



Original Article

External Validation and Refinement of Emergency Heart Failure Mortality Risk Grade Risk Model in Patients With Heart Failure in the Emergency Department

Nariman Sepehrvand, MD,^{a,b} Erik Youngson, MMath,^c Jeffrey A. Bakal, PhD,^c
Finlay A. McAlister, MD, MSc,^{c,d} Brian H. Rowe, MD, MSc,^e and
Justin A. Ezekowitz, MBCh, MSc^{a,f,g}

^a Canadian VIGOUR Centre, University of Alberta, Edmonton, Alberta, Canada

^b Department of Medicine, University of Alberta, Edmonton, Alberta, Canada

^c Patient Health Outcomes Research and Clinical Effectiveness Unit, University of Alberta, Edmonton, Alberta, Canada

^d Division of General Internal Medicine, Department of Medicine, University of Alberta, Edmonton, Alberta, Canada

^e Department of Emergency Medicine and School of Public Health, University of Alberta, Edmonton, Alberta, Canada

^f Division of Cardiology, Department of Medicine, University of Alberta, Edmonton, Alberta, Canada

^g Mazankowski Alberta Heart Institute, Edmonton, Alberta, Canada

ABSTRACT

Background: Emergency Heart Failure Mortality Risk Grade (EHMRG) assesses the risk of death within 7 days of emergency department (ED) presentation for patients with acute heart failure (AHF). We aimed to externally validate and refine the EHMRG model in patients who presented to the ED with AHF.

Methods: We performed a cohort study using administrative data for all ambulance-transported patients from Alberta (2012-2016) presenting to the ED with a primary diagnosis of AHF.

Results: Among 6708 patients with AHF, the 7-day mortality was 0.0%, 0.8%, 1.6%, 4.0%, 4.2%, and 12.0% across EHMRG risk categories (1-4, 5A and 5B). The EHMRG score had a c-index of 0.73 (95% confidence interval [CI], 0.71-0.76) for 7-day mortality and 0.71 (95% CI, 0.70-0.73) for 30-day mortality, but lower c-statistics for other outcomes (0.61-0.67). The inclusion of natriuretic peptides to the EHMRG model improved prediction (Net Reclassification Improvement, 0.268; 95% CI, 0.173-0.363; $P < 0.01$) for 7-day mortality, as did the

RÉSUMÉ

Introduction : Le Emergency Heart Failure Mortality Risk Grade (EHMRG) permet d'évaluer le risque de mortalité dans les 7 jours qui suivent la consultation au service des urgences (SU) des patients souffrant d'insuffisance cardiaque aiguë (ICA). Nous avons pour objectif de valider et d'affiner à l'externe le modèle EHMRG chez les patients qui se présentaient au SU en raison d'une ICA.

Méthodes : Nous avons réalisé une étude de cohorte à partir des données administratives de tous les patients de l'Alberta transportés par ambulance (2012-2016) au SU en raison d'un diagnostic principal d'ICA.

Résultats : Parmi les 6708 patients souffrant d'ICA, la mortalité dans les 7 jours était de 0,0 %, de 0,8 %, de 1,6 %, de 4,0 %, de 4,2 % et de 12,0 % dans toutes les catégories de risque EHMRG (1-4, 5A et 5B). Le score EHMRG montrait un C-index (indice de concordance) de 0,73 (intervalle de confiance [IC] à 95 %, 0,71-0,76) pour la mortalité dans les 7 jours et de 0,71 (IC à 95 %, 0,70-0,73) pour la mortalité dans les

Heart failure (HF) is a major public health problem, with frequent emergency department (ED) visits and hospital admissions, and increased mortality.¹ Given the high burden of acute heart failure (AHF), the paucity of effective therapies,

and wide practice variation with respect to disposition, risk-stratification tools for use in the ED to predict the short-term mortality risk have been developed using cohort studies.²⁻⁴ Despite promise, none are in widespread clinical use, and few incorporate currently available biomarkers in widespread use such as natriuretic peptides (NPs).⁵ One high-quality score, derived and validated in a population-based setting with good discriminative properties, is the Emergency Heart Failure Mortality Risk Grade (EHMRG), which was developed to identify a patient's 7-day mortality risk post-ED presentation.³ However, it is unclear whether the inclusion of NPs or a commonly accepted 5-level triage tool (Canadian Triage Acuity Score)⁶ would add to the score's

Received for publication February 12, 2019. Accepted March 1, 2019.

Ethics Statement: The research reported in this paper adhered to relevant ethical guidelines.

Corresponding author: Dr Justin A. Ezekowitz, The Canadian VIGOUR Centre, 2-132 Li Ka Shing Center for Health Innovation, Edmonton, Alberta T6G 2E1, Canada. Tel.: +1-780-492-0592; fax: +1-780-492-9486.

E-mail: jae2@ualberta.ca

See page 130 for disclosure information.

addition of the Canadian Triage and Acuity Scale (Net Reclassification Improvement, 0.111; 95% CI, 0.005-0.218; $P = 0.04$).

