Multiple Biomarkers for Risk Prediction in Chronic Heart Failure

Bonnie Ky, MD, MSCE; Benjamin French, PhD; Wayne C. Levy, MD; Nancy K. Sweitzer, MD, PhD; James C. Fang, MD; Alan H.B. Wu, PhD; Lee R. Goldberg, MD, MPH; Mariell Jessup, MD; Thomas P. Cappola, MD, ScM

Background—Prior studies have suggested using a panel of biomarkers that measure diverse biological processes as a prognostic tool in chronic heart failure. Whether this approach improves risk prediction beyond clinical evaluation is unknown.

Methods and Results—In a multicenter cohort of 1513 chronic systolic heart failure patients, we measured a contemporary biomarker panel consisting of high-sensitivity C-reactive protein, myeloperoxidase, B-type natriuretic peptide, soluble fms-like tyrosine kinase receptor-1, troponin I, soluble toll-like receptor-2, creatinine, and uric acid. From this panel, we calculated a parsimonious multimarker score and assessed its performance in predicting risk of death, cardiac transplantation, or ventricular assist device placement in comparison to an established clinical risk score, the Seattle Heart Failure Model (SHFM). During a median follow-up of 2.5 years, there were 317 outcomes: 187 patients died; 99 were transplanted; and 31 had a ventricular assist device placed. In unadjusted Cox models, patients in the highest tertile of the multimarker score had a 13.7-fold increased risk of adverse outcomes compared with the lowest tertile (95% confidence interval, 8.75–21.5). These effects were independent of the SHFM (adjusted hazard ratio, 6.80; 95% confidence interval, 4.18–11.1). Addition of the multimarker score to the SHFM led to a significantly improved area under the receiver operating characteristic curve of 0.803 versus 0.756 (P=0.003) and appropriately reclassified a significant number of patients who had the outcome into a higher risk category (net reclassification improvement, 25.2%; 95% confidence interval, 14.2–36.2%; P<0.001).

Conclusions—In ambulatory chronic heart failure patients, a score derived from multiple biomarkers integrating diverse biological pathways substantially improves prediction of adverse events beyond current metrics. (Circ Heart Fail. 2012;5:183-190.)

Key Words: biomarkers ■ chronic heart failure

Heart failure is a major public health burden that accounts for at least 290,000 deaths in the United States alone each year.1 There is substantial variation in the severity and prognosis of heart failure, ranging from mild disease that is easily managed with neurohormonal blockade to advanced illness requiring therapy with mechanical support or heart transplantation.2 Accurate assessment of prognosis is critical to guide clinical management and to identify high-risk patients who should be considered for advanced therapy. Although established predictors such as the New York Heart Association class, left ventricular ejection fraction, natriuretic peptide levels, and risk scores such as the Seattle Heart Failure Model (SHFM) exist, they do not fully explain the risk of adverse outcomes in ambulatory chronic heart failure patients.2

Clinical Perspective on p 190

The progression of heart failure is complex and is driven by multiple biological processes, including inflammation, oxidative stress, neurohormonal activation, vascular remodeling, myocyte injury, and renal impairment.3 As such, there has been growing interest in the measurement of a diverse biomarker profile, reflective of the underlying biology of heart failure, as a means to risk-stratify patients...
and improve our understanding of the underlying pathophysiology.4 We therefore evaluated the predictive utility of 8 biomarkers, reflective of diverse biological pathways in heart failure: high-sensitivity C-reactive protein (hsCRP) (inflammation), uric acid and myeloperoxidase (MPO) (oxidative stress), B-type natriuretic peptide (BNP) (neurohormonal activation), soluble fms-like tyrosine kinase receptor-1 (sFlt-1) (vascular remodeling), troponin I (TnI) (myocyte injury), soluble toll-like receptor-2 (ST2) (myocyte stress), and creatinine (renal function) in a multicenter cohort of 1513 ambulatory chronic heart failure patients. We hypothesized that a biomarker score summarizing the activity of multiple pathways implicated in heart failure would improve our ability to classify risk of adverse outcomes (cardiac transplantation, ventricular assist device [VAD] placement, or death) compared with a validated clinical risk prediction algorithm, the SHFM.

Methods

Study Population

The Penn Heart Failure Study is a National Heart, Lung, and Blood Institute–sponsored multicenter cohort study of outpatients with chronic heart failure recruited from referral centers at the University of Pennsylvania (Philadelphia, PA), Case Western University (Cleveland, OH), and the University of Wisconsin (Madison, WI).5,6 The primary inclusion criterion is a clinical diagnosis of heart failure as determined by a heart failure specialist. The resultant cohort spans a full spectrum of heart failure severity ranging from mild disease to severe disease requiring advanced therapies.5,6 Participants are excluded if they have a noncardiac condition resulting in an expected mortality of <6 months as judged by the treating physician, or if they were unable to provide consent.

At time of study entry, detailed clinical data were obtained using standardized questionnaires administered to the patient and physician, with verification through medical records. Blood samples were obtained at enrollment, processed, and stored at −80°C until time of assay. Follow-up events including all-cause mortality and cardiac transplantation were prospectively ascertained every 6 months through patient contact and verified through death certificates, transplants were prospectively ascertained every 6 months as judged by the treating physician, or if they were unable to provide consent.

