JACC
PUBLISHED BY ELSEVIER ON BEHALF OF THE AMERICAN
COLLEGE OF CARDIOLOGY FOUNDATION

VOL. 85, NO. 23, 2025

Predicting Heart Failure Outcomes Using Patient-Reported Health Status

Real-World Validation of the KCCQ-12

Bassim R. El-Sabawi, MD,[a] Mohammed A. Al-Garadi, PhD,[b] Bryan D. Steitz, PhD,[b] Amy M. Perkins, MS,[c] Allison B. McCoy, PhD,[b] Donald G. Sengstack, MS,[b] Robert A. Greevy, Jr, PhD,[c] Emily H. Burnell, BA,[d] Kelly H. Schlendorf, MD,[a] JoAnn Lindenfeld, MD,[a] Lynne W. Stevenson, MD,[a] Deepak K. Gupta, MD, MSCI,[a] John A. Spertus, MD, MPH,[e] Sean P. Collins, MD,[f,g] Michael E. Matheny, MD, MPH, MS,[b,g] Justin M. Bachmann, MD, MPH[a,g]

ABSTRACT

BACKGROUND The Kansas City Cardiomyopathy Questionnaire-12 (KCCQ-12), a patient-reported outcome measure for adults with heart failure, is associated with hospitalizations and mortality in clinical trials. Curated data sets from controlled trials differ substantially from pragmatic data collected from real-world settings, however, and few data exist on the KCCQ-12's predictive utility in clinical practice.

OBJECTIVES This study sought to evaluate the predictive utility of the KCCQ-12 for hospitalizations and mortality when administered during outpatient heart failure care.

METHODS We conducted a cohort study of patients assigned the KCCQ-12 in heart failure clinics from July 2019 through March 2024. The primary exposure was KCCQ-12 Overall Summary (KCCQ-OS) score. The primary outcomes were 90-day hospitalization and cumulative mortality. Multivariable-adjusted associations were assessed using logistic regression and Cox proportional hazards models. Gradient boosting (XGBoost) and random survival forest machine learning models were used to evaluate KCCQ-OS feature importance in predicting 90-day hospitalizations and cumulative mortality, respectively.

RESULTS Among 4,406 patients assigned the KCCQ-12, 2,888 (66%) completed at least 1 questionnaire. The median KCCQ-OS score was 59.4 (Q1-Q3: 35.4-81.8). Patients with KCCQ-OS scores <25 had higher adjusted risks of 90-day hospitalization (OR: 3.49; 95% CI: 2.50-4.90) and cumulative mortality (HR: 3.09; 95% CI: 2.29-4.17) compared with those with scores ≥75. The KCCQ-OS score was the most important feature for predicting 90-day hospitalizations in the XGBoost model (area under the receiver-operating characteristic curve: 0.760; 95% CI: 0.706-0.811) and the most important feature for predicting cumulative mortality in the random survival forest model (C-index 0.783; 95% CI: 0.742-0.824) compared with other clinical, demographic, and laboratory variables. KCCQ-12 noncompletion was independently associated with increased 90-day hospitalization (OR: 1.72; 95% CI: 1.46-2.02) and 1-year mortality (HR: 1.52; 95% CI: 1.25-1.84) after adjusting for all variables in the primary analysis.

CONCLUSIONS In outpatient heart failure care, lower KCCQ-OS scores were strongly associated with increased hospitalizations and mortality, with the greatest risk among patients with scores <25. Noncompletion of the KCCQ-12 was itself associated with worse outcomes. The KCCQ-OS score was the dominant predictor of 90-day hospitalizations and cumulative mortality in machine learning models, supporting the KCCQ-12 as a prognostic tool in routine clinical practice. (JACC. 2025;85:2253–2266) Published by Elsevier on behalf of the American College of Cardiology Foundation.

Listen to this manuscript's audio summary by Editor Emeritus Dr Valentin Fuster on www.jacc.org/journal/jacc.

From the [a]Division of Cardiovascular Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA; [b]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, USA; [c]Department of Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee, USA; [d]Office of Population Health, Vanderbilt University Medical Center, Nashville, Tennessee, USA; [e]University of Missouri-Kansas City's Healthcare Institute for Innovations in Quality and Saint Luke's Mid America Heart Institute, Kansas City, Missouri, USA; [f]Department of Emergency Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA; and the [g]Veterans Affairs Tennessee Valley Healthcare System, Nashville, Tennessee, USA.

ISSN 0735-1097/$36.00

https://doi.org/10.1016/j.jacc.2025.03.545

ABBREVIATIONS AND ACRONYMS

AUC = area under the receiver-operating characteristic curve

BNP = B-type natriuretic peptide

EHR = electronic health record

HFpEF = heart failure with preserved ejection fraction

HFrEF = heart failure with reduced ejection fraction

KCCQ-12 = Kansas City Cardiomyopathy Questionnaire-12

LVEF = left ventricular ejection fraction

NT-proBNP = N-terminal pro-B-type natriuretic peptide

PROM = patient-reported outcome measure

RSF = random survival forests

VUMC = Vanderbilt University Medical Center

Patient-reported outcome measures (PROMs) are frequently used in clinical trials and registries to evaluate functional status and quality of life.[1] Health care systems are increasingly incorporating PROMs into routine clinical practice,[2] where they serve as a means for improving clinician-patient communication regarding health status.[3] The Kansas City Cardiomyopathy Questionnaire-12 (KCCQ-12) is a 12-item questionnaire quantifying symptoms, physical function, social limitations, and quality of life in adults with heart failure.[4] The KCCQ-12 and its longer, 23-item version (KCCQ-23)[5] have been extensively validated in clinical trials,[6-13] demonstrating responsiveness to clinical change[14] and associations with subsequent hospitalizations and mortality.[15-17] Despite its widespread use in heart failure research, studies on the KCCQ-12's performance in predicting outcomes when administered as part of routine ambulatory care are sparse.

SEE PAGE 2267

Both the patient population and pragmatic data from clinical settings are very different from the controlled environments of clinical trials.[18,19] The evaluation of predictive models incorporating pragmatic data such as the KCCQ-12 collected during routine clinical practice benefits from a dual analytic framework[20] combining traditional regression models with machine learning approaches. Machine learning approaches excel at handling the complex, nonlinear relationships inherent in real-world clinical data, where measurements are asynchronous, missingness is common, and patient populations are heterogeneous.[20-22] Machine learning models also offer improved individual risk estimation as compared to pooled risk scores,[23-25] which is particularly important for clinical decision support tools in the electronic health record (EHR). These advantages come at the cost of interpretability, however, thus traditional regression analyses complement machine learning models by providing interpretable effect sizes that give clinicians context to predictive tools such as the KCCQ-12.

