ESC HEART FAILURE
ESC Heart Failure 2022; **9**: 606–613
Published online 22 November 2021 in Wiley Online Library (wileyonlinelibrary.com) **DOI:** 10.1002/ehf2.13707

ORIGINAL ARTICLE

Risk stratification in heart failure decompensation in the community: HEFESTOS score

José-María Verdu-Rotellar[1,2,3]*, Rosa Abellana[4], Helene Vaillant-Roussel[5], Lea Gril Jevsek[6], Radost Assenova[7], Djurdjica Kasuba Lazic[8], Peter Torsza[9], Liam George Glynn[10], Heidrun Lingner[11], Jacopo Demurtas[12,13], Hans Thulesius[14,15], Miguel Angel Muñoz[1,2,3] and HEFESTOS group†

[1]*Gerencia Territorial de Barcelona, Institut Català de la Salut, Barcelona, Spain;* [2]*Unitat de Suport a la Recerca de Barcelona, Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain;* [3]*School of Medicine, Universitat Autònoma de Barcelona, Bellaterra, Spain;* [4]*Departament de Fonaments Clinics, Facultat de Medicina, Universitat de Barcelona, Barcelona, Spain;* [5]*Faculty of Medicine, UPU ACCePPT, Department of General Practice, CHU, Direction de La Recherche Clinique et de l'Innovation, Clermont Auvergne University, Clermont-Ferrand, France;* [6]*Medical Faculty, University of Maribor, Maribor, Slovenia;* [7]*Department of Urology and General Medicine, Faculty of Medicine, Medical University of Plovdiv, Plovdiv, Bulgaria;* [8]*Department of Family Medicine "Andrija Stampar" School of Public Health, School of Medicine University of Zagreb, Zagreb, Croatia;* [9]*Semmelweis University, Budapest, Hungary;* [10]*Health Research Institute and Graduate Entry Medical School, University of Limerick, Limerick, Ireland;* [11]*Hannover Medical School-Center for Public Health and Healthcare, Hannover, Germany;* [12]*Primary Care Department, Azienda Usl Toscana Sud Est, Grosseto, Italy;* [13]*Clinical and Experimental Medicine PhD Program, University of Modena and Reggio Emilia, Modena, Italy;* [14]*Department of Clinical Sciences, Lund University, Lund, Sweden; and* [15]*Department of Medicine and Optometry, Linnaeus University, Växjö, Sweden*

Abstract

Aims Because evidence regarding risk stratification predicting prognosis of patients with heart failure (HF) decompensation attended in primary care is lacking, we developed and externally validated a model to forecast death/hospitalization during the first 30 days after an episode of decompensation. The predictive model is based on variables easily obtained in primary care settings.

Methods and results HEFESTOS is a multinational study consisting of a derivation cohort of HF patients recruited in 14 primary healthcare centres in Barcelona and a validation cohort from primary healthcare in 9 other European countries. The derivation and validation cohorts included 561 and 250 patients, respectively. Percentages of women in the derivation and validation cohorts were 56.3% and 47.6% (P = 0.026), respectively. Mean age was 82.2 years (SD 8.03) in the derivation cohort, and 79.3 years (SD 10.3) in the validation one (P = 0.001). HF with preserved ejection fraction represented 72.1% in the derivation cohort and 58.8% in the validation one (P = 0.004). Mortality/hospitalization during the first 30 days after a decompensation episode was 30.5% and 26% (P = 0.225) for the derivation and validation cohorts, respectively. Multivariable logistic regression models were performed to develop a score of risk. The identified predictors were worsening of dyspnoea [odds ratio (OR): 2.5; P = 0.001], orthopnoea (OR: 2.16; P = 0.01), paroxysmal nocturnal dyspnoea (OR: 2.25; P = 0.01), crackles (OR: 2.35; P = 0.01), New York Heart Association functional class III/IV (OR: 2.11; P = 0.001), oxygen saturation $\leq$ 90% (OR: 4.98; $P < 0.001$), heart rate $>$ 100 b.p.m. (OR: 2.72; P = 0.002), and previous hospitalization due to HF (OR: 2.45; $P < 0.001$). The model showed an area under the curve (AUC) of 0.807, 95% confidence interval (CI): [0.770; 0.845] in the derivation cohort and AUC 0.73, 95% CI: [0.660; 0.808] in the validation one. No significant differences between both cohorts were observed (P = 0.08). Regarding probability of hospitalization/death, three risk groups were defined: low $<$5%, medium 5–20%, and high $>$20%. Outcome incidence was 2.7% for the low-risk group, 12.8% for medium risk, and 46.2% for high risk in the derivation cohort, and 9.1%, 12.9%, and 39.6% in the validation one.

Conclusions The HEFESTOS score, based on variables easily accessible in a community setting and validated in an external European cohort, properly predicted the risk of death/hospitalization during the first 30 days after an HF decompensation episode.

Keywords Primary care; Heart failure; Decompensation; Risk stratification

Received: 18 June 2021; Revised: 1 October 2021; Accepted: 29 October 2021
**Correspondence to: José-María Verdu-Rotellar, Unitat de Suport a la Recerca, Gerència Territorial de Barcelona, Institut Català de la Salut-IDIAP Jordi Gol Departament de Salut, Generalitat de Catalunya, Carrer Sardenya 375, Entlo, 08025 Barcelona, Spain. Tel: 0034 932 073 760; Fax: 0034 933070812. Email: verdujm@gmail.com*
†*The list of the HEFESTOS investigators is given in the Supporting Information,* Appendix S1.