Laboratory Analyses

All biomarkers were measured from banked plasma obtained at the time of study entry. sFlt-1 and MPO were measured using prototype ARCHITECT chemiluminescent microparticle-based immunoassays (Abbott Laboratories, Abbott Park, IL). BNP, TnI, hsCRP, uric acid, and creatinine were measured using standard ARCHITECT immunoassays (Abbott Laboratories, Abbott Park, IL). ST2 was measured via a high sensitivity sandwich monoclonal immunoassay (Presage ST2 Assay, Critical Diagnostics, San Diego, CA).7 Full details are provided in the online-only Data Supplement.

Seattle Heart Failure Model

The SHFM is a multivariable risk prediction scoring system that has been validated in multiple heart failure populations as a predictor of mortality, urgent cardiac transplantation, or VAD placement.8 The version of the score used in this study was the SHFM-D, abbreviated as SHFM for simplicity, and is based on clinically assessed variables including demographics, heart failure characteristics, laboratory values, and medications. The derivation and validation of the SHFM has been previously described.9

Statistical Analysis

Baseline characteristics were summarized for the entire cohort using standard descriptive statistics. A log (base 2) transformation was applied to each biomarker to normalize its distribution. Spearman correlations were calculated between each biomarker pair. Cox regression models were used to determine the unadjusted association between each individual biomarker and time to the combined outcome of all-cause death, cardiac transplantation, or VAD placement. For TnI, variables that indicated a detectable value and the log-transformed continuous value were included to estimate the risk associated with a detectable value and the risk associated with the absolute value if detectable. The proportional hazards assumption was evaluated using weighted residuals.

All biomarkers were included in a multivariable Cox model, and a parsimonious set of biomarkers was selected using a stepwise procedure based on the Akaike information criterion, in which a biomarker that reduced the Akaike information criterion was retained. To estimate a multimarker score, a leave-one-out-jackknife approach was used such that the multimarker score for each individual was calculated as a weighted combination of his/her biomarker levels, with weights determined by regression coefficients, which were estimated from the parsimonious Cox model that fit the data for all other individuals. The jackknife approach ameliorates the potential for bias when applying a predictive score to the same dataset from which it was derived and avoids arbitrarily splitting the data into derivation and validation cohorts.

In primary analyses, we compared the multimarker score with the clinically used SHFM score. To ensure a fair comparison between the multimarker score and the SHFM, we used estimates of the baseline hazard function to recalibrate the SHFM within our cohort. Kaplan-Meier curves were used to contrast the survival distribution across tertiles of the multimarker and SHFM score. Hazard ratios were estimated from an unadjusted Cox regression model and a model that adjusted for the SHFM.

To evaluate the multimarker score as a discriminator of individual risk, time-dependent receiver operating characteristic (ROC) curves were used to compare the ability of the multimarker score and the SHFM to classify patients with regard to adverse event at 1 year.10 Differences in the area under the ROC curve (AUC) were used to quantify improvements in predictive accuracy. The incremental value of the multimarker score compared with the SHFM in predicting outcomes at 1 year was determined using net reclassification improvement (NRI).12,13 which represents the difference in the number of patients moving up or down clinical risk groups, stratified according to outcome. Clinically meaningful risk categories of 0% to <10%, 10% to <20%, 20% to <50%, and ≥50% risk were defined a priori.10 Cox regression models were used to predict risk at 1 year. Because some subjects were censored before 1 year, the number of events and nonevents at 1 year was estimated from the Kaplan-Meier survival estimator.14

Bootstrap resampling was used to compute standard error estimates on which to base confidence intervals for the AUC and the NRI, as well as the Wald probability values to test whether the difference between 2 AUCs and the NRI were equal to 0.10 Estimates of the AUC, the difference between 2 AUCs, and the NRI were obtained for each resampled dataset, and the standard deviation of the estimates across 1000 resampled datasets was used as the standard error.

In a secondary analysis, we rederived the SHFM in our cohort. We entered the component variables into a Cox regression model and used a leave-one-out jackknife approach to rederive the SHFM. We then compared the multimarker score with the rederived SHFM. In addition, because the SHFM-D score has been validated only for mortality, we performed a sensitivity analysis in which we used all-cause mortality alone as the outcome and compared the multimarker score with both the SHFM and its component variables. Calibration for the SHFM was assessed using the Gini index and
Results

Baseline characteristics of the 1513 patients in the study cohort are shown in Table 1. The mean age of patients at time of enrollment was 56±15 years. Sixty-six percent of the cohort was male, 74% were white, and 22% were African American. Over a maximum follow-up period of 5 years (median, 2.5 years; interquartile range, 1.4–4.0 years), there were 317 outcomes: 187 deaths, 99 cardiac transplantations, and 31 VAD implantations.

