ranking risk, not a way to choose the operation or to adjust for baseline risk, and a measure that would judge pre-fracture frailty remains to be validated in this population.

List of abbreviations

AUC, area under the receiver operating characteristic curve; CCI, Charlson Comorbidity Index; CI, confidence interval; DCA, decision-curve analysis; HFRS, Hospital Frailty Risk Score; HR, hazard ratio; ICD-10, International Classification of Diseases, 10th Revision; IDI, integrated discrimination improvement; IQR, interquartile range; LOS, length of stay; NRI, net reclassification improvement; OR, odds ratio; ROC, receiver operating characteristic; VTE, venous thromboembolism.

Declarations

Ethics approval and consent to participate. The collection of patient information and creation of the MIMIC-IV research resource were reviewed by the Institutional Review Board of the Beth Israel Deaconess Medical Center, which granted a waiver of informed consent and approved the data sharing initiative. This secondary analysis of publicly available, de-identified data was therefore exempt from further ethics review and did not require participant consent.

Clinical trial number: not applicable.

Consent for publication. Not applicable.

Availability of data and materials. The dataset supporting the conclusions of this article is available in the PhysioNet repository, MIMIC-IV version 3.1, https://doi.org/10.13026/kpb9-mt58 (https://physionet.org/content/mimiciv/3.1/). Access is restricted to credentialed users who complete the required training (CITI) and sign a data use agreement; the authors cannot share the data directly. The analysis code is available from the corresponding author on reasonable request.

Competing interests. The authors declare no competing interests.

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

Authors' contributions. Yuanjie Li: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft. Chunyu Wang: Data curation, Investigation, Validation. Liangzhao Ma: Data curation, Investigation, Validation. Junhua Fang: Investigation, Validation. Weibing Liu: Resources, Validation, Supervision, Project administration, Writing – review & editing. Jianhong Hou: Supervision, Writing – review & editing. All authors read and approved the final manuscript.

Acknowledgements. Not applicable.

Use of AI tools. The authors used a large language model (Claude, Anthropic) to assist with drafting and language editing of the text and with writing and debugging of the analysis code. All AI-generated text and code were reviewed, verified and edited by the authors, who take full responsibility for the content of the manuscript.

Additional files.

Additional file 1: Table S1. Coding burden and incremental value of the HFRS beyond the diagnosis count.

Additional file 2: Table S2. Classification of the 109 Hospital Frailty Risk Score codes and their contribution in 1,594 hip fracture admissions.

Additional file 3: Figure S1. Distribution of the HFRS.

Additional file 4: Figure S2. Sensitivity forest plot (primary-diagnosis hip fracture, age ≥85 years).

Additional file 5: Figure S3. Decile-based calibration plots (observed vs. predicted, 30/90/365-day mortality) from models refitted in the validation cohort (apparent calibration).

Additional file 6: Figure S4. Decision-curve analysis: net benefit versus threshold probability for the HFRS, the CCI, and treat-all/treat-none strategies.

Additional file 7: Figure S5. Sensitivity forest plot (first admission only, women, men, emergency, elective).

Additional file 8: Tables S3–S6. Pre-admission Hospital Frailty Risk Score: category-specific event rates, sensitivity to the prior window, first-admission analysis and cross-tabulation.

References

1. Forssten MP, Mohammad Ismail A, Ioannidis I, Wretenberg P, Borg T, Cao Y, et al. The mortality burden of frailty in hip fracture patients: a nationwide retrospective study of cause-specific mortality. Eur J Trauma Emerg Surg. 2023;49(3):1467–1475. doi:10.1007/s00068-022-02204-6. PMID: 36571633.

2. Dyer SM, Crotty M, Fairhall N, Magaziner J, Beaupre LA, Cameron ID, et al. A critical review of the long-term disability outcomes following hip fracture. BMC Geriatr. 2016;16(1):158. doi:10.1186/s12877-016-0332-0. PMID: 27590604.

3. Johnell O, Kanis JA. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. Osteoporos Int. 2006;17(12):1726–1733. doi:10.1007/s00198-006-0172-4. PMID: 16983459.

4. Lin HS, Watts JN, Peel NM, Hubbard RE. Frailty and post-operative outcomes in older surgical patients: a systematic review. BMC Geriatr. 2016;16(1):157. doi:10.1186/s12877-016-0329-8. PMID: 27580947.

608 5. Van Heghe A, Mordant G, Dupont J, Dejaeger M, Laurent MR, Gielen E. Effects of orthogeriatric
609 care models on outcomes of hip fracture patients: a systematic review and meta-analysis. Calcif Tissue
610 Int. 2022;110(2):162–184. doi:10.1007/s00223-021-00913-5. PMID: 34591127.
611 6. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic
612 comorbidity in longitudinal studies: development and validation. J Chronic Dis. 1987;40(5):373–383.
613 doi:10.1016/0021-9681(87)90171-8. PMID: 3558716.
614 7. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet.
615 2013;381(9868):752–762. doi:10.1016/S0140-6736(12)62167-9. PMID: 23395245.
616 8. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older
617 adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci. 2001;56(3):M146–M156.
618 doi:10.1093/gerona/56.3.M146. PMID: 11253156.
619 9. Fried LP, Ferrucci L, Darer J, Williamson JD, Anderson G. Untangling the concepts of disability,
620 frailty, and comorbidity: implications for improved targeting and care. J Gerontol A Biol Sci Med Sci.
621 2004;59(3):255–263. doi:10.1093/gerona/59.3.M255. PMID: 15031310.
622 10. Gilbert T, Neuburger J, Kraindler J, Keeble E, Smith P, Ariti C, et al. Development and validation
623 of a Hospital Frailty Risk Score focusing on older people in acute care settings using electronic
624 hospital records: an observational study. Lancet. 2018;391(10132):1775–1782. doi:10.1016/S0140-
625 6736(18)30668-8. PMID: 29706364.
626 11. Shimizu A, Maeda K, Fujishima I, Kayashita J, Mori N, Okada K, et al. Hospital Frailty Risk
627 Score predicts adverse events in older patients with hip fractures after surgery: analysis of a
628 nationwide inpatient database in Japan. Arch Gerontol Geriatr. 2022;98:104552.
629 doi:10.1016/j.archger.2021.104552. PMID: 34653923.
630 12. Wong BLL, Chan YH, O'Neill GK, Murphy D, Merchant RA. Frailty, length of stay and cost in
631 hip fracture patients. Osteoporos Int. 2023;34(1):59–68. doi:10.1007/s00198-022-06553-1. PMID:
632 36197493.