Introduction

Heart failure decompensation is a major cause of hospitalization and is associated with high in-hospital mortality following discharge and increased readmission rates. Moreover, it represents one of the greatest economic and health burdens for the public healthcare system.[1–4]

Heart failure decompensation is defined as the worsening of the symptoms and characteristic signs of this disease that require unplanned visits to outpatient clinics and emergency departments, or hospitalization.[5,6]

A number of studies in different countries have developed models that predict the risk of hospitalization or death in patients with this condition. These heterogeneous models, however, present high variability and have not been widely implemented in routine clinical practice. Almost all of them included hospitalized patients, and the few that assessed patients attended in primary care or outpatient clinics were based on variables not accessible to the majority of primary healthcare professionals.[7,8] In addition, they showed a reduced predictive capacity in patients with low risk of death.[9]

Most patients presenting heart failure decompensation are attended in primary care and contact their family doctor during the month prior to being hospitalized.[10]

In the primary care setting, some laboratory tests such as natriuretic peptides and echocardiographic parameters, which could help to better stratify risk of heart failure, are not usually available.

The objective of our study was, therefore, to develop and validate a risk score based on clinical variables easily accessible in primary care to predict hospitalization/death at short term (30 days) of patients presenting heart failure decompensation.

Methods

HEFESTOS is an international, multicentre, prospective cohort study. It is focused on creating a predictive prognostic model for heart failure patients who suffer a decompensation episode initially treated in primary care.

The project was originally designed in Spain where a pilot study was carried out to test its feasibility (March–July 2015).

The protocol was discussed within the European General Research Network where family doctors from nine countries agreed to collaborate in the project. The European research team worked on standardizing methodological procedures and extended the study to their respective countries. The study protocol was written in English and translated into the languages of the participant countries (French, German, Italian, Slovenian, Croatian, Bulgarian, Hungarian, and Swedish). Ethical approval was sought and granted from each country taking part.

The derivation cohort consisted of 561 heart failure patients recruited in 14 primary care centres in Barcelona (Spain) between March 2015 and June 2019. The validation cohort included 250 primary care patients from 9 European countries (France, Ireland, Germany, Italy, Slovenia, Croatia, Bulgaria, Hungary, and Sweden) recruited between March 2017 and June 2019.

Inclusion criteria were patients aged $\geq$45 years with a diagnosis of heart failure in their medical records and attended in primary care for heart failure decompensation. Recruitment was carried out after the participants provided informed, signed consent.

Heart failure decompensation was considered as such when the patient had any of the following signs or symptoms: increased dyspnoea, unexplained weight gain, or appearance/increase of peripheral oedema.

Patients were excluded if they presented severe psychiatric illness or cognitive impairment, were unable to complete clinical examinations, or had been hospitalized in the previous 30 days due to heart failure decompensation.

The Helsinki ethics declaration was followed at all times. Participants provided informed consent, and the study protocol was approved by the research ethics committees of all the participant countries.

Data were collected using specifically designed webpage forms, and quality checks were carried out every month. Potential explanatory variables were sociodemographic data, comorbidity, clinical examination, pulse oximetry, electrocardiography, and medication. Ejection fraction assessments were collected when available. Variables were recorded at the time of study inclusion.

The outcome variables were hospitalization for causes related to heart failure decompensation and/or mortality from any cause in the 30 days after the index consultation date. This information was obtained by accessing medical records and/or by telephone contacts with patients or their relatives.

Statistical analysis

Categorical data are expressed using frequencies and percentages, and continuous ones with means and standard deviations (SD). The χ^2 test was used to study the association between outcome and categorical variables, and Student's *t*-test to assess differences between continuous variables. The associations were also evaluated in terms of odds ratio. Clinically meaningful variables showing a significant level in univariate analysis ($P < 0.05$) were thereafter included in the multivariable logistic model. A backward stepwise method was employed to identify independent risk predictors with $P < 0.05$ for inclusion or deletion. The overall performance of the model was calculated with the Brier score and Nagelkerke R^2.

DOI: 10.1002/ehf2.13707

The discriminative ability of the model was assessed by Harrell's C-index (area under the receiver-operating characteristic curve). The calibration of the model was checked using the Hosmer–Lemeshow goodness-of-fit test (by deciles of the predicted probability) and plotting the observed and predicted probabilities of the model grouped into tenths using deciles. The external validation of the predictive model was also evaluated in terms of calibration and discrimination.

In creating the prognostic risk score, each final predictor had its beta-coefficient divided by the smallest figure and then rounded to the nearest integer number.[10] The predictors of a particular patient thus ranged from 0 to 23. Analysis was performed using R software for Windows Version 4.0.3 (R project for statistical computing; Vienna, Austria).

Results

A total of 561 consecutively recruited patients were included in the derivation cohort (women = 56.3%), and 250 patients (women = 47.6%) in the validation one. Mean age was 82.2 (SD 8.03) and 79.3 (SD 10.3) years in the derivation and validation cohorts, respectively.

In the validation cohort, there was a greater percentage of men and participants were younger and presented lower comorbidity, with the exception of coronary heart disease, which was higher (*Table 1*).

