

ORIGINAL RESEARCH

Baseline and Dynamic Risk Predictors of Appropriate Implantable Cardioverter Defibrillator Therapy

Katherine C. Wu , MD; Shannon Wongvibulsin , PhD; Susumu Tao, MD, PhD; Hiroshi Ashikaga, MD, PhD; Michael Stillabower, MD; Timm M. Dickfeld, MD, PhD; Joseph E. Marine, MD; Robert G. Weiss, MD; Gordon F. Tomaselli, MD; Scott L. Zeger, PhD

BACKGROUND: Current approaches fail to separate patients at high versus low risk for ventricular arrhythmias owing to over-reliance on a snapshot left ventricular ejection fraction measure. We used statistical machine learning to identify important cardiac imaging and time-varying risk predictors.

METHODS AND RESULTS: Three hundred eighty-two cardiomyopathy patients (left ventricular ejection fraction $\leq 35\%$) underwent cardiac magnetic resonance before primary prevention implantable cardioverter defibrillator insertion. The primary end point was appropriate implantable cardioverter defibrillator discharge or sudden death. Patient characteristics; serum biomarkers of inflammation, neurohormonal status, and injury; and cardiac magnetic resonance-measured left ventricle and left atrial indices and myocardial scar burden were assessed at baseline. Time-varying covariates comprised interval heart failure hospitalizations and left ventricular ejection fractions. A random forest statistical method for survival, longitudinal, and multivariable outcomes incorporating baseline and time-varying variables was compared with (1) Seattle Heart Failure model scores and (2) random forest survival and Cox regression models incorporating baseline characteristics with and without imaging variables. Age averaged 57 ± 13 years with 28% women, 66% white, 51% ischemic, and follow-up time of 5.9 ± 2.3 years. The primary end point ($n=75$) occurred at 3.3 ± 2.4 years. Random forest statistical method for survival, longitudinal, and multivariable outcomes with baseline and time-varying predictors had the highest area under the receiver operating curve, median 0.88 (95% CI, 0.75-0.96). Top predictors comprised heart failure hospitalization, left ventricle scar, left ventricle and left atrial volumes, left atrial function, and interleukin-6 level; heart failure accounted for 67% of the variation explained by the prediction, imaging 27%, and interleukin-6 2%. Serial left ventricular ejection fraction was not a significant predictor.

CONCLUSIONS: Hospitalization for heart failure and baseline cardiac metrics substantially improve ventricular arrhythmic risk prediction.

Key Words: cardiac magnetic resonance imaging ■ heart failure ■ risk stratification ■ sudden cardiac death ■ ventricular arrhythmia

See Editorial by Tan and Ellenbogen

Identifying patients most at risk for ventricular arrhythmias and sudden cardiac death (SCD) remains a clinical challenge.¹ Current guidelines directing the use of primary prevention implanted cardioverter defibrillator (ICDs) rely on a reduced left ventricular

ejection fraction (LVEF), which has low predictive efficiency.¹ Estimated annualized rates of appropriate ICD therapy in recipients of primary prevention ICDs with ischemic and nonischemic cardiomyopathy range from 8% per year in historical randomized ICD

Correspondence to: Katherine C. Wu, MD, Johns Hopkins Medical Institutions, Division of Cardiology, Halsted 559, 600 North Wolfe Street, Baltimore, MD 21287. E-mail: kwu@jhmi.edu

Supplementary Material for this article are available at <https://www.ahajournals.org/doi/suppl/10.1161/JAHA.120.017002>

For Sources of Funding and Disclosures, see page 11.

© 2020 The Authors. Published on behalf of the American Heart Association, Inc., by Wiley. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

JAHA is available at: www.ahajournals.org/journal/jaha

CLINICAL PERSPECTIVE

What Is New?

- In primary prevention implanted cardioverter defibrillator (ICD) recipients, an interim hospitalization for heart failure identified a group at high risk for subsequent ventricular arrhythmia defined as an appropriate ICD shock.
- Among ICD recipients without heart failure hospitalizations, baseline cardiac magnetic resonance imaging metrics (specifically, left ventricle heterogeneous gray and total scar, left ventricle and left atrial volumes, and left atrial total emptying fraction) as well as serum interleukin-6 levels were the strongest predictors of subsequent appropriate ICD shock; serial left ventricle ejection fraction did not provide additional prognostic value for the arrhythmic outcome.
- Machine learning statistical methods may improve risk score development by accounting for complex and dynamic interactions among risk variables and temporally varying risk.

What Are the Clinical Implications?

- The combination of clinical heart failure course, baseline cardiac magnetic resonance imaging metrics, and levels of the inflammatory biomarker interleukin-6 can most accurately stratify subsequent high versus low ventricular arrhythmic risk and may be useful for decision-making as primary prevention ICD recipients approach elective generator change.

Nonstandard Abbreviations and Acronyms

LGE	late gadolinium enhancement
RF-SLAM	random forest survival, longitudinal and multivariate outcomes
SCD	sudden cardiac death
SI	signal intensity

trials² and older national registries³ to as low as 1% per year more recently.⁴ More precise discrimination between high versus low risk ICD candidates could improve personalized, shared decision-making regarding the risk to benefit ratio of ICD insertion.²

The use of a single snapshot LVEF measurement is a poor surrogate for the pathophysiologically complex interplay of factors that increase risk for ventricular arrhythmias. A major prerequisite for the initiation and propagation of reentrant ventricular arrhythmias is a vulnerable underlying arrhythmogenic substrate characterized by the presence, amount, and architecture

of myocardial scar and fibrosis.^{5–9} Increasingly recognized is the variable trajectory of both LVEF and the clinical course of cardiomyopathies,^{10,11} though the true impact of LVEF improvement/recovery on ventricular arrhythmic outcomes remains understudied, as does the relative importance of decompensated heart failure (HF).

We conducted the current investigation to identify factors predictive of appropriate ICD shocks in a cohort of patients who met clinical guidelines for primary prevention ICD insertion. We aimed to build on the current knowledge base by leveraging a statistical learning method based on random forests to examine the complex relationships more deeply among multiple risk predictors, both baseline and time varying, and thereby isolate those factors most important in separating patients along a spectrum of risk. Specifically, we explored how ≥ 1 HF hospitalization and temporal trends in serial LVEFs, both surrogates for cardiomyopathy trajectory, affect risk stratification. We further investigated the incremental value of these time-varying factors in conjunction with a rich set of pathophysiologically driven markers of abnormal myocardial substrate measured by cardiac magnetic resonance imaging with late gadolinium enhancement (CMR-LGE).

METHODS

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Study Participants and Design

The prospective observational registry, Left Ventricular Structural Predictors of Sudden Cardiac Death (NCT01076660), enrolled 382 patients between November 2003 and April 2015 at 3 sites: Johns Hopkins Medical Institutions (Baltimore, MD), Christiana Care Health System (Newark, DE), and the University of Maryland Medical System (Baltimore, MD); details of the study protocol and interim results have been previously described.^{12–16} The institutional review boards at all sites approved the protocol and all patients signed informed consent. Patients meeting clinical criteria for primary prevention ICD insertion based on LVEF \leq 35% were approached for enrollment; 328 patients co-enrolled in PROSE-ICD (Prospective Observational Study of Implantable Cardioverter Defibrillators, NCT00733590).¹⁵ Patients underwent CMR-LGE using 1.5-T whole body scanners (Signa CV/I, GE Healthcare, Milwaukee, WI; or Siemens Avanto, Erlangen, Germany) with a standard, uniform imaging protocol and centralized image analysis at

the Johns Hopkins core laboratory. See Data S1 for exclusions.