Conclusion: The EHMRG model exhibited moderate discriminative ability in a large population-based cohort of patients with AHF in the ED. Revision of the EHMRG score through factor inclusion and exclusion could improve the model's performance.

discriminatory properties or, alternatively, more reliably predict nonfatal end points such as future ED visits or hospitalization.

The aims of this study were (1) to externally validate the EHMRG risk model using a population-based dataset of patients who presented to the ED with AHF via ambulance and (2) to potentially refine it to improve its accuracy and simplify its use in clinical practice.

Material and Methods

In this retrospective cohort study, administrative data were used to capture information on patients who accessed Emergency Medical Services (EMS) and were transported to any ED in the province of Alberta, Canada, with a discharge diagnosis of HF between April 1, 2012, and February 29, 2016. In those who had multiple ED visits during the study period, the first visit was selected as the index ED visit and the rest were excluded. Alberta has approximately 4.2 million residents, approximately 100 acute care facilities and EDs, and an annual ED volume of more than 2.3 million visits (2016 data). The study was approved by the Health Research Ethics Board of the University of Alberta.

Data sources

Anonymized data for the study were retrieved from databases described previously and maintained by Alberta Health Services.⁷ These databases contain the data on all patient interactions with the health system. The databases are comprehensive and were linked by a unique lifetime identifier, unique to each patient in Alberta. Patients who had a primary ED diagnosis of HF and were linked to an EMS transport were included in this study. HF was defined by International Classification of Diseases 10th Revision (ICD-10) codes for HF (ICD-10 I50.x). The accuracy of using this ICD-10 code against chart audit has been validated in Alberta.⁸

The National Ambulatory Care Reporting System (NACRS) database includes all visits to the ED in Alberta. Data from NACRS were merged with inpatient data from the hospital discharge database (Discharge Abstract Database) to identify subsequent admissions to hospital, in-hospital outcomes, procedures, and other codes related to comorbid conditions, derived from previous hospitalizations. The information in these databases has been demonstrated to be highly accurate for use in the research setting.⁹ Patient comorbidities were identified using the

30 jours, mais des statistiques C plus faibles pour les autres résultats cliniques (0,61-0,67). L'inclusion de peptides natriurétiques au modèle EHMRG améliorait la prédiction (indice Net Reclassification Improvement [NRI], 0,268; IC à 95 %, 0,173-0,363; $P < 0,01$) pour la mortalité dans les 7 jours, de la même manière que l'ajout de l'Échelle canadienne de triage et de gravité (indice NRI, 0,111; IC à 95 %, 0,005-0,218; $P = 0,04$).

Conclusion : Le modèle EHMRG a montré une capacité discriminative modérée dans une vaste cohorte en population générale auprès de patients souffrant d'ICA au SU. La révision du score EHMRG par l'inclusion et l'exclusion de facteurs pourraient permettre d'améliorer la performance du modèle.

ICD-10¹⁰ from the index ED visit and any hospitalizations or ambulatory encounters in the 2 years before the index ED HF visit.

A province-wide laboratory repository, similar to a previous study,⁵ was used to retrieve the data regarding laboratory test results. The following test results were included: troponin, potassium, creatinine, B-type natriuretic peptide (BNP), and N-terminal pro B-type natriuretic peptide (NT-proBNP). For this analysis, laboratory tests were included if they occurred within 12 hours before or 48 hours after arrival to ED. On the basis of the available dataset, the majority of tests (> 90%) occurred on the same day as the ED visit.

Medication data (on metolazone use and other medications in the last 180 days before index ED visit) was collected using the Pharmaceutical Information Network database, linked by a patient's unique lifetime identifier. The Pharmaceutical Information Network database includes only medications dispensed in the community and thus does not include ED medications. For additional data regarding the heart rate, systolic blood pressure (SBP), and oxygen saturation levels, we used EMS computer-aided dispatch and the EMS ePCR system, which includes the prehospital ambulance data of patients from the first medical contact until their delivery to the destination hospital. Patient deaths were identified using the Alberta Health Care Registry, and hospitalizations and ED visits were derived from Discharge Abstract Database and NACRS, respectively.

Outcomes

The primary outcome of interest was the 7-day all-cause mortality. Other outcomes of interest included 30-day all-cause mortality, 7-day and 30-day deaths, hospitalization, rehospitalization and repeat ED visits, days alive and out of hospital at 30 days, and a composite end point of 7-day and 30-day mortality/rehospitalization. Composite outcomes including death and hospitalizations or ED visits were defined as death during the index episode (ED visit for nonadmitted patients, ED visit + hospitalization for admitted patients) or death/hospitalization/ED visit within 7 or 30 days after the end of the index episode. The outcome "hospitalizations" did not include hospital admissions associated with the index ED visit.

