Editorial

Biomarker panels for prognosis prediction in heart failure on a CHARM basis

Wouter E. Kok

Cardiology Department, Academic Medical Center, Amsterdam, the Netherlands

Correspondence to: Wouter E. Kok. Cardiology Department, Academic Medical Center, F3-149, Amsterdam, the Netherlands.
Email: w.e.kok@amc.uva.nl.

Comment on: Jackson CE, Haig C, Welsh P, *et al.* The incremental prognostic and clinical value of multiple novel biomarkers in heart failure. Eur J Heart Fail 2016;18:1491-8.

Received: 06 May 2017; Accepted: 30 May 2017; Published: 07 June 2017.
doi: 10.21037/jlpm.2017.05.09
View this article at: http://dx.doi.org/10.21037/jlpm.2017.05.09

As in other medicine fields, science in heart failure is discovering and testing more biomarkers for prognostic use. The natriuretic peptides and troponins—the most cardiac specific of the present biomarkers—have proved useful biomarkers for predicting cardiac events in heart failure patients, on top of established and validated clinical prediction markers (1,2). For those who may be less familiar with the rapidly accumulating number of biomarkers, but also for the investigators trying to sort out how we should best analyse biomarkers in heart failure, the contribution that Jackson *et al.* have made to the growing number of multimarker studies in heart failure is valuable (3). The study demonstrates that in addition to BNP, five biomarkers (mid-regional pro-adrenomedullin, high-sensitive troponin T, combined free light chains, high sensitive CRP and soluble ST2) may aid in predicting 2 to 3 years mortality in a heart failure population of 628 patients, who had a recent hospitalization for heart failure and who survived the first 4 weeks after discharge. The present commentary focuses on the biomarkers that were evaluated in this study, on the model that was used, and on the value of biomarker panels for risk prediction.

Single biomarkers for predicting prognosis, tested against the natriuretic peptides

Previous studies have demonstrated the prognostic value of single biomarkers which are less cardio-specific than natriuretic peptides and troponins, but are relevant within the detrimental pathways for heart failure (soluble ST2, copeptin, mid-regional pro-adrenomedullin, galectin-3, GDF-15, cystatin C) (4-8). Still other biomarkers may add prognostic information such as high sensitive CRP and combined free light chains but with uncertain relevance for cardiac events (7,9,10). All of these newer biomarkers have been tested individually against the natriuretic peptides and appear correlated but also add independent prognostic information to clinical prediction models (3-10). Although galectin-3 level categorization demonstrates prognostic value when used as single biomarker, and adds information on top of a natriuretic peptide in a mixed population of chronic heart failure (7), it was predictive in a population with preserved ejection fraction (11) but did not perform well in two large populations of systolic heart failure when compared to natriuretic peptides (11,12).

Background clinical prediction models for testing biomarkers in heart failure

Adding prognostic information from newer biomarkers appears difficult, and its success not only depends on correlations with other biomarkers, but also on how well a clinical baseline prediction model performs, as measured in the C-statistic (area under the curve of a prediction ROC, also C-index) (10). It does not give an individual's risk assessment, but the probability that the model will discriminate those with events and those without. The change in C-statistic is usually very small per added prognostic variable, with an average improvement of the C-statistic of 0.0036 (13). Biomarkers will predict better if the baseline model has a lower C-statistic. The addition of BNP to the Seattle Heart Failure Model (SHFM)

