Predicting 30-day Mortality for Patients with Acute Heart Failure Who Are in the Emergency Department: A Cohort Study

Óscar Miró, PhD1, Xavier Rosselló, MD2,3, Víctor Gil, MD3, Francisco Javier Martín-Sánchez, PhD4, Pere Llorens, PhD5, Pablo Herrero-Puente, PhD6, Javier Jacob, PhD7, Héctor Bueno, PhD1,8, Stuart J. Pocock, PhD2,3 on behalf of the ICA-SEMES (Acute Heart Failure of the Spanish Society of Emergency Medicine) Research Group9

1Emergency Department, Hospital Clínic, Barcelona, Catalonia, Spain
2Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, United Kingdom
3National Centre for Cardiovascular Research (CNIC), Madrid, Spain
4Emergency Department, Hospital Clínico San Carlos, Universidad Complutense, Madrid, Spain
5Emergency Department, Short-Stay Unit and Home Hospitalization, Hospital General de Alicante, Spain
6Emergency Department, Hospital Universitario Central de Asturias, Oviedo, Spain
7Emergency Department, Hospital Universitari de Bellvitge, L’Hospitalet de Llobregat, Catalonia, Spain
8Cardiology Department, Hospital Universitario 12 de Octubre, Universidad Complutense, Madrid, Spain
9Consult Appendix 1 to see the complete list of the ICA-SEMES Research Group (Research group on Acute Heart Failure of the Spanish Society of Emergency Medicine) members

Word count for abstract: 272
Word count for text: 3289

Address for correspondence:
Dr. Óscar Miró,
Emergency Department, Hospital Clínic,
Villarroel 170, 08036 Barcelona, Catalonia, Spain.
Email: omiro@clinic.cat
Phone/FAX numbers: (+34) 93.227.98.33 / 93.227.56.93
Abstract

Background: Physicians in the emergency department (ED) need additional tools to stratify patients with acute heart failure (AHF) according to risk.

Objective: To predict future mortality from data readily available on ED admission.

Design: Prospective cohort study.

Setting: 34 Spanish EDs

Participants: 4867 consecutive ED patients admitted during 2009-2011 for the derivation cohort and 3229 patients admitted in 2014 for the validation cohort.

Measurements: Candidate risk factors and 30-day mortality.

Results: We found 13 independent risk factors in the derivation cohort and combined them to form an overall score, which we call the MEESSI-AHF (Multiple Estimation of risk based on the Emergency department Spanish Score In patients with AHF) score. This score predicted 30-day mortality with excellent discrimination (c-statistic=0.836) and calibration (Hosmer-Lemeshow P = 0.99), and it provided a steep gradient in 30-day mortality across risk groups (<2% mortality for patients in the 2 lowest risk quintiles and 45% mortality in the highest risk decile). We confirmed these characteristics in the validation cohort (for example, c-statistic=0.828). Multiple sensitivity analyses failed to find important amounts of confounding or bias.

Limitations: The study was confined to a single country. Participating EDs were not selected randomly. Many patients had missing data. Measuring some risk factors was subjective.

Conclusion: This tool has excellent discrimination and calibration, and it was validated in patients different from the patients used to develop it. We think physicians can consider using this tool to inform clinical decisions as we conduct further studies to determine whether the tool enhances physician decisions and improves patient outcomes.
Primary Funding Source: Spanish Ministry of Health, Catalonia Govern, Fundació Marató-TV3.

Keywords: acute heart failure, risk score, outcomes.
Introduction

Annual hospital admissions due to acute heart failure (AHF) in Europe and the USA exceed 1 million in each region and account for most of the costs of heart failure-related care (1, 2). The emergency department (ED) has a central position in the management of AHF, since about 90% of patients with this condition attend an ED to improve their symptoms (3, 4). Once initial treatments have been administered in the ED and their effects checked, decisions are made regarding subsequent patient management: specifically does the patient need to be hospitalized or can they be discharged home with proper treatment and follow-up. As a result of a mainly subjective, empirically-driven assessment, a highly variable proportion of AHF patients is currently being directly discharged to home from ED: 16.3% in US (5), 23.9% in Spain (4), and 36.2% in Canada (6).

