Journal of the American College of Cardiology
© 2014 by the American College of Cardiology Foundation
Published by Elsevier Inc.

Vol. 63, No. 2, 2014
ISSN 0735-1097/$36.00
http://dx.doi.org/10.1016/j.jacc.2013.07.087

Biomarkers

Head-to-Head Comparison of 2 Myocardial Fibrosis Biomarkers for Long-Term Heart Failure Risk Stratification

ST2 Versus Galectin-3

Antoni Bayes-Genis, MD, PHD,*† Marta de Antonio, MD,*† Joan Vila, MSC,‡§ Judith Peñafiel, BSC,‡§ Amparo Galán, MD, PHD,|| Jaume Barallat, MD,|| Elisabet Zamora, MD, PHD,*† Agustin Urrutia, MD, PHD,*† Josep Lupón, MD, PHD*†

Badalona and Barcelona, Spain

Objectives ST2 and galectin-3 (Gal-3) were compared head-to-head for long-term risk stratification in an ambulatory heart failure (HF) population on top of other risk factors including N-terminal pro–B-type natriuretic peptide.

Background ST2 and Gal-3 are promising biomarkers of myocardial fibrosis and remodeling in HF.

Methods This cohort study included 876 patients (median age: 70 years, median left ventricular ejection fraction: 34%). The 2 biomarkers were evaluated relative to conventional assessment (11 risk factors) plus N-terminal pro–B-type natriuretic peptide in terms of discrimination, calibration, and reclassification analysis. Endpoints were 5-year all-cause and cardiovascular mortality, and the combined all-cause death/HF hospitalization.

Results During a median follow-up of 4.2 years (5.9 for alive patients), 392 patients died. In bivariate analysis, Gal-3 and ST2 were independent variables for all endpoints. In multivariate analysis, only ST2 remained independently associated with cardiovascular mortality (hazard ratio: 1.27, 95% confidence interval [CI]: 1.05 to 1.53, $p = 0.014$). Incorporation of ST2 into a full-adjusted model for all-cause mortality (including clinical variables and N-terminal pro–B-type natriuretic peptide) improved discrimination (C-statistic: 0.77, $p = 0.004$) and calibration, and reclassified significantly better (integrated discrimination improvement: 1.5, 95% CI: 0.5 to 2.5, $p = 0.003$; net reclassification index: 9.4, 95% CI: 4.8 to 14.1, $p < 0.001$). Incorporation of Gal-3 showed no significant increase in discrimination or reclassification and worse calibration metrics. On direct model comparison, ST2 was superior to Gal-3.

Conclusions Head-to-head comparison of fibrosis biomarkers ST2 and Gal-3 in chronic HF revealed superiority of ST2 over Gal-3 in risk stratification. The incremental predictive contribution of Gal-3 to existing clinical risk factors was trivial. (J Am Coll Cardiol 2014;63:158–66) © 2014 by the American College of Cardiology Foundation

Heart failure (HF), a major epidemic in Western countries, is characterized by ventricular remodeling and variable degrees of myocardial fibrosis (1,2). The prognosis of HF patients, despite contemporary evidence-based treatment, remains poor (3). There is a need to refine the variables clinicians use to correctly classify patients at risk of developing adverse events. Assessment based on signs and symptoms together with echocardiography is valuable but insufficient, and some circulating biomarkers have been identified and developed for routine use. Among these are natriuretic peptides, which provide information about myocardial stretch, and are already acknowledged in HF guidelines (4–6). Novel biomarkers reflective of other pathophysiological pathways, such as ventricular remodeling and fibrosis, are promising, but their contribution must go beyond information available from conventional assessment, which already includes natriuretic peptides.

See page 167

From the *Heart Failure Unit, Hospital Universitari Germans Trias i Pujol, Badalona, Barcelona, Spain; †Department of Medicine, Autonomous University of Barcelona, Barcelona, Spain; ‡Hospital del Mar Medical Research Institute, Barcelona, Spain; §CIBER Epidemiology and Public Health, Barcelona, Spain; and the ||Biochemistry Service, Hospital Universitari Germans Trias i Pujol, Badalona, Barcelona, Spain. ST2 assays were performed by Critical Diagnostics; N-terminal pro–B-type natriuretic peptide assays were provided by Roche Diagnostics; and galectin 3 assays were partially provided by BioMerieux. Dr. Bayes-Genis has received lecture honoraria from Roche Diagnostics and Critical Diagnostics; and he owns stock in Critical Diagnostics. Dr. de Antonio has received a research grant from the Catalan Society of Cardiology. Dr. Lupón owns stock in Critical Diagnostics. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

Manuscript received May 24, 2013; revised manuscript received July 17, 2013, accepted July 23, 2013.

Two such biomarkers are commercially available and approved by the U.S. Food and Drug Administration: soluble ST2 and galectin-3 (Gal-3). ST2 is a member of the interleukin-1 receptor family and exists in 2 forms, a transmembrane receptor (ST2L) as well as a soluble decoy receptor (ST2) (7). The ligand of ST2L is interleukin-33, which is involved in reducing fibrosis and hypertrophy in mechanically strained tissues. In in vitro and in vivo models, ST2L transduces the effects of interleukin-33, whereas excess soluble ST2 leads to cardiac fibrosis and ventricular dysfunction (8–10). Gal-3 is a soluble beta-galactosidase–binding glycoprotein released by activated cardiac macrophages (11,12). Released Gal-3 in the myocardium, via a paracrine effect, stimulates proliferation of myofibroblasts and procollagen 1 deposition (13). Both ST2 and Gal-3 are reflective of fibrosis and cardiac remodeling, key in HF pathophysiology, and strongly related to outcomes (14,15). A comparative prognostic analysis of both biomarkers using state-of-the-art statistics currently recommended for biomarker implementation has not been done. Accordingly, we performed a head-to-head evaluation of ST2 and Gal-3 in a large real-life cohort with a long-term follow-up. The value of the 2 biomarkers over conventional assessment was measured in terms of iscrimination, calibration, and reclassification analysis.