At enrollment, patients underwent a baseline history and physical examination, 12-lead electrocardiography, and fasting blood collection before implantation of a single-chamber or dual-chamber ICD, or cardiac resynchronization therapy with an ICD based on guidelines. Device programming was at the discretion of the operators. Participants were evaluated biannually and after any ICD discharge. ICDs were interrogated in person or via remote transmission. Records for interim HF hospitalizations and/or LVEF measurements were reviewed.

The primary end point was appropriate ICD shock for ventricular tachycardia above the programmed rate cutoff (generally 180 bpm) or ventricular fibrillation or definite or suspected sudden arrhythmic death after review of ICD interrogations when available, medical records, death certificates, autopsy reports, and eyewitness accounts. Hinkle-Thaler criteria were used when ICD interrogations at time of death were unavailable. All arrhythmic events were adjudicated centrally by 2 clinical cardiac electrophysiologists. A third electrophysiologist reconciled disagreements. Hospitalization for HF required confirmatory documentation of signs and symptoms and intensified HF treatment.

CMR Imaging Protocol

Steady-state free precession and post-gadolinium inversion-recovery fast gradient-echo CMR sequences were used and images were analyzed with custom research software Cinetool (GE Healthcare). Cines were quantified for left ventricle (LV) and left atrial (LA) volumes and LV mass. Two observers blinded to clinical outcome determined the presence of LGE by reviewing all cross-sections. When present, LGE was quantified into core and heterogeneous scar (gray zone) extents using published methodology.^{12,13} Core extent comprised all pixels with signal intensity (SI) >50% of maximal SI within the hyperenhanced region. Gray zone extent comprised all pixels with SI > peak SI in the normal myocardium but < 50% of maximal SI within the hyperenhanced region. Total LV scar comprised the sum of gray zone and core extents. Phasic LA volumes and function were analyzed by Multimodality Tissue Tracking software (version 6.0, Toshiba, Japan) as described¹⁶ (see Data S1).

Serum Biomarker Assessment

Patients co-enrolled in PROSE-ICD had peripheral blood analyzed at baseline for serum biomarkers of inflammation, neurohormonal activation, and myocardial infarction as detailed¹⁵ and included high-sensitivity

C-reactive protein (ALPCO Diagnostics, Salem, NH); interleukin-6 (IL-6) (R&D Systems, Minneapolis, MN); interleukin-10 (R&D Systems); tumor necrosis factor α receptor II (R&D Systems); NT-proBNP (N-terminal pro-B-type natriuretic peptide) (ALPCO Diagnostics); creatine kinase MB-fraction, myoglobin, and cardiac troponin T and I (all with Meso Scale Discovery, Rockville MD).

Comparison to Seattle Heart Failure Model Risk Scores

We calculated and compared the mortality rates and proportions of sudden deaths predicted by the Seattle Heart Failure Model (SHFM) model for ICD recipients (SHFM-D)¹⁷ and the Seattle Proportional Risk model (SPRM)¹⁸, which use a subset of demographic and clinical variables (see Data S1).

Statistical Analysis

Event data were censored at 8 years after enrollment or at time of death, first appropriate ICD firing, cardiac transplant, or LV assist device implantation or loss to follow-up. Few primary end point events accrued beyond 8 years. Continuous data are reported as mean \pm SD or as median with interquartile ranges for highly skewed distributions. We adapted the random forest survival algorithm (fast unified random forests for survival, regression, and classification, RF-SRC, R package, version 2.9.0)^{19,20} to incorporate **S**urvival with time-dependent predictors, **L**ongitudinal (repeated over time), and/or **M**ultivariate outcomes in discrete time analysis (RF-SLAM) and described the development, methodology, and approach to visualization displays in 2 recent technical papers.^{21,22} Random forest has been a standard machine learning method since its description in 2001; it combines an ensemble of predictions from a collection ("forest") of individual decision trees.²³ Each individual tree generates a prediction and the overall random forest prediction is the average vote from all the trees in the forest. Random forest uses the "wisdom of the crowd" concept to minimize prediction error by averaging a large number of relatively uncorrelated bootstrap replications of the original training data.

For RF-SLAM,^{21,22} we included (1) baseline predictors consisting of demographics, comorbidities, medications, electrophysiologic parameters, laboratory values, enrollment LVEF, and CMR imaging metrics detailed in Table 1 (including 14 CMR variables and 9 biomarkers of inflammation, neurohormonal activation and myocardial infarction); and (2) dynamic predictors comprising ≥ 1 HF hospitalization (n=329 total HF hospitalizations) and serial LVEFs, with times of occurrence/measurement relative to the enrollment date. To

Table. Patient Characteristics by Primary End Point*

	No Primary End Point (n=307)	Primary End Point (n=75)	P Value
<i>Demographics and clinical characteristics</i>			
Age, y	59 (49, 66)	58 (51, 65)	0.89
Male (%)	211 (69)	63 (84)	0.01
Race: White/Black/Other	200 (65)/ 99 (32)/ 8 (3)	51 (68)/ 21 (28)/ 3 (4)	0.66
Body surface area, m ²	2.0 (1.8, 2.2)	2.1 (1.9, 2.3)	0.02
Ischemic cardiomyopathy etiology	149 (49)	44 (59)	0.15
Years from index Myocardial infarction/ cardiomyopathy diagnosis	3.83 (5.18)	5.43 (5.61)	0.02
New York Heart Association functional class I/II/III	64 (21)/ 137 (45)/ 106 (35)	20 (27)/ 31 (41)/ 24 (32)	0.55
<i>Cardiac risk factors</i>			
Hypertension	180 (59)	44 (59)	>0.99
Hypercholesterolemia	180 (59)	45 (60)	0.93
Diabetes mellitus	85 (28)	19 (25)	0.79
Nicotine use	133 (43)	44 (59)	0.02
<i>Medication usage</i>			
Angiotensin-converting enzyme inhibitor or angiotensin receptor blocker	275 (90)	66 (88)	0.85
Beta blocker	288 (94)	68 (91)	0.48
Lipid-lowering	199 (65)	56 (75)	0.14
Antiarrhythmic drugs†	18 (6)	8 (11)	0.22
Diuretics	173 (56)	54 (72)	0.02
Digoxin	50 (16)	16 (21)	0.39
Aldosterone-inhibitor	80 (26)	21 (28)	0.85
Aspirin	215 (70)	55 (73)	0.67
<i>Electrophysiologic variables</i>			
History of atrial fibrillation	51 (17)	14 (19)	0.80
Ventricular rate, bpm,	72 (63, 83)	69 (59, 81)	0.09
QRS duration, msec,	108 (96, 140)	120 (100, 144)	0.14
Presence of left bundle branch block	79 (26)	14 (19)	0.26
Biventricular implantable cardioverter defibrillator	90 (29)	17 (23)	0.31
<i>Laboratory values/ biomarkers</i>			
Sodium, mEq/L	139 (137, 141)	137 (137, 141)	0.97
Potassium, mEq/L	4.2 (4, 4.5)	4.3 (4, 4.5)	0.91
Creatinine, mEq/L	1.0 (0.8, 1.2)	1.0 (0.9, 1.2)	0.06
Estimated glomerular filtration rate, mL/min/1.73 m ²	81 (24)	80 (21)	0.64
Blood urea nitrogen, mg/dL	18 (13, 24)	20 (13, 24)	0.35
Glucose, mg/dL	103 (91, 126)	106 (93, 114)	0.93
Hematocrit, %	40.0 (37.4, 43.2)	41.3 (37.9, 44.6)	0.03
High-sensitivity C-reactive protein, µg/mL	3.2 (1.2, 7.8)	4.6 (2.0, 10.1)	0.06
IL-6, pg/mL	1.4 (0.8, 2.8)	2.0 (1.4, 4.2)	<0.01
IL-10, pg/mL	1.4 (0.9, 2.5)	1.3 (1.0, 2.8)	0.89
Tumor necrosis factor α receptor II, pg/mL	2989 (2199, 4124)	3014 (2295, 3858)	0.68
N-terminal pro-B-type natriuretic peptide, pmol/L	1750 (706, 3070)	2065 (1300, 3450)	0.08
Cardiac troponin T, ng/mL	0 (0, 0.02)	0 (0, 0.02)	0.43
Cardiac troponin I, ng/mL	0.02 (0, 0.05)	0.02 (0, 0.07)	0.98
Creatine kinase MB-fraction, ng/mL	2.6 (1.7, 4.0)	3.2 (2.1, 4.6)	0.19
Myoglobin, ng/mL	23.5 (20.8, 28.0)	23 (20.7, 29.3)	0.91
Entry LVEF by echo, %	24±8	23±7	0.1