Although decision-making in the ED is of critical importance, emergency physicians currently are not stratifying patient by risk during this process. Some biomarkers, for example, heart-specific markers like natriuretic peptides and troponin or non-specific markers like glucose or creatinine, are associated with prognosis, but cannot by themselves predict outcomes with sufficient reliability to aid decision-making (7,8). Alternatively, several AHF risk scores have been developed (9,10), but these scores have been based on hospitalized patients thus ignoring the many AHF patients, more than a third in certain countries (6), who are entirely managed in the ED and discharged home. To our knowledge, only 3 risk scores have been developed specifically for use in the ED: 2 in Canada (the Ottawa Heart Failure Risk Scale, OHFRS, and the Emergency Heart Failure Mortality Risk Grade, EHMRG) (11,12) and 1 in United States (The Improving Heart Failure Risk Stratification in the Emergency Department, STRATIFY, scale) (13). However, some were not externally validated (OHFRS, STRATIFY), some were constructed from administrative data (OHFRS, EHMRG), some excluded a substantial portion of patients (EHMRG: palliative patients excluded; OHFRS: non-consecutive sample with multiple exclusion criteria), and some were derived from databases of limited size (OHFRS: 557 patients; STRATIFY: 1033 patients). Therefore, we believe there is a need for additional tools to help physicians in the ED stratify patients with acute heart failure (AHF) according to risk.

Methods

The Acute Heart Failure in Emergency Departments (EAHFE) Registry
The EAHFE Registry collects detailed information on consecutive patients attending 34 Spanish EDs with a final diagnosis of AHF (14,15). Hospitals participate in the EAHFE Registry voluntarily, and they include university and community hospitals, EDs with high, medium or low volume of attendances (>300, 200-300, or <200/day, respectively), and hospitals from all areas of the country. Attending emergency physicians use Framingham’s clinical diagnostic criteria (16) to identify patients for the registry. Thereafter, the diagnosis is double-checked by the principal investigator of each centre, who makes the final adjudication of AHF diagnosis based on the review of medical charts and all complementary tests done during the ED stay and any hospitalization. The diagnosis was confirmed by natriuretic peptide determinations or echocardiography (17) in the 92% of patients included in the EAHFE Registry. The only exclusion criteria to be included in the EAHFE Registry is a diagnosis of ST-elevation myocardial infarction, which occurred in approximately 3% of patients.

For every patient, data on demographics, clinical history, presentation and treatments were routinely collected on specific case record forms. Interventions, treatments and patient placements (hospital admission or discharge) were entirely based on the decision-making of the attending emergency physician. Subsequent follow-up, through telephone contact and consultation of medical records, was performed between day 31 and 90. The EAHFE Registry complies with the Declaration of Helsinki and was approved by the Ethical Committees of all participating centres, and all patients gave informed consent. Around 2% of patients fulfilling inclusion and exclusion criteria refused to participate.

Study design

During the design of the EAHFE Registry, we planned to develop a model that could stratify patients according to their risk of experiencing adverse outcomes. We wanted this model to be used as soon as possible after arrival in the ED by the first emergency physician who saw the patient using variables routinely available in most EDs. We named this model MEESSI-AHF (Multiple Estimation of risk based on the Emergency department Spanish Score In patients with AHF).

When developing the model, we selected registry patients from May 2009 and November-December 2011 for the derivation cohort and patients from January-February 2014 for the validation cohort (Figure S1). We used patients in the derivation cohort to generate a 30-day mortality risk model and we used patients from the validation cohort to measure how stable the model was.
Data analysis

We first identified over 88 candidate predictor variables (Supplemental Table S1) that described baseline demographics, medical history and status at admission and could potentially have prognostic implications. To develop the risk score, we used logistic regression (without interaction terms) with checks for non-linearity and forward stepwise variable selection with an entry criterion of \( p<0.01 \). We used multiple imputation with chained equations (18) to produce 50 imputed data sets for estimating missing values. Once we identified a predictor, we then identified a cut-off value based on our clinical information about the predictor’s value (e.g., serum potassium) or about the linear trend (e.g., serum creatinine and systolic blood pressure). In the final model, we formed each continuous variable into ordered categories to facilitate their use in practice. We measured the model’s discrimination with the c-statistic, and we measured the model’s calibration by comparing observed- versus model-derived mortality risk with the Hosmer-Lemeshow statistic. We conducted sensitivity analyses by type of hospital (university vs community), by daily ED census (low-medium vs high volume), and for alternative models that did not include values for Barthel index, NT-proBNP, or troponin (in any combination) because they can be those more frequently be lacking in certain ED or in certain circumstances. We compared our model with the EHRMG model (12) in a merged data set of both derivation and validation cohorts by comparing the areas under the ROC curves for 30-day mortality with the DeLong test. We used STATA software, version 13.1 (Stata Corp, College Station, TX, USA) for all analyses.