Mortality or hospitalization during the first 30 days after an episode of decompensation was 30.5% for the derivation cohort and 26.0% for the validation one ($P = 0.225$). Of the 561 derivation cohort patients, 450 were attended in primary care settings (56% oral treatment adjustment, 43% intravenous diuretic medication), and 111 were referred to the hospital emergency room.

Among those referred to the emergency room, 85% were hospitalized and 8.1% died in the following 30 days. Of the patients attended in primary care and not initially referred to the hospital, 12.5% were eventually admitted, and 2.2% died in the following 30 days.

Of the patients who needed to be hospitalized or died in the first 30 days after the episode, those attended in primary care settings presented worse clinical parameters [crackles, higher respiratory and heart rate, lower ejection fraction, paroxysmal dyspnoea, orthopnoea, and worse functional New York Heart Association (NYHA) functional class class]. We did not find any association with the outcomes regarding comorbidity with the exception of renal failure and chronic obstructive pulmonary disease. Such conditions were more commonly observed in patients who were hospitalized or deceased. Previous hospitalization due to heart failure decompensation (between 31 and 365 days prior to inclusion) was also related to outcome occurrence (*Table 2*). All these variables were included in the analysis to create the multivariate predictive model.

The model confirmed that previous hospitalization due to heart failure decompensation, presence of crackles, paroxysmal nocturnal dyspnoea, orthopnoea, NYHA III/IV status, worsening in NYHA functional status, having a heart rate $>$ 100 b.p.m., and oxygen saturation $\leq$ 90% were independent predictors for hospitalization/death in the first 30 days following the decompensation episode.

This predictive model demonstrated a good discrimination ability with an area under the curve (AUC) at 30 days of 0.807, 95% confidence interval (CI): [0.770; 0.845]. *Figure 1* shows the receiver-operating characteristic curve (ROC) of the predicted probabilities (black line). In addition, overall performance using the Bier score had a rating of 0.015, which was below 0.25, and a Nagelkerke's R^2 of 0.341. In terms of agreement between the predicted and observed probabilities of the risk of hospitalization/death, the χ^2 Hosmer and Lemeshow test was 9 ($P = 0.3$), indicating no evidence of poor fit. Additionally, the calibration plots of the model showed a good calibration, because the triangles lay around a 45 line of the plot with a slope of 1 [*Figure 2A*].

Table 1 Characteristics of participant heart failure patients according to the study cohorts

	[All] *N* = 811	Spain cohort *N* = 561	Europe cohort *N* = 250	OR	*P* value	*N*
Sex					0.026	811
Male	376 (46.4%)	245 (43.7%)	131 (52.4%)	Ref.		
Female	435 (53.6%)	316 (56.3%)	119 (47.6%)	0.70 [0.52;0.95]		
Age (mean, SD)	81.3 (8.89)	82.2 (8.03)	79.3 (10.3)	0.96 [0.95;0.98]	0.001	808
Diabetes mellitus	358 (44.3%)	267 (47.6%)	91 (36.7%)	0.64 [0.47;0.87]	0.005	809
Cardiac ischaemia	321 (39.8%)	191 (34.0%)	130 (52.8%)	2.17 [1.60;2.95]	0.001	807
Chronic or paroxysmal atrial fibrillation	454 (56.1%)	334 (59.5%)	120 (48.4%)	0.64 [0.47;0.86]	0.004	809
Stroke	120 (14.9%)	82 (14.6%)	38 (15.4%)	1.06 [0.69;1.61]	0.861	808
Chronic renal disease	342 (42.3%)	256 (45.6%)	86 (34.7%)	0.63 [0.46;0.86]	0.005	809
Smoking	77 (9.58%)	46 (8.2%)	31 (12.8%)	1.64 [1.00;2.65]	0.059	804
Chronic obstructive pulmonary disease	223 (27.6%)	169 (30.1%)	54 (21.8%)	0.65 [0.45;0.92]	0.018	809
Hypertension	684 (88.0%)	489 (87.2%)	195 (90.3%)	1.36 [0.83;2.33]	0.283	777

OR, odds ratio; SD, standard deviation.

Table 2 Summary description according to hospitalization or/and death during the first 30 days after an episode of decompensation in heart failure patients