13. Song Y, Skinner J, Bynum J, Sutherland J, Wennberg JE, Fisher ES. Regional variations in diagnostic practices. N Engl J Med. 2010;363(1):45–53. doi:10.1056/NEJMsa0910881. PMID: 20463332.

14. Soong JTY, Kaubryte J, Liew D, Peden CJ, Bottle A, Bell D, et al. Dr Foster global frailty score: an international retrospective observational study developing and validating a risk prediction model for hospitalised older persons from administrative data sets. BMJ Open. 2019;9(6):e026759. doi:10.1136/bmjopen-2018-026759. PMID: 31230009.

15. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM. Measuring diagnoses: ICD code accuracy. Health Serv Res. 2005;40(5 Pt 2):1620–1639. doi:10.1111/j.1475-6773.2005.00444.x. PMID: 16178999.

16. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med. 2015;162(1):55–63. doi:10.7326/M14-0697. PMID: 25560714.

17. Johnson A, Bulgarelli L, Pollard T, Gow B, Moody B, Horng S, et al. MIMIC-IV (version 3.1) [dataset]. PhysioNet; 2024. doi:10.13026/kpb9-mt58.

18. Johnson AEW, Bulgarelli L, Shen L, Gayles A, Shammout A, Horng S, et al. MIMIC-IV, a freely accessible electronic health record dataset. Sci Data. 2023;10(1):1. doi:10.1038/s41597-022-01899-x. PMID: 36596836.

19. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005;43(11):1130–1139. doi:10.1097/01.mlr.0000182534.19832.83. PMID: 16224307.

20. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29–36. doi:10.1148/radiology.143.1.7063747. PMID: 7063747.

21. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837–845. PMID: 3203132.

22. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008;27(2):157–172. doi:10.1002/sim.2929. PMID: 17569110.

23. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. Stat Med. 1997;16(9):965–980. doi:10.1002/(SICI)1097-0258(19970515)16:9<965::AID-SIM509>3.0.CO;2-O. PMID: 9160492.

24. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26(6):565–574. doi:10.1177/0272989X06295361. PMID: 17099194.

25. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Med Care. 1998;36(1):8–27. doi:10.1097/00005650-199801000-00004. PMID: 9431328.

26. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Med Care. 2009;47(6):626–633. doi:10.1097/MLR.0b013e31819432e5. PMID: 19433995.

27. Chin M, Kendzerska T, Inoue J, Aw M, Mardiros L, Pease C, et al. Comparing the Hospital Frailty Risk Score and the Clinical Frailty Scale among older adults with chronic obstructive pulmonary disease exacerbation. JAMA Netw Open. 2023;6(2):e2253692. doi:10.1001/jamanetworkopen.2022.53692. PMID: 36729458.

28. Bruno RR, Wernly B, Flaatten H, Schölzel F, Kelm M, Jung C. The hospital frailty risk score is of limited value in intensive care unit patients. Crit Care. 2019;23(1):239. doi:10.1186/s13054-019-2520-8. PMID: 31266513.

29. Gilbert T, Cordier Q, Polazzi S, Street A, Conroy S, Duclos A. Combining the Hospital Frailty Risk Score with the Charlson and Elixhauser multimorbidity indices to identify older patients at risk of poor outcomes in acute care. Med Care. 2024;62(2):117–124. doi:10.1097/MLR.0000000000001962. PMID: 38079225.

30. Iezzoni LI, Foley SM, Daley J, Hughes J, Fisher ES, Heeren T. Comorbidities, complications, and coding bias: does the number of diagnosis codes matter in predicting in-hospital mortality? JAMA. 1992;267(16):2197–2203. doi:10.1001/jama.267.16.2197. PMID: 1556797.

Figure legends

Figure 1. Cohort selection flowchart. Selection of 1,594 hospitalizations (1,445 patients) aged ≥65 years with hip fracture (ICD-10 S72.0–S72.2) from MIMIC-IV v3.1.

Figure 2. Receiver operating characteristic curves by outcome. Comparison of the HFRS, the Charlson Comorbidity Index, and the Elixhauser index (unweighted and van Walraven-weighted) for 30-, 90-, and 365-day mortality and prolonged hospital stay (>10 days).

Figure 3. Outcome rates across HFRS categories. Observed mortality (30/90/365 days) and prolonged-stay rate across low (<5), intermediate (5–15), and high (>15) HFRS categories.

Figure 4. Forest plot of ΔAUC, HFRS versus comorbidity indices. ΔAUC for the HFRS versus the Charlson Comorbidity Index and versus the unweighted Elixhauser index across all outcomes in the full cohort. Filled markers indicate $p<0.05$; open markers $p\geq0.05$.

Figure 5. Kaplan–Meier curves for 365-day survival by HFRS category (log-rank $p<0.001$).

Figure 6. The HFRS computed before admission. (A) ROC curves for 365-day mortality for the pre-admission HFRS, a Charlson index built from the same prior codes, and the index-admission HFRS (n=558). (B) AUC with 95% CI for four outcomes. (C) Pre-admission versus index HFRS; the dashed line marks identical scores. The pre-admission score was at chance for 30-day mortality (AUC 0.551) and was inferior to the Charlson index at every horizon.

Figure S1. Distribution of the HFRS. Histogram of HFRS values across 1,594 admissions; dashed lines mark the cut-offs of 5 and 15 that separate the low, intermediate and high categories.

Figure S2. Sensitivity forest plot (primary-diagnosis hip fracture, age ≥85 years).

Figure S3. Decile-based calibration plots (observed vs. predicted, 30/90/365-day mortality) from models refitted in the validation cohort (apparent calibration).

Figure S4. Decision-curve analysis: net benefit versus threshold probability for the HFRS, the CCI, and treat-all/treat-none strategies.

Figure S5. Sensitivity forest plot (first admission only, women, men, emergency, elective).

Tables

Table 1. Baseline characteristics of 1,594 older adults hospitalized with hip fracture (MIMIC-IV).

