ESC European Society of Cardiology

European Journal of Heart Failure (2025) **27**, 726–736
doi:10.1002/ejhf.3585

RESEARCH ARTICLE

Predicting survival in patients with severe heart failure: Risk score validation in the HELP-HF cohort

Mauro Chiarito[1,2*†], Davide Stolfo[3†], Alessandro Villaschi[1,2], Samantha Sartori[4], Luca Baldetti[5], Carlo Mario Lombardi[6], Marianna Adamo[6], Ferdinando Loiacono[1,2], Antonio Maria Sammartino[6], Mauro Riccardi[6], Daniela Tomasoni[6], Riccardo Maria Inciardi[6], Marta Maccallini[1,2], Gaia Gasparini[1,2], Benedetta Grossi[2,7], Stefano Contessi[3], Daniele Cocianni[3], Maria Perotto[3], Giuseppe Barone[5], Marco Merlo[3], Alberto Maria Cappelletti[8], Gianfranco Sinagra[3], Daniela Pini[1], Marco Metra[6], and Matteo Pagnesi[6]

[1]Humanitas Research Hospital IRCCS, Rozzano, Milan, Italy; [2]Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, Milan, Italy; [3]Cardiovascular Department, Azienda Sanitaria Universitaria Giuliano Isontina (ASUGI), University of Trieste, Trieste, Italy; [4]Center for Interventional Cardiovascular Research and Clinical Trials, Icahn School Medicine at Mount Sinai, New York, NY, USA; [5]Cardiac Intensive Care Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy; [6]Institute of Cardiology, ASST Spedali Civili, Department of Medical and Surgical Specialties, Radiological Sciences and Public Health, University of Brescia, Brescia, Italy; [7]Department of Chemistry, Materials and Chemical Engineering, Politecnico di Milano, Milan, Italy; and [8]IRCCS San Raffaele Scientific Institute, Milan, Italy

Received 3 September 2024; revised 10 December 2024; accepted 30 December 2024; online publish-ahead-of-print 20 January 2025

Aims Accurate selection of patients with severe heart failure (HF) who might benefit from advanced therapies is crucial. The present study investigates the performance of the available risk scores aimed at predicting the risk of mortality in patients with severe HF.

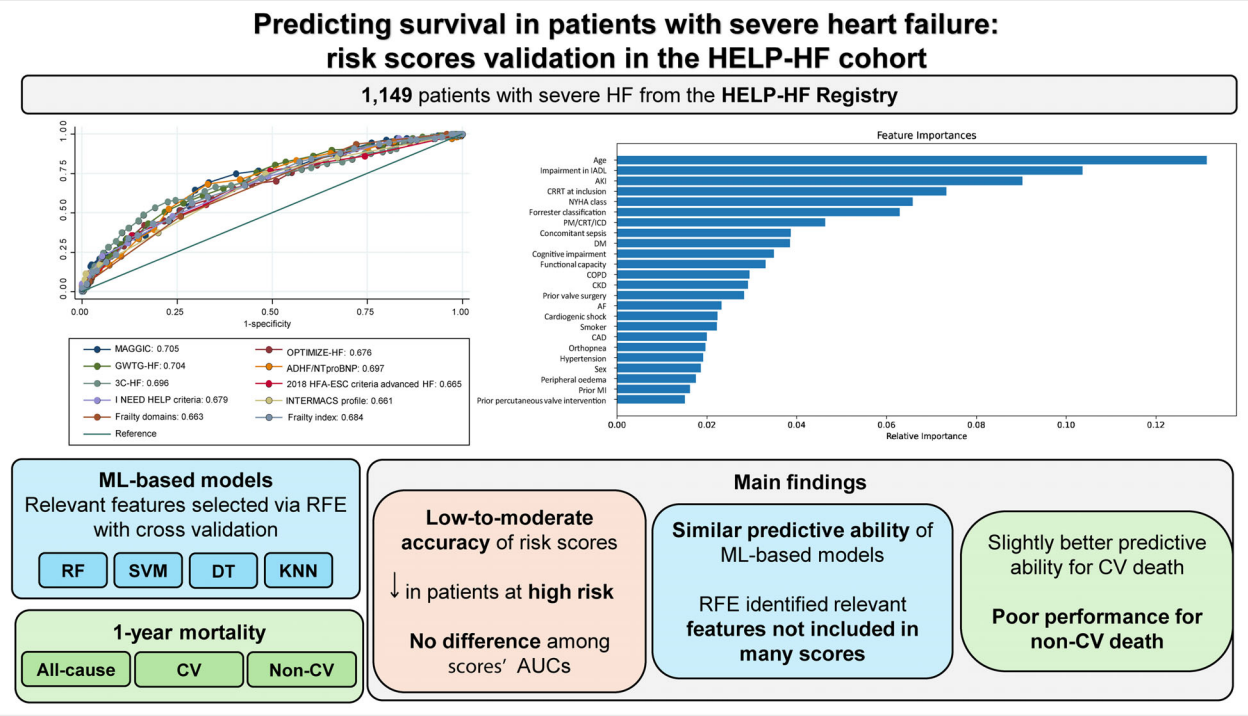
Methods and results The risk of 1-year mortality was estimated in patients with severe HF enrolled in the HELP-HF cohort according to the MAGGIC, 3-CHF, ADHF/NT-proBNP, and GWTG-HF risk scores, the number of criteria of the 2018 HFA-ESC definition of advanced HF, I NEED HELP markers, domains fulfilled of the 2019 HFA-ESC definition of frailty, the frailty index, and the INTERMACS profile. In addition, we tested the performance of different machine learning (ML)-based models to predict 1-year mortality. At 1-year follow-up, 265 patients (23.1%) died. The prognostic accuracy, tested in the subgroup of patients with completeness of all data regarding the variables included in the scores (497/1149 patients), resulted moderate for MAGGIC, GWTG-HF, and ADHF/NT-proBNP scores (area under the curve [AUC] $\geq$0.70) and only poor for the other tools. All the scores lost accuracy in estimating the rate of 1-year mortality in patients at the highest risk. Support vector machine-based model had the best AUC among ML-based models, slightly outperforming most of the tested risk scores.

Conclusion Most of the scores used to predict the risk of mortality in HF performed poorly in real-world patients with severe HF and provided inaccurate estimate of the risk of 1-year mortality in patients at the highest risk. ML-based models did not significantly outperform the currently available risk scores and their use must be validated in large cohort of patients.

*Corresponding author. Department of Biomedical Sciences, Humanitas University, Pieve Emanuele-Milan; Cardio Center, Humanitas Research Hospital IRCCS, Via Manzoni 56, 20089 Rozzano-Milan, Italy. Tel: +39 02 82247009, Email: mauro.chiarito@hunimed.eu

†Contributed equally as first co-authors.

Graphical Abstract



The predictive ability of different risk scores developed and validated to predict the risk of mortality in patients with heart failure in the HELP-HF cohort is reported and compared based on their areas under the curve. In addition to regression-based risk models, the predictive ability of different machine learning-based models was evaluated in the same cohort. Recursive feature elimination was employed to identify the most relevant features in predicting the risk of mortality. AF, atrial fibrillation; AKI, acute kidney injury; AUC, area under the curve; CAD, coronary artery disease; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; CRRT, continuous renal replacement therapy; CRT, cardiac resynchronization therapy; CV, cardiovascular; DM, diabetes mellitus; DT, decision tree; HF, heart failure; KNN, k-nearest neighbor; IADL, instrumental activity of daily living; ICD, implantable cardioverter defibrillator; MI, myocardial infarction; ML, machine learning; NYHA, New York Heart Association; PM, pacemaker; RF, random forest; RFE, recursive feature elimination; SVM, support vector machine.