Role of the funding source

This study was partially supported by competitive grants from the Institute de Salud Carlos III supported with funds from Spanish Ministry of Health, (PI10/01918, PI11/01021, PI15/01019 and PI15/00773), Fundació La Marató de TV3 (20152510), and Catalonia Govern (GRC 2009/1385 and 2014/0313). The funding source had no role in the design, conduct, and analysis of this study or in the decision to submit the manuscript for publication.

Results
The study derivation cohort comprises 4897 consecutive patients admitted to an ED with AHF during May 2009 and Nov-Dec 2011 (Figure S1). Thirty patients were excluded from analysis due to lack of follow-up, while those with censored data (48 patients with <30 days of follow-up) were included. Patients had a mean age 79.7 years, 57.1% were females, comorbidities were very frequent (83.4% had hypertension, 42.2% diabetes mellitus, 39.4% dyslipidemia, 29.9% ischemic cardiomyopathy), 89.5% had New York Heart Association (NYHA) class III-IV and 56.5% had some dependency (Barthel index <100 points) at ED arrival, and 41.5% patients had LVEF below 50%, with 52.4% of them receiving beta-blockers, 62.9% angiotensin-converter enzyme inhibitor or angiotensin-II receptor blocker, and 29.1% mineralocorticoid-receptor antagonist. Patients subsequently hospitalized (75.6%) had a median length of stay of 7 days. The rest of the characteristics of the study population are presented in Table 1.

Within 30 days of admission, 500 patients (10.3%) had died. From all of the candidate predictors, a logistic regression model was used with forward stepwise variable selection to identify the final 13 highly significant independent death predictors included into the MESSI-AHF risk score. These variables are listed in Table 1 ordered by their statistical strength of prediction (i.e. Barthel index at admission is the most highly significant) and each odds ratio is adjusted for all the other variables. Figure S2 displays the independent impact of each predictor on mortality risk based on the model in Table 1, and Table S2 shows comparison in key predictor variables in patient with and without missing values.

For any patient, one adds together their relevant risk coefficients plus the intercept coefficient in Table 1 to determine the multivariable risk score, which is the patient’s predicted log (odds) of dying within 30 days. The distribution of this risk score for all 4867 patients is shown in Figure 1. Also, the curve in Figure 1 relates a patient’s risk score to the probability of dying within 30-day of admission, which ranges from 0.005 to 0.898 with a median of 0.051. To facilitate the calculation of any patient’s risk of dying within 30 days, we have set up a website http://bernalte.cat/calculadora/ for a specific patient one enters the relevant 13 items and immediately their predicted % risk of dying within 30 days is provided.

Figure 2 shows the cumulative mortality over 30 days for patients classified into 6 risk groups: the bottom 4 quintiles and the top two deciles of the risk score’s distribution in this derivation cohort. Good discrimination of the model was achieved, with c-statistic 0.836 (95% CI 0.818 -0.853). There was a steep gradient in 30-day
mortality across risk groups: with 45% mortality for the top decile of risk and around 0.7% for the bottom quintile of risk. Similar discrimination capacity was observed in either university or community hospitals, as well as in low-medium or high-volume ED (Table 2). In this derivation cohort Figure 3(a) depicts the model goodness-of-fit, comparing observed and model-predicted 30-day mortality risk across the 6 risk groups. A useful nomenclature is as follows: low risk (first and second quintiles), intermediate risk (third and fourth quintiles), high risk (next decile) and very high risk (top decile). Sensitivity and specificity of the every risk threshold for each category plotted on a ROC curve is presented in Figure S3. Reduced models lacking Barthel index, troponin or NT-proBNP (in any combination) also showed good discriminatory capacity, ranging from 0.829 and 0.784 (Table S3). Accordingly, they have been incorporated in the website calculator.