EHMRG risk score

We used the published point coefficients for the EHMRG Risk Score.³ The following variables are used to calculate the

Table 1. Baseline characteristics in patients with and without 7-day mortality

Characteristic	Died within 7 days N = 347	Did not die within 7 days N = 6361	P value
Age, mean (SD), y	83.6 (10.2)	80.6 (10.9)	< 0.0001
Male, n (%)	176 (50.7)	2964 (46.6)	0.13
No. of ED visits in prior 6 mo, median (IQR)	1 (0-2)	1 (0-2)	0.8
CTAS score, median (IQR)	2 (2-3)	3 (2-3)	< 0.0001
ED length of stay, median (IQR), h	9.2 (5.4-17.4)	9.8 (6.5-16.9)	0.03
Admitted from index ED visit	290 (83.6)	5227 (82.2)	0.5
Comorbidities, n (%)			
Prior myocardial infarction	66 (19.0)	1252 (19.7)	0.8
Cerebrovascular disease	34 (9.8)	686 (10.8)	0.6
Peripheral vascular disease	27 (7.8)	572 (9.0)	0.4
Atrial fibrillation	152 (43.8)	2730 (42.9)	0.7
COPD	134 (38.6)	2214 (34.8)	0.14
Hypertension	203 (58.5)	4249 (66.8)	0.001
Diabetes mellitus	119 (34.3)	2542 (40.0)	0.03
Dementia	70 (20.2)	678 (10.7)	< 0.0001
Active cancer	50 (14.4)	565 (8.9)	0.0005
Charlson Comorbidity Index, mean (SD)	5.7 (3.0)	5.1 (2.6)	< 0.0001
Vital statistics at index presentation			
Systolic BP, mean (SD), mm Hg	126.6 (28.3)	143.8 (29.1)	< 0.0001
Diastolic BP, mean (SD), mm Hg	73.0 (18.8)	80.6 (19.3)	< 0.0001
Heart rate, mean (SD), beats/min	91.6 (27.0)	88.0 (24.2)	0.009
Respiratory rate, mean (SD) breaths/min	25.5 (8.9)	24.2 (8.2)	0.003
Oxygen saturation, mean (SD), %	87.4 (10.6)	91.1 (8.0)	< 0.0001
Treatments in prior 180 d, n (%)			
ACE/ARBs	196 (56.5)	4364 (68.6)	< 0.0001
β-Blockers	197 (56.8)	4143 (65.1)	0.001
MRAs	42 (12.1)	894 (14.1)	0.3
Digoxin	32 (9.2)	654 (10.3)	0.5
Loop diuretics	223 (64.3)	3852 (60.6)	0.16
Metolazone	18 (5.2)	269 (4.2)	0.4
Blood tests			
Hemoglobin, mean (SD), g/L	116.3 (21.0)	118.8 (20.5)	0.03
Sodium, mean (SD), mmol/L	137.0 (6.3)	137.2 (5.0)	0.3
Potassium, mean (SD), mmol/L	4.6 (0.9)	4.3 (0.7)	< 0.0001
Creatinine, mean (SD), mg/dL	2.0 (1.5)	1.5 (1.1)	< 0.0001
Elevated troponin, n (%)	91 (33.0)	836 (15.2)	< 0.0001
BNP, median (IQR), ng/L	1034 (615-1965)	807 (477-1444)	0.0002
NT-proBNP, median (IQR), ng/L	7547 (3165-17,608)	4310 (2098-9290)	0.0002

ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; BNP, B-type natriuretic peptide; BP, blood pressure; COPD, chronic obstructive pulmonary disease; CTAS, Canadian Triage and Acuity Scale; ED, emergency department; EHMRG, Emergency Heart failure Mortality Risk Grade; EMS, Emergency Medical Services; IQR, interquartile range; MRA, mineralocorticoid receptor antagonist; NT-proBNP, N-terminal pro B-type natriuretic peptide; SD, standard deviation. The highlighted variables are those included in the original EHMRG risk model.

EHMRG risk score: age, mode of ED presentation (ambulance vs self-presentation), SBP, heart rate, oxygen saturation, potassium concentration, creatinine concentration, troponin (greater than the upper limit of normal), presence of active cancer, and treatment with metolazone. The score reports in a range from -248 to 400+, and cut points are -49.0, -15.8, 18.0, 56.6, and 89.4, for first to fifth risk quintiles, respectively.

Comorbidity index and Canadian Triage and Acuity Scale score

The Charlson Comorbidity Index¹¹ score was calculated for all patients. The Canadian Triage and Acuity Scale (CTAS) is a mandatory element in all EDs in Canada that is used to prioritize patients to be assessed by clinicians upon arrival.⁶

Statistical analysis

Patient characteristics were summarized using proportions, means (standard deviation), and medians (interquartile range) as appropriate and compared by status of death within 7 days. We measured the ability of the EHMRG risk model to risk stratify patients and discriminate between those with and without an outcome using the area under the receiver operating characteristic curve (c-statistic), including Wald 95% confidence intervals (CIs). For EHMRG validation purposes, we provided the c-statistic considering the EHMRG score as a continuous variable.