Methods

Study population. From May 2006 to July 2010, ambulatory patients treated at a multidisciplinary HF unit were consecutively included in the study in an outpatient setting, as previously reported (16). In summary, patients were referred to the unit by cardiology or internal medicine departments and, to a lesser extent, from the emergency or other hospital departments. The principal referral criterion was HF according to the European Society of Cardiology guidelines irrespective of etiology, at least 1 HF hospitalization, or a reduced left ventricular ejection fraction (LVEF).

Blood samples were obtained by venipuncture between 9:00 AM and 12:00 AM during conventional ambulatory visits. After adequate centrifugation, the serum samples were stored at −80°C. ST2, Gal-3, and N-terminal pro–B-type natriuretic peptide (NT-proBNP) were analyzed from the same blood sample.

All participants provided written informed consent, and the study was approved by the local ethics committee. All study procedures were in accordance with the ethical standards outlined in the Declaration of Helsinki of 1975, as revised in 1983.

Follow-up and outcomes. All patients were followed up at regular pre-defined intervals, with additional visits as required in the case of decompensation, need for up-titration, or other circumstances that necessitated closer follow-up. The regular schedule of visits included a minimum of quarterly visits with nurses, biannual visits with physicians, and elective visits with geriatricians, psychiatrists, and rehabilitation physicians (14,16). Those who did not attend the regular visit were contacted by telephone.

A death was considered to be from cardiovascular origin if it was caused by: HF (decompensated HF or treatment-resistant HF, in the absence of another cause); sudden death (unexpected death, witnessed or not, of a previously stable patient with no evidence of worsening HF or any other cause of death); acute myocardial infarction (directly related in time with acute myocardial infarction, whether due to mechanic, hemodynamic, or arrhythmic complications); stroke (associated with recently appearing acute neurologic deficit); procedural (post-diagnostic or post-therapeutic procedure death); and other cardiovascular causes (e.g., rupture of an aneurysm, peripheral ischemia, or aortic dissection).

Abbreviations and Acronyms

AIC = Akaike information criterion
BIC = Bayesian information criterion
CI = confidence interval
Gal-3 = galectin-3
HF = heart failure
HR = hazard ratio
IDI = integrated discrimination improvement
IQR = interquartile range
LVEF = left ventricular ejection fraction
NRI = net reclassification improvement
NT-proBNP = N-terminal pro–B-type natriuretic peptide
ST2 = high-sensitivity soluble ST2

Five-year all-cause and cardiovascular death and the combined all-cause death or HF hospitalization were the primary endpoints. Fatal events were identified from clinical records or by reviewing the electronic clinical history of the Catalan and Spanish Health databases.

ST2 assay. Soluble ST2 was measured from banked plasma samples using a high-sensitivity sandwich monoclonal immunoassay (Presage ST2 assay, Critical Diagnostics, San Diego, California). This platform offers improved accuracy in quantifying ST2 levels, particularly at lower concentrations. The antibodies used in the Presage assay were generated from recombinant protein based on the human complementary deoxyribonucleic acid clone for the complete soluble ST2 sequence (17). The ST2 assay had a within-run coefficient of <2.5%, a total coefficient of variation of 4%, and a limit of detection of 1.31 ng/ml.

Gal-3 assay. For Gal-3 measurement, we used an enzyme-linked fluorescent assay (BioMerieux ref. 411191) on a mini-VIDAS analyzer (BioMerieux, France). The coefficient of variation for the assay was <10%, the linearity 3.3 to 100.0 ng/ml, and the limit of detection 2.4 ng/ml.

NT-proBNP assay. NT-proBNP levels were determined using an immuno-electrochemiluminescence assay on the Modular Analytics E 170 (Roche Diagnostics, Switzerland). This assay has <0.001% cross-reactivity with bioactive BNP, and in the constituent studies in this report, the assay had inter-run coefficients of variation ranging from 0.9% to 5.5% (18).

Statistical analysis. Categorical variables were expressed as percentages. Continuous variables were expressed as the mean ± SD or median (interquartile range) according to normal or skewed distribution. Survival analyses were

performed using Cox regression models. To fulfill the assumption of linearity of the covariables Gal-3, ST2, and NT-proBNP, the logarithmic function of Gal-3 and NT-proBNP, and ST2 plus the quadratic term of ST2 were used in the Cox models. Online Figure 1 shows the smoothing spline estimates for 5-year all-cause death for Gal-3 and ST2 nontransformed levels. ST2 analyses were performed per every 10 ng/ml change. The following variables were incorporated into the reference model: age; sex; LVEF (%); estimated glomerular filtration rate (ml/min/1.73 m^2); New York Heart Association functional class; presence of diabetes mellitus; ischemic etiology; hemoglobin (g/dl); serum sodium (mmol/l); beta-blocker treatment; angiotensin-converting enzyme inhibitor or angiotensin II receptor blocker treatment; and NT-proBNP level. Gal-3 and ST2 were subsequently added to this model. Log-rank tests for Kaplan-Meier survival curves were performed using Gal-3 and ST2 quartiles.

We used different measurements of performance to test the potential incremental prognostic value of these biomarkers, as follows.

DISCRIMINATION. C-statistics summarize the diagnostic discrimination. Discrimination refers to a model's ability to correctly distinguish 2 classes of outcomes. We used the index of rank correlation, Somers D, which already incorporates information of censored data. C-statistics between models were compared using the Mann-Whitney *U* test for equality concordance.

CALIBRATION. 1) The D'Agostino-Nam version of the Hosmer-Lemeshow calibration test was used to calculate a c^2 value. A model is well calibrated when predicted and observed values agree for any reasonable grouping of the observation (no statistically significant differences in the Hosmer-Lemeshow test results). 2) The Bayesian information criterion (BIC), the Akaike information criterion (AIC), and the Brier score were calculated for each model. The AIC and BIC are measures of the relative goodness of fit of a statistical model. The BIC penalizes free parameters more strongly than does the AIC. No statistical tests compare different BIC, AIC, or Brier score estimations, and lower values indicate a better model. 3) The global goodness of fit of the models was evaluated by likelihood ratio tests. A significant p value in this test means that adding a new variable to the model significantly improves the accuracy of the model.