increased the ROC by +0.03 to +0.06 depending on the baseline C-statistic (1,14). The authors of the present study used the CHARM score which is originally without BNP, and has a C-statistic of 0.75 for a 2-year occurrence of death in 24% (15). Compared to that, the SHFM without BNP has a similar C-statistic of 0.73 for a 1-year occurrence of death in 12%; the 2-year incidence of death in this model was 22% (14). The CHARM score did not have laboratory values or the use of ICD/CRT or the use of medication other than candesartan in the baseline model (15) while the SHFM includes medication, and also sodium and creatinine in the model, which have been all-time predictors of death (14). Adding laboratory values to the CHARM score was done subsequently, including creatinine, but sodium was surprisingly not retained in the final model while bilirubin levels were (16). This explains to some extent why the present study of Jackson *et al.* (3) did not include sodium as a baseline variable, but bilirubin. The authors could have added but did not add heart failure medication or ICD/CRT to the score, which can be seen as a limitation in the real world of evidence based medicine which prescribes these interventions for prognosis. However, in the study by Jackson *et al.* (3) the baseline CHARM score has BNP added and laboratory values, and as such has a C-statistic of 0.72 to begin with.

The patients that were included in the study of Jackson *et al.* were not similar to the patients in the original CHARM model, as the studied patients were recently discharged from the hospital and had a higher event rate of 46% all-cause mortality after a median follow-up of 3.2 years (3). The high event rate and longer follow-up may influence the C-statistic, and the cut points of biomarker levels used for prediction. If patients who are not predicted to die within a year will eventually die in the second year—using a biomarker that only can predict 1 year—the sensitivity of the test will decrease in the second year, as more patients die in the second year with a 'negative' biomarker result; specificity will remain the same or increase, and C-statistic may not be able to tell what is predicted in the first versus the second year. We should keep asking ourselves whether a high incident outcome is predictable or that it is inevitable, and whether we then are identifying the low risk patient better than the high risk patient by our biomarkers (7). In addition, the other discrimination statistics that we are now accustomed to use, the reclassification statistics, inform us better on how a biomarker influences prediction: in general, the reclassification occurs more often towards a lower risk than towards a higher risk, but occurs in both directions. In the study of Jackson *et al.*, we only know the net reclassification index of 33% for the presence of three or more elevated biomarkers in the 3-year observation period. It would have been of some importance to reveal the 1- and 2-year discriminatory reclassification statistics to be able to better interpret the findings of the study. In any case, some indication of the 'durability' and the altered direction of a negative and positive prediction is warranted, and was lacking in the present study.

Multimarker panels and choices how to construct them

Beyond testing individual biomarkers for risk prediction, further steps can be taken by making risk profiles of distinct classes of biomarkers (17,18). There have been other studies than the one by Jackson *et al.* predicting mortality in heart failure using multimarker panels, which have provided insight into risk assessment and how to construct the multimarker panels (19-21). Most studies have all-cause mortality as an endpoint, which is a way to describe the severity of the outcome without interference of competing risks; it is however also restricting a proper interpretation of outcome (sudden death, non-cardiac death). Jackson *et al.* selected 5 of the 9 biomarkers that demonstrated enough importance to the clinical model when added individually (3). Then they made the decision to use dichotomous cut points of biomarker levels that individually had the best combination of sensitivity and specificity, and finally they assessed the values of combinations of 1, 2, 3 to 5 elevated biomarkers (3). The result is that a panel of biomarkers has to contain at least three elevated biomarkers to significantly add risk to the clinical model, with a HR of 2.2 for mortality compared to those patients without any of the biomarkers elevated. The ultimate C-statistic of the CHARM model with BNP and bilirubin added increased from 0.721 to 0.730. This was done comparing those with at least three biomarkers to those without at least three biomarkers (no gradual risk). It was not possible to increase C-statistics to such an extent with a single biomarker. Interestingly, 20.7% of the population of the present study had no biomarker elevated beyond the used cut points and had a 3-year mortality of 18.5%. The chance of having 3 to 5 biomarkers elevated in their study was 39% (with HR 2.2), mainly (around 90% of patients) consisting of combinations of increased levels of mid-regional pro-adrenomedullin (>1,1 nmol/L) and