Keywords Heart failure • Advanced heart failure • HELP-HF • Risk scores

Introduction

Heart failure (HF) represents one of the major causes of mortality and morbidity worldwide. Different clinical scenarios are included under the diagnosis of HF, varying in terms of concomitant comorbidities and HF severity.[1] The growing prevalence of HF – affecting more than 64.3 million people worldwide – and the improved prognosis have resulted in a relevant increase in the global prevalence of severe and advanced HF.[2] This epidemiologic pattern implies important socio-economic issues: severe HF is characterized by progressive worsening of symptoms, limited quality of life, high risk of mortality, and might require a timely referral for advanced therapies, such as left ventricular assist device (LVAD) and heart transplantation.[3] In contrast, the risk of futility of invasive and expensive therapies must be considered in patients with severe HF, as they commonly present end-stage comorbidities and poor functional status.[4] Therefore, accurate selection of patients who might truly benefit from advanced therapies is crucial. Several risk scores aimed at predicting the risk of mortality in HF patients have been developed and validated in the last decades.[5–7] Nonetheless, the performance of such scores among patients with severe HF has been poorly investigated. The present study aims to evaluate the performance of different risk scores in the prediction of 1-year mortality in a cohort of patients with severe HF. In addition, we worked on the identification of specific features associated with the occurrence of mortality using different machine learning (ML)-based models, in order to improve the selection of candidates for advanced therapies and to avoid futile treatments.

Methods

Study design and study population

The present study evaluated patients enrolled in the Assessment of the I NEED HELP markers in Heart Failure (HELP-HF) registry. The design of the HELP-HF registry has been previously described.[8] Briefly, it is an observational, retrospective, multicentre registry including consecutive HF patients hospitalized for acute HF or evaluated as outpatients for chronic HF at four Italian high-volume centres between 1st January 2020 and 20th November 2021. Included patients presented at least one 'I NEED HELP' high-risk marker.[9] Institutional review board approval was waived for this registry because of its retrospective design with collection of anonymized data and without study-specific intervention.

Data collection and score selection

De-identified individual patient data on medical history, clinical presentation, echocardiography and laboratory findings, medical therapy and clinical outcomes were collected. Follow-up was performed by medical records or telephone contact. Score selection was based on the presence in the HELP-HF registry of the variables included in the main risk scores available in the literature. The MAGGIC, OPTIMIZE-HF, GWTG-HF, ADHF/NT-proBNP, and 3C-HF scores were calculated.[6,7,10–12] We excluded the BCN Bio-HF,[13] PREDICT-HF,[14] CHARM,[15] Seattle Heart Failure Model (SHFM),[16] HFSS,[17] MECKI,[18] GISSI-HF,[19] and Krakow scores,[20] due to lack of availability in the case report form of the HELP-HF registry of one or more variables needed to calculate the scores. We also evaluated the predictive ability of the 2018 HFA-ESC definition of advanced HF (based on the number of criteria fulfilled),[21] the number of I NEED HELP markers,[9] the INTERMACS profile,[22] the frailty index (FI), and the number of domains fulfilled of the 2019 HFA-ESC definition of frailty.[23] We employed a standard procedure to construct the FI, using the deficit accumulation approach (online supplementary *Table S1*).[24] Patients were also divided into two groups (FI <0.21 and ≥0.21), as previously described.[25]

Statistical analysis and machine learning-based models building workflow

Continuous variables are presented as mean ± standard deviation or median (interquartile range [IQR]). Categorical variables are presented as numbers and percentages. Follow-up was evaluated at the date of death or last available follow-up and censored at 1 year. Logistic and Cox proportional hazards regression analyses were performed to assess the association of the different scores with all-cause mortality. Results of the Cox regression analyses are reported as hazard ratio (HR) and 95% confidence interval (CI). The predictive ability of the scores was assessed by plotting receiver operating characteristic (ROC) curves from logistic regression analysis and compared via a non-parametric approach. This analysis was performed considering logistic regression models as advanced therapies for HF must be considered only in patients with life expectancy of at least 1 year with good quality of life, not considering the timing of events within 1 year,[1] making appropriate the use of logistic regression in this setting. This comparison was performed in the subgroup of patients with available data for all the variables included in the scores. In addition, each area under the curve (AUC) was separately compared with the best AUC to highlight significant differences. A sensitivity analysis including median values in case of missing information was performed. A further sensitivity analysis using the Harrell's C-index to compare the performance of the scores in predicting the risk of death considering the timing of the event was performed. The performance of the scores was assessed in the subgroup of patients enrolled during hospitalization and in those with reduced or mildly reduced ejection fraction. Observed versus model-predicted 1-year all-cause mortality was compared across quantiles of death probability estimated by each score. The calibration of the scores was tested with the Hosmer–Lemeshow goodness of fit test. In light of its limited statistical power, calibration belts were also constructed to test the relationship between estimated probabilities and observed outcome rates for each score. For the ML-based models, the number of input data was determined according to the 10 times rule: the amount of input data was ≥10 times the number of degrees of freedom (i.e. selected relevant features) of the model. Missing continuous variables were imputed using either mean or median, depending on the skewness of their distribution; in case of dichotomous missing variables, absence of data was interpreted as absence of clinical feature. Variables with more than 10% of missing data were not included. Outliers were identified using the IQR method (defined as more than 1.5 IQR below Q1 or more than 1.5 IQR above Q3) and detected outliers were trimmed and treated as missing data.[26] Relevant features selection for ML-based algorithms was performed testing recursive feature elimination with cross-validation based on random forest (RF) algorithm.[27] A single downsampling of the majority class (i.e. patients with negative outcome) was performed to obtain a balanced subpopulation (50% of patients with a negative and 50% with a positive outcome, totaling 530 patients). The ML-based model training and validation were performed with RF, support vector machine (SVM), decision tree (DT), and k-nearest neighbors (KNN). Training and validation were performed in the subgroup of patients without available data for all the variables included in the scores (training set, $n = 301$); testing was performed on the subgroup with available data for all the variables included in the scores (testing set, $n = 229$) (online supplementary *Figure Appendix S1*), to allow a reliable comparison of the predictive ability with the logistic regression-based risk scores. Importantly, there is no overlap between the training/validation and testing cohorts, allowing for an unbiased evaluation of the model's predictive capabilities. A sensitivity analysis to assess the performance of ML-based models using the entire database was performed, with a randomized split into training and testing sets, ensuring no overlap between the two cohorts. During the training phase, hyperparameter optimization was performed. Selected hyperparameters are reported in online supplementary *Table S2*. Two-sided *p*-values <0.05 were considered statistically significant. Statistical analyses were performed using Stata (version 17, Stata Corp, College Station, TX, USA), MatLab (MathWorks Inc, Natick, MA, USA) and Python (version 3.11.6, Python Software Foundation, Wilmington, DE, USA).