A series of refinements and modifications to the EHMRG model were tested by comparing the c-statistic as well as integrated discrimination improvement (IDI) and Net Reclassification Improvement (NRI) methods.¹² The NRI estimates the proportion of correct minus incorrect

Table 2. Death within 7 days and 30 days and days alive and out of hospital within 30 days by EHMGR quintiles

EHMGR quintile	EHMGR range	Patients, n	7 d deaths, n	7 d deaths, % (95% CI)	30 d deaths, n	30 d deaths, % (95% CI)	Days alive and out of hospital at 30 d, mean (95% CI)
1	≤ -49.1	237	0	0.0 (0.0-1.5)	1	0.4 (0.0-2.3)	22.7 (21.6-23.8)
2	-49.0 to -15.9	602	5	0.8 (0.3-1.9)	18	3.0 (1.8-4.7)	21.5 (20.7-22.2)
3	-15.8 to 17.9	1192	19	1.6 (1.0-2.5)	65	5.5 (4.2-6.9)	19.6 (19.0-20.2)
4	18.0-56.5	1875	75	4.0 (3.2-5.0)	189	10.1 (8.8-11.5)	17.3 (16.9-17.8)
5A	56.6-89.3	1139	48	4.2 (3.1-5.5)	149	13.1 (11.2-15.2)	15.1 (14.5-15.8)
5B	≥ 89.4	1663	200	12.0 (10.5-13.7)	428	25.7 (23.6-27.9)	11.6 (11.1-12.2)

CI, confidence interval; EHMGR, Emergency Heart failure Mortality Risk Grade. Quintiles derived from Lee et al.³

reclassifications using a modified EHMGR model after each modification over the original EHMGR model as a comparator for the outcomes of 7-day mortality. For EHMGR modification analyses, the predictive ability statistics were calculated using the individual EHMGR variables rather than the EHMGR score in the logistic models because otherwise appropriate weights would need to be derived.

Overall, the proportion of missing data was low for the components of the EHMGR score (< 2% for key variables), except for troponin (unknown for 12.7% of patients). Mean imputation was used for missing values to allow all patients to be included in the analysis. Sensitivity analyses were undertaken using a complete case analysis to determine substantial differences in the model statistics. The availability of NPs (BNP or NT-proBNP) was mutually exclusive (ie, patients did not have both BNP and NT-proBNP measured in the dataset), and 72.6% of patients had BNP or NT-proBNP measured. Because of the proportion of missing data and the mutually exclusive nature of these variables, the missing indicator method was used for evaluating whether the addition of NPs could improve the predictive ability of the EHMGR components. In addition, BNP and NT-proBNP were log-transformed because of the skewness and large variability of the values. All statistical analysis was conducted using SAS statistical software (version 9.4; SAS Institute, Inc, Cary, NC).

Results

The study cohort consisted of 6708 patients (Supplemental Fig. S1). The mean age was 80.7 years, and 3568 (53.2%) were women (Table 1); 82.2% were admitted to hospital from their index ED AHF visit. The median length of ED visit was 9.8 hours (interquartile range, 6.4-16.9), and those who were hospitalized had a median hospital length of stay of 8 days (interquartile range, 5-17). The study cohort included patients with a history of HF (3409, 50.8%) and patients with de novo HF (3299, 49.2%). The mean Charlson Comorbidity Index was 5.1 (standard deviation, 2.7). The mean EHMGR score was 51.4 (standard deviation, 62.6) and event rates were 5.2% for 7-day mortality, 12.7% for 30-day mortality, 18.8% for 30-day hospitalization/rehospitalization after discharge from their index hospital visit, and 27.8% of all patients had a repeat ED visit.

Clinical outcomes and model performance

As depicted in Table 2 and Figure 1, the 7-day mortality rate was 0.0%, 0.8%, 1.6%, 4.0%, 4.2%, and 12.0% across the EHMGR risk categories (1-4, 5A and 5B) provided in the

original derivation study.³ By using quintiles driven by our data and quintile ranges provided in another study,¹³ a similar increase in 7-day and 30-day mortality was observed across the EHMGR risk quintiles (Supplemental Tables S1 and S2; Supplemental Figs S2 and S3). The mean days alive and out of hospital at 30 days after index ED visit ranged from 22.7 days in the first EHMGR risk quintile to 11.6 days in the 5B risk category (Table 2 and Fig. 1C). An increase in the 7-day mortality was observed across EHMGR risk categories in those who were directly discharged from the ED and in those who were admitted to the hospital in their index ED visit. However, the EHMGR score performed better in risk stratifying patients among those who were hospitalized (Supplemental Table S3). The EHMGR score had a c-statistic of 0.73 (95% CI, 0.71-0.76) for 7-day mortality prediction (Table 3, Supplemental Fig. S4) and 0.71 (95% CI, 0.70-0.73) for identifying patients at risk of 30-day mortality (Supplemental Fig. S5). Testing the model against other clinical outcomes, including the composite of 30-day death, hospitalization, and rehospitalization, showed poorer results in terms of discriminatory power in the range of 0.61 to 0.67 (Table 3).

EHMGR refinement

Several scenarios to improve the discriminatory power of EHMGR model were tested, including the addition of new variables and removal of existing ones from the model. Addition of the Charlson Comorbidity Index, number of ED visits in the last 6 months, and number of hospitalization days in the prior year were shown to have no significant impact on the discriminative ability of the model (Table 4 and Supplemental Fig. S6).