Finally, we used 3229 patients recruited during Jan-Feb 2014 to validate our risk score on an external population of patients, 299 (9.26%) dying within 30 days of ED admission. Five patients of the validation cohort were excluded from analysis due to lack of follow-up, while six patients with less than 30 days follow-up were included. Comparisons for key predictor variables between derivation and validation cohorts are shown in Table S4. Distribution of the MEESSI-AHF scores is presented in Figure S4. In this validation cohort, Figure 3(b) compares the observed and model-predicted mortality in six risk groups (from lowest quintile to top decile). The model fit and extent of risk discrimination is very similar to what was found in the derivation cohort. The c-statistic in the validation cohort is 0.828 (95% CI 0.802-0.853), very similar to that achieved in model development. To check goodness of model fit, the Hosmer-Lemeshow test for the derivation cohort was P=0.99, and for the validation cohort P= 0.122. When compared with the previously developed risk score EHMRRG intended for 7-day mortality prediction (12) using a same sample of patients of the present study, the MEESSI-AHF had superior discrimination overall (c-statistic, 0.830 vs. 0.750; P<0.001; Figure S5).

Discussion

The findings we present in this study are based on a large prospective population-based cohort of consecutive AHF patients admitted to 34 hospital EDs across Spain. Patients with many types of AHF were included, except for those developing AHF during an ST-elevation myocardial infarction, and all data were recorded shortly after arrival in the ED. The 13 predictors of 30-day mortality we identified should all be promptly available in routine
clinical practice worldwide; and we have provided a web calculator [http://bernalte.cat/calculadora/] to make it easier for physicians to calculate the risk for a specific patient. Using such a calculator, emergency physicians will now be able to determine whether a patient is at high (or low) risk of dying within 30 days which, in turn, might allow for better patient management. Our score may be particularly useful in the 10% of patients at very high risk for 30-day mortality (around a 45%), as well as in the 40% of patients at low risk for 30-day mortality (<2%). Identification of both groups has important management implications. For a patient with very high risk, special attention has to be focused on ensuring that the patient and relatives are aware of the severity and, assuming they are appropriate, on prompt aggressive treatments with an emphasis on early admission to an intensive care unit. For a patient with low risk, attention should be focused on treatment that will lead to early discharge from the ED to home, which is consistent with a recent consensus about patients with <2% all-cause mortality as long as they are observed long enough in the ED (19).

In the US, the overall incidence rate of heart failure hospitalizations has declined 29.5% between 1998 and 2007 (20). We suggest that this decline could be due to better ambulatory care that avoids patient decompensation and allows proper treatment of less severe AHF episodes without hospital admission. In this sense, there is an increasing perception that more AHF patients at low risk of adverse outcomes should avoid hospitalization (4, 21), and recent consensus opinions by clinical experts advocate that approach (19,22).

Specifically, one group recommends rates of 20% to 40% direct ED discharge for patients being diagnosed with AHF (depending whether the ED lacks or possesses, respectively, a specific observation area) (19). These figures match well to patients in our low risk category (40%). Avoiding hospital admission is not only a matter of health care system efficiency improvement that could save substantial costs. Hospitalization itself could imply some potential hazards: nosocomial infection, increased errors in patient with polypharmacy, acute reactive psychosis and deteriorating functional status are quite common amongst the elderly being hospitalized. AHF patients are typically of advanced age, with a median age around 80 years in most series (4,12) (median 80 years in our cohort). However, we are not aware of any formal tools that are currently being used to aid ED risk stratification for AHF patients. Thus, some authors have argued that direct discharge of patients without objectively-based risk stratification is putting some patients at an unacceptably high risk of adverse events (6,23). This situation contrasts with improvements achieved in other high prevalent ED conditions, such as community-acquired pneumonia and acute coronary syndromes, where risk scores have
been developed (24,25) and are being widely applied to discharge less severe patients who previously would have been admitted to hospital. We believe that the **MEESSI-AHF risk score** can provide similar help in the management of patients with AHF, especially for elderly patients who are more challenging to evaluate (15).

All 13 variables we found to be predictive have been repeatedly reported as influencing the prognosis of patients with AHF (1,11-13,15,26-28). However, in our study 4 of these variables had more than 25% missing values. We adjusted for these missing values using a multiple imputation technique. Moreover, in order to match our score to what happens when real patients are in EDs, our website calculator provides a risk score even when values for Barthel index, troponin levels, and NT-ProBNP are not available and we have shown that these risk scores perform as well as the regular risk scores (Table S4).