Results

Baseline patients features and clinical presentation

Among 4753 patients with HF screened between January 2020 and November 2021, 1149 patients (24.3%) had at least one I NEED HELP high-risk marker. The mean age of the population was 75.1 ± 11.5 years and 67.3% of patients were male. At the time of enrolment, 777 patients (67.6%) were hospitalized and 372 (32.4%) were outpatients. Patients with de novo HF represented 16.3%. Median left ventricular ejection fraction was 35% (25–50%)

Table 1 Baseline characteristics

	Overall ($n = 1149$)	Survived ($n = 906$)	Dead ($n = 243$)	p-value
Age (years)	75.1 ± 11.5	77 (68–83)	80 (73–85)	<0.001
Male sex	773 (67.3)	613 (67.7)	160 (65.8)	0.592
BMI (kg/m^2)	25.7 (22.5–28.4)	26.00 (23.13–29.63)	24.44 (22.20–28.37)	0.002
New-onset HF	242 (21.1)	154 (17.0)	33 (13.6)	0.20
Time since HF diagnosis (months)	24 (1–84)	30.00 (3.00–84.00)	30.00 (5.00–96.00)	0.24
HF hospitalization(s) during last year	415 (36.1)	315 (34.8)	100 (41.2)	0.066
Type of inclusion				<0.001
Outpatient visit	372 (32.4)	332 (36.6)	40 (16.5)	
Inpatient hospitalization	777 (67.6)	574 (63.4)	203 (83.5)	
Comorbidities				
Hypertension	817 (71.1)	651 (71.9)	166 (68.3)	0.279
Diabetes	447 (38.9)	341 (37.6)	106 (43.6)	0.089
History of AF	641 (55.8)	493 (54.4)	148 (60.9)	0.070
Prior CAD diagnosis	504 (43.9)	389 (42.9)	115 (47.3)	0.221
Prior MI	380 (33.1)	280 (30.9)	100 (41.2)	0.003
Prior PCI	336 (29.2)	260 (28.7)	76 (31.3)	0.433
Prior CABG	171 (14.9)	123 (13.6)	48 (19.8)	0.016
Prior valve surgery	139 (12.1)	106 (21.7)	32 (23.2)	0.677
Prior percutaneous valve intervention				0.202
TAVR	28 (2.4)	40 (4.4)	9 (3.7)	
Mitral TEER	49 (4.3)	18 (2.0)	10 (4.1)	
Known cardiomyopathy	291 (25.3)	248 (27.2)	43 (17.7)	0.002
Peripheral artery disease	205 (17.8)	145 (16.0)	60 (24.7)	0.002
Prior stroke or TIA	173 (15.1)	131 (14.5)	42 (17.3)	0.274
COPD	266 (23.2)	196 (21.6)	70 (28.8)	0.019
Chronic kidney disease	650 (56.6)	483 (53.3)	168 (69.1)	<0.001
MCI or dementia	157 (13.7)	95 (10.5)	62 (25.5)	<0.001
ADL or IADL impairment	339 (31.3)	218 (24.1)	121 (49.8)	<0.001
NYHA functional class				<0.001
I	48 (4.2)	35 (3.9)	13 (5.3)	
II	363 (31.6)	318 (35.1)	45 (18.5)	
III	572 (49.8)	450 (49.7)	122 (50.2)	
IV	166 (14.5)	103 (11.4)	63 (25.9)	
Cardiac implantable electronic devices				0.637
Pacemaker	167 (14.5)	126 (13.9)	41 (16.9)	
ICD	183 (15.9)	145 (16.0)	38 (15.6)	
CRT-D	168 (14.6)	138 (15.2)	30 (12.3)	
CRT-P	15 (1.3)	11 (1.2)	4 (1.6)	

Values are given as mean ± standard deviation, *n* (%), or median (interquartile range).
ADL, activities of daily living; AF, atrial fibrillation; BMI, body mass index; CABG, coronary artery bypass graft; CAD, coronary artery disease; COPD, chronic obstructive pulmonary disease; CRT-D, cardiac resynchronization therapy-defibrillator; CRT-P, cardiac resynchronization therapy-pacemaker; HF, heart failure; IADL, instrumental activity of daily living; ICD, implantable cardioverter defibrillator; MCI, mild cognitive impairment; MI, myocardial infarction; NYHA, New York Heart Association; PCI, percutaneous coronary intervention; TAVR, transcatheter aortic valve repair; TEER, transcatheter edge-to-edge repair; TIA, transient ischaemic attack.

and patients with HF with reduced ejection fraction (HFrEF), mildly reduced ejection fraction, and preserved ejection fraction (HFpEF) were 56.5%, 15.0%, and 28.5%, respectively. Severe mitral regurgitation was reported in 231 patients (20.7%), while 363 (34.4%) and 482 (43.4%) patients had right ventricular dilatation and dysfunction, respectively. Baseline characteristics, data on clinical presentation and echocardiographic and laboratory findings are reported in *Table 1* and online supplementary *Tables S3* and *S4*.

Applicability and distribution of risk scores

The variables included in the risk scores and the missing rate of each variable are reported in *Table 2*. As expected, considering the inclusion criteria and the aim of the HELP-HF registry, data about INTERMACS profile, I NEED HELP markers and criteria included in the 2018 HFA-ESC definition of advanced HF were available

Table 2 Variables included in the MAGGIC, OPTIMIZE-HF, GWTG-HF, and ADHF/NT-proBNP risk scores

Variable	MAGGIC	OPTIMIZE-HF	GWTG-HF	ADHF/NT-proBNP	Missing rate, *n* (%)
Ejection fraction	X	X		X	0 (0)
Age	X	X	X		0 (0)
SBP at inclusion	X	X	X	X	19 (1.7)
BMI	X				44 (3.8)
Creatinine	X	X			29 (2.5)
NYHA class	X				0 (0)
Sex	X				0 (0)
Current smoker	X				31 (2.7)
Diabetes mellitus	X				0 (0)
COPD	X		X	X	0 (0)
First diagnosis of HF within 18 months	X				21 (1.8)
Beta-blocker therapy	X				2 (0.2)
ACEi/ARB/ARNI therapy	X				2 (0.2)
BUN			X		228 (19.8)
Sodium		X	X	X	80 (7)
Heart rate		X	X		25 (2.2)
Black race			X		0 (0)
eGFR				X	29 (2.5)
Moderate-to-severe TR				X	56 (4.9)
Haemoglobin				X	42 (3.7)
NT-proBNP/BNP				X	168 (14.6)
HF as primary cause of admission		X			0 (0)

ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; ARNI, angiotensin receptor–neprilysin inhibitor; BMI, body mass index; BNP, B-type natriuretic peptide; BUN, blood urea nitrogen; COPD, chronic obstructive pulmonary disease; eGFR, estimated glomerular filtration rate; HF, heart failure; NT-proBNP, N-terminal pro-B-type natriuretic peptide; NYHA, New York Heart Association; SBP, systolic blood pressure; TR, tricuspid regurgitation.

for all patients enrolled. Similarly, the formula to calculate the FI allowed to calculate this index in all patients. Based on variables availability, all the risk scores were concomitantly available only in 497 patients. *Table 3* shows the rate of patients with missing information regarding one or more variables for each score and excluded from the discriminative analysis.

Risk score performance

After a median follow-up of 260 days (105–390 days), the composite of death or HF hospitalization occurred in 496 patients (43.2%), 265 patients (23.1%) died, and 308 had a HF hospitalization (26.8%). Cardiovascular (CV) death and non-CV death occurred in 166 (14.5%) and 77 (7.8%) patients, respectively.

The risk of mortality was increased in patients fulfilling two, three, and four criteria of the 2018 HFA-ESC definition of advanced HF, while having only one criterion was not associated with worse prognosis (1-year all-cause mortality rate: 12.5%, 14.6%, 21.5%, 21.4%, and 40.4%; crude HR for 1-criterion increase: 1.43, 95% CI: 1.30–1.57, $p < 0.001$; $p_{\text{log-rank}} < 0.0001$). The risk of mortality was significantly higher in patients with INTERMACS profile 1–5, as compared with those with INTERMACS profile 7. Similarly, the presence of three to nine I NEED HELP criteria was associated with a progressively increased risk of mortality, as well as for increasing MAGGIC, GWTG-HF, OPTIMIZE-HF and ADHF/NT-proBNP scores, number of frailty domains, and FI. The AUC for each score is reported in *Table 3* and the ROC and Kaplan–Meier curves for each score in online supplementary *Figures S2–S11*. The predictive ability of the risk scores was poor for most of the scores, as only the MAGGIC, GWTG-HF and the ADHF/NT-proBNP had an AUC ≥0.70 (0.698 for all the three scores). ROC curves and the AUC comparisons in this subgroup of patients are reported in *Figure 1*, confirming the slightly increased predictive ability of the MAGGIC and GWTG-HF, but without significant differences among the AUCs, except for a lower predictive ability of the number of criteria of the 2018 HFA-ESC definition of advanced HF as compared with the MAGGIC and GWTG-HF scores. The results from pairwise comparison are reported in online supplementary *Table S5*. Consistent findings were observed at sensitivity analysis including median values in case of missing information (online supplementary *Figure S12*). The Harrell's C-index indicated similarly poor performance of the scores when considering the timing of events (online supplementary *Table S6*). Observed and predicted rates of mortality indicated clear escalations with increasing scores for all the risk scores, but the Hosmer–Lemeshow indicated poor fit for the MAGGIC and 3C-HF scores ($p = 0.004$ for MAGGIC, $p = 0.042$ for 3C-HF) (*Figure 2*). The calibration belts showed that, despite an underestimation of the risk for the 3C-HF score for patients at the highest scores, the confidence belts contain the bisector for expected mortality for both the scores. Of note, the prediction of mortality provided by both the models was less accurate for patients with the highest scores (*Figure 3*). Similar loss of precision was observed with all the other scores.