RECLASSIFICATION. We used the method described by Pencina et al. (19). There are 2 main statistics to assess reclassification. Integrated discrimination improvement (IDI) considers the changes in the estimated mortality prediction probabilities as a continuous variable. Net reclassification improvement (NRI) requires a previous definition of meaningful risk categories (we used tertiles for the risk of death: <13.9%, 13.9% to 30.2%, and >30.2%). NRI considers changes in the estimated mortality prediction probabilities that imply a change from 1 category to another.

Values of $p < 0.05$ from 2-sided tests were considered to indicate statistical significance. The analyses were performed using the software R statistical package (version 2.11.1, Foundation for Statistical Computing, Vienna, Austria).

Role of the funding source. Funding sources did not have a role in study design; collection, analysis, and interpretation of data; writing of the report; or the decision to submit the paper for publication.

Results

Of the 891 consecutive patients included from May 2006 to July 2010, Gal-3 and ST2 were available for 876, the final number included in this analysis. Median age was 70.2 years (interquartile range [IQR]: 60.5 to 77.2 years). Table 1 shows the baseline characteristics of the entire sample. The median follow-up time was 4.2 years (IQR: 2.6 to 6.4 years), during which 392 patients died. Follow-up for alive patients was 5.9 years (IQR: 4.1 to 6.7 years). A total of 453 HF hospitalizations were registered from 198 patients.

Table 1 Demographic and Clinical Baseline Characteristics and Treatments During Follow-Up (N = 876)

Age, yrs	70.2 (60.5–77.2)
Female	249 (28.4)
Etiology	
Ischemic heart disease	457 (52.2)
Dilated cardiomyopathy	85 (9.7)
Hypertensive	81 (9. 2)
Valvular	103 (11.8)
Other	150 (17.1)
LVEF,%	34 (26–43)
eGFR, ml/min/1.73 m^2	43.2 (29.7–59.8)
Sodium, mmol/l	139 (137–142)
Hemoglobin, g/dl	12.9 ± 1.8
NYHA functional class	
I	64 (7.3)
II	576 (65.8)
III	227 (25.9)
IV	9 (1.0)
Hypertension	534 (61.0)
Diabetes mellitus	315 (36.0)
Treatments (follow-up)	
ACEI or ARB	786 (89.7)
Beta-blocker	767 (87.6)
Spironolactone/eplerenone	342 (39.0)
Loop diuretic	742 (84.7)
Digoxin	265 (30.3)
CRT	47 (5.4)
ICD	92 (10.5)
NT-proBNP, ng/l	1,398 (529–3,016)
Galectin 3, ng/ml	16.5 (12.6–22.7)
ST2, ng/ml	38.2 (30.8–50.9)

Values are median (IQR), n (%), or mean ± SD.

ACEI = angiotensin-converting enzyme inhibitor; ARB = angiotensin II receptor blocker; CRT = cardiac resynchronization therapy; eGFR = estimated glomerular filtration rate; ICD = implantable cardioverter-defibrillator; IQR = interquartile range; LVEF = left ventricular ejection fraction; NT-proBNP = N-terminal pro–B-type natriuretic peptide; NYHA = New York Heart Association; ST2 = high-sensitivity soluble ST2.

Among cardiovascular causes of death (n = 221), refractory HF was responsible in 113 (51.1%) patients, sudden death in 45 (20.4%) patients, and acute myocardial infarction in 20 (9%) patients. Five patients were lost to follow-up and adequately censored.

Cox regression and survival. In the bivariate analysis, both biomarkers were predictors of death from all cause as continuous variables (log(Gal-3) hazard ratio [HR]: 2.69, 95% confidence interval [CI]: 2.22 to 3.27, $p < 0.001$; ST2 HR: 1.45, 95% CI: 1.32 to 1.59, $p < 0.001$) and significantly predicted cardiovascular death (log(Gal-3) HR: 2.74, 95% CI: 2.12 to 3.54, $p < 0.001$; and ST2 HR: 1.55, 95% CI: 1.31 to 1.84, $p < 0.001$). For interpretation of these HR, Gal-3 values were normalized by log transformation, whereas ST2 was normalized by adding its quadratic transformation to nontransformed ST2 levels, and ST2 analyses were performed per every 10 ng/ml change (Online Appendix). Figure 1 shows Kaplan-Meier survival curves according to Gal-3 (Fig. 1A) and ST2 (Fig. 1B) quartiles. No interaction was found between mineralocorticoid antagonists and ST2 ($p = 0.778$) or Gal-3 ($p = 0.339$).

In multivariable analysis, log(Gal-3) was independently associated only with all-cause but not with cardiovascular death (Table 2), whereas ST2 remained strongly and independently associated with both all-cause and cardiovascular death (Table 2). When high-sensitivity cardiac troponin T was included in the multivariable analysis, log(Gal-3) lost the statistical significance even for all-cause death (Online Table 1).

Both biomarkers remained independently associated with the combined endpoint (all-cause death or HF hospitalization): log(Gal-3) HR: 1.39, 95% CI: 1.06 to 1.83, $p = 0.017$; and ST2 HR: 1.18, 95% CI: 1.08 to 1.29, $p < 0.001$. When high-sensitivity cardiac troponin T was included in the multivariable analysis, log(Gal-3) lost the statistical significance (HR: 1.28, 95% CI: 0.96 to 1.70, $p = 0.088$), whereas ST2 remained statistically associated with this combined endpoint (HR: 1.19, 95% CI: 1.08 to 1.32, $p < 0.001$). Figure 2 shows Kaplan-Meier curves for the combined endpoint according to Gal-3 (Fig. 2A) and ST2 (Fig. 2B) quartiles.

Performance metrics in risk prediction models.