Our model compares favourably with other risk models. For example, our model had a c-statistic of 0.836 in the derivation cohort and 0.828 in the validation cohort, which were higher than the comparable value when we calculated the EHMRG score in 2137 patients who had all the data necessary to calculate an EHMRG score. The EHMRG model focused on a shorter-term perspective (7-day mortality) (12). We feel a longer perspective (30-day mortality) provides a better framework to create a model to aid emergency physicians. Moreover, EHMRG score excluded palliative patients (who have a higher risk of adverse events), and that could limit its generalizability. Certainly, patients only receiving palliative care are not uncommon: e.g. 10.2% of our patient had a Barthel index of 0 to 20 points (indicating complete dependence) and an additional 32.8% had a Barthel index between 21 and 60 points (indicating severe dependence) and, although not directly recorded in our study, for many of them palliative care could apply. However, we have previously demonstrated that the exclusion patients for whom palliative care could potentially apply did not significantly change the discriminatory capacity of the model (only decreased from 0.741 to 0.729) (29). Our findings, in line with previous works in this field (30), affirm that the Barthel index is a key outcome predictor, adding value to previously developed risk scores. Thus, it is important to recognize that patient frailty and dependence are key aspects that should be considered in every disease impacting on an elderly population, as it comes about AHF patients. Finally, our model has been developed using data prospectively recorded using a standardised pro forma at the time of admission to the ED, instead of using retrospective extraction from administrative reports, as was done for the EHMRG model. The latter strategy could limit reliability and completeness of data.

All the above-mentioned limitations, even with more extensive patient exclusion criteria and smaller sample,
also apply for the OHFRS model, which obtained a c-statistic of 0.77 (11). On the other hand, although the
STRATIFY score (13) was developed using data recorded prospectively, it was derived with a limited number of
cases, no external validation was done, and got moderate discriminatory capacity (c-statistic: 0.68) (13).
Therefore, for the first time, we offer a risk-model with robust data from a large-scale population-based study
to quickly assess patient prognosis.

Our study has important limitations. Some important predictors had a high number of missing values, which
we have addressed with multiple imputation techniques and sensitivity analyses. There is a possibility of a
“false positive” predictor entering the risk model after testing 88 candidate predictor variables, although use
of p<0.01 as entry criterion has minimised this risk. Some variables, e.g. Barthel index, NYHA class, association
with ACS, or low cardiac output, are partially based on subjective interpretation, but we tried to reduce this
problem by providing all research centers with a dictionary for all variables and holding meetings with all
researchers just before each recruitment phase in an attempt to minimize inconsistency. Additionally, the
precision of our model might change in the future, especially if new treatments for heart failure were able to
modify mortality, such as angiotensin II receptor blocker neprilysin inhibitors, which were not available when
this study was performed. Finally, as for any study in a single country, caution should be taken in extrapolating
findings internationally. Moreover, EDs were not randomly selected but were participants of the EAHFE
Registry, with special interest in AHF, so it is possible results could differ when applied to other EDs. Thus, we
courage others to explore validation of our risk model in other countries/regions. Nonetheless, we believe
that our model has the potential for being used widely.

In conclusion, our study demonstrates that physicians can use 13 readily available items to estimate individual
risk of 30-day mortality for patients with AHF who are admitted to the ED. With strong risk discrimination,
good model fit and external validation, this tool is now ready for clinical use. Further study is needed to
elucidate the real potential of the MEESSI-AHF risk score for enhancing physician behaviour and improving
patient outcomes. We have provided user-friendly access to a way of calculating scores for specific patients
(http://bernalte.cat/calculadora/). This tool has excellent discrimination and calibration, and it was validated
in patients different from the patients used to develop it. We think physicians can consider using this tool to
inform clinical decisions as we conduct further studies to determine whether the tool enhances physician
decisions and improves patient outcomes. We believe that this tool will be especially useful for identifying individuals at lower risk for whom further hospitalization may be not required.
Acknowledgements: All participants in the ICA-SEMES Research Group.

Conflict of interests: The authors state that they have no conflict of interests with the present work. The ICA SEMES Research Group has received unrestricted support from Otsuka, Orion Pharma and Novartis, and the present work was designed, performed, analysed and written exclusively by the authors independently of these pharmaceutical laboratories.
References


17. McMurray JJ, Adamopoulos S, Anker SD, Auricchio A, Böhm M, Dickstein K, et al. ESC Committee for Practice Guidelines. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. *Eur Heart J* 2012; **33**:1787-847. PMID: 22611136.