high-sensitive troponin T (>21.9 ng/mL), and for the third biomarker in 70% of patients either combined free light chains (>51.8 mg/L), or high sensitive CRP (>6.0 mg/L) and to a lesser extent in 47% of patients an elevated soluble ST2 of >28.4 ng/mL (3). No correlations between biomarkers were presented. The result of modeling with dichotomous cut points was compared with the result of modelling with continuous variables, but only for each of the individual markers; a similar increase in C-statistic was obtained when 1 of the 5 ultimately selected biomarkers were added to the clinical model either as continuous versus dichotomous values. There was no formal testing of a multivariate model using all individual biomarkers as continuous variables, so that we would know the interdependence of the individual variables and establish whether we should use all five biomarkers. Other multimarker studies have made a multivariable Cox regression model, from which the regression coefficients were determined of each multivariably significantly contributing biomarker and used it as weight in the multimarker panel (17,19,21). This seems more appropriate. The choice of dichotomizing biomarker levels may present a hazard for modeling, as it is the biomarker level that predicts best for this population and for this time-frame of follow-up. It may however be that for different populations and follow-up durations, say 3 to 12 months or 2 years, that cut points have to vary to keep up with risk estimation (6). For the perspective of whether a panel of biomarkers works better for a model than individual biomarkers, the study has made its point: three biomarkers tell more than one. Other arguments may however also be relevant. For instance, the variability of one measurement itself may be the cause of imperfect risk prediction (22). This may be tested in a study comparing repeated measurement of a single biomarker (to decrease variability) with the use of additional biomarkers with different pathophysiological profiles indicating the severity of heart failure, as is the current hypothesis of multimarker studies. This hypothesis was not tested by the study, but applied.

Future multimarker testing

From a point of view of risk intervention, it is probably not only relevant to exactly identify high mortality risk of patients and use this as reason to add multiple biomarkers to increase the C-statistics, because there will always be more patients at risk than there will be events (inherent imprecision). The finding that we need three or more biomarkers to improve risk categorization between high risk and 'the others' did not improve our views for categorization of these intermediate and lower risk categories. It may be that some biomarkers discriminate better between intermediate and low risk, and these are then probably the ones that add information to selected biomarkers identifying high risk.

We also need a sense of 'durability' of the predictive power of a single measurement of a biomarker, so that we know how to follow up these patients and what a possible treatment target may be at what time. A multimarker panel does not have to result in a score only, if we keep our confidence in the belief that different medications will be appropriate for different pathophysiological states in heart failure. In this respect, clustering of biomarkers and clinical findings may reveal an opportunity to address the question which patients will benefit the most from our current medications (23).

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned and reviewed by Executive Editor Zhi-De Hu (Department of Laboratory Medicine, General Hospital of Ji'nan Military Region, Ji'nan, China).

Conflicts of Interest: The author has completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/jlpm.2017.05.09).The author has no conflicts of interest to declare.

