Safety indicators in the STRONG-HF trial from a methodological perspective

Short title: STRONG-HF: methodological perspective

Xavier Rossello[1,2,3,4]

[1] *Cardiology Department, Health Research Institute of the Balearic Islands (IdISBa), Hospital Universitari Son Espases, Palma, Spain*

[2] *Health Research Institute of the Balearic Islands (IdISBa), Palma, Spain*

[3] *Facultad de Medicina, Universitat de les Illes Balears (UIB), Palma, Spain*

[4] *Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain*

Address for correspondence:

Xavier Rossello, Cardiology Department, Hospital Universitari Son Espases-IDISBA. Carretera de Valldemossa, 79, 07120 Palma, Illes Balears, Spain. Tel: +34 871 20 50 50 (ext. 64527), Fax: +34 871 20 68 68.

E-mail address: fjrossello@ssib.es

Word count: 1308 words and 14 references.

Disclosures: None.

Acute heart failure (AHF) is associated with an increased risk of post-discharge clinical outcomes.[1] In patients recently hospitalized for AHF, the Safety, Tolerability and Efficacy of Rapid Optimization, Helped by NT-proBNP Testing, of Heart Failure Therapies (STRONG-HF) trial showed that a high intensity care (HIC) strategy of rapid up-titration of guideline-directed medication (GDMT) and close follow-up reduce the risk of 180-day HF readmission or all-cause death, compared with usual care (UC).[2] Rapid GDMT up-titration involved renin angiotensin system inhibitors (angiotensin converting enzyme inhibitors, or angiotensin receptor blockers in patients intolerant to them, or angiotensin receptor-neprilysin inhibitor),[3] beta-blockers,[4] and mineralocorticoid receptor antagonists,[5] but not sodium/glucose cotransporter 2 inhibitors,[6] whilst close follow-up involved a mean of 4.8 subsequent visits.[2] Patients assigned to HIC were followed up during 180 days according to a set of pre-specified specific safety indicators which were used to guide up-titration at each follow-up visit. Most of these safety indicators have been previously described as predictors of hospital readmission and death in patients with AHF [7], and can be summarised as follows: (1) estimated glomerular filtration rate (eGFR) <30ml/min/1.73m^2, (2) serum potassium >5.0 mmol/L, (3) systolic blood pressure (SBP) <95mmHg,[8] (4) heart rate <55bpm, (5) NT-proBNP concentration >10% higher than pre-discharge values.

In this issue of the Journal of Cardiac Failure, Daniela Tomasoni *et al.* describe that among the 542 patients with AHF enrolled in the HIC arm of the STRONG-HF trial, the occurrence of any safety indicator (as per protocol) was associated with the administration of slightly lower average percentage of GDMT optimal doses (mean difference vs the HIC arm patients not reaching any safety indicator, -11.0% [95% CI -13.6 to -8.4%], P<0.001) without any impact on the occurrence of the primary outcome of 180-day HF readmission or death (HR 0.84, 95% CI 0.48 - 1.46, p=0.54).[9] Of note, the control arm was excluded from this analysis, which makes sense considering that the STRONG-HF trial was one of the reasons to update the 2021 ESC guidelines for the management of heart failure, and it has been a clinical practice changing trial.[10]

Between 1 and 6 weeks after discharge, 57.7% patients in the HIC arm met at least one safety indicator. The primary endpoint of 180-day all-cause death or HF re-admission occurred in 15.0% and 14.2% of patients with and without any safety indicator, respectively. None of the individual safety indicators, considered alone, was significantly associated with the primary endpoint, but there was some of evidence of an association between two safety indicators and the individual components of the primary outcome. SBP < 95mmHg was associated with a trend towards increased 180 days all-cause mortality (p = 0.065), and an eGFR drop to less

than 30ml/min/1.73m^2 with more HF readmissions (p = 0.020). The occurrence of a safety indicator was associated with a smaller 90-day improvement in EQ-5D VAS (adjusted mean difference -3.32 points, 95% CI -5.97 to -0.66, P=0.015).

There are several methodological issues related to this paper that deserve a thorough assessment. First, the study was powered to detect a difference between groups (HIC vs. US), and it is naturally underpowered to detect any difference in the primary outcome between subgroups within the same arm (HIC). However, the study might be well powered to detect a difference in a continuous variable (i.e., percentage of GDMT), which tend to be less challenging to detect than a difference in a binary outcome (i.e., primary endpoint). In this regard, it is very important to consider that absence of evidence is not equivalent to evidence of absence.[11] This is well reflected by the broad 95% confidence interval of the hazard ratio for any safety indicator (which includes a range between 0.48 and 1.46), as well as for each individual safety indicator. Therefore, the risk of type II error regarding the primary endpoint of this sub-study is not small.