	[All] $N = 561$	No outcome $N = 390$	Hospitalization or death $N = 171$	Odds ratio	*P* value	*N*
Demographics						
Sex						561
Male	245 (43.7%)	167 (42.8%)	78 (45.6%)	Ref.	Ref.	
Female	316 (56.3%)	223 (57.2%)	93 (54.4%)	0.89 [0.62;1.28]	0.540	
Age (mean, SD)	82.2 (8.03)	82.0 (8.02)	82.7 (8.05)	1.01 [0.99;1.03]	0.369	561
Personal history						
Diabetes mellitus	267 (47.6%)	179 (45.9%)	88 (51.5%)	1.25 [0.87;1.79]	0.226	561
Cardiac ischaemia	191 (34.0%)	130 (33.3%)	61 (35.7%)	1.11 [0.76;1.62]	0.590	561
Chronic or paroxysmal atrial fibrillation	334 (59.5%)	229 (58.7%)	105 (61.4%)	1.12 [0.77;1.62]	0.554	561
Stroke	82 (14.6%)	57 (14.6%)	25 (14.6%)	1.00 [0.59;1.65]	0.990	561
Chronic renal disease	256 (45.6%)	167 (42.8%)	89 (52.0%)	1.45 [1.01;2.08]	0.044	561
Smoking	46 (8.20%)	34 (8.72%)	12 (7.02%)	0.80 [0.38;1.54]	0.512	561
Chronic obstructive pulmonary disease	169 (30.1%)	106 (27.2%)	63 (36.8%)	1.56 [1.06;2.29]	0.023	561
Hypertension	489 (87.2%)	337 (86.4%)	152 (88.9%)	1.25 [0.73;2.24]	0.426	561
Previous hospitalization in the last year	132 (23.5%)	74 (19.0%)	58 (33.9%)	2.19 [1.46;3.29]	<0.001	561
Anamnesis and exploration						
Crackles	347 (61.9%)	209 (53.6%)	138 (80.7%)	3.60 [2.37;5.61]	<0.001	561
Heart rate > 100 b.p.m.	66 (11.8%)	34 (8.72%)	32 (18.7%)	2.41 [1.42;4.07]	0.001	561
Respiratory frequency > 24 per minute	86 (15.4%)	46 (11.9%)	40 (23.4%)	2.27 [1.41;3.63]	0.001	559
Body mass index	30.9 (6.08)	31.1 (6.19)	30.3 (5.80)	0.98 [0.95;1.01]	0.159	546
Temperature	36.0 (0.58)	36.0 (0.56)	36.1 (0.62)	1.38 [1.01;1.89]	0.046	557
Oedema	452 (80.6%)	312 (80.0%)	140 (81.9%)	1.13 [0.71;1.81]	0.613	561
Systolic blood pressure (mmHg)	136 (22.1)	135 (23.2)	137 (19.6)	1.00 [1.00;1.01]	0.374	809
Diastolic blood pressure (mmHg)	71.9 (12.3)	72.1 (11.8)	71.4 (13.4)	0.99 [0.98;1.01]	0.493	561
Ejection fraction rate						561
≥50%	381 (67.9%)	281 (72.1%)	100 (58.5%)	Ref.	Ref.	
40–49%	50 (8.91%)	30 (7.69%)	20 (11.7%)	1.87 [1.00;3.44]	0.049	
≤39%	71 (12.7%)	40 (10.3%)	31 (18.1%)	2.18 [1.28;3.67]	0.004	
Unknown	59 (10.5%)	39 (10.0%)	20 (11.7%)	1.44 [0.79;2.57]	0.228	
Signs and symptoms						
Paroxysmal nocturnal dyspnoea	160 (28.5%)	75 (19.2%)	85 (49.7%)	4.14 [2.80;6.14]	<0.001	561
Orthopnoea	272 (48.5%)	150 (38.5%)	122 (71.3%)	3.97 [2.70;5.90]	<0.001	561
NYHA						561
I or II	153 (27.3%)	135 (34.6%)	18 (10.5%)	Ref.	Ref.	
III or IV	408 (72.7%)	255 (65.4%)	153 (89.5%)	4.46 [2.68;7.83]	<0.001	
Worsening in NYHA status	416 (74.2%)	267 (68.5%)	149 (87.1%)	3.10 [1.92;5.21]	<0.001	561
Weight gain	192 (34.2%)	116 (29.7%)	76 (44.4%)	1.89 [1.30;2.74]	0.001	561
Worsening or occurrence of ankle oedemas	364 (65.1%)	242 (62.4%)	122 (71.3%)	1.50 [1.02;2.23]	0.040	559
Time elapsed since the first symptoms	9.31 (9.74)	9.80 (10.5)	8.18 (7.68)	0.98 [0.96;1.00]	0.073	553
Complementary explorations						
ECG						557
Normal	72 (12.9%)	54 (14.0%)	18 (10.6%)	Ref.	Ref.	
Abnormal	485 (87.1%)	333 (86.0%)	152 (89.4%)	1.36 [0.78;2.47]	0.279	
Oxygen saturation ≤ 90%	37 (6.61%)	11 (2.83%)	26 (15.2%)	6.09 [2.99;13.2]	<0.001	560
Treatments						
ACEi/ARB	421 (75.0%)	278 (71.3%)	143 (83.6%)	2.05 [1.31;3.30]	0.003	561
Beta-blockers	409 (72.9%)	284 (72.8%)	125 (73.1%)	1.01 [0.68;1.53]	1.000	561
Aldosterone	121 (21.6%)	74 (19.0%)	47 (27.5%)	1.62 [1.06;2.46]	0.032	561
Loop diuretics	472 (84.1%)	321 (82.3%)	151 (88.3%)	1.61 [0.96;2.82]	0.071	561

ACEi, angiotensin-converting enzyme inhibitors; ARB, angiotensin receptor blockers; ECG, electrocardiogram; NYHA, New York Heart Association; SD, standard deviation.

External validation

The AUC of the model in the validation cohort was 0.73, 95% CI: [0.660; 0.808] (*Figure 1*, grey line), which was a little lower than in derivate one although no significantly statistical differences between them were observed ($P = 0.08$). The calibration plot of the model also demonstrated good calibration between the observed and predicted probabilities with the exception of strongly predicted probabilities [*Figure 2B*].