Characteristic	Value
Hospitalizations, n	1,594
Patients, n	1,445
Age, years, median (IQR)	82 (75–88)
Female, n (%)	1,109 (69.6)
Emergency/urgent admission, n (%)	696 (43.7)
Primary diagnosis of hip fracture, n (%)	1,328 (83.3)
HFRS, median (IQR)	8.8 (5.3–13.7)
Low (<5), n (%)	346 (21.7)
Intermediate (5–15), n (%)	926 (58.1)
High (>15), n (%)	322 (20.2)
Charlson Comorbidity Index, median (IQR)	2 (1–4)
Length of stay, days, median (IQR)	4 (3–7)
30-day mortality, n (%)	146 (9.2)
90-day mortality, n (%)	278 (17.4)
365-day mortality, n (%)	488 (30.6)
Prolonged stay (>10 days), n (%)	164 (10.3)
30-day readmission, n (%)	79 (5.0)
Venous thromboembolism, n (%)	24 (1.5)

Values are n (%) unless otherwise stated; continuous variables are median (IQR). HFRS, Hospital Frailty Risk Score; IQR, interquartile range. Prolonged stay was defined as a length of stay >10 days, and readmission as a subsequent hospitalization within 30 days of discharge.

Table 2. Cross-tabulation of HFRS and Charlson Comorbidity Index: outcome rates.

Each cell reports "30-day mortality % / 365-day mortality % / prolonged stay %" with the number of patients in parentheses.

HFRS \ CCI	Low (0–1)	Moderate (2–3)	High (≥4)
Low (<5)	1.2 / 6.5 / 0.4 (n=260)	3.2 / 11.3 / 0.0 (n=62)	4.2 / 20.8 / 0.0 (n=24)
Intermediate (5–15)	4.6 / 21.4 / 4.6 (n=415)	8.5 / 30.0 / 4.8 (n=293)	14.2 / 49.1 / 10.1

			(n=218)
High (>15)	13.5 / 32.4 / 21.6 (n=74)	13.5 / 46.1 / 24.7 (n=89)	27.0 / 69.2 / 44.0 (n=159)

Table 3. Characteristics of high-risk patients missed by comorbidity-only screening.

Characteristic	CCI 0–1, HFRS >15 (n=74)	CCI 0–1, HFRS <5 (n=260)	CCI ≥4, HFRS <5 (n=24)	p (first two groups)
Age, years	84.5 (79.0–88.0)	78.0 (71.0–84.0)	80.0 (73.0–87.2)	<0.001
Female, n (%)	54 (73.0)	191 (73.5)	14 (58.3)	1.00
HFRS	18.2 (16.7–22.1)	3.5 (2.3–4.2)	3.8 (3.7–4.7)	—
Charlson index	1.0 (1.0–1.0)	0.0 (0.0–1.0)	4.5 (4.0–7.2)	—
Length of stay, days	6 (4–9)	3 (3–4)	3 (3–5)	<0.001
Length of stay >10 days, n (%)	16 (21.6)	1 (0.4)	0 (0)	<0.001
Emergency admission, n (%)	26 (35.1)	108 (41.5)	10 (41.7)	—
ICU admission, n (%)	15 (20.3)	9 (3.5)	1 (4.2)	—
30-day mortality, n (%)	10 (13.5)	3 (1.2)	1 (4.2)	<0.001
90-day mortality, n (%)	12 (16.2)	8 (3.1)	3 (12.5)	—
365-day mortality, n (%)	24 (32.4)	17 (6.5)	5 (20.8)	—
30-day readmission, n (%)	4 (5.4)	6 (2.3)	0 (0)	0.24
Non-home discharge, n (%)	64 (92.8)	220 (85.3)	22 (91.7)	0.11
Microbiology tested, n (%)	46 (62.2)	32 (12.3)	2 (8.3)	—
Positive culture, n (%)	20 (27.0)	8 (3.1)	1 (4.2)	<0.001
Haemoglobin, g/dL	10.1 (8.4–11.4)	10.8 (9.1–11.7)	10.0 (8.6–10.8)	0.13
Creatinine, mg/dL	0.8 (0.6–1.0)	0.8 (0.7–1.0)	0.9 (0.8–1.3)	0.78
Platelets, ×10⁹/L	181 (151–236)	181 (140–248)	214 (170–266)	0.68
Sodium, mmol/L	140 (136–142)	139 (137–141)	138 (137–141)	0.68
Blood urea nitrogen, mg/dL	20 (14–29)	18 (13–22)	24 (15–34)	0.014

Continuous variables are median (IQR); categorical variables are n (%) of patients with available data (non-home discharge was available for 69 and 258 patients in the first two groups). p values are from the Mann–Whitney U test for continuous variables and Fisher's exact test for categorical variables, comparing the two low-comorbidity groups. After adjustment for age, sex and admission type, the high-frailty group had an odds ratio of 10.9 (95% CI 2.7–43.5; 13 events) for 30-day mortality and 5.8 (2.8–12.1) for 365-day mortality.

Table 4. Discrimination of the HFRS versus the Charlson Comorbidity Index (DeLong test).

Outcome	Events n (%)	HFRS AUC (95% CI)	CCI AUC (95% CI)	ΔAUC	p
30-day mortality	146 (9.2)	0.726 (0.678–0.775)	0.693 (0.644–0.742)	+0.034	0.19
90-day mortality	278 (17.4)	0.729 (0.694–0.765)	0.709 (0.672–0.745)	+0.021	0.25
365-day mortality	488 (30.6)	0.725 (0.696–0.753)	0.713 (0.684–0.742)	+0.011	0.45
Prolonged stay (>10 days)	164 (10.3)	0.837 (0.799–0.876)	0.716 (0.670–0.762)	+0.122	<0.001
30-day readmission	79 (5.0)	0.615 (0.547–0.682)	0.618 (0.550–0.685)	−0.003	0.94

AUC, area under the receiver operating characteristic curve; CCI, Charlson Comorbidity Index; CI, confidence interval. ΔAUC = HFRS AUC − CCI AUC; p values are from the DeLong test for paired curves. Events n (%) is the number of events contributing to each outcome.

Table 5. Discrimination of the HFRS and three administrative comorbidity indices.