The second methodological issue is in the opposite direction, since the risk of type I error regarding the primary and secondary endpoints of this STRONG-HF sub-study is not small. The author reported 18 adjusted models, which result from the comparison of six safety indicator groups and three outcomes (the primary endpoint, and their two individual components). In this context, one of the comparisons is significant, and another one shows a trend towards significance (SBP and eGFR drop, respectively). This might be somehow expected by the rule of chance, and it might be explained by performing multiple comparisons using multiple outcomes without a hierarchical approach, or any correction for alpha spending (e.g., Bonferroni correction).[12] Whilst these two potential associations are biologically plausible, so are the associations between the other safety indicators and the three investigated clinical outcomes. A key point of a trial is about making inferences from a sample to a broader population based on the estimated treatment effect and the degree of uncertainty around it. Some degree of statistical thinking is needed to interpret the findings about SBP and eGFR, which are hypothesis-generating at best, regardless of whether they can be explained by the current scientific knowledge or not.[13]

The third methodological issue is about the individual components of the primary endpoint. These are commonly used in cardiovascular randomised controlled trials to combine evidence across 2 or more outcomes into a single primary endpoint. The danger of them is in oversimplifying the evidence by putting too much emphasis on the composite, without adequate inspection of the contribution from each individual component. In the primary

publication of the STRONG-HF trial, both components of the composite went in the same direction when assessed separately (all-cause death was a key secondary endpoint that showed a non-significant HR of 0.84, whilst heart failure hospitalisation was a pre-specified exploratory outcome that showed a significant HR of 0.56 with a p-value of 0.001).[2] However, in this secondary publication comparing the presence of any safety indicator within the intervention arm, the direction of the estimate (HR) for all-cause mortality was in the opposite direction than for heart failure hospitalisation (HR 0.66 vs. 1.64, respectively). Whilst the study was underpowered to assess the findings with more granularity, this observation is striking. It might be explained by a competing risks issue,[14] since any comparison of non-fatal outcomes ignoring all-cause mortality might be biased (e.g., if a patient dies during the 180-day period, then is no longer eligible to be re-hospitalised). Of note, the number of deaths in the intervention group was 39, whilst the number of heart failure readmissions was 47, according to the primary publication.[2] It might also be explained because of the nature of the association between safety indicators and each individual component of the primary (e.g., safety indicators can be associated with higher incidence of heart failure episodes, but with lower mortality). Given the total number of events, we can only speculate. But the latent idea is that the primary endpoint was great for the primary comparison of HIC vs UC, but perhaps it was not so great for this secondary analysis assessing the association between safety indicators and clinical outcomes within the HIC group. Within the same randomised controlled trial, a one-size-fits-all approach for clinical outcomes might not always work.

Overall, the current study highlights that patients enrolled in the HIC arm of the STRONG-HF trial meeting any safety indicator reached a slightly lower average percentage of GDMT optimal doses without experiencing an increase in the risk of the primary endpoint, compared to those without safety indicators. Regardless of the methodological exercise presented in this editorial, the current study does not provide any evidence about a potential harmful association between the presence of safety indicators and the subsequent incidence of poor outcomes. Based on the STRONG-HF trial,[10] an intensive strategy of initiation and rapid up-tritration of evidence-based treatment before discharge and during frequent and careful follow-up visits in the first six weeks following a HF hospitalisation is recommended to reduce the risk of HF rehospitalisation or death in the 2023 Focused Updated of the 2021 ESC guidelines for the management of heart failure, with a class of recommendation I and a level of evidence B.

Reference list

1. Rossello X, Bueno H, Gil V, Jacob J, Javier Martín-Sánchez F, Llorens P, Herrero Puente P, Alquézar-Arbé A, Raposeiras-Roubín S, López-Díez MP, Pocock S, Miró Ò. MEESSI-AHF risk score performance to predict multiple post-index event and post-discharge short-term outcomes. *Eur Heart J Acute Cardiovasc Care* 2021;**10**:142–152.

2. Mebazaa A, Davison B, Chioncel O, Cohen-Solal A, Diaz R, Filippatos G, Metra M, Ponikowski P, Sliwa K, Voors AA, Edwards C, Novosadova M, Takagi K, Damasceno A, Saidu H, Gayat E, Pang PS, Celutkiene J, Cotter G. Safety, tolerability and efficacy of up-titration of guideline-directed medical therapies for acute heart failure (STRONG-HF): a multinational, open-label, randomised, trial. *Lancet* 2022;**400**:1938–1952.

