Published in final edited form as:

Heart Fail Rev. 2010 July ; 15(4): 343–349. doi:10.1007/s10741-009-9144-z.

Multi-marker strategies in heart failure: clinical and statistical approaches

Larry A. Allen and

The Division of Cardiology, University of Colorado Denver, Academic Office 1, Room 7109,
12631 East 17th Avenue, Mail Stop B130, PO Box 6511, Aurora, CO 80045, USA
larry.allen@ucdenver.edu

G. Michael Felker

Duke Clinical Research Institute, 2400 Pratt St, Room 0311 Terrace Level, DUMC Box 3850,
Durham, NC 27715, USA

Abstract

Advances in genomics and proteomics promise to transform biomarker research, in which the major challenges will not be the discovery of new markers but rather the optimal selection and validation of a subgroup of clinically useful markers from the large pool of candidates. Critically, the value of new biomarkers panels will need to be assessed in the context of readily available clinical information in order to create more actionable knowledge rather than just greater complexity. Appropriate methodologies for the clinical and statistical evaluation of so called “multi-marker strategies” have not been systematically defined. Although specific criteria for the appropriate clinical and statistical evaluation of multi-marker strategies will vary based on the intended use (e.g., diagnosis vs. screening), the ultimate measure of success is the ability for a biomarker panel to both correct a meaningful portion of misclassification by standard methods (discrimination) and to improve quantification of absolute risk (calibration) in comparison to existing clinical information. Findings should be validated in an independent dataset of the representative patient population before a given multi-marker strategy can be considered for clinical use. Here, we define multi-marker strategies, summarize recent examples of biomarker combinations in heart failure, address key statistical and clinical issues, and discuss future directions for this rapidly evolving field.

Keywords

Biomarkers; Multi-marker panels; Combinations; Heart failure; Diagnosis; Screening

What is a multi-marker strategy?

Broadly defined, the term “biomarker” describes any measurable biologic parameter that provides insight into biologic processes in health or disease [1, 2]. Under this definition, blood tests, imaging studies, and physical examination findings could all be considered biomarkers. In common usage, however, the term biomarker typically refers to a quantifiable parameter that is measured from a biological sample such as blood or urine. The use of such biomarkers is now ubiquitous in clinical medicine generally and cardiovascular disease particularly. Accepted clinical uses of biomarkers include diagnosis, screening, prognosis/risk stratification, and selection or titration of therapies [3]. In addition to these

clinical uses, biomarkers may facilitate understanding of fundamental biologic processes (“bedside to bench”) and may also be useful as “surrogate endpoints” in the development and evaluation of new therapies.

Traditionally, potential biomarkers have been selected based on understanding of underlying mechanisms. Candidate markers thus identified were assessed for potential clinical utility and analytical characteristics, and subsequently validated for clinical use [4]. As reviewed elsewhere in this issue, a variety of scientific and technological advances have altered the landscape of biomarker research. The completion of the Human Genome Project and the burgeoning field of proteomics, spurred by advances in high throughput technologies such as mass spectrometry and microarrays, will result in the simultaneous identification of a vast array of potential new biomarkers for any given disease state [5, 6]. Increasingly, panels of multiple markers will be easily measured simultaneously from a single biologic sample. These developments promise to create a fundamental shift in the paradigm of biomarker research, in which the major challenges will not be the discovery of new markers but rather the selection and validation of a subgroup of clinically useful markers from the large pool of candidates. Critically, the value of these new markers will need to be assessed in the context of readily available clinical information, including existing markers and other relevant demographic and clinical variables. Appropriate methodologies for the clinical and statistical evaluation of so called “multi-marker strategies” have not been systematically defined. Careful attention to the optimal methods for evaluating new markers will be needed in order to maximize the impact of technological advances on clinical care, creating more actionable knowledge rather than just greater complexity.

Given the high prevalence and rising incidence of heart failure in combination with its extreme rates of morbidity and mortality, the development of biomarkers for diagnosis and risk assessment in heart failure will remain a priority [7]. In this article, we outline the components of an idealized multi-marker strategy, review published data on multi-marker strategies in heart failure, and discuss various statistical approaches to evaluating the additive information from new markers in the context of existing knowledge.