To address the important gap in knowledge regarding the KCCQ-12's real-world prognostic capabilities, we employed complementary regression-based and machine learning approaches to evaluate the KCCQ-12's performance in predicting hospitalizations and mortality when implemented as a standard assessment in outpatients with heart failure. This dual analytic strategy allowed us to quantify the KCCQ-12's association with outcomes through traditional effect measures while also assessing its relative importance among clinical variables in the complex environment of outpatient practice.

METHODS

METHODOLOGY STANDARDS. This pragmatic cohort study adhered to Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)[26] standards. Machine learning model development adhered to Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD)[27] standards. The study was reviewed by the Institutional Review Board of the Vanderbilt Human Subjects Research Protection Program and determined to be exempt from further review and continuing oversight.

SETTING AND DATA COLLECTION. Vanderbilt University Medical Center (VUMC) is a 1,190-bed academic health center in Nashville, Tennessee. The KCCQ-12 was implemented in VUMC Heart Failure Clinics beginning in 2019 as part of the Vanderbilt Patient-Reported Outcomes Measurement System,[2] an institutional effort to integrate PROMs into routine clinical practice. The KCCQ-12 was delivered through VUMC's Epic electronic health record (EHR, Epic Systems Corporation) 72 hours before each clinic visit via My Health at Vanderbilt,[28] the institution's MyChart patient portal (Epic Systems Corporation) (Supplemental Figure 1). Patients who had not completed the KCCQ-12 before the clinic visit were asked to complete it in clinic using electronic tablets. Scored data from the completed KCCQ-12 were available to clinicians in the EHR alongside laboratory studies, as smart phrases that could be imported into clinical notes, and as longitudinal trends (Supplemental Figure 2). Data were extracted from

The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the Author Center.

Manuscript received March 6, 2025; revised manuscript received March 28, 2025, accepted March 31, 2025.

the EHR using Structured Query Language (SQL) with Azure Data Studio (Microsoft).[29]

STUDY COHORT. The study cohort was composed of patients assigned the KCCQ-12 questionnaire during VUMC Heart Failure Clinic visits from July 2019 through March 2024. Because a minority of providers in the heart failure clinic also saw general cardiology patients, we used echocardiogram data as well as International Classification of Diseases, 10th Revision (ICD-10) claims codes to restrict the study cohort to patients with a heart failure diagnosis. We defined patients with heart failure with reduced ejection fraction (HFrEF) as those with left ventricular ejection fraction (LVEF) <50%. Patients with a most recent LVEF ≥50% and ICD-10 claims codes consistent with a heart failure diagnosis were classified as heart failure with preserved ejection fraction (HFpEF),[30] even though some may have had prior HFrEF with recovered LVEF through guideline-directed medical therapy. We excluded patients with heart transplants or ventricular assist devices. For patients completing the KCCQ-12 multiple times, the first completed KCCQ-12 was selected to maximize follow-up time.

OUTCOMES. The main outcomes were 90-day all-cause hospitalization and cumulative mortality over the entire length of follow-up, ascertained from the EHR. We evaluated hospitalizations within 90 days of KCCQ-12 completion as a binary outcome. For cumulative mortality, we calculated time to event using the number of days between KCCQ-12 completion and death with censoring applied at the last known follow-up date if the patient was still alive.

VARIABLES. Variables included in the analysis were selected a priori based on prior established associations with heart failure outcomes from a review of heart failure prediction models[17,31-33] to minimize the risk of overfitting with data-driven selection methods. The primary predictor variable was the KCCQ-12 Overall Summary (KCCQ-OS) score,[4] which aggregates 4 domains (symptom frequency, physical function, social limitations, and quality of life) as scores from 0 to 100, with higher scores indicating better health status. Covariates included demographic, clinical, and laboratory values with the value in closest temporal proximity to the KCCQ-12 completion time selected for the analysis. For laboratory studies, we queried a time window 180 days before and after completion of the KCCQ-12. Because the cohort had a combination of B-type natriuretic peptide (BNP) and N-terminal pro-B-type natriuretic peptide (NT-proBNP) levels, a *z*-score was created for a composite BNP/NT-proBNP variable by log-transforming and standardizing these values.[34] We characterized comorbidities with Elixhauser comorbidity groups by querying relevant ICD-10 codes for outpatient and inpatient visits over 1 year before KCCQ-12 completion.[35] We calculated a Social Vulnerability Index, which was designed by the Centers for Disease Control and Prevention[36] to evaluate sociodemographic advantage, using home address census tracts. We used the IterativeImputer function from the Python scikit-learn[37] package to address missing data with multiple imputation. Missingness for each variable is detailed in Supplemental Table 1.

STATISTICAL ANALYSIS. Patients assigned the KCCQ-12 were stratified by completion status, and baseline characteristics between these groups were compared using Wilcoxon rank sum tests for continuous variables and chi-square tests for categorical variables. We constructed multivariable-adjusted logistic and Cox regression models to assess the relationship between KCCQ-OS quartiles and the risk of 90-day hospitalization and cumulative mortality, respectively. Models were stratified by heart failure type (HFpEF and HFrEF) to explore associations between the KCCQ-OS score and outcomes across ejection fraction subgroups. Formal interaction testing assessed whether the relationship between KCCQ-OS scores and outcomes differed by heart failure type. This testing was performed with KCCQ-OS modeled both as categorical quartiles and as continuous variables with nonlinear splines, using likelihood ratio tests for statistical significance. We also evaluated the KCCQ-OS score as a continuous variable in logistic regression and Cox regression models incorporating natural cubic splines to account for potential nonlinear relationships with outcomes. Optimal degrees of freedom for these models were determined using the Akaike Information Criterion.[38] To enhance clinical interpretability, we calculated absolute risk estimates for both outcomes across the range of the KCCQ-OS score. For 90-day hospitalization, we predicted probabilities from the logistic regression models with 95% CIs. For mortality, we calculated 1-year mortality risks from the Cox proportional hazards models using the baseline hazard at 365 days. We generated absolute risk plots for all patients and separately for HFrEF and HFpEF subgroups.

We performed additional multivariable analyses to evaluate whether KCCQ-12 completion status independently predicted clinical outcomes. For these analyses, we included the full cohort of heart failure patients who were assigned the KCCQ-12 (both completers and noncompleters). We constructed logistic

regression models for 90-day hospitalization and Cox proportional hazards models for 1-year mortality, adjusting for the same demographic, clinical, and laboratory variables used in the primary analyses. To establish temporal consistency across both completers and noncompleters, we defined the index date for noncompleters as the check-in time for the clinic visit in which the KCCQ-12 was assigned, and for completers as the time of questionnaire completion.