3. SOLVD Investigators, Yusuf S, Pitt B, Davis CE, Hood WB, Cohn JN. Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure. *N Engl J Med* 1991;**325**:293–302.

4. Packer M, Bristow MR, Cohn JN, Colucci WS, Fowler MB, Gilbert EM, Shusterman NH. The effect of carvedilol on morbidity and mortality in patients with chronic heart failure. U.S. Carvedilol Heart Failure Study Group. *N Engl J Med* 1996;**334**:1349–1355.

5. Pitt B, Zannad F, Remme WJ, Cody R, Castaigne A, Perez A, Palensky J, Wittes J. The effect of spironolactone on morbidity and mortality in patients with severe heart failure. Randomized Aldactone Evaluation Study Investigators. *N Engl J Med* 1999;**341**:709–717.

6. Packer M, Anker SD, Butler J, Filippatos G, Pocock SJ, Carson P, Januzzi J, Verma S, Tsutsui H, Brueckmann M, Jamal W, Kimura K, Schnee J, Zeller C, Cotton D, Bocchi E, Böhm M, Choi D-J, Chopra V, Chuquiure E, Giannetti N, Janssens S, Zhang J, Gonzalez Juanatey JR, Kaul S, Brunner-La Rocca H-P, Merkely B, Nicholls SJ, Perrone S, Pina I, Ponikowski P, Sattar N, Senni M, Seronde M-F, Spinar J, Squire I, Taddei S, Wanner C, Zannad F. Cardiovascular and Renal Outcomes with Empagliflozin in Heart Failure. *New England Journal of Medicine* 2020;**383**:1413–1424.

7. Miró Ò, Rossello X, Platz E, Masip J, Gualandro DM, Peacock WF, Price S, Cullen L, DiSomma S, Oliveira Jr MT de, McMurray JJ, Martín-Sánchez FJ, Maisel AS, Vrints C, Cowie MR, Bueno H, Mebazaa A, Mueller C. Risk stratification scores for patients with acute heart failure in the Emergency Department: A systematic review. *Eur Heart J Acute Cardiovasc Care* 2020;**9**:375–398.

8. Rossello X, Bueno H, Gil V, Jacob J, Martín-Sánchez FJ, Llorens P, Herrero Puente P, Alquézar-Arbé A, Espinosa B, Raposeiras-Roubín S, Müller CE, Mebazaa A, Maggioni AP,

Pocock S, Chioncel O, Miró O. Synergistic Impact of Systolic Blood Pressure and Perfusion Status on Mortality in Acute Heart Failure. *Circ Heart Fail* 2021;**14**:312–323.

9. Tomasoni D, Davison B, Adamo M, Pagnesi M, Mebazaa A, Edwards C, Arrigo M, Barros M, Biegus J, Čelutkienė J, Čerlinskaitė-Bajorė K, Chioncel O, Cohen-Solal A, Damasceno A, Diaz R, Filippatos G, Gayat E, Kimmoun A, Lam CSP, Novosadova M, Pang PS, Ponikowski P, Saidu H, Sliwa K, Takagi K, Maaten JM Ter, Voors A, Cotter G, Metra M. Safety Indicators in Patients Receiving High-intensity Care After Hospital Admission for Acute Heart Failure: The STRONG-HF Trial. *J Card Fail* 2023.

10. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, Burri H, Butler J, Čelutkienė J, Chioncel O, Cleland JGF, Crespo-Leiro MG, Farmakis D, Gilard M, Heymans S, Hoes AW, Jaarsma T, Jankowska EA, Lainscak M, Lam CSP, Lyon AR, McMurray JJ V, Mebazaa A, Mindham R, Muneretto C, Francesco Piepoli M, Price S, Rosano GMC, Ruschitzka F, Skibelund AK, ESC Scientific Document Group. 2023 Focused Update of the 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J* 2023;**44**:3627–3639.

11. Rossello X. The p-value spectrum: from 'absence of evidence' to 'evidence of difference'. *Eur Heart J Acute Cardiovasc Care* 2023.

12. Pocock SJ, Rossello X, Owen R, Collier TJ, Stone GW, Rockhold FW. Primary and Secondary Outcome Reporting in Randomized Trials: JACC State-of-the-Art Review. *J Am Coll Cardiol* 2021;**78**:827–839.

13. Pocock SJ, McMurray JJ V, Collier TJ. Making Sense of Statistics in Clinical Trial Reports: Part 1 of a 4-Part Series on Statistics for Clinical Trials. *J Am Coll Cardiol* 2015;**66**:2536–2549.

14. Rossello X, González-Del-Hoyo M. Survival analyses in cardiovascular research, part II: statistical methods in challenging situations. *Rev Esp Cardiol (Engl Ed)* 2022;**75**:77–85.