Components of an ideal multi-marker strategy

A variety of characteristics have been proposed for the idealized biomarker in cardiovascular disease [2, 3, 8]. From an analytic perspective, an ideal biomarker should be able to be measured with precision from easily obtained biologic samples (e.g., blood rather than tissue) at a reasonable cost and in a timely manner. The ideal multi-marker platform would measure the relevant markers simultaneously from a single biologic sample, and report results in a way that integrates multiple markers’ values into simplified results (such as a risk score) that could be applied directly to clinical decision making.

Clinically, idealized characteristics for individual biomarkers or multi-marker panels vary depending on the intended use. In the diagnosis of patients with acute clinical symptoms (such as acute heart failure), markers that are highly sensitive (at the cost of lower specificity) are preferred, given that a false negative result may lead to delay in appropriate medical management in a patient with potentially life-threatening illness. Conversely, in screening for diseases of low prevalence (such as asymptomatic LV dysfunction in community populations), highly specific markers are preferred due to the prohibitive cost of subsequent testing if a large number of “false positives” are identified during screening.

Several general concepts inform the type of multi-marker strategies most likely to be clinically useful. Fundamentally, markers that reflect distinct biological processes are more likely to provide incremental clinical benefit than markers that relate to the same general physiologic phenomenon [9]. For example, in heart failure, the combination of markers of

hemodynamic stress (e.g., b-type natriuretic peptide [BNP]) and myocyte necrosis (e.g., troponin) is more likely to be clinically useful than the combination of markers reflecting similar pathophysiology (e.g., BNP and atrial natriuretic peptide [ANP]). Although this concept appears self-evident when discussing markers for which the fundamental underlying mechanisms are well understood, use of “unbiased” proteomic techniques for biomarker identification will increasingly yield potential markers with poorly described or even unknown biological functions. Optimal methods for choosing which individual biomarkers to include in a multi-marker panel have not been well established. Traditionally those few biomarkers with the strongest independent statistical association with disease or outcome have been chosen, with some emphasis on selecting biomarkers that capture distinct aspects of the disease process.

Existing multi-marker strategies in heart failure

Multi-marker strategies are increasingly used as clinical tools in the risk stratification of patients presenting with chest pain and acute coronary syndromes [10–13] and for predicting future risk of cardiovascular disease [14, 15]. In contrast, the field of multi-marker evaluation in heart failure is in its infancy. Those studies that have evaluated multi-marker strategies in heart failure have focused primarily on the prediction of prognosis, in part because current statistical techniques (multivariable modeling) are well suited to this sort of analysis. Below, we present a few representative studies from the literature to provide examples of the current state of multi-marker strategies in heart failure.

Ralli et al. [16] described the relationship between anemia, cardiac troponin I, and BNP with mortality in patients with advanced heart failure. Using a cohort of 264 patients with advanced heart failure referred to University of California Los Angeles between 1999 and 2004 for whom these biomarker data were available at the time of initial evaluation, they divided patients into groups based on the presence or absence of anemia, detectable or undetectable cardiac troponin I, and elevated or normal BNP. At 1-year, there were 46 deaths and 60 transplants. Survivors had significantly higher hemoglobin levels versus non-survivors (13.2 ± 1.7 vs. 12.0 ± 1.8 g/dl, $P < .001$), lower serum BNP (702 ± 789 vs. $1,553 \pm 1,316$ pg/ml, $P < .001$), and decreased troponin positivity (43.6% vs. 71.7%, $P < .001$). Low hemoglobin (adjusted RR 2.3, 1.1–5.0), elevated BNP (adjusted RR 17.3, 2.2–135), and detectable troponin (adjusted RR 4.0, 1.5–8.9) all remained independent predictors of 1-year mortality on multivariable analysis when considered with 22 other baseline variables (other statistically significant predictors were NYHA class IV, serum creatinine, serum sodium, albumin, and total cholesterol). The combination of these biomarkers provided added information beyond consideration of the biomarkers individually. The presence of anemia in conjunction with detectable troponin defined a group with markedly increased mortality (adjusted RR 5.3), whereas neither the presence of anemia with undetectable troponin nor the absence of anemia with detectable troponin conferred a significantly increased risk of mortality. Consequently, there was a *synergistic* interaction between anemia and troponin positivity in predicting risk ($P = .05$, $\beta = .92$). Similarly, anemia in conjunction with elevated BNP was associated with a significantly increased mortality risk (adjusted RR 10.4; 64% 1-year survival). A synergistic interaction between anemia and BNP was of borderline significance ($P = .11$, $\beta = .75$). Using these biomarkers together, patients positive for all 3 (44 of 264 patients) were at particularly high risk for death (58.4% at 1 year), nearly double the mortality rate of any other biomarker subgroup (Fig. 1). Conversely, absence of both anemia and BNP elevation identifies patients with exceptionally good prognosis (96.3% at 1 year).