Receiver operating characteristic curves and the AUC comparison for CV and non-CV death are reported in online supplementary *Figures S13* and *S14*. The AUC for CV death

Table 3 Area under the curve and 95% confidence interval and Hosmer–Lemeshow goodness of fit, number and percentage of patients with all available variables included and mean value for each score

Score	AUC	95% CI	*p*-value vs. MAGGIC	*p*-value vs. GWTG-HF	Patients with all available variables, *n* (%)	Mean score (SD)
MAGGIC	0.71	0.65–0.76		0.574	849 (73.9)	31.4 (9.1)
OPTIMIZE-HF	0.68	0.62–0.73	0.712	0.234	1031 (89.7)	35 (7.2)
GWTG-HF	0.71	0.65–0.76	0.574		879 (76.5)	41.4 (8.3)
ADHF/NT-proBNP	0.70	0.64–0.75	0.758	0.834	877 (76.3)	9.4 (4)
3C-HF	0.68	0.64–0.72	0.601	0.629	1138 (99)	25.2 (11.2)
2018 HFA-ESC definition of advanced HF	0.66	0.61–0.72	**<0.001**	**<0.001**	1149 (100)	1.9 (1.4)
I NEED HELP	0.68	0.62–0.73	0.534	0.260	1149 (100)	2.7 (1.5)
INTERMACS profile	0.66	0.60–0.72	0.726	0.768	1149 (100)	5.3 (1.5)
Frailty domains	0.66	0.61–0.71	0.173	0.052	854 (74.3)	2.3 (0.8)
Frailty index	0.69	0.63–0.74	0.386	0.178	1149 (100)	0.3 (0.1)

AUC, area under the curve; CI, confidence interval; SD, standard deviation.

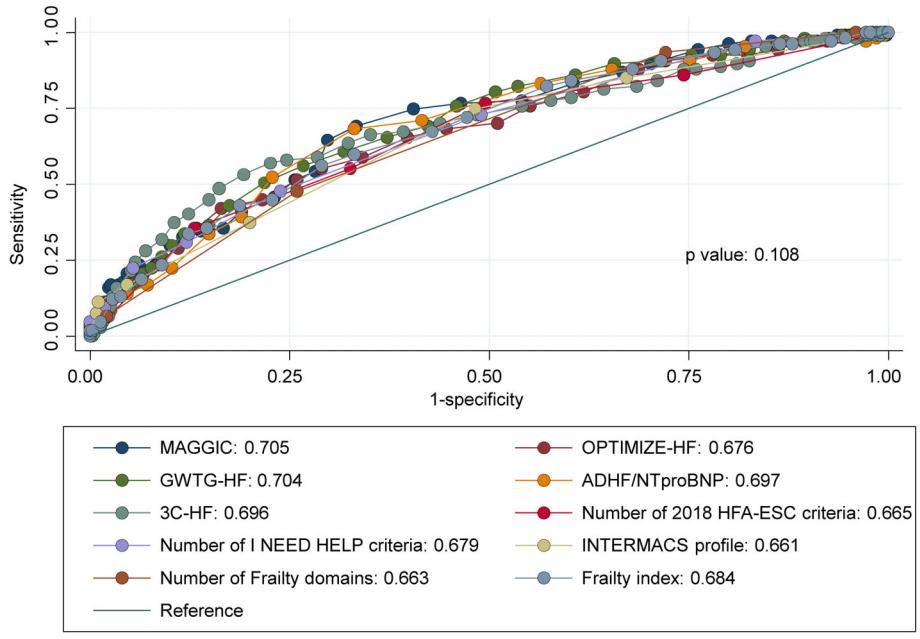


Figure 1 Comparison of the receiver operating characteristic curves of the scores evaluated in the HELP-HF cohort. All patients with completeness of data needed to calculate the scores ($n = 497$) are included in the comparative analysis of the areas under the curve (the areas under the curve of each score are reported in the box).

resulted to be higher than those related to the prediction of all-cause mortality, with the highest value (0.760) observed with the use of the MAGGIC score; conversely, the predictive ability of the scores ranged from similar to random guessing (0.525) to very poor (0.642).

Subgroup analyses

Among patients enrolled during hospitalization, the GWTG-HF showed the best predictive ability (AUC: 0.713), higher than the number of criteria of the 2018 HFA-ESC definition of advanced HF ($p = 0.023$) and of frailty domains fulfilled ($p = 0.018$). Similarly, the GWTG-HF had a better discriminative ability (AUC: 0.679) than the number of criteria of the 2018 HFA-ESC definition of advanced HF when considering only patients with reduced or mildly reduced ejection fraction ($p = 0.026$).

Machine learning-based models

The results from recursive feature elimination and the accuracy trend according to the number of selected features are reported in *Figure 4*. The more relevant features were age, impairment in

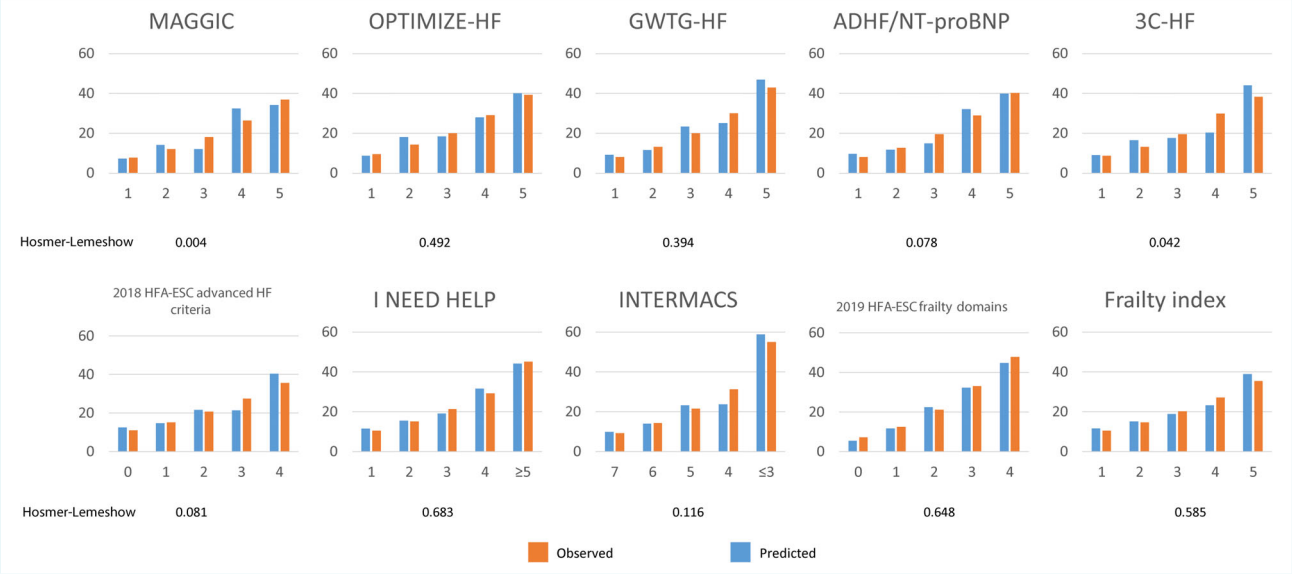


Figure 2 Observed and predicted event for each quintile or group of the testable risk scores. The rate of all-cause mortality for each quintile or group of the testable risk scores is reported with the *p*-value of the Hosmer–Lemeshow goodness of fits test below the histogram of each score.