NPs. Those who died within 7 days had higher BNP and NT-proBNP levels compared with those who lived ($P < 0.001$; Table 1). The addition of NPs (BNP or NT-proBNP) to a logistic regression model including the EHMGR score variables resulted in improved performance, as illustrated by the c-statistic increasing from 0.747 to 0.760 ($P = 0.007$) and NRI of 0.268 (95% CI, 0.173-0.363; $P < 0.0001$) for predicting 7-day death.

CTAS score. The addition of the CTAS score improved the prediction of 7-day death by increasing the c-statistic from 0.747 to 0.759 ($P = 0.002$) and an NRI of 0.111 (95% CI, 0.005-0.218; $P = 0.044$). The addition of the CTAS to a combined model of EHMGR plus NPs resulted in further improved discriminatory performance, increasing the

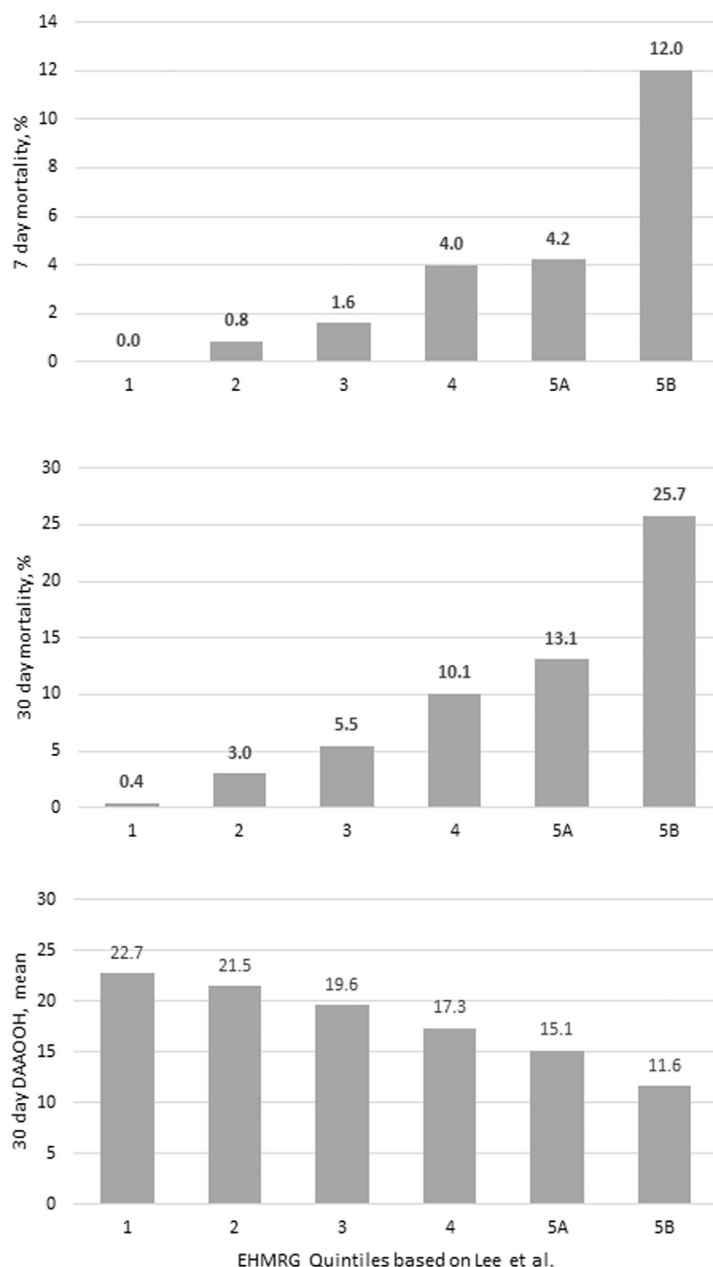


Figure 1. Death within 7 days and 30 days and days alive and out of hospital within 30 days by decile groups from Lee et al.³

c-statistic to 0.771 ($P = 0.003$) with an NRI of 0.130 (95% CI, 0.022-0.237; $P = 0.019$).

Troponin. When troponin was removed from the model, the model performed significantly worse for predicting 7-day death (c-statistic decrease from 0.747 to 0.740, $P = 0.025$; NRI = -0.269, 95% CI, -0.364 to -0.175, $P < 0.0001$).

Metolazone. Removal of the metolazone use variable did not alter the predictive ability of the model with the original EHMRG variables (c-statistic = 0.747 in both models, $P = 0.53$, NRI = -0.005, 95% CI, -0.086 to 0.076, $P = 0.93$) or the extended model that included NP and CTAS

(c-statistic = 0.771 in both models, $P = 0.59$, NRI = -0.026, 95% CI, -0.108 to 0.056, $P = 0.637$).