Another example of a multi-marker strategy in heart failure is from van Kimmenade et al. [17]. Given the highly predictive power of both NT-proBNP and parameters of renal

function relative to other variables on outcomes in multivariable analyses, they sought to study the integrative role of these two biomarkers together in a multi-marker strategy. Using 720 patients dichotomized by the median of glomerular filtration rate (GFR, cutoff 60 ml/min/1.73 m²) and NT-proBNP (cutoff 4,647 pg/ml), they found that the combination of low GFR and elevated NT-proBNP was the best predictor of 60-day mortality (odds ratio 3.46, 95% confidence interval 2.13–5.63). In the subgroup of patients with NT-proBNP below the median, prognosis was not significantly influenced by markers of renal function. Thus, considering markers of renal function in context of NT-proBNP levels in a multi-marker model adds important information to the interpretation of those markers beyond their consideration in isolation.

Other multi-marker strategies involving novel biomarkers of prognosis have recently been published. These include ST2 in concert with NT-proBNP and other serum markers [18, 19], as well as galectin-3 with NT-proBNP [20]. Generally, the statistical approaches outlined in the above examples have been replicated in these studies.

Anand et al. [21] evaluated the incremental predictive value of a variety of heart failure biomarkers to a baseline clinical risk model. In this analysis of 3,720 chronic heart failure patients from the Valsartan Heart Failure Trial (Val-HeFT), four biomarkers that had been previously demonstrated to be prognostic in heart failure were assessed: NT-proBNP, norepinephrine (NE), C-reactive protein (CRP), and troponin T. A baseline model of clinical predictors of 1-year mortality was generated, which included age, gender, NYHA class, systolic blood pressure, cholesterol, blood urea nitrogen (BUN), hemoglobin, uric acid, and ejection fraction. The c-statistic for 1-year mortality for this “clinical model” was 0.69, suggesting moderate discriminatory ability. Addition of the four candidate biomarkers to the model increased the c-statistic for the overall model to 0.73. Sequential removal of troponin T, CRP, and NE did not reduce the c-statistic significantly, suggesting that NT-proBNP alone added significant prognostic information. Notably, the c-statistic for NT-proBNP alone was 0.71, approximating the prognostic utility of the best clinical model. Thus, using the incremental change in c-statistic for the evaluation of a potential multi-marker strategy in the end simplified to a single biomarker. In contrast to the above analysis of traditional troponin T (detectable in 10.4% of the Val-HeFT population), high-sensitivity troponin T (detectable in 92% of the Val-HeFT population) was significantly additive to the base clinical model c-statistic, with and without BNP in the model [22]. This difference demonstrates how the characteristics of an assay may significantly alter the predictive performance of a biomarker.

Statistical considerations in evaluation of multi-marker strategies

Appropriate statistical techniques for assessing new markers will differ depending on the clinical purposes for which the new marker is being evaluated. In diagnostic testing and screening, the fundamental requirement is the separation of a population into two distinct groups (those with the disease and those without). For these purposes, Bayesian analyses focused on changes in pre and post test probabilities, likelihood ratios, and positive and negative predictive value are most appropriate [23]. Receiver operator characteristic (ROC) curves can be useful in determining the discriminatory capacity of a biomarker, identifying optimal cut-points for decision making, and evaluating the trade off between sensitivity and specificity. Bayesian analysis using likelihood ratios may be particularly useful for evaluating multi-marker approaches due to the ability to use them in series (e.g., post-test probability from marker #1 becomes pre-test probability for marker #2, etc.).