Scoring system

In order to build a score able to predict the risk of hospitalization/death at 30 days, we divided the model regression coefficients by the beta coefficient of males (0.35) and rounded them to the nearest integer (*Table 3*). The score ranked from 0 to 23: 0 points corresponding to female gender, no previous hospitalization, not presenting crackles or paroxysmal nocturnal dyspnoea, NYHA class I/II, without

DOI: 10.1002/ehf2.13707

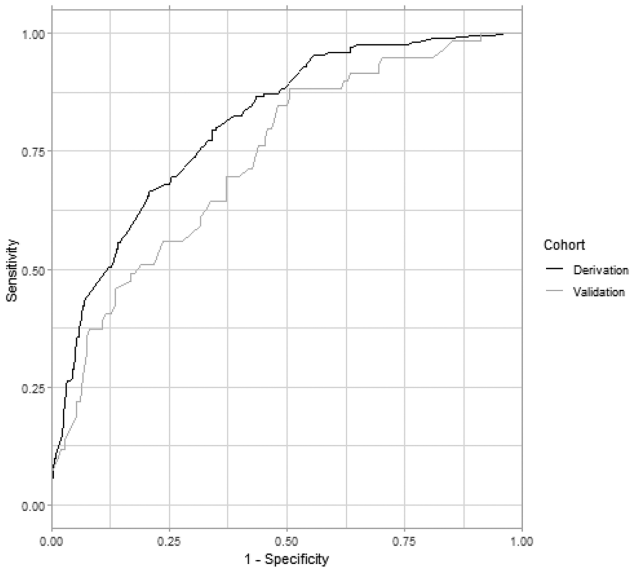
NYHA worsening, heart rate < 100 b.p.m., and oxygen saturation > 90%. The estimated risk for 0 points was 0.017, and for 23 points 0.980. The estimated risk for each score point is shown in *Figure 3*. The AUC for the score in the derivation cohort was 0.89, 95% CI: [0.78; 0.84] and 0.73, 95% CI: [0.66; 0.81] in the validation one.

In accordance with published studies,[7,8,11,12] we created three groups of risk. A <5% probability of hospitalization/death 30 days after the decompensation episode indicated patients at low risk, 5–20% medium risk, and >20% high risk. Thus, low risk for hospitalization/death corresponded to scores ≤3, medium risk 4–7, and high risk ≥8. In *Figure 3*, the predicted probabilities in days for hospitalization/death 30 days after the decompensation episode are plotted.

In the derivation cohort, cumulative incidence for low-risk patients was 2.7%, medium risk 12.8%, and high risk 46.2%. In the validation cohort, accumulated incidences were similar with the exception of the low-risk group (*Table 4*).

A URL has been created to facilitate clinicians and researchers access to the calculator: https://rabellana.shinyapps.io/HEFESTOS_Score/.

Figure 1 Receiver-operating characteristic curve of the predicted probabilities of short-term hospitalization or death in decompensated heart failure patients.



Discussion

Summary of main findings

This study presents a multinational, externally validated score, the HEFESTOS, which predicts hospitalization/death at short term (30 days) due to a heart failure decompensation episode in patients attended in primary care. It is based on clinical variables easily available for primary care professionals.

The identified independent predictors were hospitalization in the previous year, presence of crackles, paroxysmal nocturnal dyspnoea, or orthopnoea, NYHA III/IV, worsening in NYHA status, heart rate > 100 b.p.m., and oxygen saturation ≤ 90%.

This score facilitates risk stratification in community-living heart failure patients. C-statistics of 0.81, 95% CI: [0.77; 0.84] in the derivation and 0.74, 95% CI: [0.66; 0.81] in the validation cohorts, respectively, were reported. To date, these are the highest ranges of predictive scores published in studies on heart failure decompensation.[7,8]

Figure 2 Calibration plot of the predictive model for short-term heart failure hospitalization or death. (A) Derivation cohort. (B) Validation cohort. The triangles denote the mean predicted and observed event probabilities for patients grouped into tenths using deciles. The grey dashed line denotes perfect calibration. The distribution of calculated predicted probabilities is overlaid along the horizontal axis. 1 indicates patients with hospitalization or death 30 days after an episode, and 0 patients who did not undergo an event.

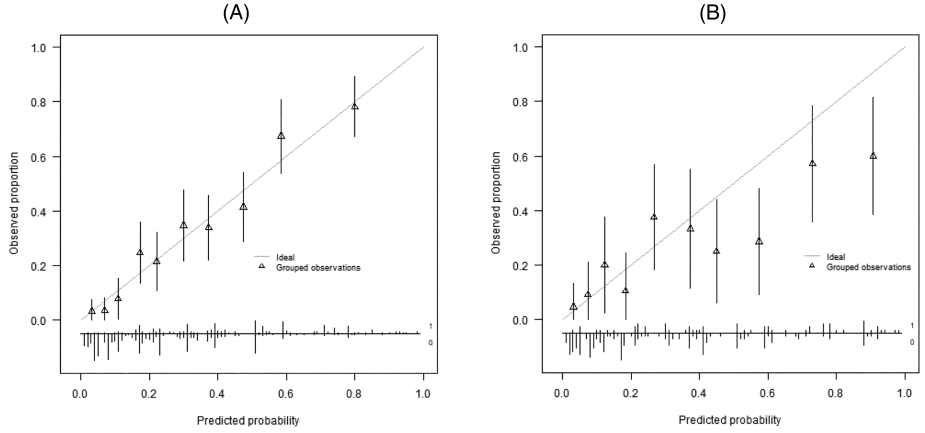
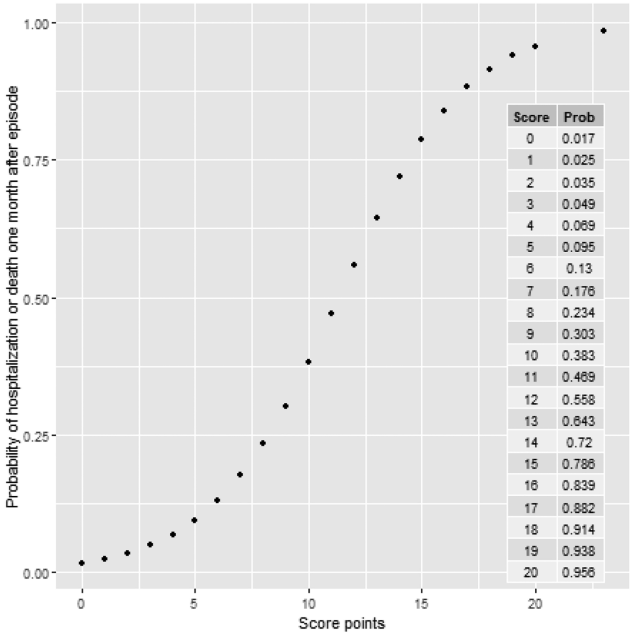


