EDITORIAL

Prognostic Assessment in Heart Failure Patients: An Unresolved Task

Evaluación pronóstica en pacientes con insuficiencia cardíaca: una tarea por resolver

ANA CAROLINA ALBA

Heart failure (HF) represents the final stage of different diseases as hypertension, coronary artery disease, myocarditis, alcoholic cardiomyopathy and Chagas' disease. Nowadays, the treatments of these conditions have improved; yet, many patients develop HF later in life with more number of comorbid disorders. Several risk factors modify the outcome of patients with HF and usually coexist at varied proportions. All these circumstances determine that the prognostic assessment of a particular HF patient is challenging. The human mind has limitations to bring together all that information and translate it into an accurate prognostic assessment, even in the case of well-informed physicians.

Our limitation as physicians to establish an accurate prognosis in many disciplines has been recognized for over 70 years. (1-4) In order to overcome this limitation, several researchers have developed many predictive models over the past 30 years to estimate the risk of future adverse events in HF patients by combining a limited number of prognostic markers. Some of these predictive models have been validated in multiple populations and have demonstrated different performance. (5)

In this issue of the Argentine Journal of Cardiology, Chirino et al. (6) have made an interesting analysis about the performance of two models to predict mortality in 704 HF patients in Argentina: the Cardiac and Comorbid Conditions - Heart Failure (3C-HF) score and the Meta-Analysis Global Group in Chronic Heart Failure (MAGGIC) score. These scorying systems were calculated using variables measured at hospital discharge or during an outpatient visit. The authors reported that both models demonstrate moderate ability to discriminate events and adequate calibration based on statistical methods. Some important points were highlighted during the discussion.

Discrimination capacity or discrimination power is the ability of a prognostic model to differentiate between groups of patients depending on whether they have a high or low risk of developing a certain event. Using the area under the curve, both predictive models show a discrimination capacity of about 0.70. In clinical practice, this means that there is 70% chance that the models will categorize patients who have experienced an event as high-risk patients and those who have not as low-risk patients. In the same sense, the 30% chance to be wrong. On the other hand, calibration is the property of a model to evaluate how close is the estimated of an event to occur in comparison to the o event rate. Both properties are associated and inter-related; however, a model with good discrimination capacity may not be well calibrated, and vice versa. Calibration is the most important quality of a model as clinical decisions are based on patients' estimated risk.

In their study, Chirino et al. (6) observed that both models overestimated the risk in low-risk patients and underestimated it in high-risk patients, thus compromising the discrimination capacity. In the same sense, the magnitude of the difference between the predicted risk and the observed risk compromises the calibration. The clinical impact of this limited performance will depend on how this model is used. For example, if a more aggressive strategy is decided for patients with an estimated risk > 5%, these patients would not be adequately identified by this model. A more aggressive strategy would be, for example, to indicate aldosterone antagonists, angiotensin-converting enzyme inhibitors or natriuretic peptides to asymptomatic patients with low left ventricular ejection fraction (LVEF), or to implant an implantable cardioverter defibrillator in patients with New York Heart Association class I and non-ischemic heart disease with reduced LVEF. In the study by Cirino et al., patients with 4% to 13% risk by the MAGGIC score had an observed risk of about 8%. The consequences of using these models would mean missing the opportunity of treating high-risk patients wrongly categorized as having low risk, or unnecessarily treating low risk patients with an aggressive

Rev Argent Cardiol 2018;86:305-306. http://dx.doi.org/10.7775/rac.v86.i5.14045

SEE RELATED ARTICLE: Rev Argent Cardiol 2018;86:322-328. http://dx.doi.org/10.7775/rac.v86.i5.13212

Address for reprints: Dr. Ana Carolina Alba. Toronto General Hospital, University Health Network. Toronto, Canada. 585 University Avenue, 6EN-246, Toronto, ON, Canada. M5G 2N2

Toronto General Hospital, University Health Network. Toronto, Canada

strategy, exposing them to side effects and wasting limited resources.

The performance of a model can be impaired by significant differences in the characteristics of the validation population. Chirino et al. emphasized that the Argentine population included in their study is older (average age: 73 years). However, these models derived from populations that are not significantly different from the one included in this study. Population samples significantly different from those included to create the original model may explain, to some extent, the limited performance of some predictive models. Yet, this does not seem to be the case in the study by Chirino et al.