Statistical analyses were conducted using R version 4.4.1 (R Foundation for Statistical Computing).[39]

MACHINE LEARNING MODELS. The focus of the study was on evaluating the real-world prognostic value of the KCCQ-OS score (as a continuous variable) rather than on extensive algorithm validation. We selected XGBoost,[40] a gradient boosting implementation, for predicting 90-day hospitalizations (a binary outcome) due to its ability to capture complex relationships between predictors and outcomes. The full data set was split into training (80%) and testing (20%) sets, stratified by outcome to maintain class distribution. XGBoost hyperparameters were optimized using 5-fold stratified cross-validation on the training set with a grid search over the following parameters: number of estimators (100, 200), maximum depth (3, 4, 5), learning rate (0.01, 0.1), and minimum child weight (1, 3). Model performance was assessed on the test set using area under the receiver-operating characteristic curve (AUC) for discrimination and Brier score for calibration, with 95% CIs calculated using 10,000 bootstrap resamples. SHapley Additive exPlanations (SHAP)[41] values were calculated to evaluate feature importance, representing the marginal contribution of each predictor to the model's output.

We used random survival forests (RSF) to predict cumulative mortality due to its ability to handle right-censored survival data and strong predictive performance in cardiovascular mortality risk prediction.[42,43] RSF is an ensemble method implemented in scikit-survival[44] that combines multiple decision trees, each trained on a bootstrap sample of the data, to improve predictive performance. The full data set was split into training (80%) and testing (20%) sets to ensure sufficient data for stable survival curve estimation. RSF hyperparameters were optimized using 5-fold cross-validation on the training set with a grid search over the following parameters: number of estimators (100, 200, 300), maximum depth (none, 5, 10), and minimum samples split (2, 5, 10). Model discrimination was assessed on the test set using Harrell's concordance index (C-index) with 95% CIs derived from 10,000 bootstrap resamples. We calculated an integrated Brier score across 10 time points to assess both discrimination and calibration. Feature importance was evaluated using permutation importance with 10 iterations, measuring the decrease in C-index when each feature's values were randomly shuffled.

All analyses were performed in Python 3.11.9[45] using scikit-learn 1.5.1,[37] scikit-survival 0.23.0,[44] and xgboost 2.1.1.[40]

RESULTS

The cohort derivation is displayed in Supplemental Figure 3. Baseline characteristics of the study cohort, stratified by KCCQ-12 completion status, are presented in **Table 1**. Of 4,406 unique patients assigned the KCCQ-12 during the study time frame, a total of 2,888 (66%) completed at least 1 questionnaire (**Central Illustration**, **Table 1**). As compared with patients who did not complete the KCCQ-12, patients who completed the questionnaire were younger (median age 65 vs 69 years; $P < 0.001$) and had a lower Social Vulnerability Index, indicating less socioeconomic disadvantage. Patients completing the questionnaire also had lower systolic blood pressure, natriuretic peptide, blood urea nitrogen, and creatinine levels.

Among the 2,888 patients who completed at least 1 KCCQ-12, 1,690 (59%) were classified as HFrEF and 1,198 (41%) as HFpEF (**Table 1**). The median KCCQ-OS score was 59.4 (35.4-81.8), indicating moderate impairment due to heart failure symptoms with a wide range of health status across the cohort. Patients with HFpEF had a lower median KCCQ-OS score compared with those with HFrEF (56.3 vs 61.5, respectively; $P < 0.01$). Interaction testing between heart failure type and KCCQ-OS scores revealed no significant interactions for either outcome, whether the KCCQ-OS score was modeled as categories (hospitalization: chi-square = 3.78, df = 3; $P = 0.286$; mortality: chi-square = 1.14, df = 3; $P = 0.768$) or with nonlinear splines (hospitalization: chi-square = 4.74, df = 2; $P = 0.094$; mortality: chi-square = 0.21, df = 2; $P = 0.899$). These findings indicate that the relationship between the KCCQ-OS score and clinical outcomes does not differ significantly between patients with HFrEF and HFpEF, supporting our combined analysis approach adjusting for ejection fraction as a continuous variable.

A total of 490 hospitalization events (17%) occurred within 90 days of KCCQ-12 completion, and 489 patients died (17%) over a median follow-up of 1.9 years (6,535 person-years of follow-up total). Regression analyses for 90-day hospitalization (**Table 2**) and mortality (**Table 3**) demonstrated that lower KCCQ-OS scores were strongly associated with worse outcomes, regardless of heart failure type

(HFrEF vs HFpEF). For 90-day hospitalization, patients with KCCQ-OS scores <25 had a markedly higher odds compared with those with scores ≥75 (OR: 3.49; 95% CI: 2.50-4.90; $P < 0.001$) with similar results for HFrEF and HFpEF subgroups (Table 2). For mortality, the lowest KCCQ-OS range (<25) predicted the highest risk compared with scores ≥75 (HR: 3.09; 95% CI: 2.29-4.17; $P < 0.001$) with similar results for HFrEF and HFpEF subgroups (Table 3).

Multivariable-adjusted absolute risk plots (Figure 1) demonstrated consistent relationships between KCCQ-OS scores and 90-day hospitalization as well as 1-year mortality. For 90-day hospitalization, risk increased steadily as KCCQ-OS scores decreased, with similar patterns observed in both HFrEF and HFpEF subgroups. Estimated 1-year mortality risk increased similarly with decreasing KCCQ-OS scores. To illustrate the clinical significance of these relationships, a patient with a KCCQ-OS score of 10 had a 22.4% risk of 90-day hospitalization and 6.8% risk of 1-year mortality as compared with 6.7% risk of 90-day hospitalization and 2.1% risk of 1-year mortality in a patient with a KCCQ-OS score of 90.

XGBoost gradient boosting models incorporating the KCCQ-12 demonstrated good predictive performance for 90-day hospitalizations, with an AUC of 0.760 (95% CI: 0.706-0.811) and Brier score of 0.129 (95% CI: 0.110-0.149). Similar predictive performance was seen in models stratified by heart failure type (Supplemental Table 2). SHAP analysis demonstrated that the KCCQ-OS score was the most important predictor of 90-day hospitalizations among all heart failure patients in the XGBoost model (Figure 2A). The KCCQ-OS score was the most important predictor in HFrEF patients and the second most important predictor in HFpEF patients, following hemoglobin (Supplemental Figure 4).

RSF models incorporating the KCCQ-OS score also demonstrated good performance for mortality risk, achieving a C-index of 0.783 (95% CI: 0.742-0.824) and integrated Brier score of 0.092 (95% CI: 0.079-0.107) in all patients. Model performance was maintained in analyses stratified by heart failure type (Supplemental Table 3). Permutation importance analysis demonstrated that the KCCQ-OS score was the most important predictor of mortality in the RSF model for all heart failure patients (Figure 2B). In models stratified by heart failure type, the KCCQ-OS score was the third most important predictor in HFrEF patients (after age and blood urea nitrogen) and the second most important predictor in HFpEF patients (after red cell distribution width-coefficient of variation) (Supplemental Figure 5).