Table 1. Baseline Characteristics of Study Participants

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All Participants (n=1513)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>Age, y, mean (SD)</td>
<td>56 (15)</td>
</tr>
<tr>
<td>Male</td>
<td>1000 (66)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1114 (74)</td>
</tr>
<tr>
<td>African American</td>
<td>330 (22)</td>
</tr>
<tr>
<td>Other</td>
<td>69 (4)</td>
</tr>
<tr>
<td><strong>Medical history and risk factors</strong></td>
<td></td>
</tr>
<tr>
<td>History of hypertension</td>
<td>877 (58)</td>
</tr>
<tr>
<td>History of diabetes</td>
<td>426 (28)</td>
</tr>
<tr>
<td><strong>Tobacco use</strong></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>138 (9)</td>
</tr>
<tr>
<td>Former</td>
<td>806 (54)</td>
</tr>
<tr>
<td>Hypercholesterolemia</td>
<td>752 (50)</td>
</tr>
<tr>
<td><strong>Heart failure characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>NYHA classification</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>258 (17)</td>
</tr>
<tr>
<td>II</td>
<td>690 (46)</td>
</tr>
<tr>
<td>III</td>
<td>444 (29)</td>
</tr>
<tr>
<td>IV</td>
<td>114 (8)</td>
</tr>
<tr>
<td>Ischemic heart failure</td>
<td>455 (30)</td>
</tr>
<tr>
<td>Systolic heart failure</td>
<td>1305 (86)</td>
</tr>
<tr>
<td>Cardiac resynchronization therapy</td>
<td>381 (25)</td>
</tr>
<tr>
<td>Defibrillator</td>
<td>636 (42)</td>
</tr>
<tr>
<td><strong>Medication use</strong></td>
<td></td>
</tr>
<tr>
<td>ACE inhibitors or ARBs</td>
<td>1314 (87)</td>
</tr>
<tr>
<td>Aldosterone antagonists</td>
<td>511 (34)</td>
</tr>
<tr>
<td>Aspirin</td>
<td>825 (55)</td>
</tr>
<tr>
<td>β-blockers</td>
<td>1328 (88)</td>
</tr>
<tr>
<td>Digoxin</td>
<td>586 (39)</td>
</tr>
<tr>
<td>Diuretics</td>
<td>1185 (78)</td>
</tr>
<tr>
<td>HMG CoA reductase inhibitors</td>
<td>756 (50)</td>
</tr>
<tr>
<td>Inotrope</td>
<td>50 (3)</td>
</tr>
<tr>
<td><strong>Clinical measures, mean (SD)</strong></td>
<td></td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td>30 (7.3)</td>
</tr>
<tr>
<td>Systolic blood pressure, mm Hg</td>
<td>114 (20)</td>
</tr>
<tr>
<td>Diastolic blood pressure, mm Hg</td>
<td>70 (12)</td>
</tr>
<tr>
<td>Estimated glomerular filtration rate, mL/min/1.73 m²</td>
<td>82 (32)</td>
</tr>
<tr>
<td>Sodium, mEq/L</td>
<td>139 (3.4)</td>
</tr>
<tr>
<td>Ejection fraction, %</td>
<td>34 (17)</td>
</tr>
<tr>
<td>Seattle Heart Failure Model score</td>
<td>−0.07 (1.0)</td>
</tr>
</tbody>
</table>

Table 1. Continued

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All Participants (n=1513)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomarkers or clinical risk scores, median (interquartile range)</td>
<td></td>
</tr>
<tr>
<td>BNP, pg/mL</td>
<td>171 (47, 576)</td>
</tr>
<tr>
<td>sFlt-1, pg/mL</td>
<td>308 (260, 381)</td>
</tr>
<tr>
<td>hsCRP, mg/L</td>
<td>0.35 (0.14, 0.89)</td>
</tr>
<tr>
<td>MPO, pmol/L</td>
<td>138 (95, 227)</td>
</tr>
<tr>
<td>ST2, ng/mL</td>
<td>27.5 (19.9, 40.8)</td>
</tr>
<tr>
<td>Tnl, Detectable, n (%)</td>
<td>973 (64)</td>
</tr>
<tr>
<td>Tnl, ng/mL</td>
<td>0.02 (0.01, 0.04)</td>
</tr>
<tr>
<td>Uric acid, mg/dL</td>
<td>7.0 (5.7, 8.8)</td>
</tr>
<tr>
<td>Creatinine, mg/dL</td>
<td>0.92 (0.76, 1.25)</td>
</tr>
</tbody>
</table>

Values are expressed as n (%) unless otherwise noted.

NYHA indicates New York Heart Association; ACE, angiotensin-converting enzyme; ARBs, angiotensin receptor blocker; BNP, B-type natriuretic peptide; sFlt-1, soluble fms-like tyrosine kinase receptor-1; hsCRP, high-sensitivity C-reactive protein; MPO, myeloperoxidase; ST2, soluble toll-like receptor-2; Tnl, troponin I.

Borgon statistic. All statistical analyses were completed using R 2.12.0, including the survival, survival ROC, and pec extension packages.

Table 1. Continued

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All Participants (n=1513)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Values are expressed as n (%) unless otherwise noted.</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Results**

Baseline characteristics of the 1513 patients in the study cohort are shown in Table 1. The mean age of patients at time of enrollment was 56±15 years. Sixty-six percent of the cohort was male, 74% were white, and 22% were African American. Over a maximum follow-up period of 5 years (median, 2.5 years; interquartile range, 1.4–4.0 years), there were 317 outcomes: 187 deaths, 99 cardiac transplantations, and 31 VAD implantations.