DISCRIMINATION. C-statistics for the prediction of all-cause death and cardiovascular death significantly increased when ST2 was incorporated into the reference model with established mortality risk factors and NT-proBNP. It did not increase for either endpoint when Gal-3 was the added biomarker (Tables 3 and 4).

The same occurred for the combined end-point (all-cause death or HF hospitalization): C-statistic 0.735 [0.711 to 0.759] for reference model, 0.742 [0. 719 to 0.765], $p = 0.033$ for the ST2 model, and 0.737 [0.713 to 0.761], $p = 0.332$ for the Gal-3 model.

CALIBRATION. The p values for the Hosmer-Lemeshow statistics indicated good calibration for all the models except for the model containing Gal-3 for all-cause mortality ($p = 0.049$). Brier scores, AIC, and BIC were lower in the models that included ST2, both for all-cause mortality (Table 3) and for cardiovascular death (Table 4). Global goodness of fit was better in models including ST2 than in the model with only established mortality risk factors, as evaluated by likelihood ratio tests for both all-cause ($p < 0.001$) (Table 3) and cardiovascular death ($p = 0.007$) (Table 4). The likelihood ratio for models including Gal-3 was nonsignificant for cardiovascular mortality ($p = 0.127$) (Table 4).

RECLASSIFICATION. IDI (risk as a continuous variable) increased significantly with the addition of ST2 to the reference model, both for all-cause (IDI: 1.5, $p = 0.003$) (Table 3) and cardiovascular death (IDI: 1.3, $p = 0.004$) (Table 4), but not with the addition of Gal-3 in any case (Tables 3 and 4). NRI (reclassification according to predefined risk categories) for all-cause death improved only after inclusion of ST2 into the full-adjusted model

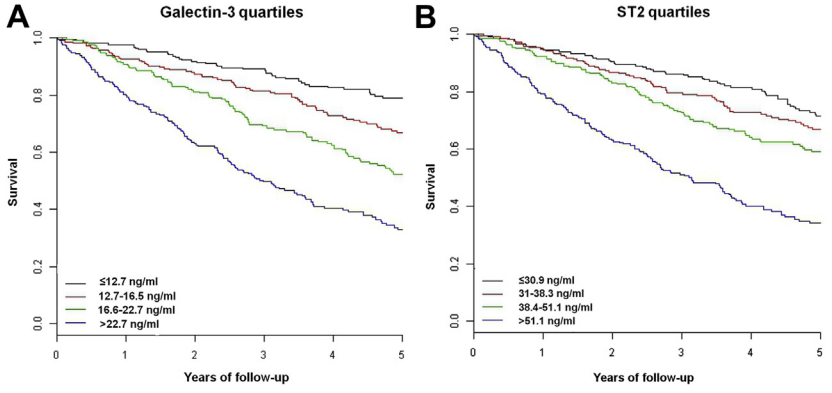

Figure 1 Kaplan-Meier Survival Curves According to Gal-3 and ST2

(A) Survival according to galectin-3 (Gal-3) quartiles. **(B)** Survival according to ST2 quartiles.

Table 2 Multivariate Cox Regression Analyses for All-Cause and Cardiovascular Mortality at 5 Years

	All-Cause Death						Cardiovascular Death					
	Gal-3			ST2			Gal-3			ST2		
	HR	95% CI	p Value	HR	95% CI	p Value	HR	95% CI	p Value	HR	95% CI	p Value
Age, yrs	1.04	1.03–1.06	<0.001	1.04	1.03–1.05	<0.001	1.04	1.03–1.06	<0.001	1.04	1.02–1.06	<0.001
Female	0.67	0.52–0.87	0.003	0.74	0.57–0.96	0.026	0.65	0.46–0.94	0.021	0.70	0.48–1.00	0.052
NYHA functional class	1.70	1.34–2.15	<0.001	1.64	1.29–2.08	<0.001	1.78	1.30–2.40	<0.001	1.83	1.34–2.50	<0.001
LVEF	1.00	1.00–1.01	0.320	1.01	1.00–1.01	0.316	1.01	1.00–1.02	0.229	1.01	1.00–1.02	0.199
Ischemic etiology of HF	1.12	0.88–1.41	0.357	1.14	0.9–1.45	0.265	1.34	0.97–1.84	0.074	1.32	0.96–1.81	0.092
Diabetes mellitus	1.17	0.93–1.46	0.179	1.16	0.93–1.45	0.187	1.28	0.95–1.7	0.102	1.25	0.93–1.69	0.143
eGFR, ml/min/1.73 m^2	1.00	1.00–1.01	0.927	1.00	0.99–1.00	0.245	1.00	0.99–1.01	0.650	1.00	0.99–1.01	0.464
Na, mmol/l	0.97	0.94–1.00	0.065	0.98	0.95–1.01	0.224	0.93	0.90–0.97	0.001	0.93	0.89–0.97	<0.001
Hb, g/dl	0.93	0.87–1.00	0.049	0.92	0.86–0.98	0.014	1.02	0.93–1.12	0.692	1.00	0.92–1.10	0.942
ACEI or ARB treatment	0.81	0.59–1.11	0.181	0.88	0.63–1.22	0.444	0.78	0.51–1.18	0.243	0.72	0.47–1.10	0.133
Beta-blocker treatment	0.52	0.39–0.69	<0.001	0.53	0.40–0.72	<0.001	0.48	0.32–0.72	<0.001	0.49	0.33–0.72	<0.001
Log(NT-proBNP)	1.12	1.15–1.41	<0.001	1.23	1.10–1.36	<0.001	1.29	1.13–1.48	<0.001	1.29	1.12–1.48	<0.001
Log(Gal-3)	1.37	1.03–1.83	0.032	—	—	—	1.35	0.92–1.98	0.127	—	—	—
ST2	—	—	—	1.23	1.12–1.36	<0.001	—	—	—	1.27	1.05–1.53	0.014

New York Heart Association as functional classes III to IV. The logarithmic functions of NT-proBNP and Gal-3 and the quadratic term of ST2 were used in the Cox models. ST2 per every 10 ng/ml change. The p value for $ST2^2$ = 0.001 for all-cause mortality and 0.024 for cardiovascular mortality. Dashes indicate non applicable.