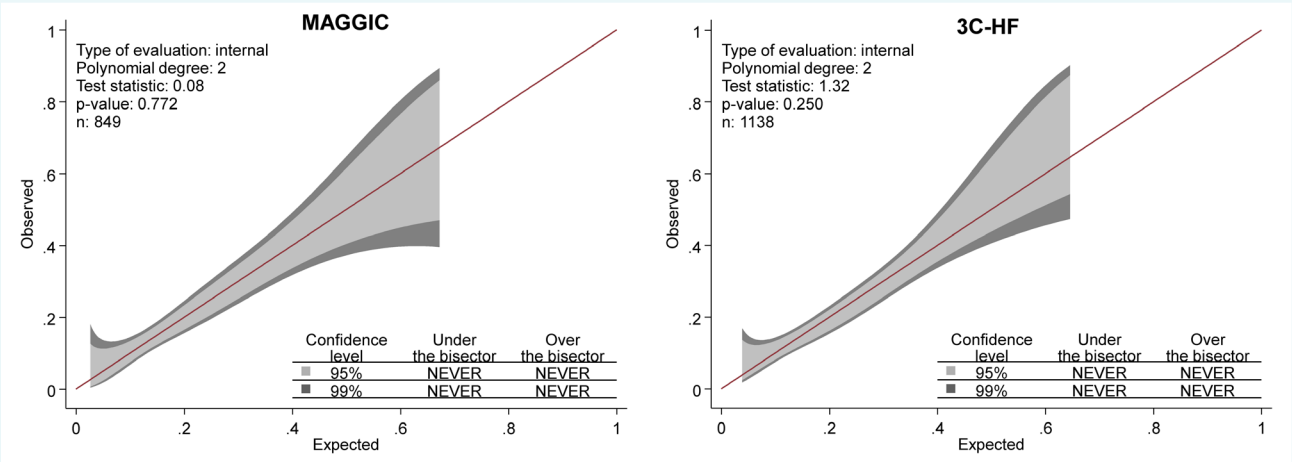


Figure 3 Calibration belts for the MAGGIC and 3C-HF scores. Calibration belts for the MAGGIC and 3C-HF scores depict the relationship between the models' fit probabilities and the observed rate of all-cause mortality.

instrumental activity in daily life, acute kidney injury, need for continuous renal replacement therapy, New York Heart Association class, and Forrester classification at enrolment. RF and SVM models exhibited the highest AUC in the training set (0.81 and 0.85, respectively), outperforming DT and KNN (0.73 and 0.70, respectively) (online supplementary *Figure S15*). The RF and SVM models showed in the testing set an AUC of 0.658 and 0.723, respectively (online supplementary *Figure S16*). Consistent findings were observed at sensitivity analysis performed after randomized split of the entire dataset in training and testing sets (online supplementary *Figure S17*).

Discussion

The main findings from the present sub-analysis from the HELP-HF registry are as follows (*Graphical Abstract*): (i) the accuracy of the risk scores in our cohort of patients with severe HF was only poor or moderate, and the estimate of mortality provided resulted to be poorly precise especially in patients at the highest risk; (ii) the scores tested performed better in the prediction of CV death than all-cause mortality, while the scores lacked of any predictive ability for non-CV death; (iii) despite the MAGGIC and GWTG-HF scores slightly outperformed the other scores in the prediction of 1-year

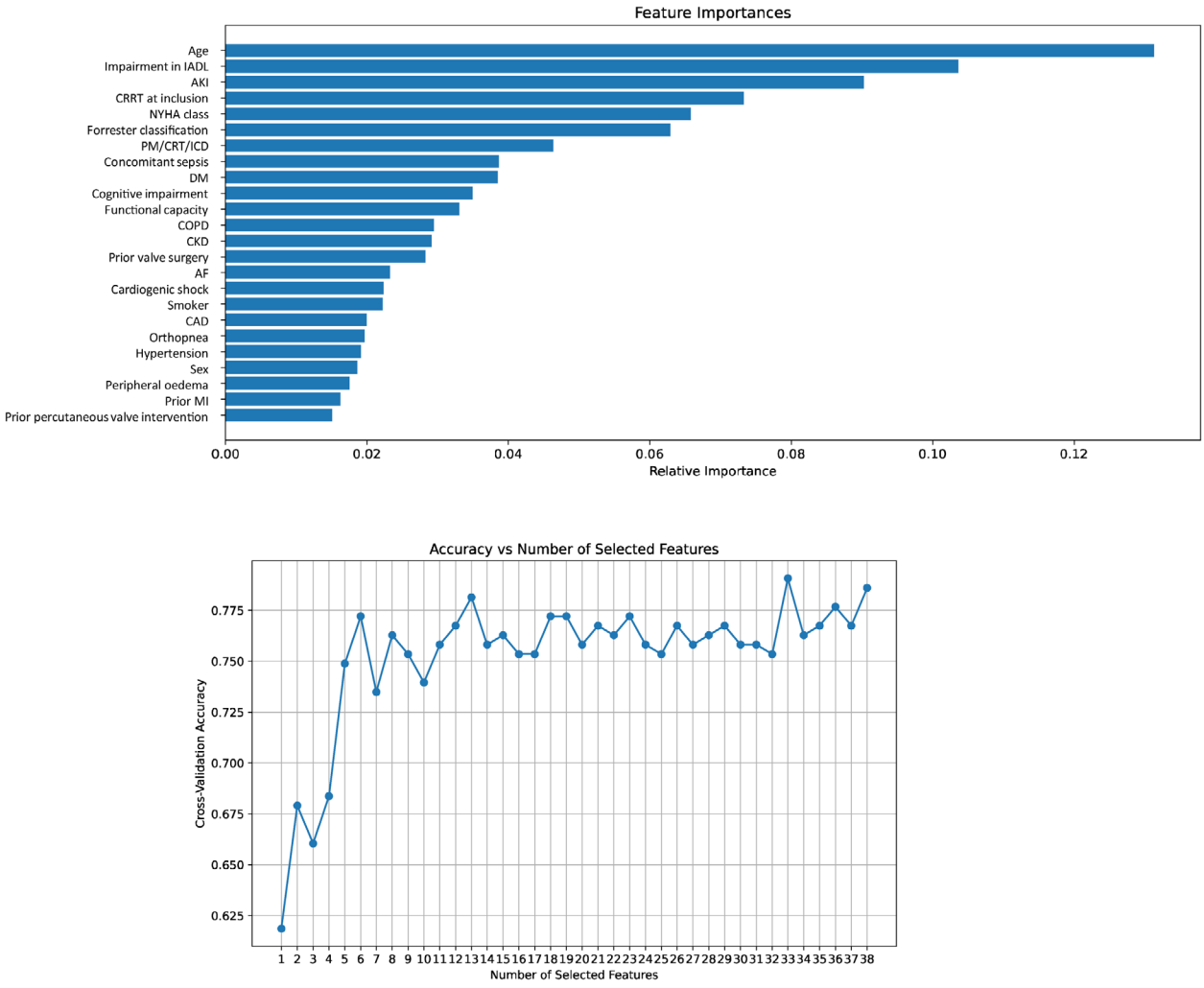


Figure 4 Feature relevance for the prediction of 1-year mortality. AF, atrial fibrillation; AKI, acute kidney injury; CAD, coronary artery disease; CRRT, continuous renal replacement therapy; CRT, cardiac resynchronization therapy; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; DM, diabetes mellitus; IADL, instrumental activity of daily living; ICD, implantable cardioverter defibrillator; MI, myocardial infarction; NYHA, New York Heart Association; PM, pacemaker.

mortality, there were no significant differences among the AUC of the testable risk scores; and (iv) ML-based models showed similar predictive ability as compared to the testable risk scores, although SVM showed a higher AUC than most of the testable risk scores.