TABLE 1 Baseline Characteristics of Patients by Kansas City Cardiomyopathy Questionnaire-12 Completion Status (N = 4,406)

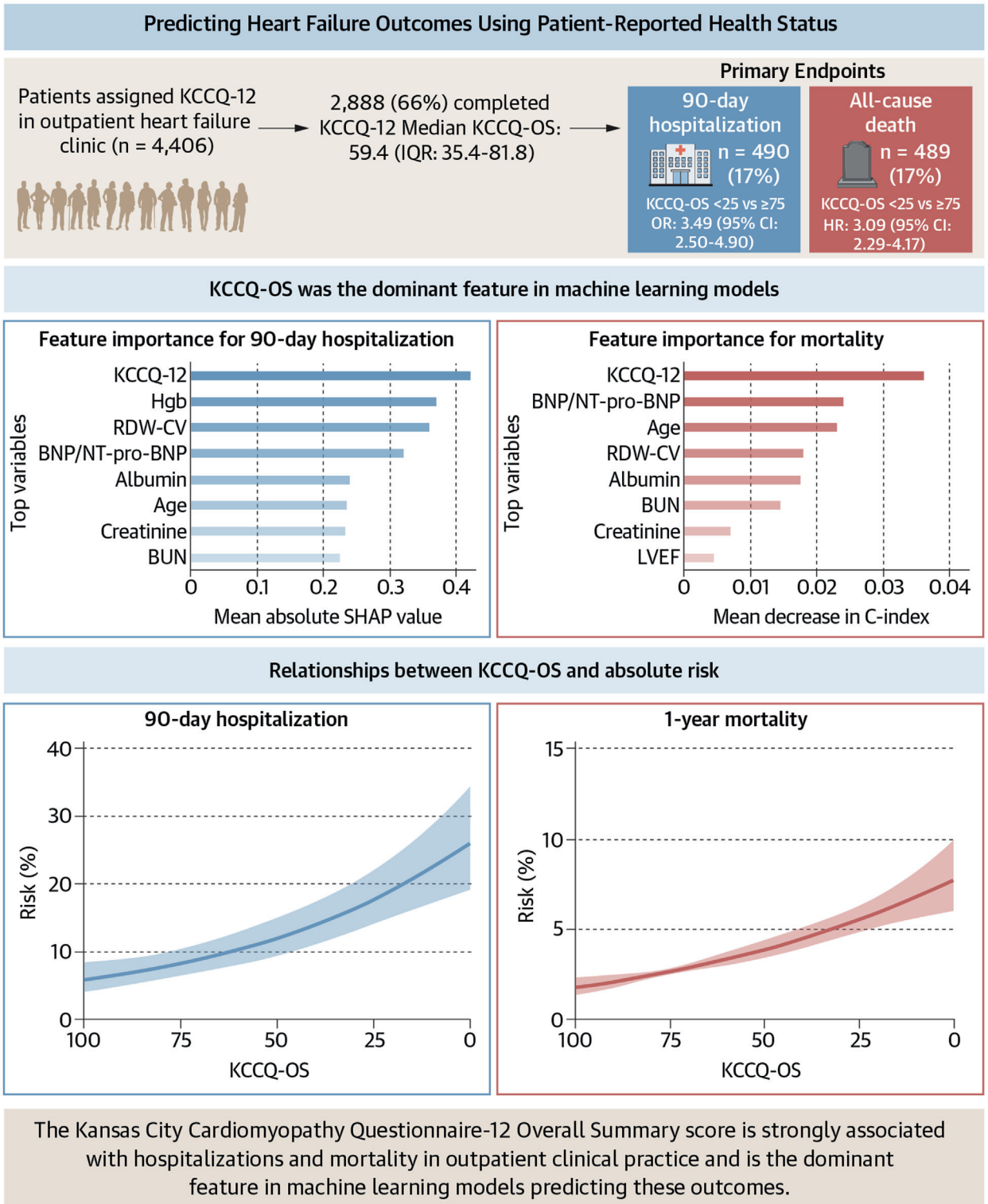
	Completed (n = 2,888)	Not Completed (n = 1,518)	*P* Value
Demographics			
Age, y	65 (54-73)	69 (59-79)	<0.001
Sex			
Female	1,203 (41.7)	630 (41.5)	0.95
Male	1,685 (58.3)	888 (58.5)	
Race			
Black	572 (19.8)	329 (21.7)	0.11
Other	65 (2.3)	24 (1.6)	
White	2,203 (76.3)	1,125 (74.1)	
Social Vulnerability Index	0.4 (0.2-0.7)	0.5 (0.3-0.8)	<0.001
KCCQ-12 scores			
KCCQ-12 Overall Summary score	59.4 (35.4-81.8)	n/a	n/a
KCCQ-12 Physical Limitation score	50 (33.3-83.3)	n/a	n/a
KCCQ-12 Symptom Frequency score	70.8 (44.4-91.7)	n/a	n/a
KCCQ-12 Quality of Life score	50 (25-75)	n/a	n/a
KCCQ-12 Social Limitation score	58.3 (33.3-87.5)	n/a	n/a
Clinical variables			
Heart failure type			
HFpEF	1,198 (41.5)	654 (43.1)	0.32
HFrEF	1,690 (58.5)	864 (56.9)	
LVEF, HFpEF, %	60 (55-65)	60.8 (55.4-67)	<0.001
LVEF, HFrEF, %	33.8 (24.5-42.3)	33 (23.8-42.3)	0.72
Systolic blood pressure, mm Hg	120 (108-134)	123 (110-139)	<0.001
Body mass index, kg/m^2	29.7 (25.7-35.3)	28.4 (24.4-33.5)	<0.001
Tobacco use			
Never	1,364 (47.2)	644 (42.4)	0.002
Quit	1,121 (38.8)	576 (37.9)	
Yes	261 (9)	179 (11.8)	
Diabetes	493 (17.1)	278 (18.3)	0.32
Hospitalized in year before completing KCCQ-12	1,194 (41.3)	n/a	n/a
Laboratory studies			
Albumin, g/dL	4.2 (4-4.4)	4.1 (3.8-4.3)	<0.001
Blood urea nitrogen, mg/dL	27 (18-47)	31 (19-51)	<0.001
Creatinine, mg/dL	1.4 (1.1-2.1)	1.5 (1.1-2.3)	0.013
Hemoglobin, g/dL	13.4 (9.9-15)	12.7 (9.9-14.6)	<0.001
MCV, fL	94 (90-98)	95 (90-98)	<0.001
RDW-CV, %	15 (13.7-17.3)	15.4 (14-17.7)	<0.001
BNP/NT-proBNP composite, *z*-score	0 (-0.8-0.7)	0.2 (-0.6-0.9)	<0.001
BNP, pg/mL	286 (81-763)	385 (97-840)	<0.001
NT-proBNP, pg/mL	789 (345-2,597)	989 (541-4,083)	<0.001
Events			
Median follow-up time, y	1.9 (0.8-3.6)	1.5 (0.7-2.3)	<0.001
90-d hospitalization	490 (17)	431 (28.4)	<0.001
1-y mortality	214 (7.4)	224 (14.8)	<0.001
Cumulative mortality	489 (16.9)	333 (21.9)	<0.001

Values are median (Q1-Q3) or n (%).