CI = confidence interval; Gal-3 = galectin-3; Hb = hemoglobin; HF = heart failure; HR = hazard ratio; other abbreviations as in Table 1.

(ST2 NRI: 9.4, $p < 0.001$; Gal-3 NRI: 0.7, $p = 0.649$) (Table 3). For cardiovascular death, NRI improved mainly for alive patients when ST2 was added to the reference model (NRI alive: 4.6, $p < 0.001$). Gal-3 did not improve but worsened the NRI for deceased patients (NRI deceased: −4.2, $p = 0.047$) (Table 4). Direct comparison of ST2 and Gal-3 models revealed that ST2 significantly improved reclassification over Gal-3 (Table 5).

Again, when we considered the combined endpoint (all-cause death or HF hospitalization), Gal-3 did not improve reclassification (IDI: 0.3, 95% CI: −0.1 to 0.8, $p = 0.157$; NRI: 0.6, 95% CI: −3.1 to 4.3, $p = 0.739$), whereas ST2 significantly improved both reclassification metrics (IDI: 1.2, 95% CI: 0.4 to 1.9, $p = 0.002$; NRI: 5.4, 95% CI: 0.7 to 10.2, $p = 0.024$).

The addition of high-sensitivity cardiac troponin T in the baseline model did not change the significant value of ST2 in discrimination and reclassification metrics (Online Tables 2 to 4).

Discussion

This study highlights the importance of assessing the true value of emerging cardiac fibrosis biomarkers above and beyond clinical risk factors and natriuretic peptides particularly in light of the newly obtained ST2 and Gal-3 American College of Cardiology/American Heart Association class II recommendation for determination of prognosis in chronic HF (20). ST2 and Gal-3 were directly compared, and our findings demonstrate that: 1) both ST2 and Gal-3 were associated with an increased risk of all-cause mortality, but only ST2 with cardiovascular mortality; and 2) ST2 significantly refined discrimination and reclassification analysis, whereas Gal-3 had negligible effects on performance metrics in risk-prediction models.

The independent prognostic value of ST2 and Gal-3 was examined on top of 11 classical risk factors (age, sex, New York Heart Association functional class, estimated glomerular filtration rate, LVEF, diabetes mellitus, sodium, hemoglobin, ischemic etiology of HF, angiotensin-converting enzyme inhibitor/angiotensin II receptor blocker treatment, and beta-blocker treatment) plus NT-proBNP. Previously, results with Gal-3 concerning outcome prediction have been conflicting whenever natriuretic peptides are incorporated into the analysis. In a short series of 232 patients, Lok et al. (21) found that Gal-3 was a significant predictor of mortality even after adjusting for NT-proBNP. By contrast, Felker et al. (22) and Gullestad et al. (23), in large series of ambulatory HF patients with long-term follow-up, found that Gal-3 was significantly predictive of long-term outcomes only in univariate analysis; this association did not persist after adjustment for other predictors, especially NT-proBNP. On the side of ST2, in all studied cohorts with or without additional biomarkers, including natriuretic peptides, ST2 unambiguously emerged as a cardinal HF risk stratifier (14,24–29). In the current study, the 2 biomarkers remained

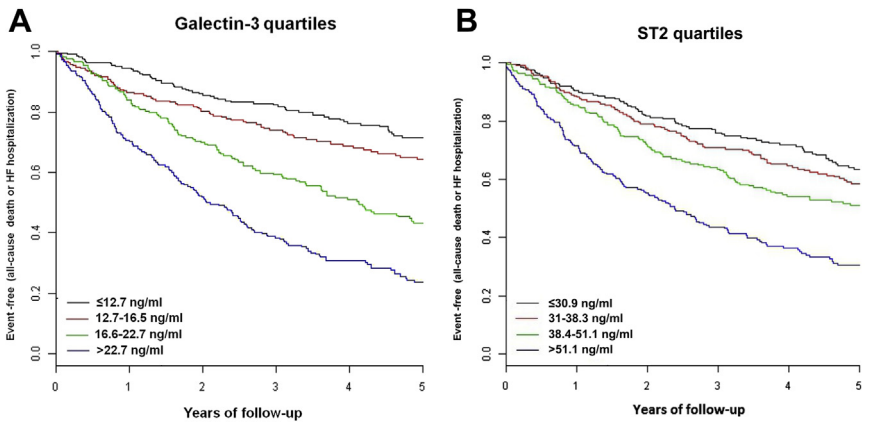

Figure 2 Kaplan-Meier Curves for the Combined Endpoint (All-Cause Mortality or HF Hospitalization) According to Gal-3 and ST2

(A) Event-free survival according to galectin-3 (Gal-3) quartiles. **(B)** Event-free survival according to ST2 quartiles. HF= heart failure.

as independent variables for all-cause mortality, but only ST2 was retained in the subgroup of cardiovascular mortality. Our data indicates that every 10 ng/ml increase in ST2 is associated to ~20% increase in risk.

