Prognostic Scores for Risk Stratification in Patients with Acute Heart Failure

Escalas de estratificación del pronóstico en pacientes con falla cardíaca aguda

John J. Sprockel, Laura Alfaro, Jennifer Cifuentes, Mónica Jiménez, Rafael A. Baron, Walter G. Chaves

ABSTRACT

Background: Heart failure is a common, progressive and potentially lethal condition. An adequate risk assessment in acute cases might optimize care and resources.

Objective: The aim of this study was to apply three of the existent risk scores to predict in-hospital mortality in patients with acute heart failure in order to understand their performance.

Methods: This prognostic study included patients hospitalized by the department of internal medicine with diagnosis of decompensated heart failure between February 2010 and March 2013. Three risk scores were used to evaluate in-hospital mortality: the ADHERE decision tree, the OPTIMIZE-HF score and the GWTG-HF score. Contingency tables were constructed using the performance of the different cutoff points and the ROC curves of each score; then, the areas under the curve were calculated.

Results: Score results were calculated in 379 patients. Mean age was 75 years, 54% were women, 79% had hypertension, and 24% were diabetics. The area under the curve for the three risk scores was 0.63 (95% CI, 0.53-0.73) for OPTIMIZE-HF, 0.57 (95% CI, 0.49-0.65) for GWTG-HF and 0.58 (95% CI, 0.47-0.68) for ADHERE.

Conclusion: In a population of patients hospitalized for decompensated heart failure, the ADHERE, OPTIMIZE and GWTG-HF risk scores showed poor discrimination ability to predict in-hospital mortality.

Key words: Heart Failure - Risk - Prognosis - Mortality

RESUMEN

Introducción: La insuficiencia cardíaca es una condición frecuente, progresiva y potencialmente mortal. La evaluación correcta del riesgo de los desenlaces en los casos agudos permitiría la optimización de la atención y de los recursos.

Objetivo: Aplicar tres de las escalas existentes para la evaluación del riesgo de muerte intrahospitalaria en pacientes con falla cardíaca aguda a fin de conocer su desempeño.

Material y métodos: Estudio con diseño de pronóstico en el que se incluyeron pacientes hospitalizados por el servicio de medicina interna con diagnóstico de falla cardíaca descompensada entre febrero de 2010 y marzo de 2013. Se aplicaron tres escalas de riesgo para evaluar la mortalidad intrahospitalaria: árbol de decisiones ADHERE, OPTIMIZE-HF y GWTG-HF. Se construyeron tablas de contingencia mediante el cálculo del desempeño para los diferentes puntos de corte, así como las curvas ROC de cada escala y luego se calcularon las áreas bajo la curva.

Resultados: Se calculó el resultado de las escalas en 379 pacientes, de los cuales el 54% eran mujeres; la mediana de edad fue de 75 años, el 79% eran hipertensos y el 24% eran diabéticos. El área bajo la curva del OPTIMIZE-HF fue de 0,63 (IC 95% 0,53-0,73), la del GWTG-HF fue de 0,57 (IC 95% 0,49-0,65) y la del ADHERE fue de 0,58 (IC 95% 0,47-0,68).

Conclusion: En una población de pacientes hospitalizados por falla cardíaca descompensada, las escalas de riesgo ADHERE, OPTIMIZE y GWTG-HF mostraron una pobre capacidad de discriminación del riesgo de muerte intrahospitalaria.

Palabras clave: Falla cardiaca - Riesgo - Pronóstico - Mortalidad

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>Area under the curve</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquartile range</td>
</tr>
<tr>
<td>LVEF</td>
<td>Left ventricular ejection fraction</td>
</tr>
<tr>
<td>NT-proBNP</td>
<td>N-terminal pro B-type natriuretic peptide</td>
</tr>
</tbody>
</table>


Accepted: 10/30/2016
Address for reprints: John Sprockel Díaz - Hospital de San José- Fundación Universitaria de Ciencias de la Salud - Cra. 19 No. 8ª - 32 Pbx: 3538100 - e-mail: jsprockel@fucusalud.edu.co