Median biomarker levels are provided in Table 1. Among the 8 biomarkers, there were moderate correlations between BNP and sFlt-1 (R=0.54), BNP and ST2 (R=0.40), and BNP and TnI (R=0.55) (online-only Data Supplement Figure I). There were also moderate correlations between sFlt-1 and TnI (R=0.43) and between creatinine and uric acid (R=0.51) (P<0.001 for all comparisons).

Consistent with previously published studies, each of the 8 biomarkers was individually associated with the combined risk of all-cause death, cardiac transplantation, or VAD implantation. Creatinine and uric acid showed nonlinear associations with outcome (online-only Data Supplement Figure II) and were modeled using quadratic terms. A multivariable Cox model including all biomarkers together was fit, and a parsimonious set was selected, based on a stepwise model selection procedure (online-only Data Supplement Table I). The following 7 markers remained in the multimarker score: BNP, sFlt-1, hsCRP, ST2, Tnl (detectable versus not), uric acid, and creatinine (Table 2).

Figure 1 displays the risk of transplant or VAD-free survival according to tertiles of the multimarker or SHFM score. In unadjusted models, patients in the middle tertile of multimarker score had a 4.69-fold increased risk of adverse outcomes, and those in the highest tertile had a 13.7-fold elevated risk. These findings remained robust after adjustment for the SHFM (Table 2). Hazard ratios were similar in subgroups defined by systolic (n=1305) or diastolic (n=198) heart failure, but our power to detect subtle differences between these subgroups was limited.
The AUC for the multimarker score was 0.798 at 1 year (95% confidence interval [CI], 0.763–0.833), which indicated a strong ability to discriminate individual risk, and which was superior to that of the SHFM (AUC, 0.756; 95% CI, 0.717–0.795). As shown in Figure 2, addition of the multimarker score to the SHFM significantly improved predictive accuracy compared with the SHFM (AUC, 0.803; 95% CI, 0.769–0.837; \( P < 0.001 \)). Risk reclassification analyses revealed that use of the multimarker score in combination with the SHFM significantly improved classification of 1-year risk compared with the SHFM (NRI, 24.1%; 95% CI, 11.6–36.7%; \( P < 0.001 \)) (Table 3). The improvement was limited to patients who had an adverse event (NRI, 25.2%; 95% CI, 14.2–36.2%; \( P < 0.001 \)), signifying a better ability to classify patients at higher risk.

We obtained similar results in secondary analyses comparing the multimarker score with the SHFM components. The AUC for the SHFM components was 0.776 (95% CI, 0.735–0.818) at 1 year (online-only Data Supplement Figure III). Addition of the multimarker score to the SHFM components significantly improved predictive accuracy compared with the SHFM components (AUC, 0.811; 95% CI, 0.771–0.850; \( P < 0.019 \)). In addition, the multimarker score and SHFM components in combination significantly improved classification of 1-year risk compared with the SHFM components (NRI, 12.2%; 95% CI, 0.6–23.8%; \( P = 0.039 \)) (online-only Data Supplement Table II).

### Table 2: Association of Multimarker Score With Risk of All-Cause Death, Cardiac Transplantation, or Ventricular Assist Device Placement

<table>
<thead>
<tr>
<th>Risk Category†</th>
<th>Kaplan-Meier 1-Y Risk Estimate (95% CI)</th>
<th>Unadjusted Hazard Ratio (95% CI)</th>
<th>P Value</th>
<th>Adjusted* Hazard Ratio (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>1.6% (0.5%, 2.7%)</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>Moderate</td>
<td>6.4% (4.3%, 8.6%)</td>
<td>4.69 (2.92, 7.54)</td>
<td>(&lt;0.001</td>
<td>3.50 (2.17, 5.67)</td>
<td>(&lt;0.001</td>
</tr>
<tr>
<td>(-0.51 \leq \text{Multimarker score} &lt; 0.52)</td>
<td>26% (22%, 30%)</td>
<td>13.7 (8.75, 21.5)</td>
<td>(&lt;0.001</td>
<td>6.80 (4.18, 11.1)</td>
<td>(&lt;0.001</td>
</tr>
<tr>
<td>High</td>
<td>26% (22%, 30%)</td>
<td>13.7 (8.75, 21.5)</td>
<td>(&lt;0.001</td>
<td>6.80 (4.18, 11.1)</td>
<td>(&lt;0.001</td>
</tr>
</tbody>
</table>

CI indicates confidence interval.

*Adjusted for Seattle Heart Failure Model score.

†Bottom tertile of multimarker score defined as low risk, middle tertile defined as moderate risk, and top tertile defined as high risk.