Index	30-day mortality, AUC (95% CI)	90-day mortality, AUC (95% CI)	365-day mortality, AUC (95% CI)	Prolonged stay (>10 d), AUC (95% CI)
Hospital Frailty Risk Score	0.726 (0.678–0.775)	0.729 (0.694–0.765)	0.725 (0.696–0.753)	0.837 (0.799–0.876)
Charlson Comorbidity Index	0.693 (0.644–0.742)	0.709 (0.672–0.745)	0.713 (0.684–0.742)	0.716 (0.670–0.762)
Elixhauser index, unweighted count	0.698 (0.648–0.747)	0.691 (0.654–0.728)	0.700 (0.671–0.729)	0.809 (0.768–0.850)
Elixhauser index, van Walraven weights	0.752 (0.705–0.799)	0.724 (0.688–0.760)	0.728 (0.700–0.757)	0.791 (0.749–0.834)

The Elixhauser index was derived from ICD-10 codes with the Quan algorithm, both as an unweighted count of 30 comorbidity categories and with van Walraven weights. DeLong comparisons: at 30 days, HFRS vs. Charlson p=0.19, HFRS vs. unweighted Elixhauser p=0.16, Charlson vs. unweighted Elixhauser p=0.81; at 90 days, HFRS vs. unweighted Elixhauser p=0.019; at 365 days all p>0.08. For prolonged stay, DeLong comparisons were HFRS vs. CCI p<0.001, HFRS vs. unweighted Elixhauser p=0.12, and HFRS vs. weighted Elixhauser p=0.018; comparisons involving the weighted Elixhauser index and mortality were not formally tested.

Table 6. Incremental value of the HFRS (base model = age + sex + Charlson Comorbidity Index).

Outcome	OR per point (95% CI)	ΔAUC	p	IDI (95% CI)	NRI (95% CI)
30-day mortality	1.10 (1.07–1.13)	+0.041	0.005	+0.029 (0.007–0.052)	+0.218 (0.062–0.374)

90-day mortality	1.10 (1.07–1.13)	+0.029	0.001	+0.042 (0.019–0.065)	+0.136 (0.042–0.230)
365-day mortality	1.10 (1.07–1.12)	+0.026	<0.001	+0.039 (0.020–0.059)	+0.124 (0.054–0.195)
Prolonged stay (>10 days)	1.22 (1.18–1.26)	+0.118	<0.001	+0.148 (0.103–0.192)	+0.340 (0.239–0.442)
30-day readmission	1.07 (1.03–1.11)	+0.001	0.96	+0.011 (−0.005–0.027)	+0.153 (0.013–0.294)

IDI and NRI p-values: all IDI p<0.05 except readmission (p=0.17); all NRI p<0.05 except readmission (p=0.032, borderline).

Table 7. Reverse incremental value, Cox survival, and sensitivity analyses.

Reverse incremental value (base = age + sex + HFRS; adding Charlson)

Outcome	CCI OR per point	ΔAUC	p
30-day mortality	1.17	+0.028	0.003
90-day mortality	1.22	+0.031	<0.001
365-day mortality	1.31	+0.036	<0.001
Prolonged stay (>10 days)	1.14	+0.004	0.30

Cox proportional-hazards (365-day mortality, Breslow)

Model	Variable	HR (95% CI)	p
Unadjusted	HFRS per point	1.10 (1.09–1.12)	<0.001
Adjusted (age, sex, CCI)	HFRS per point	1.07 (1.06–1.09)	<0.001
Adjusted (age, sex, CCI)	Charlson per point	1.17 (1.13–1.21)	<0.001
Adjusted (age, sex, CCI)	Intermediate vs. low	2.98 (2.03–4.39)	<0.001
Adjusted (age, sex, CCI)	High vs. low	4.99 (3.33–7.50)	<0.001

Sensitivity analyses (HFRS AUC vs. CCI AUC for prolonged stay)

Subgroup (n)	Prolonged stay
Primary diagnosis hip fracture (1328)	0.862 vs. 0.721
Age ≥85 years (616)	0.798 vs. 0.634
First admission only (1445)	0.850 vs. 0.722
Women (1109)	0.836 vs. 0.722
Men (485)	0.840 vs. 0.684
Emergency admission (696)	0.857 vs. 0.718
Elective admission (898)	0.819 vs. 0.710

OR, odds ratio; HR, hazard ratio; CI, confidence interval; CCI, Charlson Comorbidity Index; AUC, area under the receiver operating characteristic curve. In the reverse analysis, ΔAUC is the gain from adding the CCI to a base model of age, sex and HFRS. Cox models are for 365-day mortality. Sensitivity analyses report the AUC for prolonged stay of the HFRS versus the CCI in each subgroup.

Table 8. Incremental value of the HFRS beyond age, sex and routine admission laboratory values.

Metric	30-day mortality	90-day mortality	365-day mortality
AUC, age and sex	0.596	0.623	0.634
AUC, age, sex and laboratory values	0.679	0.686	0.689
AUC, plus HFRS	0.756	0.750	0.753
ΔAUC (p)	+0.078 (0.0006)	+0.064 (<0.001)	+0.064 (<0.001)
OR per HFRS point (95% CI)	1.11 (1.07–1.14)	1.11 (1.08–1.14)	1.12 (1.09–1.15)
Likelihood-ratio χ^2 (p)	37.8 (<0.001)	60.5 (<0.001)	92.2 (<0.001)
IDI (95% CI)	0.034 (0.013–0.065)	0.055 (0.030–0.089)	0.076 (0.047–0.107)
NRI (95% CI)	0.328 (0.178–0.584)	0.357 (0.199–0.542)	0.360 (0.208–0.466)

Laboratory values were haemoglobin (minimum), serum creatinine (maximum), platelet count (minimum) and serum sodium (maximum) within the first 24 hours of admission. Complete-case analysis (n=1,113; 146 deaths at 30 days). Confidence intervals for IDI and NRI were obtained from 500 bootstrap resamples; p values for ΔAUC are from the DeLong test.

Table 9. Composition of the Hospital Frailty Risk Score in 1,594 hip fracture admissions.

Panel A. Points contributed by code class.

Code class	Codes occurring, n	Mean points per admission	Share of all points, %	Admissions with ≥1 code, n (%)
Acute event, injury or complication	51	6.56	65.6	1,594 (100)
Chronic or pre-existing condition	32	2.69	26.9	1,332 (83.6)
Geriatric syndrome, function or care status	17	0.75	7.5	630 (39.5)

Panel B. Discrimination of each component alone (AUC).

Outcome	Total HFRS	Acute	Chronic	Chronic + geriatric	Diagnosis count
30-day mortality	0.726	0.726	0.606	0.622	0.748
365-day mortality	0.725	0.699	0.653	0.668	0.727
Prolonged stay (>10 days)	0.837	0.848	0.611	0.668	0.881

Panel C. Sensitivity of the classification.