Hospital de San José - Fundación Universitaria de Ciencias de la Salud. Bogotá, Colombia
INTRODUCTION
Heart failure is a common, progressive and potentially lethal condition. As Gheorghiade et al. noted, the incidence of acute heart failure is very similar to that of acute myocardial infarction (the clinical condition with the highest mortality worldwide), with similar in-hospital mortality. Yet, after discharge, heart failure-related mortality is between four and five times greater. (1) A registry performed by the Hospital de San José during a 4-year period (2010-2013), including data of 462 patients, reported an in-hospital mortality of 8.9% and 30-day mortality of 13.8% for decompensated heart failure, (2) while in-hospital mortality of ST-segment elevation myocardial infarction was 6.7% in a review of acute coronary syndromes performed between 2009 and 2010 and 9.5% in another study carried out in 2012. (4)

United States data indicate that mortality rates at 30 days, 1 year and 5 years after hospitalization for heart failure are 10.4%, 22 % and 42.3%, respectively; (5) adjusted mortality at 1 year decreased from 31.7% in 1999 to 29.6% in 2008 (p <0.001). (6) Recent data of the cohort in Olmsted County, Minnesota, reveal that although the age– and sex–adjusted incidence of heart failure declined from 316 per100,000 in 2000 to 219 per 100,000 in 2010 (a rate reduction of 37.5%), mortality among the new cases of heart failure remained constant at 5 years after the diagnosis, around 52.6%, and was frequently ascribed to noncardiovascular causes (54.3%). (7) Over the past years, length of hospital stay has decreased, but has been associated with greater 30-day mortality after hospitalization and greater readmission rates due to heart failure. (8) By 2012, the total cost attributable to heart failure was estimated in 30.7 billion dollars, and 68% was due to direct medical costs. (9)

Several prognostic markers have been evaluated in decompensated heart failure. Clinical markers are: hypertension, (10), elevated heart rate, (11), third heart sound and elevated jugular venous pressure (12, 13); lab tests: renal function, (14) hyponatremia, (15), anemia, (16) cardiac troponins, (10, 17) natriuretic peptides; (18) imaging tests: cardiothoracic ratio, (19) ejection fraction, (20), pulmonary hypertension; (21) functional class: 6-minute walk test, (22), peak exercise oxygen consumption, (23); hemodynamic parameters: pulmonary capillary wedge pressure; and use of neuro-modulating drugs, among others. Despite these correlations, individual predictors are rarely sufficient to provide accurate risk estimation. (24)

Because heart failure is a complex syndrome, patients with this condition have a wide spectrum of mortality risk. Patients at high risk of mortality should receive high rates of drug treatment to maximize the benefit, while interventions should be limited and out-patient care should be considered in those at low risk. More than 64 prediction scores have been developed to help risk assessment. (25) Nevertheless, the use of these tools in clinical practice is still very low. The most often stated barriers to apply these instruments are doubts concerning over-simplification of risk assessment, potential risk of over-treatment, or the fact that the numerical information resulting from prediction directives is often not helpful for decision-making. (26)

The aim of the present study is to evaluate the correlation of the three main risk scores to predict the risk of mortality in patients with acute heart failure: the ADHERE risk tree, (27), the OPTIMIZE-HF score, (28) and the GWTG-HF score (29) in the population of patients hospitalized in a tertiary care center in Bogotá, Colombia.

METHODS
A prognostic evaluation study was performed from a prospective cohort of patients admitted to the emergency department of Hospital San José in Bogotá, Colombia, from February 2010 to March 2013, (2) which included patients >18 years with decompensated heart failure according to the Framingham cohort diagnostic criteria, hospitalized by the department of internal medicine.

Major criteria were: paroxysmal nocturnal dyspnea, neck vein distention, pulmonary rales, cardiomegaly, acute pulmonary edema, third heart sound and hepatojugular reflex. Minor criteria were: lower limb edema, nocturnal cough, dyspnea on exertion, hepatomegaly; pleural effusion and tachycardia > 120 bpm. Presence of at least 2 major criteria or 1 major criterion and 2 minor criteria was required to enter the study. (30)