Ethical Statement: The author is accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. May HT, Horne BD, Levy WC, et al. Validation of the Seattle Heart Failure Model in a community-based heart failure population and enhancement by adding B-type natriuretic peptide. Am J Cardiol 2007;100:697-700.
2. Nagarajan V, Hernandez AV, Tang WH. Prognostic value of cardiac troponin in chronic stable heart failure: a systematic review. Heart 2012;98:1778-86.
3. Jackson CE, Haig C, Welsh P, et al. The incremental prognostic and clinical value of multiple novel biomarkers in heart failure. Eur J Heart Fail 2016;18:1491-8.
4. Weinberg EO, Shimpo M, Hurwitz S, et al. Identification of serum soluble ST2 receptor as a novel heart failure biomarker. Circulation 2003;107:721-6.
5. Neuhold S, Huelsmann M, Strunk G, et al. Comparison of copeptin, B-type natriuretic peptide, and amino-terminal pro-B-type natriuretic peptide in patients with chronic heart failure: prediction of death at different stages of the disease. J Am Coll Cardiol 2008;52:266-72.
6. von Haehling S, Filippatos GS, Papassotiriou J, et al. Mid-regional pro-adrenomedullin as a novel predictor of mortality in patients with chronic heart failure. Eur J Heart Fail 2010;12:484-91.
7. Lok DJ, Klip IT, Lok SI, et al. Incremental prognostic power of novel biomarkers (growth-differentiation factor-15, high-sensitivity C-reactive protein, galectin-3, and high-sensitivity troponin-T) in patients with advanced chronic heart failure. Am J Cardiol 2013;112:831-7.
8. Dupont M, Wu Y, Hazen SL, et al. Cystatin C identifies patients with stable chronic heart failure at increased risk for adverse cardiovascular events. Circ Heart Fail 2012;5:602-9.
9. Jackson CE, Haig C, Welsh P, et al. Combined Free Light Chains Are Novel Predictors of Prognosis in Heart Failure. JACC Heart Fail 2015;3:618-25.
10. Wang TJ. Multiple biomarkers for predicting cardiovascular events: lessons learned. J Am Coll Cardiol 2010;55:2092-5.
11. de Boer RA, Lok DJ, Jaarsma T, et al. Predictive value of plasma galectin-3 levels in heart failure with reduced and preserved ejection fraction. Ann Med 2011;43:60-8.
12. Gullestad L, Ueland T, Kjekshus J, et al. The predictive value of galectin-3 for mortality and cardiovascular events in the Controlled Rosuvastatin Multinational Trial in Heart Failure (CORONA). Am Heart J 2012;164:878-83.
13. Ouwerkerk W, Voors AA, Zwinderman AH. Factors influencing the predictive power of models for predicting mortality and/or heart failure hospitalization in patients with heart failure. JACC Heart Fail 2014;2:429-36.
14. Levy WC, Mozaffarian D, Linker DT, et al. The Seattle Heart Failure Model: prediction of survival in heart failure. Circulation 2006;113:1424-33.
15. Pocock SJ, Wang D, Pfeffer MA, et al. Predictors of mortality and morbidity in patients with chronic heart failure. Eur Heart J 2006;27:65-75.
16. Allen LA, Felker GM, Pocock S, et al. Liver function abnormalities and outcome in patients with chronic heart failure: data from the Candesartan in Heart Failure: Assessment of Reduction in Mortality and Morbidity (CHARM) program. Eur J Heart Fail 2009;11:170-7.
17. Wang TJ, Gona P, Larson MG, et al. Multiple biomarkers for the prediction of first major cardiovascular events and death. N Engl J Med 2006;355:2631-9.
18. Braunwald E. Biomarkers in heart failure. N Engl J Med 2008;358:2148-59.
19. Ky B, French B, Levy WC, et al. Multiple biomarkers for risk prediction in chronic heart failure. Circ Heart Fail 2012;5:183-90.
20. Lassus J, Gayat E, Mueller, et al. Incremental value of biomarkers to clinical variables for mortality prediction in acutely decompensated heart failure: the Multinational Observational Cohort on Acute Heart Failure (MOCA) study. Int J Cardiol 2013;168:2186-94.
21. Richter B, Koller L, Hohensinner PJ, et al. A multi-biomarker risk score improves prediction of long-term mortality in patients with advanced heart failure. Int J Cardiol 2013;168:1251-7.
22. Wu AH, Wians F, Jaffe A. Biological variation of galectin-3 and soluble ST2 for chronic heart failure: implication on interpretation of test results. Am Heart J 2013;165:995-9.
23. Ahmad T, Pencina MJ, Schulte PJ, et al. Clinical implications of chronic heart failure phenotypes defined by cluster analysis. J Am Coll Cardiol 2014;64:1765-74.

doi: 10.21037/jlpm.2017.05.09
Cite this article as: Kok WE. Biomarker panels for prognosis prediction in heart failure on a CHARM basis. J Lab Precis Med 2017;2:25.