SUPPLEMENTAL MATERIAL
Expanded Methods

Exclusion criteria

Exclusion criteria comprised contraindications to CMR, New York Heart Association (NYHA) functional class IV, acute myocarditis, acute sarcoidosis, infiltrative disorders (e.g. amyloidosis), congenital heart disease, or hypertrophic cardiomyopathy. Renal insufficiency was an exclusion for contrast (n=42) with creatinine clearance, CrCl<30 ml/min after July 2006; CrCl<60 ml/min after February 2007, though non-contrast images for LV and LA structure and function were used.

Quantification of phasic LA volumes and function

Endocardial and epicardial borders of the LA were planimetered. A pixel-to-pixel matching technique automatically tracks voxels during the cardiac cycle. Maximum LA volume (Vmax; LA volume at end-systole before mitral valve opening), minimum LA volume (Vmin; LA volume at end-diastole immediately after mitral valve closure), and pre-atrial contraction LA volume (VpreA; LA volume before atrial contraction) were measured using the LA volume curve. LA functional metrics were calculated as: LA total emptying fraction (LAef) = (Vmax-Vmin)×100% / Vmax; LA passive ef = (Vmax-VpreA)×100% / Vmax; and LA active ef = (VpreA-Vmin) ×100% / VpreA.

RF-SLAM methodology

A key feature of RF-SLAM is the creation of counting process information units (CPIUs) to account for time-varying covariates. A person’s event hazard is assumed to be constant within each CPIU but this strategy then allows predictor variables to change from one interval to the next. The estimated hazard is (a Bayesian analogue of) the total events divided by the total person-time at risk. People can contribute more or fewer person-times. For example, persons who are censored or have an event during the interval contribute fewer person-times than someone who completes the interval without an event. If no person completes the full interval, the constant hazard is based upon the evidence from the earlier part of the interval.

After partitioning the follow-up time into CPIUs, we use a Poisson regression splitting criterion that does not impose the proportional hazards assumption that the predictors have a common effect across the entire follow-up period. Bootstrapping is performed at the patient-level and the predictions for each CPIU for each person are used to generate discrete-time piece-wise constant hazard functions for each bootstrap replication. Variable selection and imputation of missing covariates are inherent to the RF methods used here. For time-varying covariates, the values for those covariates are only recorded for that interval if they were measured during that interval of time. If values were not measured during the interval, then they were set to not available (NA, missing). These NA values were then imputed during the RF approaches using adaptive tree imputation. We used default values of 1000 as the number of trees, 10% of the number of CPIUs as the minimum terminal node size, and the number of variables to test at each potential split point as the square root of the number of predictors in the model.

We compared RF-SLAM with both baseline and time-varying covariates and RF-SLAM with baseline covariates only.
RF-SLAM partial dependence plots

The goal of a partial dependence plot is to display the relationship of the predicted risk of the outcome with a single or set of predictor(s) controlling for the others in the model. In this application, we have estimated partial dependence plots using linear regression with the RF-SLAM out-of-bag predicted values as the outcome and a natural spline of each set of predictors with 3 degrees of freedom that interacts with an indicator variable for having had $\geq 1$ interval HF hospitalizations. The model was estimated with complete cases only and using multiple imputation by chained equations (MICE) that were not qualitatively different.

RFS methodology

To compare to the RF-SLAM methods, we also constructed RFS models. We used default settings for the number of trees (1000), node size (15), and number of variables to try at each potential split (square root of the number of predictors in the model). After building a random forest with the RFS approach, the survival and cumulative hazard estimates are obtained. Although the RFS method does not provide piecewise-constant hazard predictions, we developed an approach to obtain the discrete-time hazard estimates to facilitate comparisons between the methods. Specifically, the survival predictions were obtained from the RFS method and a smooth curve was fitted to the predictions to obtain an estimate of the survival function. Afterwards, the value of the derivative of the log of the estimated survival function was obtained every half-year (i.e. 0.5, 1, 1.5 years, etc.) to derive comparable hazard estimates to the RF-SLAM approach.