Patients with the following conditions were excluded: decompensated diabetes mellitus, urgent dialysis, diagnosis of Child class C cirrhosis, acute liver failure, nephrotic syndrome, hypovolemic shock secondary to upper gastrointestinal bleeding, septic shock, or end-stage cancer documented

| Table 1. Demographic characteristics of the population |
|----------|-----------|-----------|
|          | n (%)     |           |
| Age in years, median (IQR) | 75 (65-82) |           |
| Female sex | 205 (54)  |           |
| **Comorbidities**          |           |           |
| COPD                    | 162 (42)  |           |
| Hypertension           | 301 (79)  |           |
| Coronary artery disease | 76 (20)   |           |
| Type 2 diabetes mellitus | 94 (24)   |           |
| Chronic kidney failure  | 55 (14)   |           |
| **Laboratory tests**      |           |           |
| Creatinine levels, median (IQR) mg/dL | 1 (0.8-1.4) |           |
| Sodium, median (IQR) mg/dL | 139 (135-142) |           |
| BUN, median (IQR) mg/dL   | 24 (18-35) |           |
| Troponin I, median (IQR) μg/dL | 0.04 (0.015-0.09) |           |
| Hemoglobin, median (IQR) mg/dL | 13.9 (11.7-15.5) |           |
| NT-ProBNP >4,630         | 197 (51)   |           |
| LVEF, <40%               | 176 (46.4) |           |

in the clinical record or subsequently confirmed.

The information was entered into a specially designed data collection form and the socio-demographic variables, clinical variables, concomitant conditions, history of hospitalizations and physical examination data were recorded. Blood samples were obtained to measure blood urea nitrogen, creatinine levels, N-terminal pro B-type natriuretic peptide (NT-proBNP) (5600 Integrated Immunodiagnostic System), hemoglobin levels, sodium and troponin I (VITROS Immunodiagnostic System). An electrocardiogram was taken to all the patients at admission. Left ventricular ejection fraction (LVEF) was calculated from the echocardiograms performed at the department of cardiology of the institution or from previous studies performed within 3 months. All the patients were followed up until discharge. In-hospital mortality was documented and 30-day mortality was investigated by telephone calls or by consulting the Affiliate Unique Registry (RUAF).

Statistical analysis
All the calculations were performed using STATA 13 software package. Continuous variables are expressed as mean ± standard deviation and categorical variables as absolute and relative frequencies. Data for the calculation of the different scores was available in 379 patients of the original cohort of 462. A nonparametric estimation of the area under the ROC curve (AUC) was performed for each of the scores evaluated as well as the calculation of the receiver operation characteristics for the different cutoff points according to the percentiles (OPTIMIZE and GWTG-HF) or risk classification (ADHERE decision tree). A p value < 0.05 was considered statistically significant.

The number of in-hospital deaths was divided by the best cutoff point for NT pro-BNP (4630 pg/dL) found in the original paper. (2)

Ethical considerations
This study was approved by the Human Research Ethics Committee of the Fundación Universitaria de Ciencias de la Salud and the Hospital San José. As the study was considered to be of negligible risk, an informed consent was not necessary. The original study received funding from the internal call for proposals N 4-2009.

RESULTS
Between February 2010 and March 2013, 485 patients with diagnosis of decompensated heart failure un-

<table>
<thead>
<tr>
<th>Cutoff point</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>LR+</th>
<th>LR-</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100%</td>
<td>0.0%</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>93.8%</td>
<td>3.8%</td>
<td>0.97</td>
<td>1.67</td>
</tr>
<tr>
<td>3</td>
<td>81.3%</td>
<td>38.3%</td>
<td>1.32</td>
<td>0.49</td>
</tr>
<tr>
<td>4</td>
<td>34.4%</td>
<td>84.7%</td>
<td>2.25</td>
<td>0.77</td>
</tr>
<tr>
<td>5</td>
<td>3.1%</td>
<td>97.7%</td>
<td>1.35</td>
<td>0.99</td>
</tr>
<tr>
<td>6</td>
<td>0.0%</td>
<td>99.4%</td>
<td>0.00</td>
<td>1.01</td>
</tr>
<tr>
<td>7</td>
<td>0.0%</td>
<td>99.7%</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>&gt; 7</td>
<td>0.0%</td>
<td>100%</td>
<td>-</td>
<td>1.00</td>
</tr>
</tbody>
</table>
derwent screening, and data from 379 patients was available to calculate the scores. Demographics of the series are depicted in Table 1. Median age was 75 [interquartile range (IQR: 65-82)] years, 54% of patients were women, 79% had hypertension, 42% chronic obstructive pulmonary disease and 20% coronary artery disease. Less than half of the patients (46%) had LVEF <40%.