‡Multimarker score calculated as \( 0.21 \times \log_2 \text{BNP (pg/mL)} + 0.176 \times \log_2 \text{sFlt-1 (pg/mL)} + 0.067 \times \log_2 \text{hsCRP (mg/dL)} + 0.274 \times \log_2 \text{ST2 (ng/mL)} + 0.772 \times \text{detectable TnI (versus not)} - 1.718 \times \log_2 \text{uric acid (mg/dL)} + 0.298 \times \log_2 \text{creatinine (mg/dL)} - 0.138 \times (\log_2 \text{creatinine})^2 \), where \( \log_2 \) denotes log (base 2) transformation.

Figure 1. Event-free survival according to multimarker score category and tertiles of Seattle Heart Failure Model (SHFM) (multimarker score calculated as described in Table 2). Kaplan-Meier curves illustrate the incidence of all-cause death, cardiac transplantation, or ventricular assist device placement among Penn Heart Failure Study participants according to multimarker score category tertiles (\( P < 0.001 \) by log-rank test) (A) and SHFM score tertiles (\( P < 0.001 \) by log-rank test) (B).
We also obtained similar results with all-cause mortality as the outcome. The SHFM was well-calibrated in this cohort for the outcome of all-cause mortality with a predicted versus observed 1-year survival of 93.7% versus 94.0% ($P = 0.21$). The AUC for the multimarker score was 0.808 (95% CI, 0.762–0.854); for the SHFM, 0.761 (95% CI, 0.708–0.813); and for SHFM components, 0.749 (95% CI, 0.690–0.808).

Addition of the multimarker score improved predictive accuracy compared with the SHFM (AUC, 0.809; 95% CI, 0.763–0.854; $P = 0.038$) and compared with the SHFM components (AUC, 0.790; 95% CI, 0.735–0.846; $P = 0.065$). The multimarker score also improved classification of 1-year risk of all-cause mortality in combination with the SHFM (NRI, 18.0; 95% CI, 2.4%–33.6%; $P = 0.023$) and in combination with the SHFM components (NRI, 13.7%; 95% CI, −3.7% to 31.1%; $P = 0.12$). Although point estimates were similar, power was reduced in our sensitivity analysis due to the decreased event rate for all-cause mortality compared with the combined outcome of mortality, transplantation, and VAD implantation.

**Discussion**

In 2008, Braunwald classified circulating biomarkers into categories based on their pathophysiologic effects in heart failure and hypothesized that multiple biomarkers in combination would provide a valuable means for risk stratification. Our findings strongly support this hypothesis. In a multicenter cohort of 1513 ambulatory heart failure patients, we found that a multimarker score composed of 7 biomarkers, reflective of diverse biological axes, was a strong predictor of risk and significantly improved the prediction of outcomes compared with the most commonly used clinical risk score in heart failure, the SHFM. Patients in the highest multimarker tertile had a nearly 14-fold unadjusted risk of death, transplant, or VAD placement compared with the lowest tertile, and this risk remained nearly 7-fold after adjustment for the SHFM. The multimarker score showed a substantial ability to discriminate individual patient risk at 1 year (AUC, 0.798) that was again superior to the SHFM.

**Table 3. Reclassification of 1-Year Risk of All-Cause Death, Cardiac Transplantation, or Ventricular Assist Device Placement With the Addition of Multimarker Score to SHFM Score**

<table>
<thead>
<tr>
<th>Reclassification</th>
<th>SHFM Score</th>
<th>SHFM Score Plus Multimarker Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events, n (%)†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0% to &lt;10%</td>
<td>22 (50%)</td>
<td>19 (43%)</td>
</tr>
<tr>
<td>10% to &lt;20%</td>
<td>4 (7%)</td>
<td>6 (11%)</td>
</tr>
<tr>
<td>20% to &lt;50%</td>
<td>3 (7%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>50–100%</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Nonevents, n (%)†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0% to &lt;10%</td>
<td>764 (88%)</td>
<td>93 (11%)</td>
</tr>
<tr>
<td>10% to &lt;20%</td>
<td>121 (36%)</td>
<td>154 (46%)</td>
</tr>
<tr>
<td>20% to &lt;50%</td>
<td>11 (9%)</td>
<td>29 (23%)</td>
</tr>
<tr>
<td>50–100%</td>
<td>0 (0%)</td>
<td>2 (2%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Net Reclassification</th>
<th>Up</th>
<th>Down</th>
<th>NRI</th>
<th>95% CI</th>
<th>$P$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events†</td>
<td>56</td>
<td>12</td>
<td>25.2%</td>
<td>(14.2%, 36.2%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Nonevents†</td>
<td>177</td>
<td>163</td>
<td>−1.1%</td>
<td>(−5.6%, 3.4%)</td>
<td>0.64</td>
</tr>
<tr>
<td>Total</td>
<td>24.1%</td>
<td>(11.6%, 36.7%)</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SHFM indicates Seattle Heart Failure Model score; CI, confidence interval; and NRI, net reclassification index.

*Multimarker score calculated as described in Table 2.

†Number of events and nonevents estimated from 1-year Kaplan-Meier risk estimates; row percentages may not sum to 100% due to estimation.
score to the SHFM appropriately reclassified a large proportion (24.1%) of patients as higher risk. These findings support the usefulness of multiple biomarkers as part of an algorithm for assessing prognosis in heart failure.