For markers used in prognostication or risk stratification, multivariable modeling techniques (such as logistic regression or Cox proportional hazards analysis) are traditionally used to

assess the contribution of individual predictors to the overall risk of adverse outcomes. The amount of risk associated with elevation of a given biomarker is typically reported as a hazard ratio (for Cox proportional hazards models) or an odds ratio (for logistic regression models). Numerically, the value for such ratios is dependent on the increment reported (e.g., hazard ratio of x per y change in a continuous variable), making comparison between studies difficult. Standardized hazard ratios (i.e., the hazard associated with a change of 1 standard deviation in a continuous variable) provide some degree of standardization and simplify comparisons between studies, but are rarely used. Dichotomization of biomarker data (e.g., above vs. below the median in a given population) may simplify analysis but sacrifices statistical power. Additional complexity is introduced by the fact that biomarker values are frequently non-normally distributed and often require transformation (such as log transformation) before they conform to parametric assumptions. Subsequent difficulties involving interpretation of transformed variables in the clinical setting can render such analyses effectively useless to the practicing clinician.

Hazard ratios and P -values may be deceptive measures of the true clinical value of a given biomarker. For example, studies with very large sample sizes may generate highly significant P -values for markers with only modest association with risk (i.e., hazard ratio close to 1), leading to a potential overestimation of the importance or utility of the markers in question. Alternatively, a high hazard or odds ratio may not necessarily correlate with a high degree of clinical utility for all indications. Even markers with very high hazard ratios for a given outcome often have substantial overlap in the distribution of values between those patients with or without the disease of interest. Such a scenario requires a trade off between sensitivity (negative predictive value) and specificity (positive predictive value). BNP in the diagnosis of acute dyspnea provides an example of this phenomenon. In the Breathing Not Properly study, $\text{BNP} > 100$ pg/ml was very strongly correlated with the diagnosis of heart failure in multivariable modeling (odds ratio 29.6, $P < 0.001$) [24]. Despite this extremely high odds ratio, the specificity (76%) and positive predictive value (79%) of BNP at this cut-off were modest. As noted above, this necessary trade off between sensitivity and specificity must be informed by an understanding of the clinical implications of both false negative and false positive results.

A critical statistical issue in evaluating new biomarkers in heart failure (or any disease) is the assessment of the “value added” to existing data by the new marker in question. The most relevant question is not “Is this a good marker?”, but rather “Does the addition of this marker improve current clinical decision making to the degree that the added cost/complexity is justified?” For multi-variable models, changes in the measures of overall model performance (e.g., the c-statistic for logistic regression) can assess the incremental improvements in model performance with each added marker (Fig. 2). This approach has been suggested as the optimal means for assessing the added value of new markers to clinical prediction in the context of available information [25].

However, others have argued the pitfalls of reliance on c-statistics for assessment of the predictive value of a new biomarker [26]. The c-statistic only provides the discriminatory ability of the biomarker (its ability to classify patients into those with and without disease/outcome), which is often the important characteristic for diagnostic testing but which may not be the key factor in risk modeling; the c-statistic does not make any comment on the calibration of a biomarker (its ability to accurately quantify absolute risk). Additionally, the use of ROC curves to determine added value may be less directly applicable for survival models that have a time to event component (such as Cox proportional hazards models—widely used in heart failure studies). Prognostic discrimination analyses may be performed by converting the outcome to a dichotomous variable at a fixed point in time (e.g., 1-year mortality), but with resultant loss of statistical power when switching from time to event

analysis to logistic regression [21, 27]. Complex statistical techniques have been developed to approximate ROC curves for Cox models, but they are not widely utilized [28, 29].