Scheme	Chronic share of points, %	Acute share of points, %	HFRS ≥5 but non-acute <5, %	Admissions with non-acute score >15, n
Primary (19 ambiguous codes as assigned)	26.9	65.6	66.0	0
All 19 ambiguous codes counted as chronic	40.1	55.8	52.4	13

All 19 ambiguous codes excluded	30.8	64.5	73.2	0

Non-acute refers to the sum of chronic and geriatric points. The largest single acute contribution was the index fracture code (S72), present in every admission and worth 1.4 points by definition of the cohort. The maximum non-acute score observed in the primary classification was 14.3 points, 0.7 points below the >15 threshold. The 19 ambiguous codes were N39, L97, K26, M80, R69, K59, R13, R29, Z50, U80, B95, B96, X59, Y95, R31, R33, R94, R02 and G45; their individual assignments and the number of admissions carrying each code are listed in the deposited analysis files. Code counts in Panel A refer to the 100 codes that occurred in this cohort; across all 109 codes, 56 are acute, 34 chronic and 19 geriatric or care status (Additional file 2: Table S2). Excluding the fracture code itself, which every admission carries by definition, acute codes still contributed 60.0% of the remaining points, and that component reproduced the same discrimination (30-day AUC 0.726).

Table 10. The Hospital Frailty Risk Score computed before the index admission.

Panel A. Cohort and the two scores

Characteristic	**Value**
Index admissions with prior ICD−10 diagnoses in the window	558
Unique patients	495
Prior admissions per patient, median	1
Days from the last prior admission to the fracture, median	108
Pre-admission HFRS, median (IQR)	11.0 (6.3–17.9)
Index-admission HFRS, median (IQR)	10.0 (6.5–14.7)
Admissions in which the pre-admission score exceeded the index score, n (%)	323 (57.9)
Spearman ρ between the two scores	0.28
Share of index-score points already coded before admission, median (IQR)	34.5% (17.8–51.1)
Index HFRS >15 already >15 before admission, n (%)	60/131 (45.8)
Pre-admission CCI, median	3

Panel B. Discrimination (AUC, 95% CI)

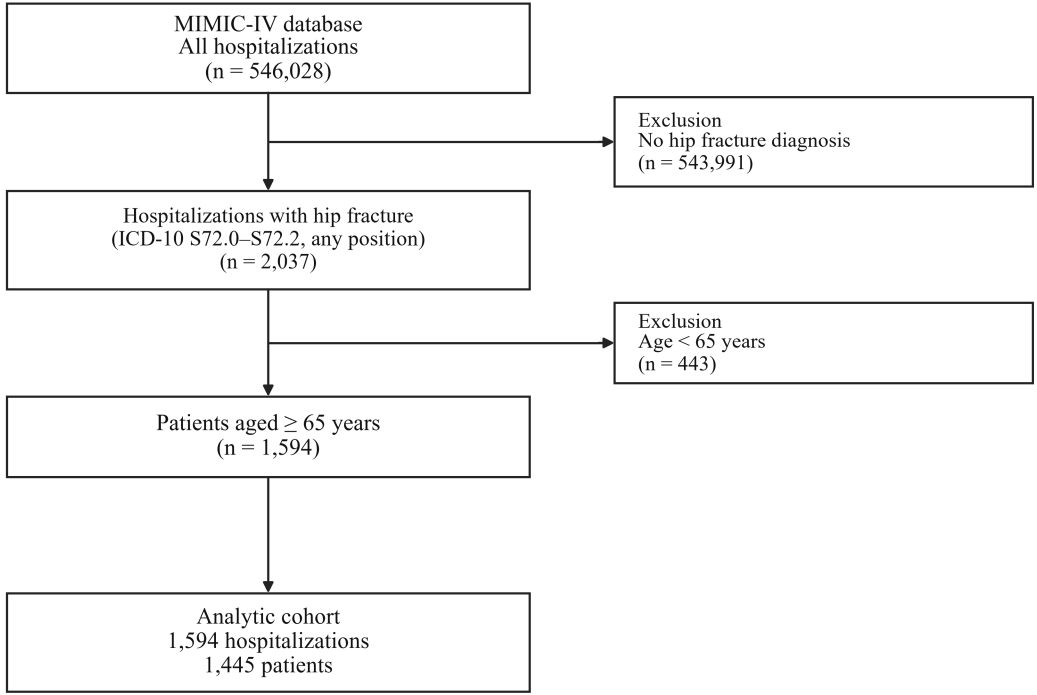
Outcome	**Events**	**Pre-admission HFRS**	**Pre-admission CCI**	**Index HFRS**	**Δ vs. pre-CCI**	**p**
30-day mortality	58	0.551 (0.471–0.631)	0.649 (0.569–0.729)	0.674 (0.595–0.754)	−0.098	0.018
90-day	106	0.598	0.674	0.710	−0.075	0.015