Table 3 Multivariable predictors to predict hospitalization or death 30 days after an episode of decompensation in heart failure patients

	Odds ratio (95% CI)	Beta coefficient	*P* value	Points (additive score)	Risk groups
(Intercept)		−4.050	<0.001		Low risk
Male	1.43 (0.93; 2.17)	0.350	0.107	1	Score ≤ 3 points
Previous hospitalization[a]	2.45 (1.51; 3.99)	0.895	<0.001	3	
Crackles	2.35 (1.45; 3.89)	0.856	0.001	2	
Paroxysmal nocturnal dyspnoea	2.25 (1.41; 3.59)	0.809	0.001	2	Medium risk
Orthopnoea	2.16 (1.36; 3.46)	0.770	0.001	2	Scores = 4–7 points
NYHA class III–IV	2.11 (1.17; 3.94)	0.745	0.016	2	
Worsening in NYHA status	2.50 (1.46; 4.45)	0.917	0.001	3	High risk
Heart rate > 100 b.p.m.	2.72 (1.46; 5.11)	1.002	0.002	3	Scores ≥ 8
Oxygen saturation ≤ 90%	4.98 (2.16; 12.18)	1.606	<0.001	5	

CI, confidence interval; NYHA, New York Heart Association.
[a]Due to heart failure decompensation (between 31 and 365 days prior to inclusion).

Figure 3 Predicted probabilities of hospitalization or death 30 days after an episode of decompensation in heart failure patients.



Comparison with existing literature

In comparison with other authors, the participants in this real-world, observational study were older, there was a higher proportion of women, and a greater number of heart failure patients with preserved ejection fraction.[7,8] All of which concurs with the characteristics of patients treated in the community.[1]

Most scores predicting heart failure decompensation events have been performed in hospitalized patients and included variables not available in routine primary care practice. The discrimination capacity of such models varies between 0.54 and 0.86.[8] Systematic reviews[7,8] have considered that, among the studies analysed, only three of the scores[13–15] were accurate and had been properly validated. The predicted outcome most frequently used was 30 day mortality.

Table 4 Number of hospitalizations or deaths 30 days after episode of decompensation in heart failure patients, among the total of patients in each stratum, and cumulative incidence, according to derivation and validation cohorts and risk score groups

Risk score groups (score points)	Derivation cohort Number of patients (cumulative incidence)	Validation cohort Number of patients (cumulative incidence)
Low risk (≤3 points)	2/73 (2.7%)	2/27 (9.09%)
Medium risk (between 4 and 7 points)	21/163 (12.8%)	7/54 (12.9%)
High risk (≥8 points)	148/322 (46.22%)	50/126 (39.68%)

Among the scores that predict short-term hospitalization, the RENDISCORE study[12] elaborated a model to predict readmission at 1 month and was based on previously hospitalized patients, some of whom were recruited in primary care. It included natriuretic peptides, symptoms and signs of left heart failure, and glomerular filtration rate < 60 mg/dL. The hospital readmission percentage was 3.1%, which was lower than our findings. This may be explained by patients being incorporated at hospital discharge and, as a consequence, clinically stable at the inclusion date.[12]

The MEESSI model[15] was developed and validated to stratify risk of death of heart failure patients 1 month after being attended in a hospital emergency room. For patients with characteristics similar to those of our participants, it had, however, low predictive capacity.

The general triage scales in emergency departments such as the Canadian Triage Acuity Scale (CTAS), Manchester Triage System (MTS), and Triage Andorran Model - Triage Spanish System (MAT-SET) have not demonstrated usefulness for the prediction of hospitalization or death at 30 days in patients with heart failure decompensation.[16]

Thibodeau and Drazner[17] showed that clinical examination using variables such as those included in our model is crucial

in the management of heart failure patients. Such evaluations provide crucial prognostic information and may help to guide decision-making.