Heart failure is one of the most common chronic diseases worldwide, encumbered by impaired functional capacity and activities of daily living, worsening symptoms, and high rate of mortality. Nonetheless, survival of patients with HF gradually increased in the last decades.[28] Despite these improvements, many patients still progress to severe stages of HF, characterized by frequent HF hospitalization and poor life expectancy.[3] Patients with severe HF are often unable to receive or up-titrate neurohormonal modulating therapy, but might benefit from LVAD or heart transplantation. However, the decision to consider advanced therapies must take into account the costs of such interventions and the limited availability of resources. Indeed, international guidelines recommend to consider advanced therapies only in patients with at least 1 year of life expectancy with good quality of life.[1,29] Hence, these critical decisions must be based on the presence and severity of comorbidities – extremely common in patients with severe HF[30] – and on life expectancy. Therefore, a precise evaluation of the risk of mortality is essential to define the most appropriate treatment plans for each patient, in order to improve resource allocation, avoiding futile procedures and better informing patients and caregivers about the possible benefits of advanced therapies or palliative care.

In this setting, the use of validated risk scores might be helpful in identifying patients at increased risk of mortality. We evaluated the applicability and the discriminative and predictive ability of some of the major HF risk prediction tools in patients affected by severe HF. In line with other reports evaluating different cohorts of patients with HF,[31,32] we were able to calculate only a limited number of validated risk scores. In previous studies, the authors overcame

missing data with multiple imputation analysis or imputing the default values of the score web-calculators or the median values in their population.[33,34] To provide a reliable estimate of the applicability of the risk models, we calculated the risk scores only in patients with all the necessary variables available. Other risk scores could not be evaluated, given the lack of data derived from cardiopulmonary test, such as the HFSS[17] or the MECKI scores,[18] information on laboratory values and drugs, such as the BCN Bio-HF[13] or the PREDICT-HF scores,[14] or on electrocardiographic and conduction disturbance data, included in the CHARM score,[15] confirming how many prognostic scores cannot be easily applied in daily practice due to the difficulty to retrieve all required variables, even in a registry that, although retrospective, was focused on patients with severe HF. An ideal tool should include only features evaluated during routine clinical practice and should be easily calculated at the bedside. As a result, the application in daily clinical practice of these scores is extremely rare, as confirmed by a recent report from the ESC HF Long-Term Registry Investigators, which showed that less than 1% of ambulatory HF patients received a prognostic estimate by their physician.[31] The lack of user-friendly tools to calculate some of these scores, together with their poor performance in some subgroups of patients with HF, such as those with severe HF, might explain their limited use.[21] For instance, a recent analysis from the LEVO-D registry on 403 patients receiving intermittent inotropic support with levosimendan as destination therapy, showed a suboptimal discrimination and calibration of the MAGGIC and BCN Bio-HF scores, with an underestimation of the risk of mortality.[34] An external validation of the MAGGIC risk score from the Swedish Heart Failure Registry reported an underestimation of the risk of mortality among patients at the highest estimated risk.[35] Likewise, the MAGGIC, OPTIMIZE-HF, and GWTG-HF scores have shown a suboptimal discriminative ability in patients waiting for heart transplantation too.[36] In line with these findings, we reported only a moderate or poor discriminative ability of the tested risk scores, showing an AUC $\geq$0.70 only for the MAGGIC, GWTG-HF and the ADHF/NT-proBNP scores. Similarly, the calibration of these scores in our cohort resulted to be inaccurate, with large CIs shown by calibration belts at the top of the risk for all the tested risk scores. Lastly, an ideal score to select patients who could truly benefit from advanced HF therapy should be able to properly predict the risk of non-CV death, as the risk of CV death is significantly modified by heart transplantation or LVAD implantation. The extremely poor performance of the tested scores in the prediction of non-CV death, coupled with their inadequate calibration for patients at the highest risk, remarks the need of specific scores for patients with severe HF. Several aspects might explain why available risk scores perform poorly in contemporary patients with severe HF. First, risk models rarely consider device therapy (none of those tested in our cohort). Second, score development is often based on studies that date back before the introduction of neurohormonal modulating and device therapy. Third, none of the available scores take into account the aetiology of HF, despite several studies highlighted how the benefits of drugs and device vary according to HF aetiology,[37–40] and ischeamic etiology carries an increased risk of worse prognosis across all the spectra of HF severity.[4] Fourth, the cohorts used to develop most of the scores included only a minority of patients with HFpEF, which are commonly older, more often female and affected by atrial fibrillation and chronic kidney disease, and present lower frequency of underlying coronary artery disease than those with HFrEF.[41] Lastly, and most importantly, there is paucity of scores specifically developed and validated in patients with severe HF, and those available require variables not routinely collected and do not include data on comorbidities, malnutrition, frailty, and daily life impairment. The prognostic impact of these aspects is well established[23,42,43] and highlighted by the relevance of impairment in instrumental activity in daily life found in our cohort at recursive feature elimination analysis. As a result, international guidelines are not consistent as regards to the recommendation about the use of risk scores: while the American Heart Association/American College of Cardiology guidelines give a 2a class of recommendation to the use of validated multivariable risk scores in ambulatory or hospitalized patients with HF,[29] European guidelines barely mention these tools,[1] also in light of the suboptimal discriminative ability shown.[34–36] Moreover, regression techniques assume a linear and homogeneous relationship between variables and outcomes, an assumption that could hamper their performance in complex diseases such as severe HF.[44,45] ML-based methods could potentially overcome these limitations: the model based on SVM showed better discriminative ability than most of the validated risk scores tested in our cohort, remarking how ML-based models might perform better than scores based on regression techniques in presence of tangled interplay between baseline features and clinical outcome. Nonetheless, ML-based models present different and equally relevant issues, such as limited interpretability, which is a major limitation in clinical practice.[46]

Study limitations

Our results should be interpreted in light of some limitations. First, the retrospective nature of our study did not allow us to collect variables included in other validated risk models and to assess their performance in the setting of severe HF. Moreover, missing data were noticed for some of the variables included in the scores investigated, despite the registry was specifically designed for patients with severe HF; a prospective study should be designed to address whether systematic recording of variables might improve score performance. Nonetheless, in light of the low upper CIs (0.76 for the score with the highest AUC, the MAGGIC) it seems unlikely that the performance of the scores tested would be significantly ameliorated by completeness of data. Second, in a limited number of cases we had to use a proxy of the variable included in the scores: for instance, the 3C-HF score considers diabetes mellitus with organ damage, while our dataset collected data only on diabetes mellitus status. Third, the disappointingly low AUCs for non-CV death must be interpreted in light of the low rate of non-CV death as compared with CV death (8% vs. 15%), but remark how these scores are inadequate to identify those patients who could truly benefit from advanced therapies. Fourth, our population included only patients evaluated at four high-volume, tertiary care Italian hospitals, and patients' features might vary in cohorts of different ethnicity and nationality. Therefore, future large studies including patients from other

countries and different ethnicity are necessary to confirm our findings.