Statistical versus clinical significance

There is little agreement on what degree of change in the overall model performance would represent a clinically meaningful improvement. Critically, there may be significant discrepancy between what represents a *statistically significant* improvement in model performance, and what represents a *clinically meaningful* addition to decision making. In interpreting clinical trial results, for example, metrics are available to assist clinicians in applying risk reductions to individual patients, such as the number needed to treat (NNT) to avoid one adverse outcome. Such constructs are not readily available for the statistical analysis of model performance, resulting in a potential disconnect between statistical and clinical significance. Finally, methodology such as the change in the model c-statistic is substantially influenced by the order in which variables are added to the model. The addition of any variable to a model already containing many strong predictors of outcome is unlikely to result in a dramatic change in the c-statistic. For example, in the analysis from Val-HeFT discussed above [21], inclusion of NT-proBNP as the initial variable in the model would have resulted in the conclusion that many other known powerful predictors (such as age, ejection fraction, BUN, etc.) were of little prognostic importance. Such a conclusion would obviously be inconsistent with a large body of data and clinical experience. This example makes the point that statistical significance is only one property of potential new markers to be considered, along with availability, ease of interpretation, and cost. Novel markers that provide significant advantages in ease of use or low cost may be adopted even if they are not statistically superior to existing markers.

As a general principle, models perform better in their derivation dataset than they do in independent datasets, and this may be particularly true for multi-marker data where individual markers are significantly correlated with each other. Validation of multi-marker tools in independent datasets is, therefore, a critical but infrequently reported step in evaluating multi-markers strategies in clinical medicine. Such validation not only improves confidence in the estimates of marker performance, but also extends the generalizability of observations by demonstrating their utility in different populations [27].

Finally, the use of multivariable modeling, while statistically powerful, often results in models that are too complex to be clinically useful. Such models may be simplified by using dichotomous (rather than continuous) assessment of each variable, but such an approach sacrifices statistical power. A variety of such risk scores that incorporate biomarkers data are present in the literature, but adoption of these models into routine clinical practice has been limited [30–34].

Future directions

For the immediate future, multi-marker strategies are likely to develop downstream of the identification and validation of individual biomarkers. To date, development in this area has been characterized by slow acceptance of a relative few biomarkers into routine clinical practice after extensive testing, followed by evaluation of pairs or triplets of well-validated markers together in simple multi-marker strategies. In the genomic era, this step-wise approach will become increasingly impractical due to the large number of new markers to be evaluated.

Central to the development of multi-marker strategies is the tension between maximizing total information provided by a multi-marker panel and minimizing complexity in analysis and interpretation of the results. In order to limit complexity, most existing strategies have

favored relatively simple approaches. As the field moves forward, one option is to take advantage of advances in the field of computation biology by employing artificial neural networks to analyze complex changes in multiple biomarkers simultaneously [35–37]. The main advantage of neural networks over traditional statistical techniques is that the model does not have to be explicitly defined before beginning the analysis. Neural networks can be trained to recognize the relevant patterns in data (linear and non-linear), whereas a statistical model requires prior knowledge of the relationships between the factors under investigation. Neural networks are also better able to combine both ordinal and nominal data, retaining statistical power.

Not only do neural networks promise a method for interpreting highly complex panels of biomarkers in relation to diagnosis and prognosis, they may also be better equipped to identify hidden domains that underlie the pathophysiology of the disease. Certain biomarkers may have greater relative value in multi-marker applications than on their own. For example, the independent adjusted likelihood or hazards ratio of a biomarker may be lower than many other biomarkers, but if the biomarker identifies a distinct biochemical mechanism underlying the disease process then it may have relatively greater incremental value when added to an overall model. Traditional statistical approaches that first evaluate biomarkers individually and only later assess value in combination with other markers may potentially filter out important information, suggesting the benefits of more “unbiased” approaches in designing multi-marker strategies. Exploratory factor analysis provides a statistical approach for elucidating the mechanistic structure, which underlies heart failure within a patient population [38]. If a handful of domains could be identified through factor analysis, then key combinations of biomarkers, which optimally associated with each domain, could be chosen for inclusion into a multi-marker panel. Theoretically, this would lead to an improved strategy for biomarker selection, which would optimize predictive performance.