(Continues)

Table. Continued

	No Primary End Point (n=307)	Primary End Point (n=75)	P Value
<i>CMR structural and functional indices</i>			
LVEF, %	27.8±10.3	25.1±8.8	0.04
LV end-diastolic volume index, ml/m ²	115.6 (94.3, 141.2)	128.9 (101, 156.9)	0.02
LV end-systolic volume index, ml/m ²	80.2 (63.6, 110)	100 (68.4, 124.7)	0.01
LV mass index, ml/m ²	71.8 (58.4, 85.3)	76.8 (64.8, 95.7)	0.02
LA maximal volume index, ml/m ²	40.6 (31.1, 57.6)	43.4 (32.4, 63.4)	0.13
LA minimal volume index, ml/m ²	22.8 (16.2, 39.7)	29.0 (20.4, 47.2)	0.01
LA preatrial volume index, ml/m ²	32.8 (24.7, 48.6)	38.1 (27.2, 58.8)	0.04
LA total emptying fraction, %	40.4 (27.7, 49.4)	33.1 (20.4, 45.3)	<0.01
LA passive emptying fraction, %	13.3 (7.6, 21.1)	13.3 (7.3, 19.0)	0.54
LA active emptying fraction, %	28.4 (17.4, 38.3)	21.7 (11.7, 31.1)	<0.01
<i>CMR hyperenhancement</i>			
Late gadolinium enhancement present, %	176 (66)	56 (86)	<0.01
Gray zone, grams	4.5 (0, 14.4)	13.3 (3.1, 22.6)	<0.01
Core, grams	7.2 (0, 19.7)	18.5 (3.7, 24.0)	<0.01
Total scar, grams	14.0 (0, 36.7)	30.5 (7.0, 50.2)	<0.01
<i>Other clinical outcomes</i>			
All-cause mortality (%)	99 (32)	41 (55)	<0.01
Time to death (years)	6.8±3.3	7.5±3.3	0.14

CMR indicates cardiac magnetic resonance; IL-10, interleukin-10; IL-6, interleukin-6; LA, left atrium; LV, left ventricle; LVEF, left ventricle ejection fraction. (Parts of the table were included in²¹ and ²² [https://medinform.jmir.org/2020/6/e15791/] and reprinted here with permission. Both are open-access articles distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium.)

*Results shown are number (%), median (interquartile range), or mean (SD).

†Amiodarone was the only antiarrhythmic prescribed for the indication of atrial arrhythmias.

incorporate time-varying covariates, we created counting process information units, a preprocessing step that partitions the multiple time-varying events and other variables for each person into user-specified bins of time. We prespecified bins at 6-month intervals to reflect biannual clinic visits. By using counting process information units, we accounted for the timeline of prior events and temporal variability in covariates. Further details of RF-SLAM are summarized in Data S1.

We used the Hosmer-Lemeshow statistic²⁴ to test the null hypothesis that the RL-SLAM predictions are well calibrated to the observed rates of events. To compare RF-SLAM with other approaches, we constructed random forest survival and Cox proportional hazards models including baseline covariates with and without imaging metrics (see Data S1). Model performance was evaluated using time-dependent area under the receiver-operator characteristic curves (AUC), an extension to the case of longitudinal data with time-varying predictors.²⁵ Eight models were compared (see Table S1).

To improve the interpretation of the RF-SLAM model results, we summarized its predictions using a single summary regression tree.²² The summary tree was grown until it explained $\geq 95\%$ of the variation in the RF-SLAM predictions. We also calculated variable

dependence plots to visualize how the predictions vary with the main categories of covariates (HF hospitalizations, imaging variables, and IL-6) selected into the summary tree, controlling for the others in the model. Further details are provided in Data S1.

Analyses were conducted with R version 3.6.3 (https://www.r-project.org). Statistical significance was defined as P value < 0.05 .

RESULTS

Baseline Characteristics

Baseline cohort characteristics are shown in Table 1, with all covariates included in the statistical models. At a median follow-up of 5.8 years, the primary end point, appropriate shocks or sudden arrhythmic death, occurred in 75 patients (incidence rate of 3.7 per 100 person-years) after enrollment with 2 deaths and 73 shocks. There were 140 all-cause deaths (incidence rate of 5.3 per 100 person-years). Among patients without a primary event, the incidence rate of death was 4.7 per 100 person-years versus 7.3 in those with an event (cumulative hazard ratio 1.5; 95% CI, 1.1-2.2; $P<0.02$). Patients with the primary end point were more

frequently male, smokers, and had higher body surface areas and longer durations of cardiomyopathy. All patients were well treated medically for HF with similar regimens except for higher diuretic use among those with events. Concentrations of IL-6 were higher among those with primary events. Most CMR indices were abnormal in the primary event group with lower LVEF, higher LV volumes and mass, and larger scar extents. LA minimal and preatrial contraction volumes were larger among those with events and total and active LA emptying fractions were lower.

SHFM and SPRM Model Predictions

Mean SHFM-D-predicted life expectancy was similar for patients with versus without events (14.7 ± 6.1 versus 15.1 ± 6.0 years) translating into similarly low risk scores for all-cause death (Table S2). The SPRM-predicted proportion of mortality attributable to SCD was slightly higher in those eventually experiencing a primary event versus those who did not (0.62 versus 0.57, $P=0.05$, Table S3).

HF Hospitalizations and LVEF Measurements

One hundred and forty patients experienced at least 1 HF hospitalization (incidence rate of 6 per 100 person-years). Among the 75 primary event patients, 41 (55%) had ≥ 1 preceding HF hospitalization (incidence rate of 5.2 per 100 person-years). Among the 307 event-free patients, 99 (32%) had ≥ 1 interval HF hospitalization during follow-up (incidence rate of 9.4 per 100 person-years). The cumulative unadjusted hazard ratio for HF among those with versus without a primary event was 1.8; 95% CI, 1.2-2.6; $P=0.002$. 91% of the cohort underwent ≥ 1 subsequent LVEF measurement at 2 years after enrollment; 73% had a second LVEF at 3.7 years; and 53% had a third LVEF at 5.2 years. All LVEFs were used in RF-SLAM.