Lab tests showed median creatinine level of 1mg/

<table>
<thead>
<tr>
<th>Percentiles of the OPTIMIZE score</th>
<th>NT-proBNP &lt;4,630 pg/dl</th>
<th>NT-proBNP ≥4,630 pg/dl</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;3, n (%)</td>
<td>2 (33.3)</td>
<td>4 (66.7)</td>
</tr>
<tr>
<td>≥3, n (%)</td>
<td>8 (30.8)</td>
<td>18 (69.2)</td>
</tr>
</tbody>
</table>

**Table 3.** Performance of in-hospital deaths according to the best cutoff point for natriuretic peptide in the OPTIMIZE score

**NT-pro BNP:** N-terminal pro B-type natriuretic peptide.

<table>
<thead>
<tr>
<th>Author and date</th>
<th>Model evaluated</th>
<th>Outcome</th>
<th>Database</th>
<th>Number of patients</th>
<th>Number of deaths</th>
<th>Deaths %</th>
<th>C (or AUC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fonarow, 2005 (27)</td>
<td>ADHERE CARD</td>
<td>In-hospital mortality</td>
<td>ADHERE</td>
<td>Development: 33,046</td>
<td>NA</td>
<td>NA</td>
<td>0.687</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Validation: 32,229</td>
<td>NA</td>
<td>NA</td>
<td>0.668</td>
</tr>
<tr>
<td>Abraham, 2008 (28)</td>
<td>OPTIMIZE-HF risk score</td>
<td>In-hospital mortality</td>
<td>OPTIMIZE-HF</td>
<td>Development: 37,548</td>
<td>1.217</td>
<td>3.24</td>
<td>0.753</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Validation: 937</td>
<td>28</td>
<td>2.99</td>
<td>0.756</td>
</tr>
<tr>
<td>Peterson, 2010 (29)</td>
<td>GWTG-HF risk score</td>
<td>In-hospital mortality</td>
<td>GWTG-HF</td>
<td>Development: 181,830</td>
<td>4.649</td>
<td>2.56</td>
<td>0.746</td>
</tr>
</tbody>
</table>

**Development**

| Auble, 2007 (34) | ADHERE CARD | In-hospital mortality (and at 30 days) | Pennsylvania Health Care Cost Containment Council database of demographic and administrative variables and Cardinal Health Information Companies-MediQual Systems Atlas Severity of Illness System database | 32,160 | 1,432 | 4.45 | 0.68 (0.66) |
|                  | EFFECT |         |          | 33,533 | 1,498 | 4.47 | 0.74 (0.73) |
|                  | BWH |         |          | 33,533 | 1,498 | 4.47 | 0.61 (0.59) |

**Validation**

| Austin, 2010 (35) | EFFECT-HF mortality model | In-hospital mortality | EFFECT HF | 15,844 | NA | NA | 0.772 |
| ADHERE Logistic regression model | ADHERE CARD | In-hospital mortality |          | 0.747 |
| ADHERE Logistic regression model | OPTIMIZE-HF risk score | In-hospital mortality (and at 90 days) | Multicenter | 701 | 53 | 7.60 | 0.758 |
| ADHERE Logistic regression model | GWTG-HF risk score | In-hospital mortality | West Tokyo Heart Failure (Wet-HF) registry | 1876 | 68 | 3.62 | 0.763 |
| Shiraishi, 2016 (37) | GWTG-HF risk score | In-hospital mortality | Hospital San José | 379 | 32 | 8.44 | 0.58 |
| Present study | ADHERE CARD | In-hospital mortality |                       | 0.63 |
|                | OPTIMIZE-HF risk score |                       |                       | 0.57 |
One of the main concerns in patients with heart failure is the risk of death, which varies significantly among different populations. The risk score, the ability of the score to predict risk, when natriuretic peptide levels are added to the model most commonly evaluated. (34-36) In two cas-
it was very poor. The ADHERE decision tree was the
formance is hardly acceptable, while in the present study
validation studies (34-37). In most cases, the perfor-
veloped the risk scores evaluated (27-29) and of the
performance of scores in our population.
The AUC evaluating the discrimination of each
score to predict in-hospital mortality was as follows:
0.63 (95% CI, 0.53-0.73) for the OPTIMIZE score, 0.57
(95% CI, 0.49-0.65) for the GWTG-HF score and 0.58
(95% CI, 0.47-0.68) for the ADHERE decision tree
(Figure 1).
The analysis of each decile of the OPTIMIZE-HF
score showed that decile 3 had the highest discrimi-
nation capacity, with a sensitivity of 81%, a specific-
ity of 38% and an accuracy of 41% (Table 2). For the
GWTG-HF score, the second decile had the highest discrimina-
tion ability, with a sensitivity of 94% and a specific-
ity of 16%.
The AUC for NT pro-BNP was 0.63 (95% CI, 0.54-
0.73), and after dividing the number of deaths by the
best cutoff point for NT pro-BNP (4,630 pg/dL) the
likelihood of death was two times higher, indepen-
dently of the result of the OPTIMIZE score (Table 3).