BNP = B-type natriuretic peptide; MCV = mean corpuscular volume; NT-proBNP = N-terminal pro-B-type natriuretic peptide; RDW-CV = red cell distribution width-coefficient of variation.

Patients who did not complete the KCCQ-12 had significantly higher rates of adverse events compared with those who completed the questionnaire (90-day hospitalization: 28.4% vs 17.0%; $P < 0.001$ and 1-year

CENTRAL ILLUSTRATION **Predictive Utility of the Kansas City Cardiomyopathy Questionnaire in Outpatient Clinical Practice**



El-Sabawi BR, et al. JACC. 2025;85(23):2253-2266.

The Kansas City Cardiomyopathy Questionnaire-12 Overall Summary score is strongly associated with hospitalizations and mortality in outpatient clinical practice and is the dominant feature in machine learning models predicting these outcomes. BNP = B-type natriuretic peptide; BUN = blood urea nitrogen; Hgb = hemoglobin; KCCQ-OS = Kansas City Cardiomyopathy Questionnaire-Overall Summary; LVEF = left ventricular ejection fraction; NT-proBNP = N-terminal pro-B-type natriuretic peptide; RDW-CV = red cell distribution width-coefficient of variation; SHAP = SHapley Additive exPlanations.

TABLE 2 Association of the Kansas City Cardiomyopathy Questionnaire-12 (KCCQ-12) With 90-Day Hospitalization (N = 2,888)

	All Patients		HFrEF		HFpEF	
90-d Hospitalization	OR (95% CI)	*P* Value	OR (95% CI)	*P* Value	OR (95% CI)	*P* Value
KCCQ-12 scores						
KCCQ-OS score 50-74, fair to good	1.46 (1.06-2.02)	<0.05	1.86 (1.21-2.89)	<0.01	1.09 (0.66-1.80)	0.73
KCCQ-OS score 25-49, poor to fair	2.31 (1.70-3.15)	<0.001	2.88 (1.89-4.44)	<0.001	1.94 (1.23-3.10)	<0.01
KCCQ-OS score 0-24, very poor to poor	3.49 (2.50-4.90)	<0.001	3.69 (2.33-5.88)	<0.001	3.37 (2.03-5.64)	<0.001
Demographics						
Age, per 10-y increase	1.12 (1.02-1.22)	<0.05	1.05 (0.93-1.18)	0.42	1.24 (1.07-1.43)	<0.01
Female	0.88 (0.69-1.11)	0.27	1.17 (0.85-1.61)	0.33	0.62 (0.43-0.87)	<0.01
Race, Black	0.94 (0.70-1.27)	0.70	0.95 (0.64-1.39)	0.79	0.90 (0.55-1.43)	0.65
Race, other	0.78 (0.37-1.53)	0.50	0.65 (0.21-1.63)	0.40	1.14 (0.37-3.08)	0.81
Social Vulnerability Index, per 1-SD increase	1.05 (0.93-1.18)	0.43	1.02 (0.87-1.19)	0.81	1.09 (0.92-1.30)	0.32
Clinical variables						
Left ventricular ejection fraction, per 5% increase	1.04 (1.00-1.07)	<0.05	1.06 (0.98-1.13)	0.13	0.97 (0.86-1.10)	0.64
Body mass index, per 5 kg/m^2 increase	1.06 (0.98-1.14)	0.13	0.98 (0.88-1.08)	0.65	1.17 (1.05-1.31)	<0.01
Systolic blood pressure, per 10-mm Hg increase	1.02 (0.96-1.08)	0.53	1.04 (0.96-1.12)	0.30	1.00 (0.92-1.08)	0.99
Tobacco, quit	0.92 (0.73-1.15)	0.46	0.85 (0.62-1.16)	0.30	1.02 (0.73-1.44)	0.89
Tobacco, yes	1.04 (0.69-1.55)	0.84	0.98 (0.58-1.60)	0.92	1.27 (0.63-2.44)	0.49
Hospitalization in prior year	1.23 (0.97-1.55)	0.08	1.32 (0.96-1.81)	0.08	1.07 (0.75-1.53)	0.71
Diabetes	1.12 (0.85-1.47)	0.41	1.02 (0.69-1.48)	0.92	1.24 (0.83-1.84)	0.29
Laboratory studies						
Sodium, per 5-mEq/L increase	1.34 (1.12-1.61)	<0.01	1.39 (1.08-1.80)	<0.05	1.36 (1.05-1.78)	<0.05
Blood urea nitrogen, per 10-mg/dL increase	1.05 (0.99-1.11)	0.08	1.07 (0.99-1.15)	0.09	1.04 (0.95-1.13)	0.39
Creatinine, per 0.5-mg/dL increase	1.03 (0.99-1.07)	0.16	1.00 (0.95-1.06)	0.93	1.05 (0.99-1.11)	0.12
Hemoglobin, per 1-g/dL decrease	1.06 (1.01-1.12)	<0.05	1.02 (0.96-1.09)	0.46	1.13 (1.04-1.23)	<0.01
Mean corpuscular volume, per 5-fL increase	1.18 (1.08-1.30)	<0.001	1.23 (1.08-1.41)	<0.01	1.14 (0.99-1.32)	0.06
RDW-CV, per 1% increase	1.13 (1.09-1.18)	<0.001	1.14 (1.08-1.20)	<0.001	1.13 (1.06-1.20)	<0.001
Albumin, per 0.5-g/dL increase	1.15 (1.00-1.33)	<0.05	1.16 (0.96-1.41)	0.12	1.18 (0.95-1.47)	0.14
BNP/NT-proBNP composite, per 1-SD increase	1.30 (1.13-1.49)	<0.001	1.33 (1.11-1.61)	<0.01	1.34 (1.08-1.66)	<0.01

Abbreviations as in Table 1.

mortality: 14.8% vs 7.4%; $P < 0.001$) (Table 1). These differences persisted in multivariable-adjusted models, where KCCQ-12 noncompletion was associated with a higher odds of 90-day hospitalization (OR: 1.72; 95% CI: 1.46-2.02; $P < 0.001$) and increased risk of 1-year mortality (HR: 1.52; 95% CI: 1.25-1.84; $P < 0.001$) after adjusting for the variables in the primary analyses. These findings indicate that the difference in event rates between completers and noncompleters cannot be fully explained by the observed differences in the covariates.