Sensitivity analysis

We repeated the analysis on complete cases without data imputation, and the results remained similar (data available upon request). For the model refinements described, all models were well calibrated as suggested by large P values in the Hosmer–Lemeshow goodness-of-fit test (Supplemental Table S5). Because the c-statistics increased with an increase in the number of predictors included in the model, we measured the Akaike Information Criterion, which showed larger decreases in Akaike Information Criterion for the

Table 3. EHMRG model's discriminatory power for predicting different outcomes

Outcome	C-index - EHMRG score	C-index - EHMRG variables
Death within 7 d	0.732 (0.706-0.759)	0.747 (0.721-0.774)
Death/hospitalization within 7 d	0.667 (0.650-0.684)	0.674 (0.658-0.691)
Death/hospitalization/re-ED visits within 7 d	0.632 (0.616-0.648)	0.639 (0.623-0.654)
Death within 30 d	0.714 (0.696-0.733)	0.728 (0.710-0.746)
Death/hospitalization within 30 d	0.636 (0.622-0.650)	0.642 (0.628-0.656)
Death/hospitalization/re-ED visits within 30 d	0.609 (0.596-0.623)	0.617 (0.604-0.631)

ED, emergency department; EHMRG, the Emergency Heart failure Mortality Risk Grade.

improved models in accordance with the *c*-statistic, NRI, and IDI results (Supplemental Table S5).

Discussion

Clinicians treating patients with AHF in the ED often need to assess the risk of future untoward events, with only a short observation period and limited available data. Our study has several findings relevant to this situation. First, we found moderate to high discriminative power for the EHMRG risk model in predicting 7-day and 30-day mortality but poorer performance in predicting the risk of 30-day death/rehospitalization/repeat ED visits. These findings highlight the need for further research to develop and refine models that can accurately predict which patients are likely to return to the ED. Second, in our study, the addition of NPs or CTAS score to the model enhanced its performance; further refinement is important for all models as new predictors become available. Finally, although the EHMRG model performs less well in the absence of troponin component, it performs acceptably in the absence of “treatment with metolazone” component. Because metolazone is uncommonly used, this variable could easily be dropped with increased sensibility and no loss in other psychometric properties.

Although most patients presenting to the ED with AHF are admitted to the hospital, approximately one third of all patients are treated and discharged home directly from the ED.¹⁴ The proportion admitted was higher in our study (82%) because we focused only on those patients transported by ambulance. ED crowding is a growing problem, and the pressure to avoid admission of patients is intense. There is a compelling need for effective, easy-to-apply models to risk stratify patients with AHF in the ED setting.

The available AHF risk models are mostly derived from the studies that were focused on hospitalized patients, rather than ED patients.¹⁵⁻¹⁷ These hospital-based models may not have the same utility in the ED setting, and thus other models have been developed. The Ottawa Heart Failure Risk Scale was derived in a small-sized sample of low-risk patients with AHF to predict 30-day death or 14 day nonfatal events (*c*-statistic = 0.75) and requires some variables (eg, 3-minute walk test results), which may not be readily available for those trying to perform retrospective comparative effectiveness studies.⁴ The STRATIFY risk model was developed from a prospective cohort of 1033 patients who presented at a few EDs in the United States with AHF and aimed to identify patients at a high risk of 30-day adverse events (eg, death, cardiopulmonary resuscitation, mechanical cardiac support, intubation, emergent dialysis, coronary revascularization, or

acute coronary syndrome). The inclusion of some components in this composite end point and the unwieldy size of the model (13 variables) have raised questions about the applicability of the STRATIFY model in the risk stratification of patients with HF. Both the Ottawa Heart Failure Risk Scale and the STRATIFY decision tool were constructed in cohorts with modest sample size and need further external validation.

The EHMRG was derived and validated retrospectively in a population-based setting (12,591 patients) and reported a *c*-statistic of 0.81 in the derivation dataset and 0.83 in the validation cohort. Subsequently, it was explored by a group of investigators in Spain using the data of Epidemiology of Acute Heart Failure in Emergency Departments (EAHFE) study, with a *c*-statistic of 0.74 (95% CI, 0.68-0.79) for 7-day mortality;¹⁸ our results are similar.

Although the timeframe of 7 days was used for the development of EHMRG, other time frames (5 days, 14 days, and 30 days) and end points (rehospitalization or initial hospitalization or ED visit) have been proposed. Our study showed that the EHMRG risk score, despite being developed for the outcome of 7-day mortality, is a useful tool to identify patients at risk of death by 30 days. Nevertheless, it performed poorly in predicting other 30-day clinical events (composite of death, hospitalization for those discharged from the ED and rehospitalization for those who were admitted at the index ED visit, and repeat ED visits).

We tested the addition of NPs to the original EHMRG model. NPs were not incorporated in the EHMRG model, because the testing was not common at the time in the ED settings that were used for EHMRG risk model development.³ Alberta has provided province-wide access to NP testing in all Alberta EDs since 2012;⁵ thus, we could explore the addition of NPs. Our results show that the model has a superior performance with the incorporation of NP (eg, BNP/NT-proBNP) results. The Ottawa Risk Score was also shown to have a slightly higher *c*-statistic when BNP was included in the model (*c* = 0.77 vs 0.75), although no NRI or IDI was calculated.⁴ This is a predictable and expected observation, considering the literature supporting the prognostic value of NPs in predicting outcomes in AHF.¹⁹

In our study, the addition of CTAS, a widely used triage tool developed, endorsed, and used across Canada and internationally,⁶ to the EHMRG score added to the model's discriminatory ability. The CTAS is an easy-to-use 5-level tool that determines the priority for patients needing to be seen in the ED. This improved discrimination happened both when the CTAS was added to the original model and when added to the EHMRG plus NP model.