Assessing risk of future adverse events is a fundamental task in clinical medicine. Research in community-based cohorts has led to the development of clinical scores that estimate the risk of new-onset cardiovascular disease, the most widely used being the Framingham risk score. These scores are simple, easily calculated, and have been a major advance for preventive medicine. Attempts to improve risk prediction by adding a diverse panel of biomarkers and genetic variants to clinical risk scores is an active area of research, with mixed results, depending on the study population.

Once cardiovascular disease is established, however, approaches to assessing risk of adverse events (ie, prognosis) are less well defined although still critically important as an objective means to guide treatment strategy. The SHFM is the most widely validated clinical risk score to predict prognosis in chronic heart failure and is easily calculated from readily available clinical data and well-calibrated in our cohort but has been shown to underestimate risk and overestimate survival. Biomarkers can improve on the use of primarily clinical data (ie, SHFM) and offer advantages of rapid availability (within minutes to hours), reproducibility, and quantitative insight into the underlying disease mechanisms. A number of previous studies have investigated the prognostic utility of 1 or 2 biomarkers in chronic heart failure. Several of these have demonstrated associations with outcome but their clinical value remains a matter of debate. As recently reviewed, one explanation for the slow translation of biomarkers to clinical use relates to overly optimistic risk estimates in initial publications. An alternative explanation is that individual biomarkers may be insufficient to assess multisystem disorders. By contrast, our findings demonstrate that a diverse multimarker panel, which provides a more comprehensive signature of the underlying pathophysiology, offers substantial ability to determine individual patient prognosis.

These findings support the concept of a multimarker tool for prognosis, but identifying an “optimal” panel of biomarkers for assessing heart failure, or any other disease, is a formidable task. We chose biomarkers that quantify known biological abnormalities in heart failure that affect outcome, including myocyte injury (TnI), neurohormonal activation (BNP), myocyte stress (ST2), vascular growth and remodeling (sFlt-1), inflammation (CRP), oxidative stress (uric acid, MPO), and renal dysfunction (creatinine). There are certainly other biomarkers that were not included, such as galectin-3, adrenomedullin, adiponectin, GDF-15, or high-sensitivity TnI. We also found substantial correlations among many of the biomarkers measured, strongly suggesting that there is no single “optimal panel.” Rather, there are probably a number of biomarker panels that may perform equivalently, and choosing which to deploy in clinical practice will depend on factors such as cost, ease of assay, and potential therapeutic implications. As technology for proteomic and metabolomic profiling continues to advance, unbiased biomarker screens will undoubtedly reveal a large number of candidates for risk assessment in heart failure, and this challenge will be magnified.

We acknowledge the strengths and limitations of our study. First, our population was recruited from heart failure referral centers with broad inclusion criteria. As a result, the study cohort has a high prevalence of comorbid conditions and a broad spectrum of heart failure. The high burden of comorbidity parallels the complexity of heart failure patients in clinical practice and stands in contrast to biomarker studies in clinical trial populations, which typically have fewer comorbid conditions. Furthermore, the broad spectrum of heart failure enabled biomarker assessment over the full range of disease, from mild to severe. Second, the multimarker score was derived within our dataset, and as such, there is a potential to have results appear more optimistic for this score. To address this concern, we performed rigorous analyses that included the implementation of a leave-one-out jackknife estimation approach, which avoids self-influence, arbitrary splitting of the data, and provides a highly robust measure to assess the validity of our findings. Our supplemental analyses using the SFHM and its components and using death as the primary outcome yielded qualitatively similar findings. Although we recalibrated the SHFM for our composite end point for risk reclassification analyses, recalibration did not affect estimation of the AUC (discrimination). Finally, our NRI estimates may be sensitive to the number and selection of cut-points.

We caution that the most rigorous comparison is to assess both the biomarker score and the SHFM in completely independent cohorts, emphasizing the need for future research. We focused on a select panel of markers, but it is certainly plausible that there remain additional markers that could have further improved prognosis. The optimal panel of markers, the change in these markers over time, and how these changes might help guide therapeutic interventions remain to be defined, as does the cost-effectiveness of such an approach are necessary studies before its clinical application.

In summary, we demonstrate that a multimarker score composed of 7 circulating biomarkers showed a substantial ability to predict risk in chronic heart failure and outperformed the most widely validated clinical risk score in heart failure, the SHFM. These findings strongly support ongoing efforts to use unbiased proteomic and metabolomic technologies to develop signatures of heart failure progression.

Sources of Funding
Dr Ky was supported by the National Institutes of Health (NIH)/Clinical and Translational Science Award KL1 RR024132, NIH K23 HL095661-01, and the Heart Failure Society of America Research Fellowship Award. This work was also supported by NIH HL088577 (Dr Cappola). Assay support was provided by Abbott Diagnostics and Critical Diagnostics. Neither the funding organizations, Abbott...
Disclosures

Dr Cappola reports receiving research support from Abbott Diagnostics. Drs Ky and Cappola are coinventors on a pending patent application for the use of sFlt-1 as a biomarker in heart failure. Dr Levy received research support for the SHFM from NIH, General Electric, Epocrates, and HeartWare and honoraria from Glaxo SmithKline.