DISCUSSION

Among patients with heart failure seen in the ambulatory setting, the risks of hospitalization within 90 days and mortality were both approximately 3-fold higher in those with KCCQ-OS scores <25 with similar findings in analyses stratified by heart failure type. Machine learning models incorporating the KCCQ-OS score demonstrated good prognostic performance for both 90-day hospitalizations and cumulative mortality. The KCCQ-OS score was the dominant predictor of both hospitalizations and mortality as compared with other variables in these models. Noncompletion of the KCCQ-12 was independently associated with a higher risk of hospitalizations and mortality. Collectively, these findings highlight the potential of the KCCQ-12 to identify high-risk heart failure outpatients who may benefit from closer monitoring and more intensive management.

Our findings align with previous studies that evaluated the prognostic value of the KCCQ in heart failure populations. One study found that KCCQ-23 Overall Summary scores <25 (indicating severe symptoms) were associated with increased hospitalizations and mortality among 505 HFrEF outpatients[16] and another found that the KCCQ-12 Overall Summary score was associated with increased mortality and readmissions among inpatients with heart failure,[46] though both studies used hospitalizations and mortality as a composite outcome. The magnitude of risk we observed (approximately 3-fold higher risk of both 90-day hospitalization and mortality for KCCQ-OS scores <25 vs ≥75) is comparable to gradients

TABLE 3 Association of the Kansas City Cardiomyopathy Questionnaire-12 (KCCQ-12) With Cumulative Mortality (N=2,888)

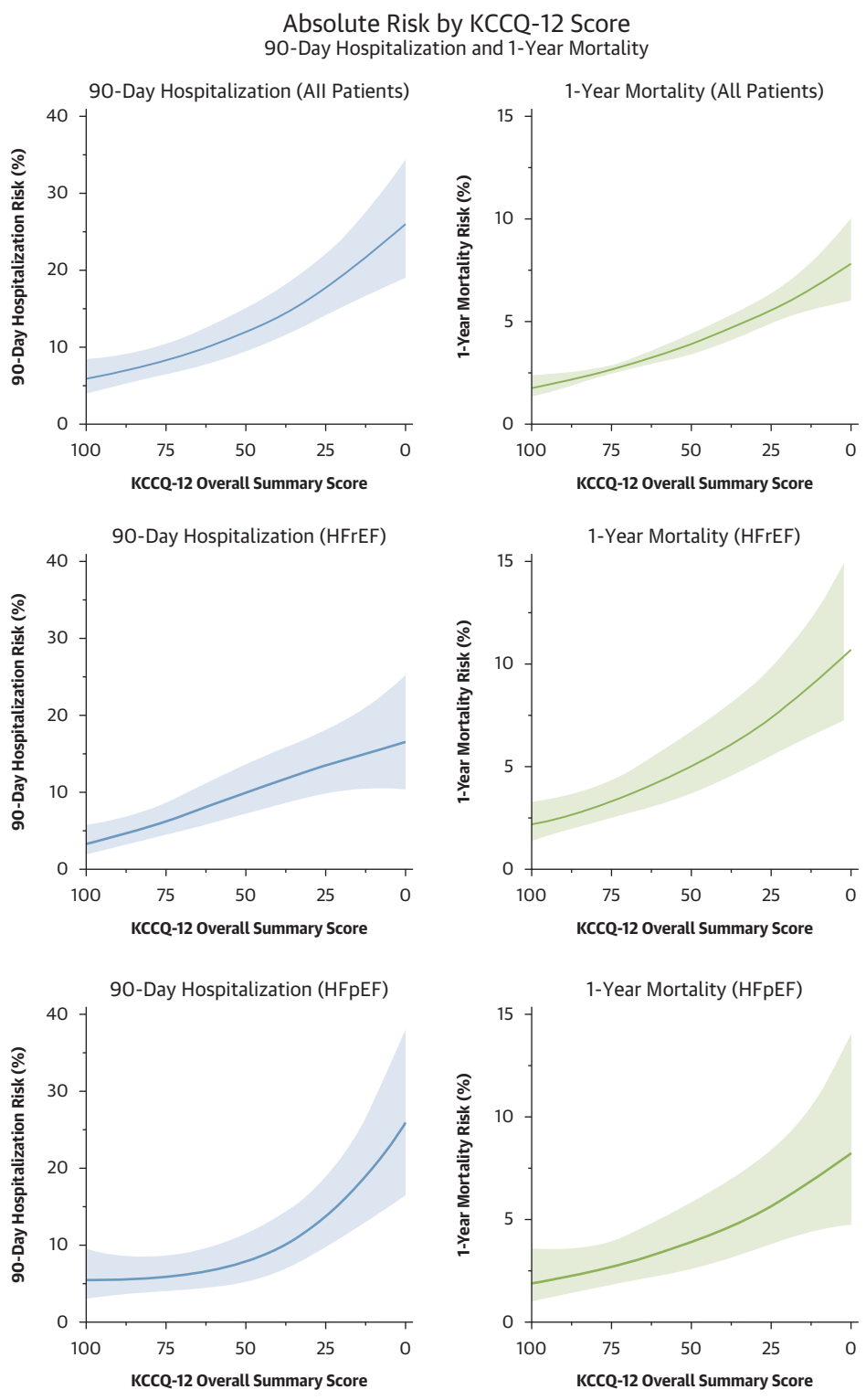
	All Patients		HFrEF		HFpEF	
Cumulative Mortality	HR (95% CI)	*P* Value	HR (95% CI)	*P* Value	HR (95% CI)	*P* Value
KCCQ-12 scores						
KCCQ-OS score 50-74, fair to good	1.69 (1.26-2.26)	<0.001	1.66 (1.15-2.37)	<0.01	1.68 (1.01-2.80)	<0.05
KCCQ-OS score 25-49, poor to fair	2.01 (1.50-2.69)	<0.001	2.28 (1.59-3.27)	<0.001	1.67 (1.00-2.79)	<0.05
KCCQ-OS score 0-24, very poor to poor	3.09 (2.29-4.17)	<0.001	3.47 (2.40-5.02)	<0.001	2.81 (1.66-4.77)	<0.001
Demographics						
Age, per 10-y increase	1.43 (1.32-1.56)	<0.001	1.51 (1.36-1.68)	<0.001	1.37 (1.18-1.58)	<0.001
Sex, female	1.23 (1.01-1.50)	<0.05	1.17 (0.91-1.51)	0.21	1.52 (1.08-2.14)	<0.05
Race, Black	1.20 (0.93-1.54)	0.15	1.17 (0.87-1.59)	0.30	1.22 (0.77-1.95)	0.40
Race, other	0.63 (0.33-1.18)	0.15	0.49 (0.21-1.10)	0.08	0.85 (0.29-2.47)	0.77
Social Vulnerability Index, per 1-SD increase	0.98 (0.89-1.08)	0.65	1.04 (0.93-1.18)	0.49	0.88 (0.75-1.04)	0.13
Clinical variables						
Left ventricular ejection fraction, per 5% increase	0.94 (0.92-0.97)	<0.001	0.97 (0.92-1.02)	0.27	0.90 (0.79-1.01)	0.08
Body mass index, per 5 kg/m^2 increase	0.98 (0.92-1.06)	0.67	1.05 (0.96-1.15)	0.30	0.93 (0.82-1.05)	0.24
Systolic blood pressure, per 10-mm Hg increase	0.97 (0.93-1.02)	0.26	0.97 (0.91-1.04)	0.39	0.97 (0.90-1.05)	0.50
Tobacco, quit	1.17 (0.96-1.42)	0.11	1.21 (0.94-1.55)	0.14	1.17 (0.84-1.63)	0.37
Tobacco, yes	1.40 (1.01-1.93)	<0.05	1.37 (0.92-2.02)	0.12	1.86 (1.00-3.47)	<0.05
Hospitalization in prior year	1.22 (1.00-1.48)	<0.05	1.24 (0.97-1.57)	0.08	1.05 (0.73-1.51)	0.78
Diabetes	1.17 (0.94-1.46)	0.17	1.05 (0.79-1.40)	0.74	1.42 (0.98-2.06)	0.06
Laboratory studies						
Sodium, per 5-mEq/L increase	0.87 (0.74-1.02)	0.09	0.74 (0.61-0.91)	<0.01	1.17 (0.90-1.53)	0.24
Blood urea nitrogen, per 10-mg/dL increase	1.06 (1.01-1.11)	<0.05	1.06 (1.00-1.13)	<0.05	1.04 (0.96-1.13)	0.35
Creatinine, per 0.5-mg/dL increase	1.02 (0.98-1.05)	0.30	1.01 (0.97-1.05)	0.67	1.03 (0.98-1.08)	0.29
Hemoglobin, per 1-g/dL decrease	1.00 (0.96-1.04)	0.98	0.98 (0.93-1.03)	0.51	1.05 (0.97-1.15)	0.25
Mean corpuscular volume, per 5-fL increase	1.13 (1.05-1.22)	<0.01	1.10 (1.00-1.21)	<0.05	1.20 (1.05-1.36)	<0.01
RDW-CV, per 1% increase	1.11 (1.07-1.14)	<0.001	1.10 (1.06-1.14)	<0.001	1.13 (1.07-1.19)	<0.001
Albumin, per 0.5-g/dL increase	0.76 (0.67-0.86)	<0.001	0.77 (0.66-0.91)	<0.01	0.74 (0.60-0.91)	<0.01
BNP/NT-proBNP composite, per 1-SD increase	1.34 (1.19-1.52)	<0.001	1.37 (1.16-1.61)	<0.001	1.38 (1.13-1.68)	<0.01