Regarding the variables included in our model, hospital admission in the previous year has been shown in multiple studies to be a risk factor for hospitalization.[1,18,19] NYHA functional class III or IV[15] and reduced oxygen saturation[12–15,19–21] as well as elevated heart rates[13,14,22] have also been included in other models. In addition, symptoms of left ventricle overload have been reported to be predictive for both heart failure diagnosis and hospitalization.[12,15]

The elevated mean age of patients included in our population may explain the differences in prognosis as this variable was not statistically significant. Because of the narrow interval of age, and most patients being older than 75 years, these differences were not expected unlike other studies with a wider age range.

An international consensus document on heart failure management has provided guidelines from an emergency medical perspective. It has been suggested that not more than 2% of patients discharged for a heart failure decompensation should die in the first 30 days and that the rate of readmissions should be lower than 10%. The lack of risk stratification of these patients before decision-making has, however, been identified as one of the possible reasons for poorer outcomes.[11]

Most patients who were hospitalized or died had been initially referred by primary care professionals to the emergency room. Nevertheless, a considerable percentage of those not referred at the moment of initial evaluation finally needed to be admitted to hospital (12.5%) and 2% died. These figures clearly exceeded the benchmarks in the consensus recommendations.[11] We consider that the proposed score may improve prioritization with respect to referring patients to hospital and intensify clinical follow-up of those at moderate or high risk (more than 3 points in the score).

Strengths and limitations

Given the characteristics of the population attended in primary care, such as advanced age, high proportion of women, and predominance of preserved ejection fraction, the external validity of this score is applicable in settings and geographical areas with similar characteristics to those of the present study. Validation studies will be required to apply the score in other populations with different characteristics to the patients included in the present study. Nevertheless, the high proportion of women, unlike other previously published scoring systems, is a strength of our study.

Many countries participated; there may therefore be differences in protocols and healthcare system characteristics with respect to criteria in managing heart failure decompensation. Nevertheless, such variability could also be seen as a strength, because we aimed to validate the model in differing countries and medical systems. It is possible that we included patients presenting more symptomatic decompensations and neglected those with less severe decompensations that might affect the prediction score.

Given the characteristics of an observational study carried out in real clinical practice in primary care and the considerable number of researchers involved, it has been difficult to ensure strictly consecutive recruitment. Whilst this could imply a selection bias, it does not, however, affect the validity of the model.

Conclusions

The HEFESTOS score is a usable tool based on variables easily collected in primary care. It stratifies the risk of death or hospitalization in heart failure patients presenting decompensation.

Conflict of interest

None declared.

Funding

This work was supported by the Carlos III Institute of Health, Ministry of Economy and Competitiveness (Spain) (Instituto de Salud Carlos III), awarded on the 2014 call under the Health Strategy Action 2013–2016, within the National Research Program oriented to Societal Challenges, within the Technical, Scientific and Innovation Research National Plan 2013–2016, co-funded with European Union ERDF funds (PI14/00583) and The European General Practice Research Network (EGPRN) (HP/2015.036).

Supporting information

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

Data S1. Supporting information.