Conclusions

Although specific criteria for the appropriate clinical and statistical evaluation of multi-marker strategies will vary based on the intended use (e.g., diagnosis vs. screening), a few general statements can be made.

1. Candidate biomarkers should be shown to have statistical association with disease or outcome that is independent of traditional risk factors and other well-known biomarkers.
2. Utility of new markers should be evaluated with a focus on the incremental information added compared to methods of determining diagnosis or prognosis. The ultimate yardstick of success for a multi-marker strategy is its ability to correct a meaningful portion of misclassification by standard methods (discrimination). At the same time, attention should be paid to how well the biomarker improves quantification of absolute risk (calibration).
3. Findings should be validated in an independent dataset of the representative patient population before a given strategy can be considered for clinical use.

Future development of multi-marker approaches has the potential to improve heart failure care along the entire spectrum of disease, by improving screening, simplifying diagnosis, clarifying prognosis, and tailoring treatment. No single biomarker can accomplish all these goals in isolation, suggesting that multi-markers approaches are likely to become increasingly prevalent in heart failure care. Careful consideration of optimal approaches to the clinical and statistical evaluation of such multi-marker strategies will be critical in order

to maximize the clinical impact of the advances in technology and improve the care of patients with heart failure.

References

1. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther.* 2001; 69:89–95. doi:10.1067/mcp.2001.113989. [PubMed: 11240971]
2. Vasan RS. Biomarkers of cardiovascular disease: molecular basis and practical considerations. *Circulation.* 2006; 113:2335–2362. doi:10.1161/CIRCULATIONAHA.104.482570. [PubMed: 16702488]
3. Mark DB, Felker GM. B-type natriuretic peptide—a biomarker for all seasons? *N Engl J Med.* 2004; 350:718–720. doi: 10.1056/NEJMe038233. [PubMed: 14960748]
4. Morrow DA, de Lemos JA. Benchmarks for the assessment of novel cardiovascular biomarkers. *Circulation.* 2007; 115:949–952. doi:10.1161/CIRCULATIONAHA.106.683110. [PubMed: 17325253]
5. Donahue MP, Marchuk DA, Rockman HA. Redefining heart failure: the utility of genomics. *J Am Coll Cardiol.* 2006; 48:1289–1298. doi:10.1016/j.jacc.2006.05.062. [PubMed: 17010784]
6. Arab S, Gramolini AO, Ping P, et al. Cardiovascular proteomics: tools to develop novel biomarkers and potential applications. *J Am Coll Cardiol.* 2006; 48:1733–1741. doi:10.1016/j.jacc.2006.06.063. [PubMed: 17084242]
7. Braunwald E. Biomarkers in heart failure. *N Engl J Med.* 2008; 358:2148–2159. doi:10.1056/NEJMra0800239. [PubMed: 18480207]
8. Morrow DA, de Lemos JA, Sabatine MS, Antman EM. The search for a biomarker of cardiac ischemia. *Clin Chem.* 2003; 49:537–539. doi:10.1373/49.4.537. [PubMed: 12651803]
9. Morrow DA, Braunwald E. Future of biomarkers in acute coronary syndromes: moving toward a multimarker strategy. *Circulation.* 2003; 108:250–252. doi:10.1161/01.CIR.0000078080.37974.D2. [PubMed: 12876133]
10. Newby LK, Storrow AB, Gibler WB, et al. Bedside multimarker testing for risk stratification in chest pain units: The chest pain evaluation by creatine kinase-MB, myoglobin, and troponin I (CHECKMATE) study. *Circulation.* 2001; 103:1832–1837. [PubMed: 11294799]
11. James SK, Lindback J, Tilly J, et al. Troponin-T and N-terminal pro-B-type natriuretic peptide predict mortality benefit from coronary revascularization in acute coronary syndromes: a GUSTO-IV substudy. *J Am Coll Cardiol.* 2006; 48:1146–1154. doi:10.1016/j.jacc.2006.05.056. [PubMed: 16978997]
12. Sabatine MS, Morrow DA, de Lemos JA, et al. Multi-marker approach to risk stratification in non-ST elevation acute coronary syndromes: simultaneous assessment of troponin I, C-reactive protein, and B-type natriuretic peptide. *Circulation.* 2002; 105:1760–1763. doi:10.1161/01.CIR.0000015464.18023.0A. [PubMed: 11956114]
13. Westerhout CM, Fu Y, Lauer MS, et al. Short- and long-term risk stratification in acute coronary syndromes: the added value of quantitative ST-segment depression and multiple biomarkers. *J Am Coll Cardiol.* 2006; 48:939–947. doi:10.1016/j.jacc.2006.04.085. [PubMed: 16949483]
14. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation.* 1998; 97:1837–1847. [PubMed: 9603539]
15. Ridker PM, Rifai N, Rose L, Buring JE, Cook N. Comparison of C-reactive protein and low-density lipoprotein cholesterol levels in the prediction of first cardiovascular events. *N Engl J Med.* 2002; 347:1557–1565. doi:10.1056/NEJMoa021993. [PubMed: 12432042]
16. Ralli S, Horwich TB, Fonarow GC. Relationship between anemia, cardiac troponin I, and B-type natriuretic peptide levels and mortality in patients with advanced heart failure. *Am Heart J.* 2005; 150:1220–1227. doi:10.1016/j.ahj.2005.01.049. [PubMed: 16338262]
17. van Kimmenade RR, Januzzi JL Jr, Baggish AL, et al. Amino-terminal pro-brain natriuretic peptide, renal function, and outcomes in acute heart failure: redefining the cardiorenal interaction? *J Am Coll Cardiol.* 2006; 48:1621–1627. doi:10.1016/j.jacc.2006.06.056. [PubMed: 17045898]