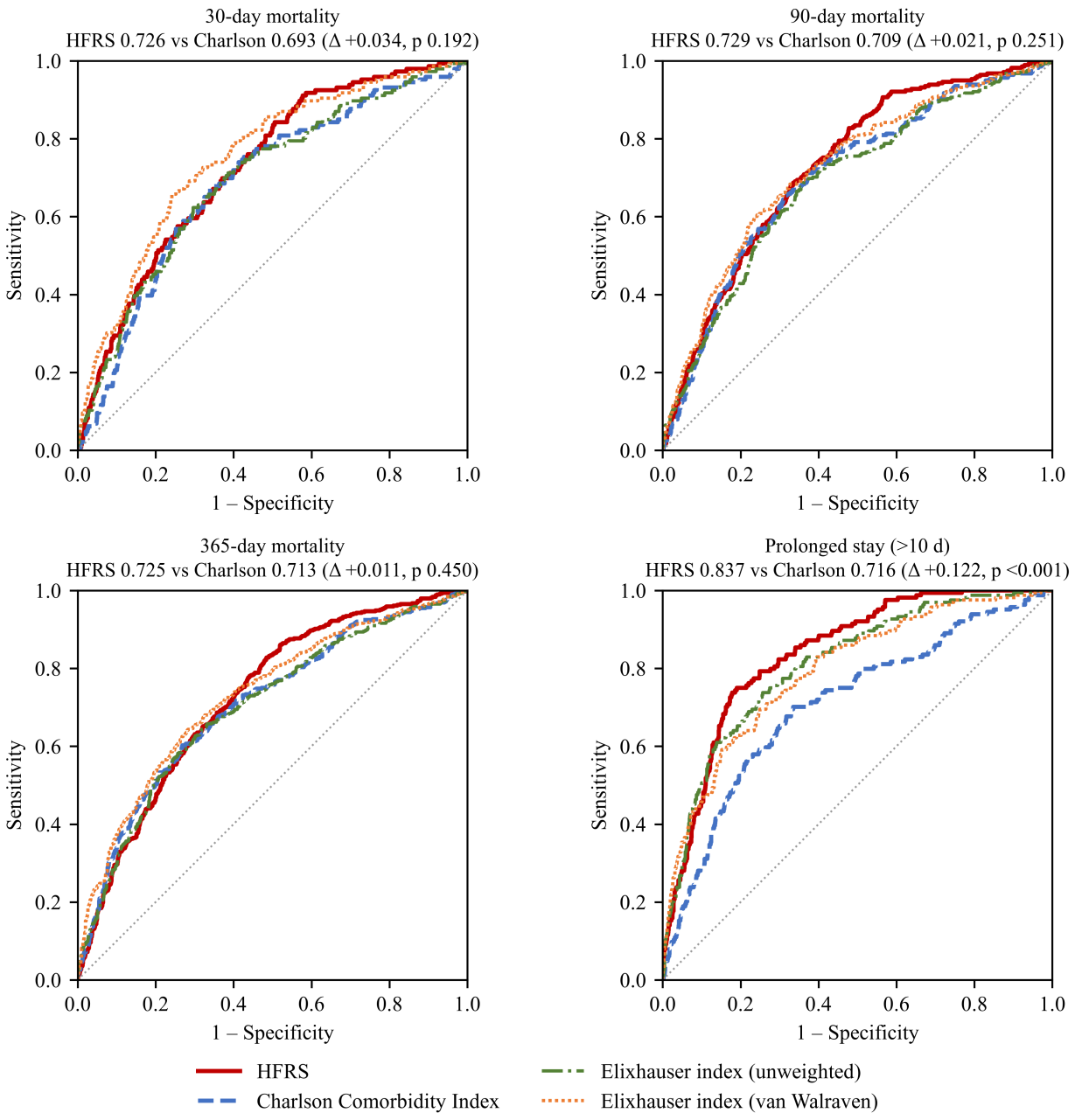
mortality		(0.536–0.660)	(0.613–0.734)	(0.651–0.769)		
365-day mortality	204	0.594 (0.544–0.643)	0.683 (0.636–0.731)	0.686 (0.639–0.733)	−0.090	0.001
Prolonged stay (>10 d)	60	0.622 (0.543–0.702)	0.652 (0.573–0.731)	0.789 (0.718–0.859)	−0.029	0.487

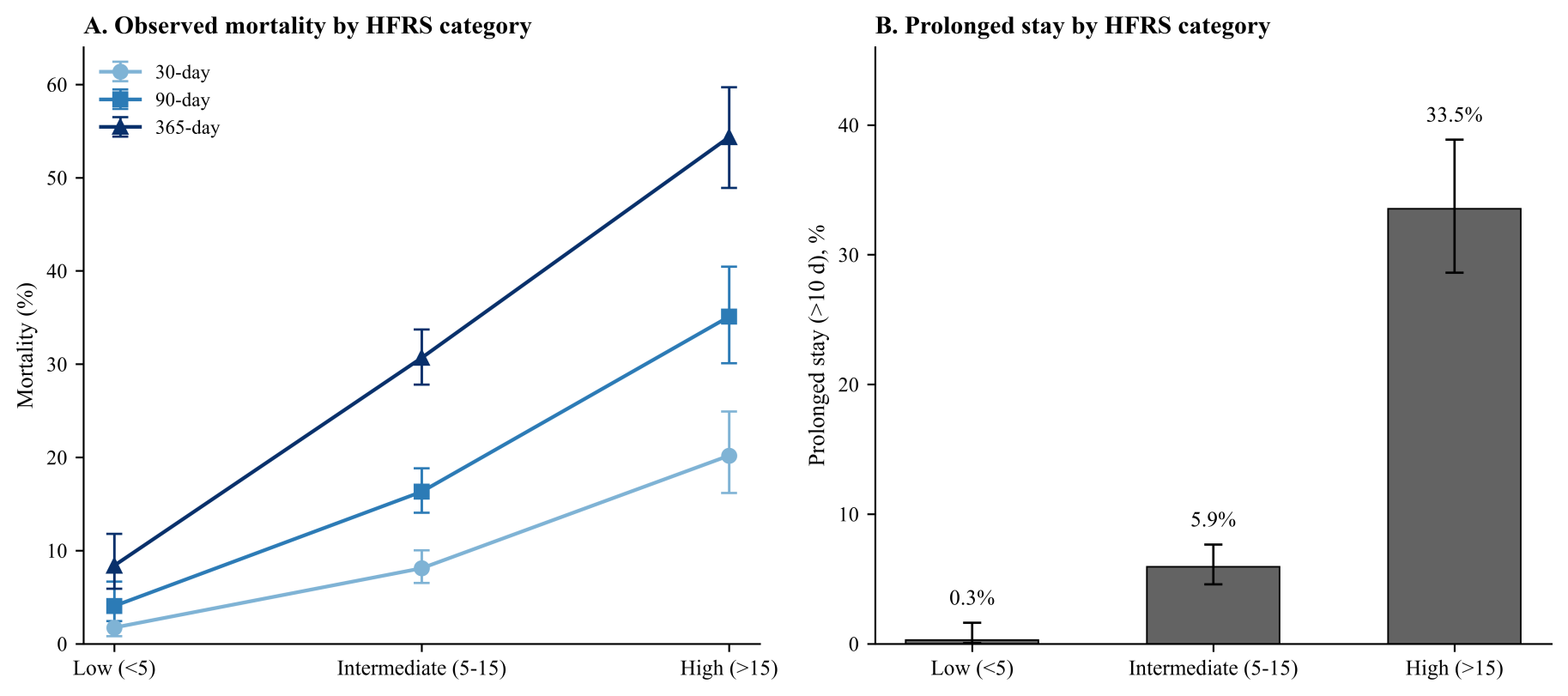
Panel C. Incremental value of the pre-admission HFRS over age, sex and the pre-admission CCI