Diagnostic Performance of the Models

For primary end point prediction, the diagnostic performances over time for all 8 models (Table S1) are shown in Figure 1. Inclusion of baseline imaging metrics improved the performance of each method. With only baseline, nonimaging covariates, the AUC for RF-SLAM at 5 years was estimated to be 0.73 (95% CI, 0.58-0.86). Adding imaging to baseline covariates improved the estimated AUC for RF-SLAM to 0.78 (95% CI, 0.63-0.89). With the further addition of time-varying indices of ≥ 1 HF hospitalization and serial LVEFs to baseline covariates including imaging, RF-SLAM model performance was highest with estimated AUC of 0.88 (95% CI, 0.75-0.96). In

comparison, the AUC for SPRM was 0.57 and SHFM at 5 years was 0.53.

RF-SLAM Prediction Model Results

Calibration of RF-SLAM was excellent ($\chi^2=9.94$, $P=0.26$) with no significant difference between predicted and observed event rates and no evidence of systematic bias (eg, overprediction at one extreme and underprediction at the other). A summary tree representing aggregate results from all bootstrapped trees for RF-SLAM that included baseline imaging and nonimaging risk factors and time-dependent covariates is shown in Figure 2. It depicts the top 7 predictor variables and the decision rules at each node for prediction of outcomes at 5 years following enrollment. These top variables accounted for 96% of the total variation in the predicted values from RF-SLAM. HF hospitalizations accounted for 67% of the total variation. Imaging metrics, namely, LV diastolic volume, total and gray zone scar extents, indexed LA minimum volume, and total LA emptying fractions, together accounted for another 27% of the total variation. Interleukin-6 concentration accounted for 2%. Patients can be separated along a gradient of ensuing yearly primary event risk ranging from low to high. Notably, serial LVEFs did not significantly add to the prediction when HF hospitalizations and imaging metrics were already incorporated.

Figure 3 complements the results shown in Figure 2 by depicting the relationships between each set of top predictors and the primary end point. The 3 panels show the relationship with increasing numbers of HF hospitalizations (Figure 3, panel A); the collective effect of the 5 imaging variables stratified by HF status and controlling for IL-6 level (Figure 3, panel B); and the effect of IL-6, conditional on the rest of the variables (Figure 3, panel C). The HF dependence is strongest with any interval HF hospitalization (≥ 1) being more associated with risk of the end point with an increase of roughly 10 events per 100 person-years in absolute event risk (range of 4 to 14% or 3.5-fold). HF hospitalizations beyond the first one did not further add significantly to the prediction model. The risk association of imaging variables is strongest among people without interval HF hospitalizations. Together, the imaging variables explain roughly a 3-fold risk of the primary end point between the range of 2-6 events per 100 person-years. There is a relatively weak relationship with IL-6.

DISCUSSION

Our principal finding is that a HF hospitalization significantly predicts subsequent life-threatening ventricular arrhythmia following ICD insertion. Larger

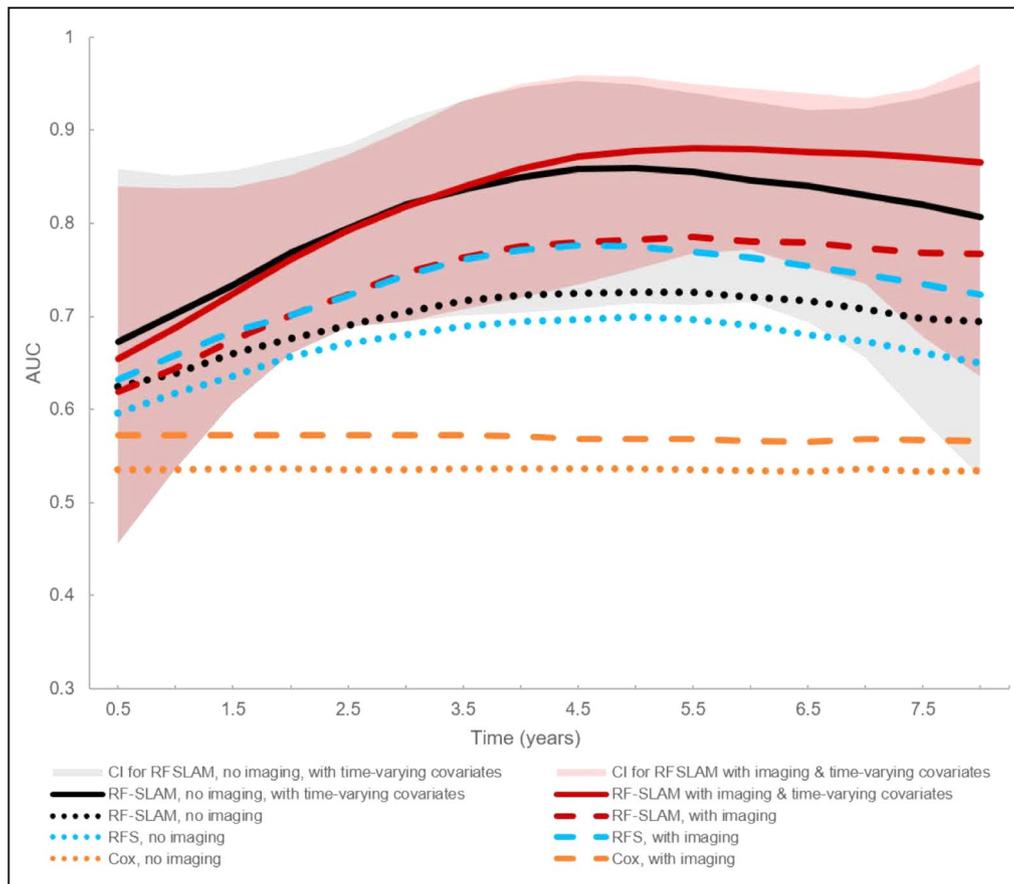


Figure 1. Median area under the curve (AUC) performances for predicting the primary end point for each of the 8 models.

Models incorporating only baseline covariates are shown as dotted or dashed lines. The 95% CIs for the AUCs over time for random forest statistical method for survival, longitudinal, and multivariable outcomes (RF-SLAM) incorporating time-varying covariates with (pink shaded area) and without (gray shaded area) imaging are also shown. RF-SLAM with both imaging and time-varying covariates (dark red solid line) had the highest AUC. RFS indicates random forest survival method.

CMR-derived LV and LA volumes, larger total scar and gray zone extents, and lower LA emptying fractions and serum IL-6 concentrations were the top baseline variables and contributed to risk prediction particularly among patients without HF. When clinical HF events and baseline CMR metrics and IL-6 levels are already included as covariates, serial LVEFs did not add significantly to the prediction model and no LVEF threshold could be identified above which risk is reduced. Identifying temporal changes to covariates as contributors to ventricular arrhythmic prediction is novel and pathophysiologically relevant because of the highly variable progression of cardiac disease and risk factors. Our findings may apply to risk assessment at the time of ICD generator end of life to inform decision-making regarding the risk/benefit of replacement and warrant further investigation.

An interim HF hospitalization²⁶ has been identified previously as a risk factor for subsequent appropriate ICD shocks. In the ICD arm of the MADIT-II (Multicenter

Automatic Defibrillator Intervention Trial II), interim HF hospitalization was associated with a 2.5-fold increased hazard ratio for subsequent appropriate ICD therapy.²⁶ We also found a marked 3.5-fold increase in relative risk for a HF hospitalization with subsequent ICD firing generally occurring not immediately afterwards but on average 2.4±1.7 years later. This suggests that clinical HF decompensation is a marker for more chronic adverse mechanical and/or electrical remodeling that is proarrhythmic. Our results further highlight and emphasize the importance of clinical HF instability rather than trajectories in LVEF in predicting increased propensity for arrhythmias.