18. Januzzi JL Jr, Peacock WF, Maisel AS, et al. Measurement of the interleukin family member ST2 in patients with acute dyspnea: results from the PRIDE (Pro-Brain Natriuretic Peptide Investigation of Dyspnea in the Emergency Department) study. *J Am Coll Cardiol.* 2007; 50:607–613. doi:10.1016/j.jacc.2007.05.014. [PubMed: 17692745]
19. Rehman SU, Martinez-Rumayor A, Mueller T, Januzzi JL Jr. Independent and incremental prognostic value of multi-marker testing in acute dyspnea: results from the ProBNP Investigation of Dyspnea in the Emergency Department (PRIDE) study. *Clin Chim Acta.* 2008; 392:41–45. doi: 10.1016/j.cca.2008.03.002. [PubMed: 18387360]
20. van Kimmenade RR, Januzzi JL Jr, Ellinor PT, et al. Utility of amino-terminal pro-brain natriuretic peptide, galectin-3, and apelin for the evaluation of patients with acute heart failure. *J Am Coll Cardiol.* 2006; 48:1217–1224. doi:10.1016/j.jacc.2006.03.061. [PubMed: 16979009]
21. Anand IS, Rector T, Latini R, Masson S, Maggioni AP, Cohn JN. Do biomarkers add prognostic information to routine measures of the severity of heart failure? *J Am Coll Cardiol.* 2006; 47(4A):A66–A67. doi:10.1016/j.jacc.2005.12.054.
22. Latini R, Masson S, Anand IS, et al. Prognostic value of very low plasma concentrations of troponin T in patients with stable chronic heart failure. *Circulation.* 2007; 116:1242–1249. doi: 10.1161/CIRCULATIONAHA.106.655076. [PubMed: 17698733]
23. Diamond GA, Denton TA, Berman DS, Cohen I. Prior restraint: a Bayesian perspective on the optimization of technology utilization for diagnosis of coronary artery disease. *The American Journal of Cardiology.* 1995; 76:82–86. doi:10.1016/S0002-9149(99)80809-1. [PubMed: 7793413]
24. Maisel AS, Krishnaswamy P, Nowak RM, et al. Rapid measurement of b-type natriuretic peptide in the emergency diagnosis of heart failure. *N Engl J Med.* 2002; 347:161–167. doi: 10.1056/NEJMoa020233. [PubMed: 12124404]
25. Kattan MW. Judging new markers by their ability to improve predictive accuracy. *J Natl Cancer Inst.* 2003; 95:634–635. [PubMed: 12734304]
26. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation.* 2007; 115:928–935. doi:10.1161/CIRCULATIONAHA.106.672402. [PubMed: 17309939]
27. Baggish AL, Siebert U, Lainchbury JG, et al. A validated clinical and biochemical score for the diagnosis of acute heart failure: the ProBNP Investigation of Dyspnea in the Emergency Department (PRIDE) Acute Heart Failure Score. *Am Heart J.* 2006; 151:48–54. doi:10.1016/j.ahj.2005.02.031. [PubMed: 16368291]
28. Zheng Y, Cai T, Feng Z. Application of the time-dependent roc curves for prognostic accuracy with multiple biomarkers. *Biometrics.* 2006; 62:279–287. doi:10.1111/j.1541-0420.2005.00441.x. [PubMed: 16542256]
29. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics.* 2005; 61:92–105. doi:10.1111/j.0006-341X.2005.030814.x. [PubMed: 15737082]
30. Morrow DA, Antman EM, Charlesworth A, et al. TIMI risk score for ST-elevation myocardial infarction: a convenient, bedside, clinical score for risk assessment at presentation: an intravenous nPA for treatment of infarcting myocardium early II trial substudy. *Circulation.* 2000; 102:2031–2037. [PubMed: 11044416]
31. Levy WC, Mozaffarian D, Linker DT, et al. The Seattle Heart Failure Model: prediction of survival in heart failure. *Circulation.* 2006; 113:1424–1433. doi:10.1161/CIRCULATIONAHA.105.584102. [PubMed: 16534009]
32. Felker GM, Leimberger JD, Califf RM, et al. Risk stratification after hospitalization for decompensated heart failure. *Journal of Cardiac Failure.* 2004; 10:460–466. doi:10.1016/j.cardfail.2004.02.011. [PubMed: 15599835]
33. Fonarow GC, Adams KF Jr, Abraham WT, Yancy CW, Boscardin WJ for the Adhere Scientific Advisory Committee Study Group and Investigators. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. *J Am Med Assoc.* 2005; 293:572–580. doi:10.1001/jama.293.5.572.
34. Aaronson KD, Schwartz JS, Chen TM, Wong KL, Goin JE, Mancini DM. Development and prospective validation of a clinical index to predict survival in ambulatory patients referred for cardiac transplant evaluation. *Circulation.* 1997; 95:2660–2667. [PubMed: 9193435]