The differences in baseline mortality or in the association between prognostic factors and mortality in patients with different characteristics may have a negative impact on the performance of a prognostic model. Both models (the MAGGIC score and the 3C-HF score) were developed and now validated by Chirino et al. in patients with different types cardiac dysfunction, including patients with preserved and reduced LVEF. The underlying pathophysiology of HF patients with reduced LVEF is different from those with HF and preserved LVEF. They are associated with different phenotypes and include diverse proportions of concomitant cardiovascular diseases, such as hypertension, coronary artery disease, atrial fibrillation and pulmonary hypertension, and of non-cardiovascular diseases, as diabetes, chronic kidney failure, anemia, chronic lung disease and obesity. (7, 8) Treatment of HF with reduced or preserved LVEF is also different. The therapeutic effects of drugs prescribed in HF patients, such as beta-blockers, angiotensin-converting enzyme inhibitors or angiotensin II receptor blockers, are also different. These drugs improve survival in HF patients with reduced LVEF, but have not demonstrated to improve survival in patients with preserved LVEF. One might wonder whether the limited performance of these models might be related to the fact that these two groups of patients are essentially different, and that a predictive model may have inadequate performance in one of these populations, or in both.

There are no comparative data to assess if these predictive models, or similar models, have different performance, depending on the type of HF. The study by Chirino et al. does not provide this information.

Previous studies performed in patients with cardiovascular disease or HF have shown that physicians have limited ability to assess prognosis and that they usually overestimate patients' mortality. (9-11) Prognostic models represent a promising solution in this sense but, so far, there is little evidence on their performance in actual clinical practice.

Chirino et al. should be congratulated for this elegant study; their work demonstrates that there are still issues to be solved. The entire scientific community has the task to provide more evidence on how to define prognosis in HF patients. I invite the investigators to keep on working together building and evaluating more applicable risk stratification models and looking for the best performance possible, providing opportunities for rigorous validation.

Conflicts of interest

None declared.

(See authors' conflicts of interest forms on the website/Supplementary material).

REFERENCES

1. Sarbin TR. A Contribution to the Study of Actuarial and Individual Methods of Prediction. Am J Sociol 1943;48:593-602. http://doi.org/d7npfh

2. Sarbin TR. The logic of prediction in psycology. Psychol Rev 1944;51:210-28. http://doi.org/dczxxk

3. Goldberg LR. Simple models or simple processes? Some research on clinical judgments. Am Psychol 1968;23:483-96. http://doi.org/crtcjr

4. Dawes RM, Faust D, Meehl PE. Clinical versus actuarial judgment. Science 1989;243:1668-74. http://doi.org/bpmh5c

5. Alba AC, Agoritsas T, Jankowski M, Courvoisier D, Walter SD, Guyatt GH, Ross HJ. Risk prediction models for mortality in ambulatory patients with heart failure: a systematic review. Circ Heart Fail 2013;6:881-9. http://doi.org/f52pqb

6. Chirino Navarta DA, Rinaldi MC, Aguayo O, Palacios R, Trejo G, Leonardi MA, et al.. Usefullness of Different Mortality Risk Scores in Patients with Heart Failure. A Retrospective Observational Study. Rev Argent Cardiol 2018;86:338-44.

7. Senni M, Paulus WJ, Gavazzi A, Fraser AG, Diez J, Solomon SD, et al., New strategies for heart failure with preserved ejection fraction: the importance of targeted therapies for heart failure phenotypes. Eur Heart J 2014;35:2797-815. http://doi.org/f6mr9g

8. Ferrari R, Böhm M, Cleland JG, Paulus WJ, Pieske B, Rapezzi C, et al. Heart failure with preserved ejection fraction: uncertainties and dilemmas. Eur J Heart Fail 2015;17:665-71. http://doi.org/cvd5

9. Ambardekar AV, Thibodeau JT, DeVore AD, Kittleson MM, Forde-McLean RC, Palardy M, Mountis MM, et al. Discordant Perceptions of Prognosis and Treatment Options Between Physicians and Patients With Advanced Heart Failure. JACC Heart Fail 2017;5:663-71. http://doi.org/cvd6

10. Muntwyler J, Abetel G, Gruner C, Follath F. One-year mortality among unselected outpatients with heart failure. Eur Heart J 2002;23:1861-6. http://doi.org/bh4x62

11. Chew DP, Junbo G, Parsonage W, Kerkar P, Sulimov VA, Horsfall M, et al; Perceived Risk of Ischemic and Bleeding Events in Acute Coronary Syndrome Patients (PREDICT) Study Investigators et al; Perceived risk of ischemic and bleeding events in acute coronary syndromes. Circulation Cardiovasc Qual Outcomes 2013;6:299-308. http://doi.org/f4w3j2