Including baseline CMR variables improved the predictions for all statistical models compared with baseline variables without imaging. Although an HF hospitalization significantly improved RF-SLAM performance regardless of inclusion of the imaging variables, SCD risk remains elevated among those without HF. Among patients with a primary event, 45% did

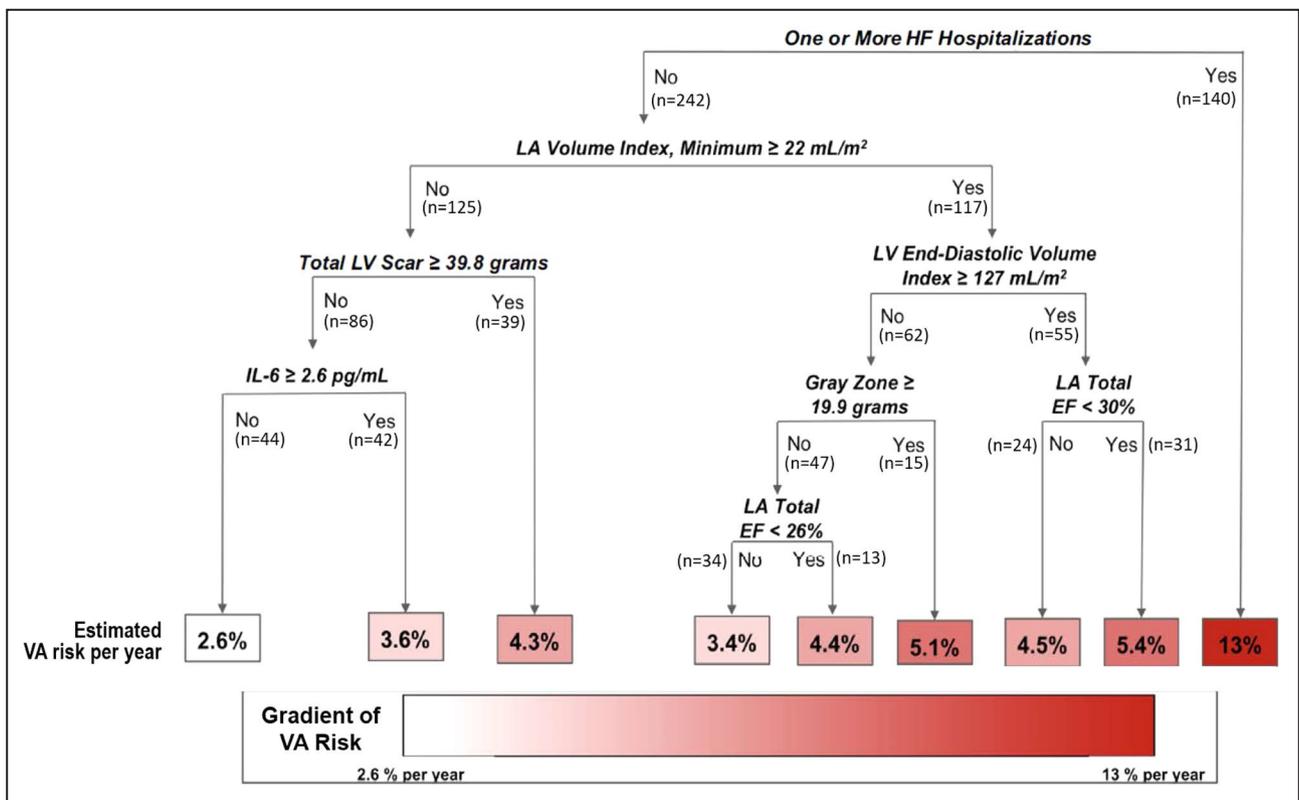


Figure 2. Summary tree of RF-SLAM depicting the top 7 predictors for the primary end point at 5 years of follow-up that accounted for > 95% of the prediction.

Decision rules at each tree node are shown in bold italics and the number of cohort patients meeting criteria at each node is noted. The annual predicted ventricular arrhythmic (VA) risk is shown at the bottom of the decision tree. The VA risk boxes are color coded according to the magnitude of the annual risk, with white corresponding to the lowest risk subgroup and dark red corresponding to the highest risk subgroup. EF indicates ejection fraction; HF, heart failure; IL, interleukin; LA, left atrium; LV, left ventricle; and RF-SLAM, random forest statistical method for survival, longitudinal, and multivariable outcomes.

not have a preceding HF hospitalization, highlighting the need for multiple highly performing risk metrics. The variable dependence plots for the imaging metrics demonstrate their additive predictive value among those without HF hospitalizations. Hence, our findings reinforce and quantify the incremental significance of the identified imaging covariates, supplementing published results. Scar and gray zone extents are important underlying pathophysiologic substrates supporting the initiation and propagation of reentrant ventricular arrhythmias.^{5,27} Severe LV chamber dilation measured by echocardiography predicted appropriate ICD shocks in a mixed ischemic and nonischemic cardiomyopathy HF cohort²⁸ and among SCD cases from a community-based study.²⁹ Chronic LV dilation resulting from adverse LV remodeling is associated with changes in cardiac ion channels and the distribution and expression of connexin proteins that may be proarrhythmic.³⁰ Here, we found that elevated indexed LV diastolic volume was an informative predictor with the advantage over prior studies of high reproducibility and decreased variability associated with CMR measurements of LV size.

The predictive importance of atrial remodeling as a marker of the cumulative, long-term effects of elevated LV filling pressures and LV wall stress has been highlighted in studies of patients with HF and reduced LVEF. Smaller, retrospective analyses using CMR^{31,32} and echocardiography (the MUSIC Study [Muerte Súbita en Insuficiencia Cardíaca])³³ suggest that lower LA emptying fractions and larger indexed LA size independently predict SCD. Our findings support the relevance of atrial abnormalities as markers of chronic diastolic dysfunction for predicting ventricular arrhythmia even among low LVEF patients. Although NTproBNP, which we measured only at baseline, was not among the top variables of importance after imaging variables were incorporated, larger LA minimum volumes were associated with higher baseline BNP levels ($P < 0.01$). LA structural indices likely better reflect chronic and cumulative elevations of LV pressure rather than single BNP measures. Further investigation of mechanisms are warranted.

Prior studies of the role of inflammation in predicting ICD therapy have been inconsistent. Some studies reported associations between increased levels