Table 4. EHMARG model performance in predicting 7-day death after trying different refinement options

Variables included in model	C-index (95% CI)	P value for C-index difference	NRI (95% CI)	NRI P value	IDI (95% CI)	IDI P value
All EHMARG variables	0.747 (0.721-0.774)	-	-	-	-	-
+ Charlson Comorbidity Index	0.749 (0.722-0.775)	0.279	0.066 (-0.041 to 0.174)	0.230	0.000243 (-0.0007 to 0.0012)	0.616
+ No. of ED visits in prior 6 mo	0.748 (0.721-0.774)	0.516	0.022 (-0.080 to 0.124)	0.696	0.000228 (-0.0003 to 0.0007)	0.382
+ No. of days hospitalized in prior year	0.751 (0.724-0.778)	0.085	0.106 (-0.007 to 0.205)	0.054	0.00109 (-0.0007 to 0.0029)	0.239
+ CTAS score	0.759 (0.733-0.785)	0.002	0.111 (0.005-0.218)	0.044	0.002956 (0.0004-0.0056)	0.026
+ BNP/NT-proBNP	0.760 (0.735-0.786)	0.007	0.268 (0.173-0.363)	< 0.0001	0.006066 (0.0032-0.0089)	< 0.0001
- Troponin	0.740 (0.713-0.766)	0.025	-0.269 (-0.364 to -0.175)	< 0.0001	-0.003451 (-0.0067 to -0.0002)	0.039
- Metolazone	0.747 (0.721-0.774)	0.530	-0.005 (-0.086 to 0.076)	0.929	-0.000005 (-0.0001 to 0.0001)	0.868
+ BNP/NT-proBNP + CTAS*	0.771 (0.746-0.797)	0.003	0.130 (0.022-0.237)	0.019	0.003529 (0.0006-0.0065)	0.019
+ BNP/NT-proBNP + CTAS - Metolazone†	0.771 (0.746-0.797)	0.587	-0.026 (-0.108 to 0.056)	0.637	-0.000002 (-0.0001 to 0.0001)	0.978

BNP, B-type natriuretic peptide; CTAS, Canadian Triage and Acuity Scale; ED, emergency department; EHMARG, the Emergency Heart failure Mortality Risk Grade; NRI, Net Reclassification Improvement; NT-proBNP, N-terminal pro B-type natriuretic peptide;

* Comparison is relative to the model including BNP/NT-proBNP.

† Comparison is relative to the model including BNP/NT-proBNP and CTAS.

Conversely, in our study, there was no association between the use of metolazone and the 7-day mortality outcome. This is consistent with the findings of the EHMARG30-ST study,²⁰ in which the use of metolazone was not associated with 30-day mortality post-ED visit. However, excluding the troponin component from the EHMARG model significantly reduced its discriminatory performance. The proposal to use high-sensitivity troponins will require future revalidation as they are used more broadly in clinical practice.

With the growing volume of evidence and expanding number of risk prediction models in clinical practice each day, it becomes paramount to find and incorporate innovative approaches to keep clinicians up to speed in using the evidence and facilitating decision-making for them. Clinical decision support tools,^{21,22} electronic alerts,^{23,24} mobile apps,²⁵ and real-time automatic calculators connected to electronic medical records have evolved significantly with the goal of using the best evidence and improving the clinical practice for our patients and healthcare professionals, and have the potential to become an inseparable part of the future's clinical practice.

Study limitations

Some strengths and limitations are noteworthy. Our study includes data from all EDs in Alberta, which makes it more representative of the overall patient population, although our results apply only to patients arriving by ambulance. The extent to which presentation by ambulance serves as a proxy for variables already in the model (eg, age) or not accounted for (eg, socioeconomic status, social isolation) that create a propensity to use ambulance services is unknown. A small number of patients had missing values on 1 or more variables; a mean imputation method for missing values was used. A complete case analysis was performed, which confirmed the lack of a difference in model performance. We lacked the data on cause-specific deaths in this study; however, all-cause mortality remains a key metric for patients with HF.

Conclusions

The EHMARG model had a moderate discriminative performance for 7-day mortality in a cohort of patients with a primary diagnosis of AHF who arrived to the ED by ambulance. The study also explored possibilities to improve the model's prediction performance by adding new variables that are already widely available in the ED setting, such as NP level and CTAS score. Future applications of AHF ED risk scores should incorporate these findings and focus on integration and real-time automatic calculation, the use of continuous rather than dichotomized variables (eg, NPs, troponin, SBP), and the use of active integration into clinical decision making.

Acknowledgements

This study is based in part on data provided by Alberta Health to Alberta Health Services Analytics Branch. The interpretation and conclusions contained are those of the researchers and do not necessarily represent the views of the Government of Alberta or Alberta Health Services. The Government of Alberta, Alberta Health, and Alberta Health Services do not express any opinion in relation to this study.