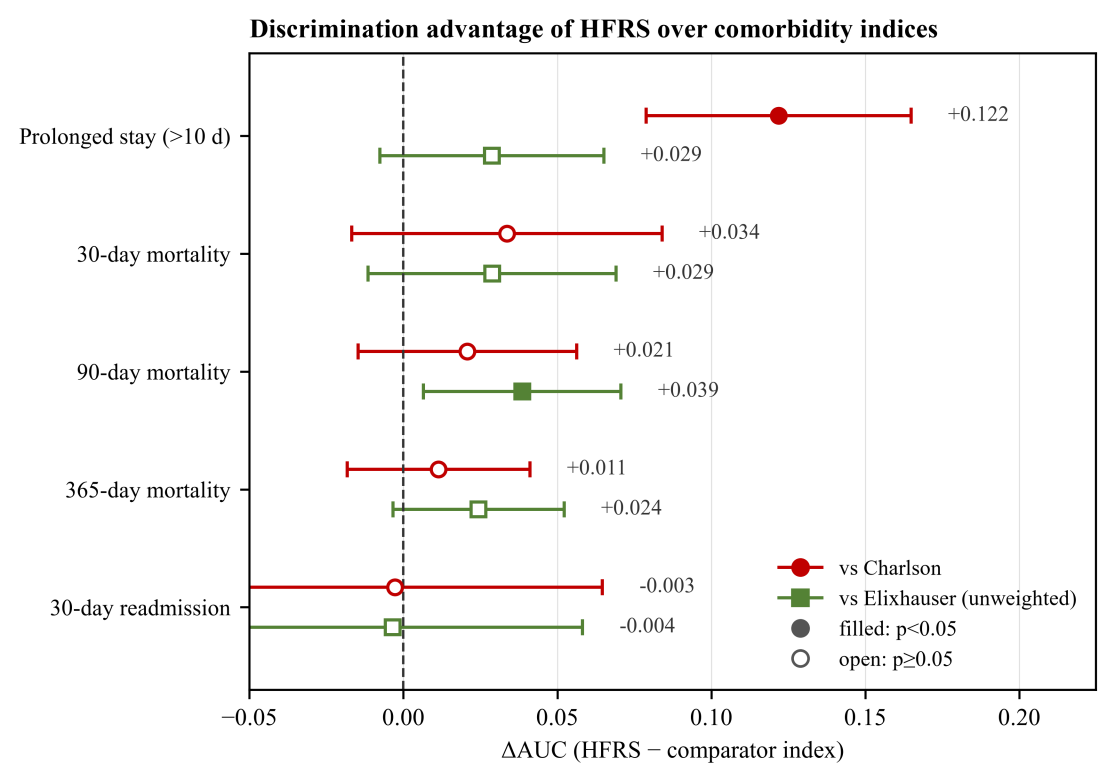
Outcome	**OR per point (95% CI)**	**AUC base → full**	**ΔAUC**	**LR p**	**IDI (95% CI)**	**NRI (95% CI)**
30-day mortality	1.01 (0.98–1.04)	0.686 → 0.684	−0.003	0.540	+0.002 (−0.012 to +0.016)	−0.011 (−0.122 to +0.099)
90-day mortality	1.03 (1.01–1.06)	0.719 → 0.718	−0.001	0.007	+0.020 (−0.012 to +0.053)	+0.024 (−0.080 to +0.129)
365-day mortality	1.01 (0.99–1.04)	0.717 → 0.714	−0.002	0.190	+0.002 (−0.009 to +0.013)	−0.014 (−0.075 to +0.047)
Prolonged stay (>10 d)	1.01 (0.98–1.04)	0.665 → 0.677	+0.011	0.358	+0.001 (−0.007 to +0.010)	−0.011 (−0.125 to +0.102)

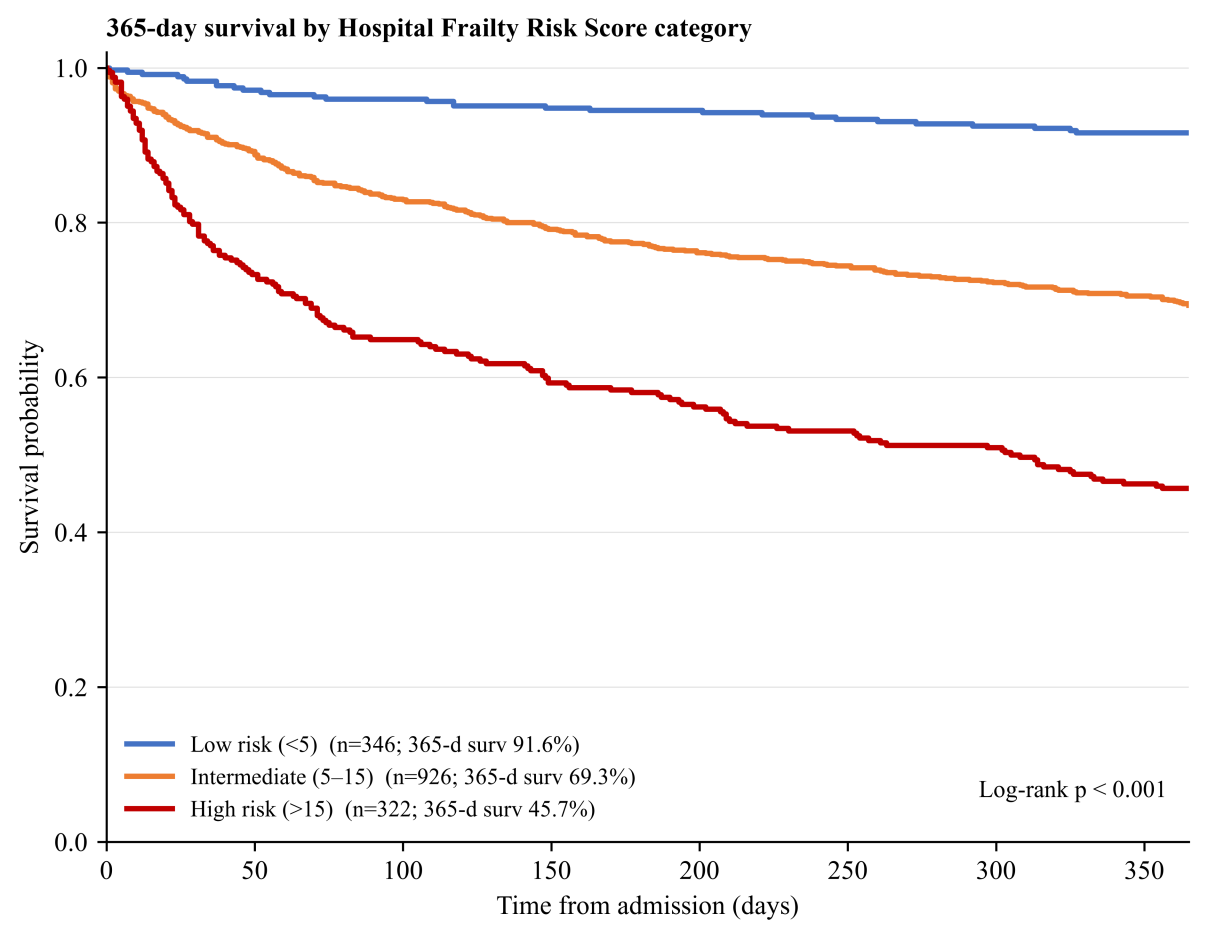
Pre-admission HFRS was computed from the unique three-character ICD−10 codes of hospitalizations in the 730 days before the index admission, using the original 109-code weights. IDI and NRI 95% CIs come from 500 clustered bootstrap resamples of patients. AUC 95% CIs use the Hanley–McNeil method.