The additional prognostic information gained by any biomarker over an established risk model needs to be determined using adequate statistical tools (30). At present, a major problem in selecting a biomarker is the proportional increase in economic burden, so any addition should be justified by adequate discrimination, calibration, and reclassification analyses (31). First, value of Gal-3 and ST2 on discrimination metrics: Gal-3 did not significantly increase discrimination (as assessed by the C-statistic) of the reference model. By contrast, incorporation of ST2 into a fully adjusted model significantly improved the C-statistic, which significantly rose up to 0.770 (p value relative to reference model, 0.004). Second, calibration of the models: The full set of calibration analyses used in this study to confirm correspondence of predicted and observed values indicated that overall, the model with ST2 is more accurate. In all models, the Hosmer-Lemeshow test was expected to be nonsignificant; yet, the model that incorporated Gal-3 was significant for all-cause mortality. The Brier score measures the average squared deviation between predicted probabilities for a set of events and their outcomes, so a lower score represents higher accuracy. Given any 2 estimated models, the model with the lower BIC, AIC, and Brier scores is preferred. In this study, the Brier score, the AIC, and the BIC were lower in the ST2 model. Third, value of the studied biomarkers on reclassification metrics: The model with ST2 significantly increased IDI and NRI for all-cause and cardiovascular mortality. Gal-3 had negligible or even deleterious effects on reclassification. Indeed, Gal-3 NRI for cardiovascular mortality reached significance in the opposite direction with a value of −4.2, which is indicative of worsening patient reclassification. Together, these main findings suggest that the pathways identified by ST2 profoundly affect risk stratification in the context of chronic HF and that the incremental predictive value of adding Gal-3 to existing clinical risk factors, particularly above and beyond NT-proBNP, is marginal.

Fibrosis is a fundamental component of the adverse structural remodeling of myocardium present in the failing heart (32). Replacement fibrosis appears at sites of previous cardiomyocyte necrosis to preserve the structural integrity of the myocardium, but not without adverse functional consequences. Increased stress or injury to the myocardium due to acute myocardial infarction, uncontrolled hypertension, and other forms of myocyte damage can contribute to fibrosis and cardiac remodeling. Responses to acute and chronic damage can involve recruitment of immune cells to the myocardium; production of cell signaling proteins from local perycites, mast cells, and macrophages, resulting in activation of resident fibroblasts and myofibroblasts; and the deposition of procollagen into the extracellular matrix, which is irreversibly cross linked to collagen-generating cardiac fibrosis. A multitude of regulators are involved in the pathophysiology of cardiac fibrosis and include ST2 and Gal-3. Given the limited benefit of Gal-3 observed in our cohort of ambulatory chronic HF patients, in which remodeling and fibrosis may be at an advanced stage, it is conceivable that Gal-3 could have a more prominent role in earlier stages of fibrosis pathobiology and ventricular remodeling. Indeed, recent studies found that higher levels of Gal-3 are associated with increased risk for new-onset HF in apparently healthy people (33); in addition, plasma Gal-3 is elevated in patients admitted with acute myocardial infarction and reduced ejection fraction at baseline (34). Gal-3 may be a modest complement to other HF biomarkers by providing an "upstream" signal of myocardial fibrotic state. Nevertheless, much remains to be clarified about Gal-3 at different stages of HF. Nativi et al. (35) recently reported

Table 3 Performance of the Models for All-Cause Mortality at 5 Years

	Reference Model	Model With Gal-3	Model With ST2
Discrimination			
C-statistic	0.757	0.760	0.770
	(0.733 to 0.782)	(0.735 to 0.785)	(0.746 to 0.793)
	Reference	p = 0.143	p = 0.004
Calibration			
H-L	Chi-square: 8.6	Chi-square: 16.9	Chi-square: 14.2
	p = 0.48	p = 0.049	p = 0.12
Brier score	0.171	0.170	0.165
AIC	4,020	4,016	4,003
BIC	4,077	4,078	4,070
Likelihood ratio	Reference	p = 0.032	p < 0.001
Reclassification			
IDI		0.2 (−0.2 to 0.6)	1.5 (0.5 to 2.5)
	Reference	p = 0.288	p = 0.003
NRI—all		0.7 (−2.4 to 3.9)	9.4 (4.8 to 14.1)
	Reference	p = 0.649	p < 0.001
NRI—deceased		−0.1 (−2.6 to 2.4)	4.4 (0.9 to 7.9)
	Reference	p = 0.929	p = 0.014
NRI—alive		0.8 (−1.2 to 2.9)	5.0 (2.0 to 8.1)
	Reference	p = 0.143	p = 0.001

Values are n or n (95% CI) unless otherwise indicated. Reference model includes age, female, ischemic etiology of heart failure, LVEF, NYHA functional class, diabetes mellitus, eGFR, ACEI or ARB treatment, beta-blocker treatment, sodium, hemoglobin, NT-proBNP. Model with Gal-3: Reference model + Gal-3. Model with ST2: Reference model + ST2. All p values versus the reference model.

AIC = Akaike information criterion; BIC = Bayesian information criterion; H-L = Hosmer-Lemeshow test; IDI = integrated discrimination improvement; NRI = net reclassification improvement; other abbreviations as in Tables 1 and 2.

that serum Gal-3 levels stay elevated despite replacement of diseased myocardium and reversal of HF state with heart transplant. These findings suggest that Gal-3 is a systemic biomarker rather than being specific to HF. By contrast, ST2 measurement provides a strong serologic overview of the cumulative myocardial fibrotic process and ultimately is a relevant addition to the predictive ability of the practicing clinician.

Because progressive cardiac fibrosis is a central aspect in the progression of cardiac dysfunction as well as the primary substrate for lethal arrhythmias and sudden death, it is intuitive that a blood marker of cardiac fibrosis would

Table 4 Performance of the Models for Cardiovascular Mortality at 5 Years

	Reference Model	Model With Gal-3	Model With ST2
Discrimination			
C-statistic	0.776	0.778	0.783
	(0.745 to 0.807)	(0.747 to 0.809)	(0.753 to 0.813)
	Reference	p = 0.288	p = 0.04
Calibration			
H-L	Chi-square: 10.2	Chi-square: 5.3	Chi-square: 14.7
	p = 0.33	p = 0.81	p = 0.1
Brier score	0.127	0.127	0.125
AIC	2,251	2,250	2,245
BIC	2,308	2,312	2,311
Likelihood ratio	Reference	p = 0.127	p = 0.007
Reclassification			
IDI		0.2 (−0.3 to 0.6)	1.3 (0.4 to 2.1)
	Reference	p = 0.447	p = 0.004
NRI—all		−4.2 (−8.8 to 0.5)	2.4 (−2.5 to 7.2)
	Reference	p = 0.078	p = 0.344
NRI—deceased		−4.2 (−8.3 to −0.1)	−2.3 (−6.2 to 1.6)
	Reference	p = 0.047	p = 0.254
NRI—alive		<0.1 (−1.8 to 1.8)	4.6 (2.1 to 7.2)
	Reference	p = 0.998	p < 0.001

Values are n or n (95% CI) unless otherwise indicated. Models as defined in Table 3. All p values versus the reference model.