F.A.M. had full access to all the data used for this study and takes responsibility for the analysis and interpretation.

Funding Sources

This study was supported by operating grants from the Canadian Institutes of Health Research and Alberta Innovates — Health Solutions.

Disclosures

None of the funding agencies had any input into design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. J.A.E. and F.A.M. receive salary support from Alberta Innovates — Health Solutions. F.A.M. is supported by the University of Alberta/Capital Health Chair in Cardiovascular Outcomes Research. B.H.R. is supported by a salary award as a Tier I Canada Research Chair in Evidence-based Emergency Medicine from the CIHR and the Government of Canada (Ottawa, ON). B.H.R. is also the Scientific Director of the Emergency Strategic Clinical Network within AHS. N.S. received a graduate studentship from Alberta Innovates — Health Solutions.

References

- Tran DT, Ohinmaa A, Thanh NX, et al. The current and future financial burden of hospital admissions for heart failure in Canada: a cost analysis. *CMAJ Open* 2016;4:e365-70.
- Collins SP, Jenkins CA, Harrell FE Jr, et al. 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;3:737-47.
- Lee DS, Stitt A, Austin PC, et al. Prediction of heart failure mortality in emergent care: a cohort study. *Ann Intern Med* 2012;156:767-75. w-261, w-262.
- Stiell IG, Clement CM, Brison RJ, et al. 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.
- Sepehrvand N, Bakal JA, Lin M, et al. Factors associated with natriuretic peptide testing in patients presenting to emergency departments with suspected Heart Failure. *Can J Cardiol* 2016;32:986.e981-e988.
- Bullard MJ, Chan T, Brayman C, et al. Revisions to the Canadian Emergency Department Triage and Acuity Scale (CTAS) Guidelines. *CJEM* 2014;16:485-9.
- Bakal JA, McAlister FA, Liu W, Ezekowitz JA. Heart failure re-admission: measuring the ever shortening gap between repeat heart failure hospitalizations. *PLoS One* 2014;9:e106494.
- Frolova N, Bakal JA, McAlister FA, et al. Assessing the use of international classification of diseases-10th revision codes from the emergency department for the identification of acute heart failure. *JACC Heart Fail* 2015;3:386-91.
- Quan H, Parsons GA, Ghali WA. Validity of information on comorbidity derived from ICD-9-CCM administrative data. *Med Care* 2002;40:675-85.
- Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005;43:1130-9.
- Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373-83.
- Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27:157-72. discussion 207-12.
- Lee DS, Lee JS, Schull MJ, et al. Design and rationale for the Acute Congestive Heart Failure Urgent Care Evaluation: The ACUTE Study. *Am Heart J* 2016;181:60-5.
- Brar S, McAlister FA, Youngson E, Rowe BH. Do outcomes for patients with heart failure vary by emergency department volume? *Circ Heart Fail* 2013;6:1147-54.
- Fonarow GC, Adams KF Jr, Abraham WT, et al. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. *JAMA* 2005;293:572-80.
- Hsieh M, Auble TE, Yealy DM. Validation of the Acute Heart Failure Index. *Ann Emerg Med* 2008;51:37-44.
- Peterson PN, Rumsfeld JS, Liang L, 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.
- Gil V, Miro O, Schull MJ, et al. Emergency Heart Failure Mortality Risk Grade score performance for 7-day mortality prediction in patients with heart failure attended at the emergency department: validation in a Spanish cohort. *Eur J Emerg Med* 2018;25:169-77.
- Januzzi JL Jr, Sakhujia R, O'Donoghue M, et al. Utility of aminoterminal pro-brain natriuretic peptide testing for prediction of 1-year mortality in patients with dyspnea treated in the emergency department. *Arch Intern Med* 2006;166:315-20.
- Greig D, Austin PC, Zhou L, et al. Ischemic electrocardiographic abnormalities and prognosis in decompensated heart failure. *Circ Heart Fail* 2014;7:986-93.
- Wai JM, Aloeos C, Mowrey WB, et al. Using clinical decision support through the electronic medical record to increase prescribing of high-dose parenteral thiamine in hospitalized patients with alcohol use disorder. *J Subst Abuse Treat* 2019;99:117-23.
- Khan S, Richardson S, Liu A, et al. Improving provider adoption with adaptive clinical decision support surveillance: an observational study. *JMIR Hum Factors* 2019;6:e10245.
- Sepehrvand N, Moosavi-Toomatari B. *Venous Thromboembolism & Fuzzy-Based Electronic Alerts*. Saarbrücken, Germany: Lambert Academic Publishing, 2012.
- Sepehrvand N, Pakdel FG, Rahimi-Rad MH, et al. Practice guidelines and clinical risk assessment models: is it time to reform? *BMC Med Inform Decis Mak* 2011;11:63.
- Patzner RE, Basu M, Larsen CP, et al. iChoose Kidney: a clinical decision aid for kidney transplantation versus dialysis treatment. *Transplantation* 2016;100:630-9.

Supplementary Material

To access the supplementary material accompanying this article, visit *CJC Open* at <https://www.cjopen.ca> and at <https://doi.org/10.1016/j.cjco.2019.03.003>.