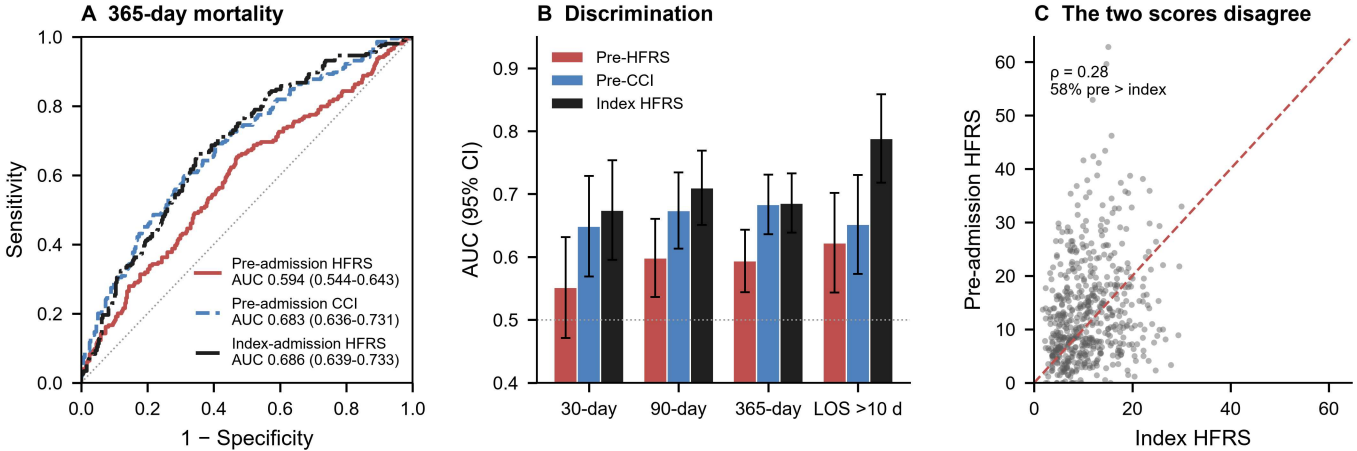

MIMIC-IV database
All hospitalizations
(n = 546,028)
Exclusion
No hip fracture diagnosis
(n = 543,991)
Hospitalizations with hip fracture
(ICD-10 S72.0–S72.2, any position)
(n = 2,037)
Exclusion
Age < 65 years
(n = 443)
Patients aged ≥ 65 years
(n = 1,594)
Analytic cohort
1,594 hospitalizations
1,445 patients


30-day mortality
HFRS 0.726 vs Charlson 0.693 (Δ +0.034, p 0.192)
90-day mortality
HFRS 0.729 vs Charlson 0.709 (Δ +0.021, p 0.251)
365-day mortality
HFRS 0.725 vs Charlson 0.713 (Δ +0.011, p 0.450)
Prolonged stay (>10 d)
HFRS 0.837 vs Charlson 0.716 (Δ +0.122, p <0.001)
Sensitivity
1 – Specificity
0.0
0.2
0.4
0.6
0.8
1.0
HFRS
Charlson Comorbidity Index
Elixhauser index (unweighted)
Elixhauser index (van Walraven)


A. Observed mortality by HFRS category
30-day
90-day
365-day
Mortality (%)
0
10
20
30
40
50
60
Low (<5)
Intermediate (5-15)
High (>15)
B. Prolonged stay by HFRS category
Prolonged stay (>10 d), %
0
10
20
30
40
0.3%
5.9%
33.5%
Low (<5)
Intermediate (5-15)
High (>15)


Discrimination advantage of HFRS over comorbidity indices
Prolonged stay (>10 d)
30-day mortality
90-day mortality
365-day mortality
30-day readmission
+0.122
+0.029
+0.034
+0.029
+0.021
+0.039
+0.011
+0.024
-0.003
-0.004
vs Charlson
vs Elixhauser (unweighted)
filled: p<0.05
open: p≥0.05
−0.05
0.00
0.05
0.10
0.15
0.20
ΔAUC (HFRS − comparator index)


365-day survival by Hospital Frailty Risk Score category
Survival probability
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
250
300
350
Time from admission (days)
Low risk (<5) (n=346; 365-d surv 91.6%)
Intermediate (5–15) (n=926; 365-d surv 69.3%)
High risk (>15) (n=322; 365-d surv 45.7%)
Log-rank p < 0.001


A 365-day mortality
Sensitivity
1 − Specificity
Pre-admission HFRS
AUC 0.594 (0.544-0.643)
Pre-admission CCI
AUC 0.683 (0.636-0.731)
Index-admission HFRS
AUC 0.686 (0.639-0.733)
B Discrimination
AUC (95% CI)
Pre-HFRS
Pre-CCI
Index HFRS
30-day
90-day
365-day
LOS >10 d
C The two scores disagree
ρ = 0.28
58% pre > index
Pre-admission HFRS
Index HFRS

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1TableS1.docx
- Additionalfile8TablesS3S6.docx
- Additionalfile3FigureS1.png
- Additionalfile6FigureS4.png
- Additionalfile5FigureS3.png
- Additionalfile4FigureS2.png
- Additionalfile7FigureS5.png
- Additionalfile2TableS2.docx