Abbreviations as in Tables 1 to 3.

Table 5 Direct Comparison of Performance for All-Cause and Cardiovascular Mortality at 5 Years of Models Containing Gal-3 and ST2

	All-Cause Mortality		Cardiovascular Mortality	
	Gal-3 vs. ST2		Gal-3 vs. ST2	
Discrimination				
C-statistic	0.760	0.770	0.778	0.783
	(0.735 to 0.785)	(0.746 to 0.793)	(0.747 to 0.809)	(0.753 to 0.833)
	p = 0.035		p = 0.254	
Calibration				
H-L	Chi-square: 16.9	Chi-square: 14.2	Chi-square: 5.3	Chi-square: 14.7
	p = 0.049	p = 0.12	p = 0.81	p = 0.1
Brier score	0.170	0.165	0.127	0.125
AIC	4,016	4,003	2,250	2,245
BIC	4,078	4,070	2,312	2,311
Reclassification				
IDI		1.3 (0.2 to 2.4)		1.1 (0.1 to 2.1)
	Reference	p = 0.019	Reference	p = 0.029
NRI—all		7.8 (2.5 to 13.1)		4.5 (–0.4 to 9.4)
	Reference	p = 0.004	Reference	p = 0.074
NRI—deceased		3.4 (–0.9 to 7.6)		0.5 (–3.6 to 4.6)
	Reference	p = 0.118	Reference	p = 0.800
NRI—alive		4.5 (1.3 to 7.7)		3.9 (1.4 to 6.5)
	Reference	p = 0.005	Reference	p = 0.002

Values are n or n (95% CI) unless otherwise indicated. All models include age, female, ischemic etiology of heart failure, LVEF, NYHA functional class, Diabetes mellitus, eGFR, ACEI or ARB treatment, beta-blocker treatment, sodium, hemoglobin, NT-proBNP. All p values versus the reference model. Reference model = clinical factors + Gal-3; model with ST2 = clinical factors + ST2.

Abbreviations as in Tables 1 to 3.

be independently associated with cardiovascular mortality. This study shows that increased serum levels of ST2 were not only predictive of all-cause mortality but also of cardiovascular mortality. A previous study has already demonstrated the value of ST2 in predicting sudden cardiac death in ambulatory patients with mild-to-moderate chronic HF and left ventricular systolic dysfunction (36). Those authors found that the prognostic value of ST2 was independent of other clinical variables and, importantly, complementary to NT-proBNP. At present, no single test reliably predicts sudden death in patients with HF (37), but the combination of ST2 and NT-proBNP markedly improved risk stratification to identify high- and low-risk patients; this fact may have an important impact on clinical decision making, particularly for delineating optimal preventive strategies.

Study limitations. First, whether serial measurements of both biomarkers at pre-defined time points would have improved risk stratification was not incorporated into the design and is beyond the scope of the present report. Second, with regard to imaging techniques, ultrasounds were primarily used to characterize ventricular remodeling, and cardiac magnetic resonance imaging was not routinely performed or available to all patients. Finally, the population was a general HF population treated at a specific and multidisciplinary HF unit in a tertiary care hospital; most patients were referred from the cardiology department and thus were relatively young men with HF of ischemic etiology and reduced LVEF. As such, these results cannot necessarily be extrapolated to a global HF population. The low use of implantable cardioverter-defibrillators in this consecutive cohort is representative of HF management in Mediterranean countries. It is possible that more widespread use of implantable cardioverter-defibrillators might change our findings. We must also acknowledge that the estimation of effect size from adding biomarker measurements to the clinical model is limited.

Conclusions

The head-to-head comparison of 2 new-generation fibrosis biomarkers revealed that ST2 is an important addition to established risk factors, whereas the additive value of Gal-3 was trivial. The incorporation of ST2 into clinical practice for the prediction of all-cause and cardiovascular mortality should be readily contemplated by the practicing clinician. Further studies should confirm whether this superiority of ST2 is present at all stages of the HF continuum.

Reprint requests and correspondence: Dr. Antoni Bayes-Genis, Department of Cardiology, Hospital Germans Trias I Pujol, Crta. Canyet s/n, Badalona, Barcelona 08916, Spain. E-mail: grupicrec@gmail.com.

REFERENCES

1. Creemers EE, Pinto YM. Molecular mechanisms that control interstitial fibrosis in the pressure-overloaded heart. Cardiovasc Res 2011; 89:265–72.