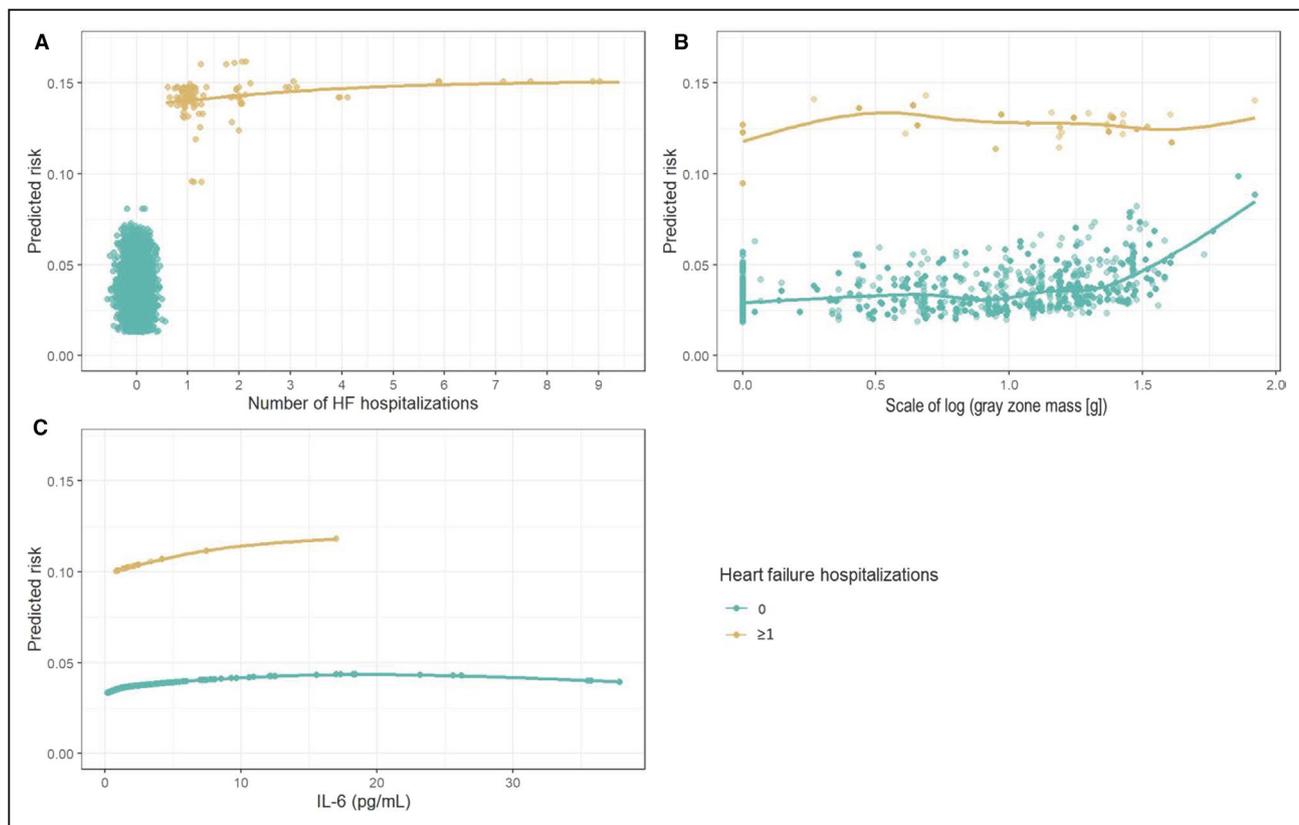


Figure 3. Variable dependence plots calculated from the RF-SLAM predicted values, stratified by interim HF hospitalization. (A) shows an individual's risk of the primary end point as a function of the number of interim HF hospitalizations. (B) shows the collective effect of all 5 imaging variables, stratified by HF status, holding IL-6 constant, and plotted against the scale of the imaging variable gray zone mass, selected because it best illustrates the collective effect of all imaging variables and reflects the largest range of effects. (C) is the risk attributable to IL-6, controlling for all of the other variables. The dependence plots can be used to ascertain a person's risk given his/her HF status along the gradient of imaging results (here gray zone mass) and ≥1 interval HF hospitalization or by IL-6 level and HF status. HF indicates heart failure; IL, interleukin; and RF-SLAM, random forest statistical method for survival, longitudinal, and multivariable outcomes.

of proinflammatory biomarkers and appropriate ICD shocks.^{15,34,35} Cytokines such as IL-6 may modulate expression and/or function of ion channels through direct cardiomyocyte effects to promote arrhythmogenesis.³⁴ Although we found a very small contribution of IL-6 level accounting for only 2% of the variance in prediction, it was a top predictor combined with HF hospitalizations and CMR measures. Prior statistical approaches that suboptimally controlled for variable interactions may explain the equivocal associations between inflammatory marker levels and SCD.

Published studies of primary prevention ICD recipients with and without cardiac resynchronization therapy suggest that although an improved LVEF reduces the relative risk of SCD outcomes, absolute risk remains elevated at 3-5% per year.^{36,37} Our results further build upon prior findings by suggesting the lack of significant incremental value of serial LVEFs when HF hospitalization and CMR indices are already considered and reinforce the imperfect predictive capability of LVEF for SCD, particularly at the individual patient

level.^{1,5} Although we did not impose any threshold values to define LVEF recovery, RF-SLAM methodology is able to detect a meaningful threshold, if present.

Risk scores such as SHFM-D and SPRM^{17,18} use simple demographic and clinical characteristics. Although most effective at stratifying patients based on low probability of benefit from ICDs but who are at increased risk for nonarrhythmic mortality,³⁸ they have limited discrimination of differences in appropriate ICD therapy.³⁹ Although such risk models reduce the number of eligible ICD candidates, they fail to distinguish those at lowest risk for ventricular arrhythmia in whom an ICD might be safely deferred that comprise >25% of potential candidates. This observation is relevant because, as the rates of both total and arrhythmic mortality decrease with advances in atherosclerotic disease and HF treatment and concerns about rising health costs intensify, it will be crucial to balance the risk/benefits of primary prevention ICDs by targeting those with the highest risk of ventricular arrhythmia most likely to

benefit paired with an acceptably low risk of noncardiac mortality.

Finally, our approach demonstrates how machine learning methods can improve ventricular arrhythmic risk prediction and interpretation. Compared to traditional regression methods, advantages of RF-SLAM include easy accommodation of time-varying predictors such as HF; missing data common to clinical research and practice; key predictor selection from a potentially large number, most of which are unimportant; and nonlinear effects including interactions among predictors. With a time-to-event outcome, the most important potential interaction comprises the follow-up time duration. In many chronic diseases, the major predictors change over the course of the process. For example, we have shown how dramatically the occurrence of a HF hospitalization increases the risk of a primary event. RF-SLAM also emphasizes depiction of time-varying risk performance (AUC) on an interval-to-interval (eg, biannual clinic visit) basis rather than aggregately at 1 time point. This may better inform clinical decision-making when the quality of predictions changes throughout the disease process.

Visualization as a method for explaining RF-SLAM predictions is increasingly important in clinical applications because black-box predictions lacking explanation raise concerns for clinical translation. With a single classification tree summarizing nearly all of the information in the RF predictions, the user can better understand the most important predictors that comprise the tree and how they interact with one another in a more clinically intuitive manner. The tree methods utilized here can be implemented on a smart phone or other simple calculating device. A second visualization approach with clinical utility is the variable dependence plot, scaled to functional dependence of the relationship of the absolute risk of the primary end point on each predictor or set of predictors of interest, controlling for the other values by computationally assigning them their average values. In this case, stratifying by interval HF hospitalization makes it easier to see the association of risk with the imaging variables and IL-6 within the subgroup without HF. Finally, because our algorithm is built using the open-source R-environment, it is readily adaptable to incorporate new data and for applications to other areas.

We recognize our study's limitations. Our study was by design observational with a long enrollment phase and ICD programming parameters were not prescriptive. There was a slight decline over time in appropriate ICD therapies but a greater decline in inappropriate therapies (Table S4) likely influenced by published studies triggering changes in device programming. The relatively small number of primary events limits the

power to detect small deviations in risk predictions. Our results require independent validation using a larger number of arrhythmic events from other cohorts, particularly to refine and improve the identification of sufficiently low-risk patients in whom an ICD could be deferred. However, the ability to identify extremes and gradations of arrhythmic risk as we do here may enhance the design of future randomized controlled trials by identifying specific risk strata. Although we did not incorporate serial CMR-LGE scans to track temporal changes in cardiac structure and tissue characterization, an ongoing study addresses this (NCT00733590, clinicaltrials.gov). Serial LVEFs were not mandated, similar to other studies that investigated the prognostic value of recovered LVEF on SCD outcomes but did not account for CMR covariates as we do here.^{37,40} The clinically obtained LVEFs were not centrally interpreted. The LVEF findings require further confirmation. Similarly, future investigations could focus on the incremental predictive value of serial measurements of other variables, which is beyond the scope here. We did not investigate other statistical machine learning approaches. Optimizing visualization approaches to explaining machine learning-based model results remain a work in-progress. Genetic information was not available.