References


27. Ioannidis JPA, Panagiotou OA. Comparison of effect sizes associated with biomarkers reported in highly cited individual articles and in subsequent meta-analyses. JAMA. 2011;305:2200–2210.


---

**CLINICAL PERSPECTIVE**

The progression of heart failure is complex and is driven by multiple biological processes. We hypothesized that a biomarker score summarizing the activity of multiple heart failure associated pathways would improve our ability to estimate prognosis. In a multicenter cohort of 1513 chronic systolic heart failure patients, we measured a contemporary biomarker panel consisting of: high-sensitivity C-reactive protein, myeloperoxidase, B-type natriuretic peptide, soluble fms-like tyrosine kinase receptor-1, troponin I, soluble toll-like receptor-2, creatinine, and uric acid. From this panel, we calculated a parsimonious multimarker score and assessed its performance in predicting risk of death, cardiac transplantation, or ventricular assist device placement. Over a median follow-up of 2.5 years, we determined that the multimarker score was strongly associated with a risk of adverse outcomes, and this effect was independent of the Seattle Heart Failure Model. Furthermore, addition of the multimarker score to the Seattle Heart Failure Model significantly improved discriminative ability and reclassified 25% of the patients into more appropriate, higher-risk categories. These findings strongly support the concept of a multimarker tool for prognosis in chronic human heart failure.
Multiple Biomarkers for Risk Prediction in Chronic Heart Failure

Circ Heart Fail. 2012;5:183-190; originally published online February 23, 2012;
doi: 10.1161/CIRCHEARTFAILURE.111.965020
Circulation: Heart Failure is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2012 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-3289. Online ISSN: 1941-3297

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circheartfailure.ahajournals.org/content/5/2/183

Data Supplement (unedited) at:
http://circheartfailure.ahajournals.org/content/suppl/2012/02/23/CIRCHEARTFAILURE.111.965020.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Heart Failure can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Heart Failure is online at:
http://circheartfailure.ahajournals.org//subscriptions/
Supplemental Material

Biomarker Assays

ST2 Assay

ST2 was measured from banked plasma via a highly sensitive sandwich monoclonal immunoassay (Presage™ ST2 assay, Critical Diagnostics, New York, New York). This platform offers improved accuracy in quantifying ST2 levels, particularly at lower concentrations. The antibodies used in the Presage™ assay were generated from a recombinant protein based upon the human cDNA clone for the complete soluble ST2 sequence. The intra- and inter-assay coefficients of variation (CV) were less than 4.0% and 2.5%, respectively. The lower limit of detection of ST2 was 2 ng/ml and the upper limit was 200 ng/ml.

sFlt-1

sFlt-1 was measured using prototype ARCHITECT immunoassays (Abbott Laboratories, Abbott Park, IL). The sFlt-1 immunoassay measures both free and bound sFlt-1, with an assay range of 15 to 50,000 pg/ml. The intra- and interassay coefficients of variation (CV) ranged from 1.3% to 5.2% and 1.9% to 5.9%, respectively.

BNP

BNP was measured using the ARCHITECT® BNP chemiluminescent microparticle immunoassay (Abbott Laboratories, Abbott Park, IL) as previously described. The assay range was from 10 to 5,000 pg/ml. The intra- and interassay CV ranged from 0.9% to 5.6% and 1.7% to 6.7%, respectively.
TnI

TnI was measured using a chemiluminescent microparticle immunoassay on the ARCHITECT® system (Abbott Laboratories, Abbott Park, IL). The limits of detection were <0.01 ng/ml. The total CV was \( \leq 10\% \).

hsCRP

hsCRP was measured using the MULTIGENT Vario assay on the ARCHITECT® system (Abbott Laboratories, Abbott Park, IL), which is a latex immunoassay. The lower limit of detection is 0.01 mg/dL. The total CV was \( \leq 6\% \).

MPO

MPO was measured using a chemiluminescent microparticle immunoassay on the ARCHITECT® system (Abbott Laboratories, Abbott Park, IL) using a 2-step immunoassay. The total CV was \( \leq 10\% \).

Uric Acid

The uric acid assay was based on the methods of Trivedi and Kabasakalian and measured on the ARCHITECT® platform (Abbott Laboratories, Abbott Park, IL). The intra- and interassay CV were <0.5% and <1.6%, respectively.

Creatinine

Creatinine was measured on the ARCHITECT® MULTIGENT platform (Abbott Laboratories, Abbott Park, IL). The intra- and interassay CV were \( \leq 4\% \).
### Supplementary Table 1: Association of biomarkers with risk of all-cause death, cardiac transplantation, or VAD placement