Abbreviations as in Table 1.

reported in prior work. For example, the KCCQ-12 validation study[4] found an approximate 3-fold difference for KCCQ-OS scores <25 compared with ≥75 in 6-month composite event rates for death and cardiovascular hospitalization among stable and acute HF recovery patients. Our analyses detail strong associations between the KCCQ-OS score and 90-day hospitalizations and mortality, each analyzed as a separate endpoint and across subgroups of heart failure type.

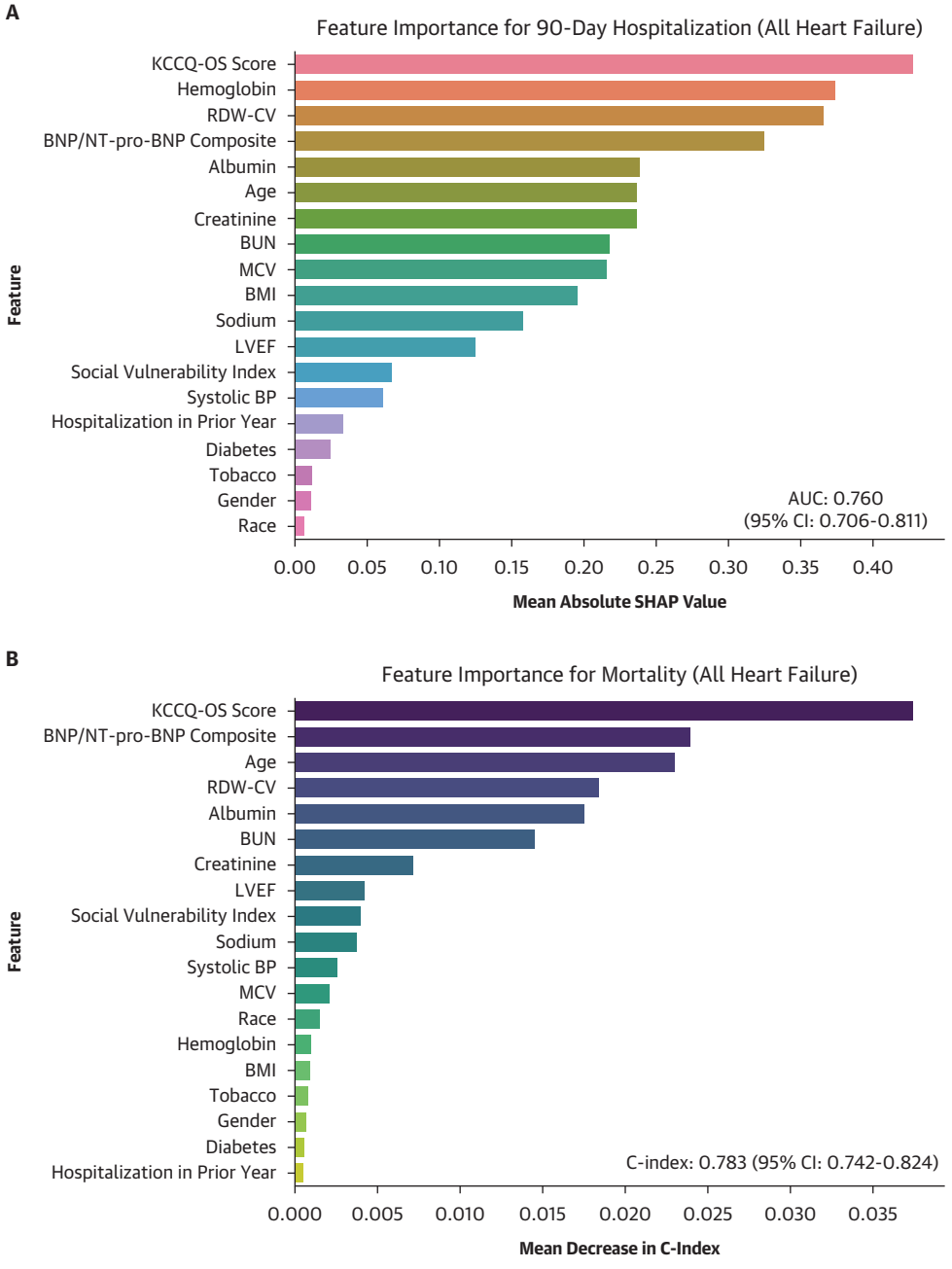
Building on the significant associations between KCCQ-OS scores and hospitalizations and mortality in traditional multivariable regression models, we found that the KCCQ-OS score emerged as 1 of the most important features in machine learning models predicting these outcomes. A study of HFpEF patients using data from the TOPCAT (Treatment of Preserved Cardiac Function Heart Failure with an Aldosterone Antagonist) randomized controlled trial found that the feature importance of the KCCQ-23 Overall Summary score in predicting hospitalizations and mortality (AUC increment 0.62)[17] was comparable to that of other variables such as blood urea nitrogen level (AUC increment 0.60) and age (AUC increment 0.70). Our finding that the KCCQ-OS score had greater feature importance than all other demographic and clinical variables when collected in a nontrial setting was novel and likely due to several fundamental differences between clinical trial and pragmatic data. First, pragmatic cohorts often capture a broader spectrum of disease severity than clinical trials. For example, the median KCCQ-OS score in our cohort (59.4) was lower than baseline scores reported in several landmark heart failure trials such as PARADIGM-HF (Prospective Comparison of ARNI [Angiotensin Receptor-Neprilysin Inhibitor] With ACEI [Angiotensin-Converting-Enzyme Inhibitor] to Determine Impact on Global Mortality and Morbidity in Heart Failure),[12] PARAGON-HF (Efficacy and Safety of LCZ696 Compared to Valsartan, on Morbidity and Mortality in Heart Failure Patients With Preserved Ejection Fraction),[11] and EMPEROR-Preserved (Empagliflozin Outcome Trial in Patients With Chronic Heart Failure With Preserved Ejection Fraction)[8] (mean KCCQ scores 72.3, 71.4, and 68.9, respectively). Second, pragmatic cohorts incorporate latency for various data elements.