2. Fan D, Takawale A, Lee J, Kassiri Z. Cardiac fibroblasts, fibrosis and extracellular matrix remodeling in heart disease. Fibrogenesis Tissue Repair 2012;5:15.
3. McMurray JJ, Stewart S. Epidemiology, aetiology, and prognosis of heart failure. Heart 2000;83:596–602.
4. McMurray JJ, Adamopoulos S, Anker SD, et al., for the ESC Committee for Practice Guidelines. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: the Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Eur Heart J 2012;33:1787–847.
5. Jessup M, Abraham WT, Casey DE, et al. 2009 focused update: ACCF/AHA guidelines for the diagnosis and management of heart failure in adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2009;53:1343–82.
6. Doust JA, Pietrzak E, Dobson A, Glasziou P. How well does B-type natriuretic peptide predict death and cardiac events in patients with heart failure: systematic review. BMJ 2005;330:625.
7. Januzzi JL Jr. ST2 as a cardiovascular risk biomarker: from the bench to the bedside. J Cardiovasc Transl Res 2013;6:493–500.
8. Weinberg EO, Shimpo M, De Keulenaer GW, et al. Expression and regulation of ST2, an interleukin-1 receptor family member, in cardiomyocytes and myocardial infarction. Circulation 2002;106:2961–6.
9. Sanada S, Hakuno D, Higgins LJ, Schreiter ER, McEnzie ANJ, Lee RT. IL-33 and ST2 comprise a critical biomechanically induced and cardioprotective signaling system. J Clin Invest 2007;117:1538–49.
10. Chackerian AA, Oldham ER, Murphy EE, Schmitz J, Pflanz S, Kastelein RA. IL-1 receptor accessory protein and ST2 comprise the IL-33 receptor complex. J Immunol 2007;179:2551–5.
11. Dumic J, Dabelic S, Flögel M. Galectin-3: an open-ended story. Biochim Biophys Acta 2006;1760:616–35.
12. Krzéslak A, Lipinska A. Galectin-3 as a multifunctional protein. Cell Mol Biol Lett 2004;9:305–28.
13. Sharma UC, Pokgarel S, van Brakel TJ, et al. Galectin-3 marks activated macrophages in failure-prone hypertrophied hearts and contributes to cardiac dysfunction. Circulation 2004;110:3121–8.
14. Bayes-Genis A, de Antonio M, Galán A, et al. Combined use of high-sensitivity ST2 and NTproBNP to improve the prediction of death in heart failure. Eur J Heart Fail 2012;14:32–8.
15. 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.
16. Zamora E, Lupón J, Vila J, et al. Estimated glomerular filtration rate and prognosis in heart failure: value of the Modification of Diet in Renal Disease Study-4, chronic kidney disease epidemiology collaboration, and Cockroft-Gault formulas. J Am Coll Cardiol 2012;59:1709–15.
17. Dieplinger B, Januzzi JL, Steinmair M, et al. Analytical and clinical evaluation of a novel high-sensitivity assay for measurement of soluble ST2 in human plasma—the Presage ST2 assay. Clin Chim Acta 2009;409:33–40.
18. Januzzi JL, van Kimmenade R, Lainchbury J, et al. NT-proBNP testing for diagnosis and short-term prognosis in acute decompensated heart failure: an international pooled analysis of 1256 patients: the International Collaborative of NT-proBNP Study. Eur Heart J 2006;27:330–7.
19. Pencina MJ, D'Agostino RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med 2011;30:11–21.
20. Yancy CW, Jessup M, Bozkurt B, et al., for the ACCF/AHA Task Force Members. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2013;62:1495–539.
21. Lok DJ, Van Der Meer P, de la Porte PW, et al. Prognostic value of galectin-3, a novel marker of fibrosis, in patients with chronic heart failure: data from the DEAL-HF study. Clin Res Cardiol 2010;99:323–8.
22. Felker GM, Fiuzat M, Shaw LK, et al. Galectin-3 in ambulatory patients with heart failure: results from the HF-ACTION study. Circ Heart Fail 2012;5:72–8.
23. 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.
24. Ky B, French B, McCloskey K, et al. High-sensitivity ST2 for prediction of adverse outcomes in chronic heart failure. Circ Heart Fail 2011;4:180–7.
25. Mueller T, Dieplinger B, Gegenhuber A, Poelz W, Pacher R, Haltmayer M. Increased plasma concentrations of soluble ST2 are predictive for 1-year mortality in patients with acute destabilized heart failure. Clin Chem 2008;54:752–6.
26. Broch K, Ueland T, Nymo SH, et al. Soluble ST2 is associated with adverse outcome in patients with heart failure of ischaemic aetiology. Eur J Heart Fail 2012;8:268–77.
27. Bayes-Genis A, Pascual-Figal D, Januzzi JL, et al. Soluble ST2 monitoring provides additional risk stratification for outpatients with decompensated heart failure. Rev Esp Cardiol 2010;63:1171–8.
28. Rehman SU, Mueller T, Januzzi JL Jr. Characteristics of the novel interleukin family biomarker ST2 in patients with acute heart failure. J Am Coll Cardiol 2008;52:1458–65.
29. Hlatky MA, Greenland P, Arnett DK, et al., for the American Heart Association Expert Panel on Subclinical Atherosclerotic Diseases and Emerging Risk Factors and the Stroke Council. Criteria for evaluation of novel markers of cardiovascular risk: a scientific statement from the American Heart Association. Circulation 2009;119:2408–16.
30. Giannesi D. Multimarker approach for heart failure management: perspectives and limitations. Pharmacol Res 2011;64:11–24.
31. Lupón J, de Antonio M, Galán A, et al. Combined use of the novel biomarkers high-sensitivity troponin T and ST2 for heart failure risk stratification vs conventional assessment. Mayo Clin Proc 2013;88:234–43.
32. Mann DL. Mechanisms and models in heart failure: a combinatorial approach. Circulation 1999;100:999–1008.
33. Ho JE, Liu C, Lyass A, et al. Galectin-3, a marker of cardiac fibrosis, predicts incident heart failure in the community. J Am Coll Cardiol 2012;60:1249–56.
34. Weir RAP, Petrie CJ, Murphy A, et al. Galectin-3 and cardiac function in survivors of acute myocardial infarction. Circ Heart Fail 2013;6:492–8.
35. Nativi JJ, Kremers WK, Hasin T, et al. Galectin-3, a biomarker of poor prognosis in heart failure remains unchanged after heart transplantation. Paper presented at: AHA Congress; November 4–6, 2012; Los Angeles, CA.
36. Pascual-Figal DA, Ordoñez-Llanos J, Tornel PL, et al., for the MUSIC Investigators. Soluble ST2 for predicting sudden cardiac death in patients with chronic heart failure and left ventricular systolic dysfunction. J Am Coll Cardiol 2009;54:2174–9.
37. Lane RE, Cowie MR, Chow AW. Prediction and prevention of sudden cardiac death in heart failure. Heart 2005;91:674–80.

Key Words: biomarkers ■ heart failure ■ myocardial fibrosis ■ remodeling ■ survival.

APPENDIX

For supplemental tables and a figure, please see the online version of this article.