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Univariable HR (95% CI)</th>
<th>Multivariable HR (95% CI)</th>
<th>Parsimonious HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>log₂ BNP – pg/ml</td>
<td>1.46 (1.39, 1.54)</td>
<td>1.23 (1.15, 1.32)</td>
<td>1.23 (1.15, 1.32)</td>
</tr>
<tr>
<td>log₂ sFlt-1 – pg/ml</td>
<td>1.68 (1.54, 1.84)</td>
<td>1.17 (0.99, 1.38)</td>
<td>1.19 (1.02, 1.40)</td>
</tr>
<tr>
<td>log₂ hsCRP – mg/l</td>
<td>1.26 (1.19, 1.34)</td>
<td>1.06 (0.99, 1.14)</td>
<td>1.07 (0.99, 1.14)</td>
</tr>
<tr>
<td>log₂ MPO – pmol/l</td>
<td>1.20 (1.09, 1.32)</td>
<td>1.04 (0.93, 1.15)</td>
<td>—</td>
</tr>
<tr>
<td>log₂ ST2 – ng/ml</td>
<td>2.05 (1.84, 2.30)</td>
<td>1.32 (1.14, 1.52)</td>
<td>1.31 (1.14, 1.52)</td>
</tr>
<tr>
<td>TnI – ng/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detectable (vs. not)</td>
<td>3.79 (2.58, 5.70)</td>
<td>2.00 (1.32, 3.04)</td>
<td>2.16 (1.48, 3.18)</td>
</tr>
<tr>
<td>log₂ TnI</td>
<td>1.14 (1.07, 1.22)</td>
<td>1.04 (0.96, 1.12)</td>
<td>—</td>
</tr>
<tr>
<td>log₂ Uric Acid – mg/dl</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear</td>
<td>0.12 (0.02, 0.99)</td>
<td>0.17 (0.02, 1.36)</td>
<td>0.18 (0.02, 1.42)</td>
</tr>
<tr>
<td>Quadratic</td>
<td>1.65 (1.15, 2.36)</td>
<td>1.36 (0.95, 1.95)</td>
<td>1.35 (0.94, 1.93)</td>
</tr>
<tr>
<td>log₂ Creatinine – mg/dl</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear</td>
<td>2.69 (2.08, 3.47)</td>
<td>1.30 (0.97, 1.76)</td>
<td>1.31 (0.97, 1.76)</td>
</tr>
<tr>
<td>Quadratic</td>
<td>0.78 (0.67, 0.90)</td>
<td>0.87 (0.73, 1.03)</td>
<td>0.87 (0.74, 1.03)</td>
</tr>
</tbody>
</table>

**Abbreviations:** HR, Hazard ratio; CI, Confidence interval

log₂ denotes log (base 2) transformation
Supplementary Table 2: Reclassification of 1-year risk of all-cause death, cardiac transplantation, or VAD placement with the addition of multimarker score to the individual SHFM components

<table>
<thead>
<tr>
<th>Reclassification</th>
<th>SHFM score plus multimarker score</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHFM score</td>
<td>0%-&lt;10%</td>
</tr>
<tr>
<td>Events, no. (row %)†</td>
<td></td>
</tr>
<tr>
<td>0%-&lt;10%</td>
<td>26 (63%)</td>
</tr>
<tr>
<td>≥10%-%&lt;20%</td>
<td>6 (14%)</td>
</tr>
<tr>
<td>≥20%-&lt;50%</td>
<td>0</td>
</tr>
<tr>
<td>≥50%-100%</td>
<td>0</td>
</tr>
<tr>
<td>Non-events, no. (row %)†</td>
<td></td>
</tr>
<tr>
<td>0%-&lt;10%</td>
<td>856 (91%)</td>
</tr>
<tr>
<td>≥10%-%&lt;20%</td>
<td>87 (35%)</td>
</tr>
<tr>
<td>≥20%-&lt;50%</td>
<td>2 (1%)</td>
</tr>
<tr>
<td>≥50%-100%</td>
<td>0</td>
</tr>
</tbody>
</table>

Net reclassification

<table>
<thead>
<tr>
<th>Up</th>
<th>Down</th>
<th>NRI</th>
<th>95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events†</td>
<td>43</td>
<td>21</td>
<td>13.2</td>
<td>(2.1%, 24.3%)</td>
</tr>
<tr>
<td>Non-events†</td>
<td>140</td>
<td>127</td>
<td>-1.0</td>
<td>(-4.3%, 2.3%)</td>
</tr>
<tr>
<td>Total</td>
<td>127</td>
<td>127</td>
<td>12.2</td>
<td>(0.6%, 23.8%)</td>
</tr>
</tbody>
</table>

Abbreviations: SHFM, Seattle Heart Failure Model score; CI, confidence interval; NRI, net reclassification index

†Number of events and non-events estimated from 1-year Kaplan-Meier risk estimates; row percentages may not sum to 100% due to estimation
Supplementary Figure 1: Pairwise Spearman rank correlation between biomarkers

The Spearman rank correlation coefficients are represented in each ellipsoid. The shape of the ellipsoid and intensity of the color reflects the degree of correlation. P < 0.001 for all.
Supplementary Figures 2A and 2B: Fitted values for uric acid and creatinine

Nonlinearity of relationship between uric acid and creatinine and risk of all-cause death, cardiac transplantation, or VAD implantation. Solid lines show estimated association. Shaded regions demonstrate pointwise confidence regions.
Supplementary Figure 3: ROC curves for events at 1 year

ROC curves comparing the ability of the individual components of the SHFM score and the multimarker score to correctly classify patients who died, required cardiac transplantation, or VAD placement by 1 year of followup.
References