Conclusion

All the available risk scores for prediction of mortality in HF showed poor or moderate discriminative ability in our cohort of patients with severe HF and lacked accuracy in estimating the risk of mortality in patients at the highest risk. ML-based models might be considered to improve risk prediction in this setting, but their use still needs to be validated in large cohorts of patients. Our findings remark the need for specific risk prediction tools for patients with severe HF, to help physicians, patients and relatives in the best choice between advanced therapies and palliative care.

Supplementary Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Acknowledgement

Open access funding provided by BIBLIOSAN.

Conflict of interest: D.S. reports personal fees from Novartis, Merck, GSK and Acceleron, outside the submitted work. M. Merlo reports personal fees from Pfizer, Novartis, Novo Nordisk and Vifor pharma, outside of the present work. G.S. reports consulting fees from Novartis, Impulse Dynamics and Biotronik, and speaker and honoraria from Novartis, Bayer, AstraZeneca, Boston Scientific, Vifor Pharma, Menarini and Akcea Therapeutics, outside the submitted work. M. Metra received personal consulting honoraria of minimal amount from Abbott, Amgen, Bayer, Edwards Therapeutics, LivaNova and Vifor Pharma for participation in advisory board meetings and executive committees of clinical trials. All the other authors have nothing to disclose.

References

1. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, *et al.* 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). With the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur J Heart Fail* 2022;**24**:4–131. https://doi.org/10.1002/ejhf.2333
2. Groenewegen A, Rutten FH, Mosterd A, Hoes AW. Epidemiology of heart failure. *Eur J Heart Fail* 2020;**22**:1342–1356. https://doi.org/10.1002/ejhf.1858
3. Dunlay SM, Roger VL, Killian JM, Weston SA, Schulte PJ, Subramaniam AV, *et al.* Advanced heart failure epidemiology and outcomes: A population-based study. *JACC Heart Fail* 2021;**9**:722–732. https://doi.org/10.1016/j.jchf.2021.05.009
4. Villaschi A, Pagnesi M, Stolfo D, Baldetti L, Lombardi CM, Adamo M, *et al.* Ischemic etiology in advanced heart failure: Insight from the HELP-HF registry. *Am J Cardiol* 2023;**204**:268–275. https://doi.org/10.1016/j.amjcard.2023.07.114
5. Mehra MR, Nayak A, Morris AA, Lanfear DE, Nemeh H, Desai S, *et al.* Prediction of survival after implantation of a fully magnetically levitated left ventricular assist device. *JACC Heart Fail* 2022;**10**:948–959. https://doi.org/10.1016/j.jchf.2022.08.002
6. Pocock SJ Ariti CA McMurray JJV Maggioni A Køber L Squire IB, *et al.*; Meta-Analysis Global Group in Chronic Heart Failure. Predicting survival in heart failure: A risk score based on 39 372 patients from 30 studies *Eur Heart J* 2013 **34** 1404 1413. https://doi.org/10.1093/eurheartj/ehs337
7. Scrutinio D, Guida P, Ammirati E, Oliva F, Frigerio M. Long-term prognostic implications of the ADHF/NT-proBNP risk score in patients admitted with advanced heart failure. *J Heart Lung Transplant* 2016;**35**:1264–1267. https://doi.org/10.1016/j.healun.2016.07.007
8. Pagnesi M, Lombardi CM, Chiarito M, Stolfo D, Baldetti L, Loiacono F, *et al.* Prognostic impact of the updated 2018 HFA-ESC definition of advanced heart failure: Results from the HELP-HF registry. *Eur J Heart Fail* 2022;**24**:1493–1503. https://doi.org/10.1002/ejhf.2561
9. Baumwol J. 'I Need Help' – A mnemonic to aid timely referral in advanced heart failure. *J Heart Lung Transplant* 2017;**36**:593–594. https://doi.org/10.1016/j.healun.2017.02.010
10. Abraham WT Fonarow GC Albert NM Stough WG Gheorghiade M Greenberg BH *et al.*; OPTIMIZE-HF Investigators and Coordinators. Predictors of in-hospital mortality in patients hospitalized for heart failure: Insights from the Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF) *J Am Coll Cardiol* 2008 **52** 347 356 https://doi.org/10.1016/j.jacc.2008.04.028
11. Peterson PN, Rumsfeld JS, Liang L, Albert NM, Hernandez AF, Peterson ED, *et al.* A validated risk score for in-hospital mortality in patients with heart failure from the American Heart Association Get With the Guidelines program. *Circ Cardiovasc Qual Outcomes* 2010;**3**:25–32. https://doi.org/10.1161/CIRCOUTCOMES.109.854877
12. Senni M, Parrella P, De MR, Cottini C, Böhm M, Ponikowski P, *et al.* Predicting heart failure outcome from cardiac and comorbid conditions: The 3C-HF score. *Int J Cardiol* 2013;**163**:206–211. https://doi.org/10.1016/j.ijcard.2011.10.071
13. Lupón J, de Antonio M, Vila J, Peñafiel J, Galán A, Zamora E, *et al.* Development of a novel heart failure risk tool: The Barcelona bio-heart failure risk calculator (BCN Bio-HF calculator). *PLoS One* 2014;**9**:e85466. https://doi.org/10.1371/journal.pone.0085466
14. Simpson J, Jhund PS, Lund LH, Padmanabhan S, Claggett BL, Shen L, *et al.* Prognostic models derived in PARADIGM-HF and validated in ATMOSPHERE and the Swedish Heart Failure Registry to predict mortality and morbidity in chronic heart failure. *JAMA Cardiol* 2020;**5**:432–441. https://doi.org/10.1001/jamacardio.2019.5850
15. Pocock SJ, Wang D, Pfeffer MA, Yusuf S, McMurray JJV, Swedberg KB, *et al.* Predictors of mortality and morbidity in patients with chronic heart failure. *Eur Heart J* 2006;**27**:65–75. https://doi.org/10.1093/eurheartj/ehi555
16. Levy WC, Mozaffarian D, Linker DT, Sutradhar SC, Anker SD, Cropp AB, *et al.* The Seattle Heart Failure Model: Prediction of survival in heart failure. *Circulation* 2006;**113**:1424–1433. https://doi.org/10.1161/CIRCULATIONAHA.105.584102
17. Aaronson KD, Schwartz JS, Chen TM, Wong KL, Goin JE, Mancini DM. Development and prospective validation of a clinical index to predict survival in ambulatory patients referred for cardiac transplant evaluation. *Circulation* 1997;**95**:2660–2667. https://doi.org/10.1161/01.cir.95.12.2660
18. Agostoni P Corrà U Cattadori G Veglia F Gioia R la Scardovi AB *et al.*, MECKI Score Research Group. Metabolic exercise test data combined with cardiac and kidney indexes, the MECKI score: A multiparametric approach to heart failure prognosis *Int J Cardiol* 2013 **167** 2710 2718. https://doi.org/10.1016/j.ijcard.2012.06.113
19. Barlera S Tavazzi L Franzosi MG Marchioli R Raimondi E Masson S, *et al.*; GISSI-HF Investigators. Predictors of mortality in 6975 patients with chronic heart failure in the Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardico-Heart Failure trial: Proposal for a nomogram *Circ Heart Fail* 2013 **6** 31 39. https://doi.org/10.1161/CIRCHEARTFAILURE.112.967828
20. Dziewięcka E, Gliniak M, Winiarczyk M, Karapetyan A, Wiśniowska-Śmiałek S, Karabinowska A, *et al.* Mortality risk in dilated cardiomyopathy: The accuracy of heart failure prognostic models and dilated cardiomyopathy-tailored prognostic model. *ESC Heart Fail* 2020;**7**:2455–2467. https://doi.org/10.1002/ehf2.12809
21. Crespo-Leiro MG, Metra M, Lund LH, Milicic D, Costanzo MR, Filippatos G, *et al.* Advanced heart failure: A position statement of the Heart Failure Association of the European Society of Cardiology. *Eur J Heart Fail* 2018;**20**:1505–1535. https://doi.org/10.1002/ejhf.1236
22. Stevenson LW, Pagani FD, Young JB, Jessup M, Miller L, Kormos RL, *et al.* INTERMACS profiles of advanced heart failure: The current picture. *J Heart Lung Transplant* 2009;**28**:535–541. https://doi.org/10.1016/j.healun.2009.02.015
23. Vitale C, Jankowska E, Hill L, Piepoli M, Doehner W, Anker SD, *et al.* Heart Failure Association of the European Society of Cardiology position paper on frailty in patients with heart failure. *Eur J Heart Fail* 2019;**21**:1299–1305. https://doi.org/10.1002/ejhf.1611
24. Butt JH, Dewan P, Jhund PS, Anand IS, Atar D, Ge J, *et al.* Sacubitril/valsartan and frailty in patients with heart failure and preserved ejection fraction. *J Am Coll Cardiol* 2022;**80**:1130–1143. https://doi.org/10.1016/j.jacc.2022.06.037
25. Sanders NA, Supiano MA, Lewis EF, Liu J, Claggett B, Pfeffer MA, *et al.* The frailty syndrome and outcomes in the TOPCAT trial. *Eur J Heart Fail* 2018;**20**:1570–1577. https://doi.org/10.1002/ejhf.1308
26. Smiti A. A critical overview of outlier detection methods. *Comput Sci Rev* 2020;**38**:100306. https://doi.org/10.1016/j.cosrev.2020.100306