35. Atienza F, Martinez-Alzamora N, De Velasco JA, Dreiseitl S, Ohno-Machado L. Risk stratification in heart failure using artificial neural networks. *Proc AMIA Symp.* 2000; 3:2–6.
36. Ortiz J, Ghefter CG, Silva CE, Sabbatini RM. One-year mortality prognosis in heart failure: a neural network approach based on echocardiographic data. *J Am Coll Cardiol.* 1995; 26:1586–1593. doi:10.1016/0735-1097(95)00385-1. [PubMed: 7594090]
37. Penny W, Frost D. Neural networks in clinical medicine. *Med Decis Making.* 1996; 16:386–398. doi:10.1177/0272989X9601600409. [PubMed: 8912300]
38. Goodman E, Dolan LM, Morrison JA, Daniels SR. Factor analysis of clustered cardiovascular risks in adolescence: obesity is the predominant correlate of risk among youth. *Circulation.* 2005; 111:1970–1977. doi:10.1161/01.CIR.0000161957.34198.2B. [PubMed: 15837951]

Fig. 1. Multi-marker prediction of 1-year mortality in chronic heart failure based on dichotomization of three continuous variables (hemoglobin, BNP, and troponin I). From Ralli et al. [14], with permission

Fig. 2. Methodology for displaying sequential added value of individual biomarkers for risk prediction in acute coronary syndromes using the change in the c-index from logistic regression models. From Westerhaut et al. [13], with permission