In conclusion, our proof-of-concept study highlights the potential to improve individualized risk assessment for ventricular arrhythmias by incorporating both a temporally varying risk factor and baseline covariates. Specifically, a set of 7 predictors consisting of HF hospitalization, CMR variables, and IL-6 levels parsimoniously accounted for the vast majority of subsequent appropriate ICD shocks. Our results support the importance of the complex interplay of pathophysiologically driven markers of underlying myocardial substrate, temporal changes in clinical HF status, and systemic inflammation in identifying increased risk for ventricular arrhythmias.

ARTICLE INFORMATION

Received April 10, 2020; accepted August 11, 2020.

Affiliations

From the Department of Medicine, Division of Cardiology, Johns Hopkins University School of Medicine, Baltimore, MD (K.C.W., S.T., H.A., J.E.M., R.G.W.); Department of Biomedical Engineering and School of Medicine, Johns Hopkins University, Baltimore, MD (S.W., H.A.); The Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, Baltimore, MD (R.G.W.); Christiana Care Health Systems Inc, Newark, DE (M.S.); Department of Medicine, University of Maryland Medical Systems, Baltimore, MD (T.M.D.); Albert Einstein College of Medicine and Montefiore Medicine, Bronx, NY (G.F.T.); and Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD (S.L.Z.).

Acknowledgments

We thank Ms. Emily Scott, MS, for her expert assistance with Figure 3 computations and graphical presentation.

Sources of Funding

This work was supported by the National Institutes of Health (R01HL103812 to KCW; R01HL132181 to GFT and KCW; 5T32GM007309 and F30HL142131 to SW).

Disclosures

None.

Supplementary Materials

Data S1

Tables S1–S4

References 41–46

REFERENCES

- Myerburg RJ, Goldberger JJ. Sudden cardiac arrest risk assessment: Population science and the individual risk mandate. *JAMA Cardiol.* 2017;2:689–694.
- Zeitler EP, Al-Khatib SM, Friedman DJ, Han JY, Poole JE, Bardy GH, Bigger JT, Buxton AE, Moss AJ, Lee KL, et al. Predicting appropriate shocks in patients with heart failure: patient level meta-analysis from SCD-HeFT and MADIT II. *J Cardiovasc Electrophysiol.* 2017;28:1345–1351.
- Greenlee RT, Go AS, Peterson PN, Cassidy-Bushrow AE, Gaber C, Garcia-Montilla R, Glenn KA, Gupta N, Gurwitz JH, Hammill SC, et al. Device therapies among patients receiving primary prevention implantable cardioverter-defibrillators in the Cardiovascular research network. *J Am Heart Assoc.* 2018;7:e008292. DOI: 10.1161/JAHA.117.008292.
- Sabbag A, Suleiman M, Laish-Farkash A, Samania N, Kazatsker M, Goldenberg I, Glikson M, Beinart R. Contemporary rates of appropriate shock therapy in patients who receive implantable device therapy in a real-world setting: From the Israeli ICD Registry. *Heart Rhythm.* 2015;12:2426–2433.
- Wu KC. Sudden cardiac death substrate imaged by magnetic resonance imaging: From investigational tool to clinical applications. *Circ Cardiovasc Imaging.* 2017;10:e005461.
- Aliot EM, Stevenson WG, Almendral-Garrote JM, Bogun F, Calkins CH, Delacretaz E, Della Bella P, Hindricks G, Jais P, Josephson ME, et al. EHRA/HRS expert consensus on catheter ablation of ventricular arrhythmias: developed in a partnership with the European Heart Rhythm Association (EHRA), a Registered Branch of the European Society of Cardiology (ESC), and the Heart Rhythm Society (HRS); in collaboration with the American College of Cardiology (ACC) and the American Heart Association (AHA). *Heart Rhythm.* 2009;6:886–933.
- Roberts WC, Siegel RJ, McManus BM. Idiopathic dilated cardiomyopathy: analysis of 152 necropsy patients. *Am J Cardiol.* 1987;60:1340–1355.
- de Bakker JM, Stein M, van Rijen HV. Three-dimensional anatomic structure as substrate for ventricular tachycardia/ventricular fibrillation. *Heart Rhythm.* 2005;2:777–779.
- Peters NS, Wit AL. Myocardial architecture and ventricular arrhythmogenesis. *Circulation.* 1998;97:1746–1754.
- Wilcox JE, Fonarow GC, Ardehali H, Bonow RO, Butler J, Sauer AJ, Epstein SE, Khan SS, Kim RJ, Sabbah HN, et al. "Targeting the heart" in heart failure: myocardial recovery in heart failure with reduced ejection fraction. *JACC. Heart Failure.* 2015;3:661–669.
- Nijst P, Martens P, Mullens W. Heart failure with myocardial recovery - the patient whose heart failure has improved: what next? *Prog Cardiovasc Dis.* 2017;60:226–236.
- Schmidt A, Azevedo CF, Cheng A, Gupta SN, Bluemke DA, Foo TK, Gerstenblith G, Weiss RG, Marban E, Tomaselli GF, et al. Infarct tissue heterogeneity by magnetic resonance imaging identifies enhanced cardiac arrhythmia susceptibility in patients with left ventricular dysfunction. *Circulation.* 2007;115:2006–2014.
- Wu KC, Gerstenblith G, Guallar E, Marine JE, Dalal D, Cheng A, Marban E, Lima JA, Tomaselli GF, Weiss RG. Combined cardiac magnetic resonance imaging and C-reactive protein levels identify a cohort at low risk for defibrillator firings and death. *Circ Cardiovasc Imaging.* 2012;5:178–186.
- Zhang Y, Guallar E, Weiss RG, Stillabower M, Gerstenblith G, Tomaselli GF, Wu KC. Associations between scar characteristics by cardiac magnetic resonance and changes in left ventricular ejection fraction in primary prevention defibrillator recipients. *Heart Rhythm.* 2016;13:1661–1666.
- Cheng A, Zhang Y, Blasco-Colmenares E, Dalal D, Butcher B, Norgard S, Eldadah Z, Ellenbogen KA, Dickfeld T, Spragg DD, et al. Protein biomarkers identify patients unlikely to benefit from primary prevention implantable cardioverter defibrillators: findings from the Prospective Observational Study of Implantable Cardioverter Defibrillators (PROSE-ICD). *Circ Arrhythm Electrophysiol.* 2014;7:1084–1091.
- Tao S, Ashikaga H, Ciuffo LA, Yoneyama K, Lima JAC, Frank TF, Weiss RG, Tomaselli GF, Wu KC. Impaired left atrial function predicts inappropriate shocks in primary prevention implantable cardioverter-defibrillator candidates. *J Cardiovasc Electrophysiol.* 2017;28:796–805.
- Levy WC, Lee KL, Hellkamp AS, Poole JE, Mozaffarian D, Linker DT, Maggioni AP, Anand I, Poole-Wilson PA, Fishbein DP, et al. Maximizing survival benefit with primary prevention implantable cardioverter-defibrillator therapy in a heart failure population. *Circulation.* 2009;120:835–842.
- Shadman R, Poole JE, Dardas TF, Mozaffarian D, Cleland JG, Swedberg K, Maggioni AP, Anand IS, Carson PE, Miller AB, et al. A novel method to predict the proportional risk of sudden cardiac death in heart failure: derivation of the Seattle Proportional Risk Model. *Heart Rhythm.* 2015;12:2069–2077.
- Ishwaran H, Kogalur UB. Random forests for survival, regression, and classification (RF-SRC), R-package version 2.5.0. 2017.
- Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *Ann. Appl. Stat.* 2008;2:841–860.
- Wongvibulsin S, Wu KC, Zeger SL. Clinical risk prediction with random forests for survival, longitudinal, and multivariate (RF-SLAM) data analysis. *BMC Med Res Methodol.* 2019;20:1.
- Wongvibulsin S, Wu KC, Zeger SL. Improving clinical translation of machine learning approaches through clinician-tailored visual displays of "black box" algorithms. *JMIR Medical Informatics.* 2020;8:e15791.
- Breiman L. Random Forests. *Mach Learn.* 2001;45:5–32.
- Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol.* 1982;115:92–106.
- Bansal A, Heagerty PJ. A tutorial on evaluating the time-varying discrimination accuracy of survival models used in dynamic decision making. *Medical Decis Making.* 2018;38:904–916.
- Singh JP, Hall WJ, McNitt S, Wang H, Daubert JP, Zareba W, Ruskin JN, Moss AJ, Investigators M-I. Factors influencing appropriate firing of the implanted defibrillator for ventricular tachycardia/fibrillation: findings from the Multicenter Automatic Defibrillator Implantation Trial II (MADIT-II). *J Am Coll Cardiol.* 2005;46:1712–1720.
- Disertori M, Rigoni M, Pace N, Casolo G, Mase M, Gonzini L, Lucci D, Nollo G, Ravelli F. Myocardial fibrosis assessment by Ige is a powerful predictor of ventricular tachyarrhythmias in ischemic and non-ischemic lv dysfunction: A meta-analysis. *JACC Cardiovasc Imaging.* 2016;9:1046–1055.
- Aleong RG, Mulvahill MJ, Halder I, Carlson NE, Singh M, Bloom HL, Dudley SC, Ellinor PT, Shalaby A, Weiss R, et al. Left ventricular dilatation increases the risk of ventricular arrhythmias in patients with reduced systolic function. *J Am Heart Assoc.* 2015;4: 10.1161/JAHA.114.001566.
- Narayanan K, Reinier K, Teodorescu C, Uy-Evanado A, Aleong R, Chugh H, Nichols GA, Gunson K, London B, Jui J, et al. Left ventricular diameter and risk stratification for sudden cardiac death. *J Am Heart Assoc.* 2014;3: 10.1161/JAHA.114.001193.
- Aiba T, Tomaselli GF. Electrical remodeling in the failing heart. *Curr Opin Cardiol.* 2010;25:29–36.
- Lydell CP, Mikami Y, Homer K, Peng M, Cornhill A, Rajagopalan A, Arasaratnam P, Cowan K, Roberts A, Sumner C, et al. Left atrial function using cardiovascular magnetic resonance imaging independently predicts life-threatening arrhythmias in patients referred to receive a primary prevention implantable cardioverter defibrillator. *Can J Cardiol.* 2019;35:1149–1157.
- Rijnierse MT, Kamali Sadeghian M, Schuurmans Stekhoven S, Biesbroek PS, van der Lingen AC, van de Ven PM, van Rossum AC, Nijveldt R, Allaart CP. Usefulness of left atrial emptying fraction to predict ventricular arrhythmias in patients with implantable cardioverter defibrillators. *Am J Cardiol.* 2017;120:243–250.
- Bayes-Genis A, Vazquez R, Puig T, Fernandez-Palomeque C, Fabregat J, Bardaji A, Pascual-Figal D, Ordóñez-Llanos J, Valdes M, Gabarrus A, et al. Left atrial enlargement and NT-proBNP as predictors of