27. Michelucci U. Feature importance and selection. *Fundamental Mathematical Concepts for Machine Learning in Science*. Cham: Springer International Publishing; 2024. p229–242.
28. Jones NR, Roalfe AK, Adoki I, Hobbs FDR, Taylor CJ. Survival of patients with chronic heart failure in the community: A systematic review and meta-analysis. *Eur J Heart Fail* 2019;**21**:1306–1325. https://doi.org/10.1002/ejhf.1594
29. Heidenreich PA, Bozkurt B, Aguilar D, Allen LA, Byun JJ, Colvin MM, *et al.* 2022 AHA/ACC/HFSA Guideline for the management of heart failure: A report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation* 2022;**145**:e895–e1032. https://doi.org/10.1161/CIR.0000000000001063
30. Conrad N, Judge A, Canoy D, Tran J, Pinho-Gomes AC, Millett ERC, *et al.* Temporal trends and patterns in mortality after incident heart failure: A longitudinal analysis of 86 000 individuals. *JAMA Cardiol* 2019;**4**:1102–1111. https://doi.org/10.1001/jamacardio.2019.3593
31. Canepa M, Fonseca C, Chioncel O, Laroche C, Crespo-Leiro M, Coats A, *et al.*; ESC HF Long-Term Registry Investigators. Performance of prognostic risk scores in chronic heart failure patients enrolled in the European Society of Cardiology Heart Failure Long-Term Registry. *JACC Heart Fail* 2018;**6**:452–462. https://doi.org/10.1016/j.jchf.2018.02.001
32. Masè M, Rossi M, Setti M, Barbati G, Teso MV, Ribichini FL, *et al.* Applicability and performance of heart failure prognostic scores in dilated cardiomyopathy: The real-world experience of an Italian referral center for cardiomyopathies. *Int J Cardiol* 2024;**396**:131562. https://doi.org/10.1016/j.ijcard.2023.131562
33. Codina P, Lupón J, Borrellas A, Spitaleri G, Cediel G, Domingo M, *et al.* Head-to-head comparison of contemporary heart failure risk scores. *Eur J Heart Fail* 2021;**23**:2035–2044. https://doi.org/10.1002/ejhf.2352
34. Codina P, Dobarro D, Juan-Bagudá J de, Frutos F de, Lupón J, Bayes-Genis A, *et al.*; Spanish LEVO-D registry Collaborators. Heart failure risk scores in advanced heart failure patients: Insights from the LEVO-D registry. *ESC Heart Fail* 2023;**10**:2875–2881. https://doi.org/10.1002/ehf2.14400
35. Sartipy U, Dahlström U, Edner M, Lund LH. Predicting survival in heart failure: Validation of the MAGGIC heart failure risk score in 51 043 patients from the Swedish Heart Failure Registry. *Eur J Heart Fail* 2014;**16**:173–179. https://doi.org/10.1111/ejhf.32
36. Nguyen LS, Coutance G, Ouldamar S, Zahr N, Brechot N, Galeone A, *et al.* Performance of existing risk scores around heart transplantation: Validation study in a 4-year cohort. *Transpl Int* 2018;**31**:520–530. https://doi.org/10.1111/tri.13122
37. Ruwald MH, Solomon SD, Foster E, Kutyifa V, Ruwald AC, Sherazi S, *et al.* Left ventricular ejection fraction normalization in cardiac resynchronization therapy and risk of ventricular arrhythmias and clinical outcomes: Results from the Multicenter Automatic Defibrillator Implantation Trial with Cardiac Resynchronization Therapy (MADIT-CRT) trial. *Circulation* 2014;**130**:2278–2286. https://doi.org/10.1161/CIRCULATIONAHA.114.011283
38. Balmforth C, Simpson J, Shen L, Jhund PS, Lefkowitz M, Rizkala AR, *et al.* Outcomes and effect of treatment according to etiology in HFrEF: An analysis of PARADIGM-HF. *JACC Heart Fail* 2019;**7**:457–465. https://doi.org/10.1016/j.jchf.2019.02.015
39. Metra, M, Jankowska, EA, Pagnesi, M, Anker, SD, Butler, J, Dorigotti, F, *et al.*; AFFIRM-AHF Investigators. Impact of ischaemic aetiology on the efficacy of intravenous ferric carboxymaltose in patients with iron deficiency and acute heart failure: Insights from the AFFIRM-AHF trial. *Eur J Heart Fail* 2022;**24**(10):1928–1939. https://doi.org/10.1002/ejhf.2630
40. Linde C Abraham WT Gold MR Daubert C; REVERSE Study Group. Cardiac resynchronization therapy in asymptomatic or mildly symptomatic heart failure patients in relation to etiology: Results from the REVERSE (REsynchronization reVErses Remodeling in Systolic left vEntricular Dysfunction) study *J Am Coll Cardiol* 2010 **56** 1826 1831. https://doi.org/10.1016/j.jacc.2010.05.055.
41. Borlaug BA. Evaluation and management of heart failure with preserved ejection fraction. *Nat Rev Cardiol* 2020;**17**:559–573. https://doi.org/10.1038/s41569-020-0363-2
42. Villaschi A, Chiarito M, Pagnesi M, Stolfo D, Baldetti L, Lombardi CM, *et al.* Frailty according to the 2019 HFA-ESC definition in patients at risk for advanced heart failure: Insights from the HELP-HF registry. *Eur J Heart Fail* 2024;**26**:1399–1407. https://doi.org/10.1002/ejhf.3234
43. Pagnesi M, Serafini L, Chiarito M, Stolfo D, Baldetti L, Inciardi RM, *et al.* Impact of malnutrition in patients with severe heart failure. *Eur J Heart Fail* 2024;**26**:1585–1593. https://doi.org/10.1002/ejhf.3285
44. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;**15**:361–387. https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4
45. Concato J, Peduzzi P, Holford TR, Feinstein AR. Importance of events per independent variable in proportional hazards analysis. I. Background, goals, and general strategy. *J Clin Epidemiol* 1995;**48**:1495–1501. https://doi.org/10.1016/0895-4356(95)00510-2
46. Chiarito M, Luceri L, Oliva A, Stefanini G, Condorelli G. Artificial intelligence and cardiovascular risk prediction: All that glitters is not gold. *Eur Cardiol* 2022;**17**:e29. https://doi.org/10.15420/ecr.2022.11