References

1. Farré N, Vela E, Clèries M, Bustins M, Cainzos-Achirica M, Enjuanes C, Moliner P, Ruiz S, Verdú-Rotellar JM, Comín-Colet J. Real world heart failure epidemiology and outcome: a population-based analysis of 88,195 patients. *PLoS ONE* 2017; **12**: e0172745.
2. Heidenreich PA, Albert NM, Allen LA, Bluemke DA, Butler J, Fonarow GC, Ikonomidis JS, Khavjou O, Konstam MA, Maddox TM, Nichol G, Pham M, Piña IL, Trogdon JG. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. *Circ Heart Fail* 2013; **6**: 606–619.
3. Farré N, Vela E, Clèries M, Bustins M, Cainzos-Achirica M, Enjuanes C, Moliner P, Ruiz S, Verdú-Rotellar JM, Comín-Colet J. Medical resource use and expenditure in patients with chronic heart failure: a population-based analysis of 88 195 patients. *Eur J Heart Fail* 2016; **18**: 1132–1140.
4. Shoaib A, Farag M, Nolan J, Rigby A, Patwala A, Rashid M, Kwok CS, Perveen R, Clark AL, Komajda M, Cleland JGF. Mode of presentation and mortality amongst patients hospitalized with heart failure? A report from the First Euro Heart Failure Survey. *Clin Res Cardiol* 2019; **108**: 510–519.
5. Gheorghiade M, Zannad F, Sopko G, Klein L, Piña IL, Konstam MA, Massie BM, Roland E, Targum S, Collins SP, Filippatos G, Tavazzi L, International Working Group on Acute Heart Failure Syndromes. Acute heart failure syndromes: current state and framework for future research. *Circulation* 2005; **112**: 3958–3968.
6. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJ, Falk V, González-Juanatey JR, Harjola VP, Jankowska EA, Jessup M, Linde C, Nihoyannopoulos P, Parissis JT, Pieske B, Riley JP, Rosano GM, Ruilope LM, Ruschitzka F, Rutten FH, van der Meer P, Authors/Task Force Members; Document Reviewers. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur J Heart Fail* 2016; **18**: 891–975.
7. Michaud AM, Parker SIA, Ganshorn H, Ezekowitz JA, McRae AD. Prediction of early adverse events in emergency department patients with acute heart failure: a systematic review. *Can J Cardiol* 2017; **34**: 168–179.
8. Miró Ò, Rossello X, Platz E, Masip J, Gualandro DM, Peacock WF, Price S, Cullen L, DiSomma S, de Oliveira MT Jr, McMurray JJ, Martín-Sánchez FJ, Maisel AS, Vrints C, Cowie MR, Bueno H, Mebazaa A, Mueller C, Study Group on Acute Heart Failure of the Acute Cardiovascular Care Association of the European Society of Cardiology. Risk stratification scores for patients with acute heart failure in the Emergency Department: a systematic review. *Eur Heart J Acute Cardiovasc Care* 2020; **9**: 375–398.
9. Miró Ò, Gil V, Rosselló X, Martín-Sánchez FJ, Llorens P, Jacob J, Herrero P, Herrera Mateo S, Richard F, Escoda R, Fuentes M, Martín Mojarro E, Llauger L, Bueno H, Pocock S. Patients with acute heart failure discharged from the emergency department and classified as low risk by the MEESSI score (multiple risk estimate based on the Spanish emergency department scale): prevalence of adverse events and predictability. *Emergencias* 2019; **31**: 5–14.
10. Sullivan LM, Massaro JM, D'Agostino RB Sr. Presentation of multivariate data for clinical use: the Framingham Study risk score functions. *Stat Med* 2004; **23**: 1631–1660.
11. Miró Ò, Levy PD, Möckel M, Pang PS, Lambrinou E, Bueno H, Hollander JE, Harjola VP, Diercks DB, Gray AJ, DiSomma S, Papa AM, Collins SP. Disposition of emergency department patients diagnosed with acute heart failure: an international emergency medicine perspective. *Eur J Emerg Med* 2017; **24**: 2–12.
12. Alvarez-García J, Ferrero-Gregori A, Puig T, Vázquez R, Delgado J, Pascual-Figal D, Alonso-Pulpón L, González-Juanatey JR, Rivera M, Worner F, Bardají A, Cinca J, Investigators of the Spanish Heart Failure Network (REDINSCOR). A simple validated method for predicting the risk of hospitalization for worsening of heart failure in ambulatory patients: the Redin-SCORE. *Eur J Heart Fail* 2015; **17**: 818–827.
13. Lee DS, Stitt A, Austin PC, Stukel TA, Schull MJ, Chong A, Newton GE, Lee JS, Tu JV. Prediction of heart failure mortality in emergent care: a cohort study. *Ann Intern Med* 2012; **156**: 767–775.
14. Stiell IG, Clement CM, Brison RJ, Rowe BH, Borgundvaag B, Aaron SD, Lang E, Calder LA, Perry JJ, Forster AJ, Wells GA. A risk scoring system to identify emergency department patients with heart failure at high risk for serious adverse events. *Acad Emerg Med* 2013; **20**: 17–26.
15. Miró Ò, Rossello X, Gil V, Martín-Sánchez FJ, Llorens P, Herrero-Puente P, Jacob J, Piñera P, Mojarro EM, Lucas-Imbernón FJ, Llauger L, Agüera C, López-Díez MP, Valero A, Bueno H, Pocock SJ, ICA-SEMES Research Group. Analysis of how emergency physicians' decisions to hospitalize or discharge patients with acute heart failure match the clinical risk categories of the MEESSI-AHF scale. *Ann Emerg Med* 2019; **74**: 204–215.
16. Miró Ò, Tost J, Herrero P, Jacob J, Martín-Sánchez FJ, Gil V, Fernández-Pérez C, Escoda R, Llorens P, ICA-SEMES Research Group. Short-term predictive capacity of two different triage systems in patients with acute heart failure: TRICA-EAHFE study. *Eur J Emerg Med* 2016; **23**: 435–441.
17. Thibodeau JT, Drazner MH. The role of the clinical examination in patients with heart failure. *JACC Heart Fail* 2018; **6**: 543–551.
18. Garcia-Gutierrez S, Quintana JM, Antón-Ladislao A, Gallardo MS, Pulido E, Rilo I, Zubillaga E, Morillas M, Onaindia JJ, Murga N, Palenzuela R, Ruiz JG, AHFRS Group. Creation and validation of the acute heart failure risk score: AHFRS. *Intern Emerg Med* 2017; **12**: 1197–1206.
19. Lee DS, Schull MJ, Alter DA, Austin PC, Laupacis A, Chong A, Tu JV, Stukel TA. Early deaths in patients with heart failure discharged from the emergency department: a population-based analysis. *Circ Heart Fail* 2010; **3**: 228–235.
20. Jacob J, Miró Ò, Herrero P, Martín-Sánchez FJ, Gil V, Tost J, Aguirre A, Escoda R, Alquézar A, Andueza JA, Llorens P, en representación del grupo ICA-SEMES. Predicting short-term mortality in patients with acute exacerbation of chronic heart failure: the EAHFE-3D scale. *Med Intensiva* 2016; **40**: 348–355.
21. Collins SP, Jenkins CA, Harrell FE Jr, Liu D, Miller KF, Lindsell CJ, Naftilan AJ, McPherson JA, Maron DJ, Sawyer DB, Weintraub NL, Fermann GJ, Roll SK, Sperling M, Storrow AB. Identification of emergency department patients with acute heart failure at low risk for 30-day adverse events: the STRATIFY decision tool. *JACC Heart Fail* 2015; **3**: 737–747.
22. Greig D, Austin PC, Zhou L, Tu JV, Pang PS, Ross HJ, Lee DS. Ischemic electrocadiographic abnormalities and prognosis in decompensated heart failure. *Circ Heart Fail* 2014; **7**: 986–993.