- sudden cardiac death in patients with heart failure. *Eur J Heart Fail*. 2007;9:802–807.
34. Lazzarini PE, Capecchi PL, El-Sherif N, Laghi-Pasini F, Boutjdir M. Emerging arrhythmic risk of autoimmune and inflammatory cardiac channelopathies. *J Am Heart Assoc*. 2018;7: 10.1161/JAHA.118.010595.
 35. Streitner F, Kuschyk J, Veltmann C, Brueckmann M, Streitner I, Brade J, Neumaier M, Bertsch T, Schumacher B, Borggreve M, et al. Prospective study of interleukin-6 and the risk of malignant ventricular tachyarrhythmia in ICD-recipients—a pilot study. *Cytokine*. 2007;40:30–34.
 36. Smer A, Saurav A, Azzouz MS, Salih M, Ayan M, Abuzaid A, Akinapelli A, Kanmanthareddy A, Rosenfeld LE, Merchant FM, et al. Meta-analysis of risk of ventricular arrhythmias after improvement in left ventricular ejection fraction during follow-up in patients with primary prevention implantable cardioverter defibrillators. *Am J Cardiol*. 2017;120:279–286.
 37. Adabag S, Patton KK, Buxton AE, Rector TS, Ensrud KE, Vakil K, Levy WC, Poole JE. Association of implantable cardioverter defibrillators with survival in patients with and without improved ejection fraction: Secondary analysis of the Sudden Cardiac Death in Heart Failure Trial. *JAMA cardiology*. 2017;2:767–774.
 38. Buxton AE. Sudden death in ischemic heart disease - 2017. *Int J Cardiol*. 2017;237:64–66.
 39. van der Heijden AC, van Rees JB, Levy WC, van der Bom JG, Cannegieter SC, de Bie MK, van Erven L, Schalij MJ, Borleffs CJ. Application and comparison of the FADES, MADIT, and SHFM-D risk models for risk stratification of prophylactic implantable cardioverter-defibrillator treatment. *Europace*. 2017;19:72–80.
 40. Ruwald MH, Solomon SD, Foster E, Kutyla V, Ruwald AC, Sherazi S, McNitt S, Jons C, Moss AJ, Zareba W. Left ventricular ejection fraction normalization in cardiac resynchronization therapy and risk of ventricular arrhythmias and clinical outcomes: results from the Multicenter Automatic Defibrillator Implantation Trial With Cardiac Resynchronization Therapy (MADIT-CRT) trial. *Circulation*. 2014;130:2278–2286.
 41. Tao S, Ciuffo LA, Lima JAC, Wu KC, Ashikaga H. Quantifying left atrial structure and function using single-plane tissue-tracking cardiac magnetic resonance. *Magn Reson Imaging*. 2017;42:130–138.
 42. Ishwaran H, Kogalur UB. Random survival forests for r. *R News*. 2007;7:25–31.
 43. Ishwaran H, Kogalur UB. Fast unified random forests for survival, regression, and classification (rf-src). R-package version 2.9.3. 2020.
 44. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011;30:377–399.
 45. van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in r. *J Stat Softw*. 2011;45:1–67.
 46. Heagerty PJ, Zheng Y. Survival model predictive accuracy and roc curves. *Biometrics*. 2005;61:92–105.

SUPPLEMENTAL MATERIAL

Data S1.

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^{16, 41}

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) = $(V_{\max} - V_{\min}) \times 100\% / V_{\max}$; LA passive ef = $(V_{\max} - V_{\text{preA}}) \times 100\% / V_{\max}$; and LA active ef = $(V_{\text{preA}} - V_{\min}) \times 100\% / V_{\text{preA}}$.

RF-SLAM methodology^{21, 22}

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.^{20, 21, 42, 43} 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 plots⁴⁴

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 ≥ 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^{20, 42, 43}

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 methodology⁴⁵

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.^{22, 25, 46}

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).

	Cox regression	RFS	RFSLAM*
Covariates included in each model			
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

(median, IQR)			
---------------	--	--	--

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%