**DISCUSSION**

Risk assessment is essential in patients with heart
failure for an adequate management of this popula-
tion. A study performed in 2002 documented that phy-
cicians overestimated the probability of severe comp-
lications in patients with acute heart failure, which
might be related with the overutilization of resources. (31) A survey performed in 2008 among 1,450 geri-
atriicians, cardiologists, internists and family practi-
tioners revealed that less than 25% of respondents
believed that they could accurately predict death in
heart failure patients. (32)

In another study, over half of the patients admitted
for heart failure to an acute care facility were low risk,
(33) a finding that is similar and even higher in our
results (> 80%). For this reason, the cases with the
highest scores may be underrepresented, affecting the
possibility of obtaining reliable information about the
performance of scores in our population.

Table 4 presents the results of the studies that de-
veloped the risk scores evaluated (27-29) and of the
validation studies (34-37). In most cases, the perfor-
ance is hardly acceptable, while in the present study
it was very poor. The ADHERE decision tree was the
model most commonly evaluated. (34-36) In two cas-
es, when natriuretic peptide levels are added to the
risk score, the ability of the score to predict risk in-
creases significantly. (36, 37)

Although the results seem to indicate that the
OPTIMIZE score has a better performance, the AUC
overlaps with the other two scores evaluated and,
thus, we cannot confirm that this score was the best.
In any case, the AUC range obtained would mean
that the use of any of these scores would not be useful
in practice. Yet, it is not possible to know if it would
be useful for the negative cases or if the addition of
natriuretic peptide would improve discrimination.

It is highly possible that our setting might have
factors different from those explored in the different
scores that would affect mortality, probably making mortality of the lower scores higher than the mortal-
ity found in the studies in which the scores were de-
veloped or validated.

**Limitations**

Among the limitations of this study, we must mention
the low number of events, which prevents consider-
ing our study as a real validation of these scores. The
intermediate and high score levels are underrepre-
sented, a fact that contrasts with the higher mortal-
ity compared with other trials, except for the study by
Scrutinio (36) who reported a mortality rate similar
to our findings.

**CONCLUSION**

In a population of patients hospitalized for decom-
penated heart failure, the ADHERE, OPTIMIZE
and GWTG-HF risk scores showed poor discrimina-
tion to predict in-hospital mortality. The population
of patients with decompensated heart failure gathers
around the lower risk scores. The results obtained
seem to indicate that use of these scores in our envi-
ronment would provide little information for decision-
making.

**Conflicts of interest**

None declared.

(See authors’ conflicts of interest forms in the web-
site/Supplementary material).

**REFERENCES**

MA, et al. Acute heart failure syndromes: current state and frame-
org/10.1161/01.CIR.0000182956.39024.42
2. Chaves WG, Diaztagle JJ, Sprockel JJ, Hernández JI, Benavidez
J, Henao D y cols. Factores asociados a mortalidad en pacientes con
http://www.scielo.org.co/pdf/ami/v39n4/v39n4a03.pdf
clínica y tratamiento de los pacientes con síndrome coronario agudo
4. Sprockel JJ, Diaztagle JJ, Chaves WG, Henao JC, Simon CJ, Afa-
nador D y cols. Calidad de la atención de los síndromes coronarios
agudos: implementación de una ruta crítica. Rev Colom Cardiol

554 ARGENTINE JOURNAL OF CARDIOLOGY / VOL 84 Nº 6 / DECEMBER 2016