Cox regression methodology

For Cox regression models, a subset of the possible predictor variables was chosen using forward variable selection; multiple imputation was used to impute missing data.

AUC and confidence intervals

The AUC and 95% confidence intervals were obtained from 500 bootstrap iterations using the time-dependent AUC (https://cran.r-project.org/web/packages/risksetROC/index.html) as described previously.

Comparison to Seattle Heart Failure Model risk scores

We derived the SHFM-D scores using covariates consisting of: age, gender, NYHA class, weight, LVEF, systolic blood pressure, ischemic/nonischemic etiology, medication use (angiotensin converting enzyme inhibitor or angiotensin receptor blocker; beta-blocker; statin; allopurinol; aldosterone blocker; diuretic type and dose; hemoglobin; total cholesterol; and serum sodium); and accounting for presence of LBBB, QRS-duration, and device type (ICD or CRT-D).

SPRM scores were derived from the following covariates: age, gender, NYHA class, diabetic status, use of digoxin, body mass index, LVEF, systolic blood pressure, and serum sodium and creatinine levels.

Both SHFM-D and SPRM metrics were compared between those with and without the primary endpoint using Cox proportional hazards regression.
Expanded Results

Table S1. Statistical models compared (n=8).

| Covariates included in each model                                      | Cox regression | RFS | RFSLAM* |
|-----------------------------------------------------------------------|----------------|-----|---------|
| Baseline non-CMR variables only                                       | X              | X   | X       |
| Baseline non-CMR and CMR variables                                   | X              | X   | X       |
| Time-varying with baseline non-CMR variables only                    |                |     | X       |
| Time-varying with baseline non-CMR and CMR variables                 |                |     | X       |

*RF-SLAM incorporates CPIUs and Poisson log likelihood in all iterations.

Table S2. Predicted mortality by Seattle Heart Failure Model for ICD patients (SHFM-D).

|                                | No primary event (n=307) | Primary event (n=75) | p-value |
|--------------------------------|--------------------------|----------------------|---------|
| 1 year predicted mortality (%) | 3.5 ± 3.6                | 3.5 ± 2.8            | 0.34    |
| 2 year predicted mortality (%) | 6.9 ± 6.7                | 6.8 ± 5.3            | 0.42    |
| 5 year predicted mortality (%) | 17.1 ± 14.0              | 17.5 ± 12.2          | 0.39    |
| Mean life expectancy (years)  | 15.1 ± 6.0               | 14.7 ± 6.1           | 0.38    |
| 5 year SHFM score (median, IQR)| -0.38 (-0.9-0.26)        | -0.22 (-0.89-0.20)   | 0.39    |

Table S3. Predicted proportion of sudden deaths by Seattle Proportional Risk Model (SPRM).

|                                | No primary event (n=307) | Primary event (n=75) | p-value |
|--------------------------------|--------------------------|----------------------|---------|
| SPRM-predicted proportion of SCD (median, IQR) | 0.57 (0.50-0.67) | 0.62 (0.52-0.69) | 0.054 |
| Predicted ICD hazard ratio     | 0.66 (0.54-0.77)         | 0.60 (0.51-0.73)    | 0.06    |
Table S4. Incidence rates (per year) by enrollment phase.

| Enrollment years       | 2003-2006 (n=148) | 2007-2011 (n=131) | 2012-2015 (n=103) |
|------------------------|-------------------|-------------------|-------------------|
| **Primary endpoint events** |                   |                   |                   |
| 0-3 year incidence rate | 5.4%              | 3.7%              | 2.6%              |
| 3-5 year incidence rate | 2.9%              | 1.0%              | 4.0%              |
| >5 year incidence rate  | 3.3%              | 4.1%              | 3.2%              |
| **Inappropriate ICD firings** |                   |                   |                   |
| 0-3 year incidence rate | 5.6%              | 3.8%              | 0.8%              |
| 3-5 year incidence rate | 5.3%              | 1.0%              | 0.4%              |
| >5 year incidence rate  | 1.9%              | 0.6%              | 1.4%              |