FIGURE 1 **Association of the Kansas City Cardiomyopathy Questionnaire-12 With Hospitalizations and Mortality**



Absolute risk estimates for 90-day hospitalization and 1-year mortality across the range of Kansas City Cardiomyopathy Questionnaire-12 Overall Summary (KCCQ-OS) scores. Probabilities were derived from multivariable-adjusted logistic regression models for hospitalization and Cox proportional hazards models for mortality, incorporating restricted cubic splines to account for nonlinear relationships. Shaded areas represent 95% CIs. Analyses were conducted in all heart failure patients (N = 2,888) and stratified by heart failure phenotype (heart failure with reduced ejection fraction [HFrEF], n = 1,690; heart failure with preserved ejection fraction [HFpEF], n = 1,198).

FIGURE 2 **Feature Importance for Hospitalizations and Mortality in Machine Learning Models**



Feature importance is shown using SHAP (SHapley Additive exPlanations) values for the gradient boosting model's (XGBoost) 90-day hospitalization predictions (A) and permutation importance for the random survival forest model's cumulative mortality predictions (B). Higher values indicate a greater impact on model predictive accuracy in all patients with heart failure (N = 2,888). AUC = area under the receiver-operating characteristic curve; BMI = body mass index; BNP = B-type natriuretic peptide; BP = blood pressure; BUN = blood urea nitrogen; KCCQ-12 = Kansas City Cardiomyopathy Questionnaire-12; LVEF = left ventricular ejection fraction; MCV = mean corpuscular volume; NT-pro-BNP = N-terminal B-type natriuretic peptide; RDW-CV = red cell distribution width-coefficient of variation.

This contrasts with clinical trials, in which participants often have blood draws and vital signs collected at the same time as the KCCQ-12 during study visits. Last, pragmatic cohorts have higher proportions of missing data as compared with highly curated clinical trial data sets, necessitating multiple imputation or similar methods. All these considerations may increase the predictive utility of PROMs, such as the KCCQ-12, in routine clinical care.

One additional, important difference between pragmatic cohorts incorporating the KCCQ-12 (or any PROM) as compared with clinical trials is potential nonresponse bias. Capturing the perspectives of as many patients as possible is important as patients who do not complete PROMs tend to be older, less healthy, and less advantaged compared with those who do.[47] This pattern was also apparent in the current study. Our 66% completion rate for the KCCQ-12 aligns with PROM completion rates in prior studies, which have generally ranged from 50% to 80% in PROMs deployed in routine clinical practice.[48] There were statistically significant differences in comorbidity burden between patients who completed the KCCQ-12 compared to those who did not, though these differences were relatively small in magnitude (**Table 1**). Both 90-day hospitalizations and mortality were significantly higher in the KCCQ-12 noncompletion group, however, and noncompletion of the questionnaire was independently associated with worse outcomes after multivariable adjustment. In examining the association between KCCQ-12 noncompletion and outcomes, it is essential to acknowledge the role of healthy cohort bias, as patients who completed the KCCQ-12 likely had better health status and were more advantaged in ways that were not measured. However, these findings are consistent with a substantial body of literature demonstrating that PROM noncompletion signals increased clinical risk. In a population-based study of >120,000 patients with cancer, patients who never completed symptom questionnaires had a 52% higher mortality risk.[49] Similarly, patients who failed to return mailed outcome surveys after total knee arthroplasty[50] and gastric fundoplication[51] had significantly worse functional status and symptom burden when eventually assessed. The relationship between PROM noncompletion and adverse outcomes across multiple clinical domains, including heart failure, suggests that missing patient-reported data may represent an important signal warranting clinical attention rather than merely representing a methodological limitation. Future research should explore whether targeted outreach to PROM noncompleters might identify high-risk patients who could benefit from more intensive monitoring.

Although the current study focused on the prognostic capabilities of the KCCQ-12, the instrument is both a predictor of clinical outcomes and a patient-centered outcome itself. Interventions designed to improve KCCQ-OS scores have the potential to engage a "virtuous cycle"[52] in which health status impacts hospitalization rates and health care resource use, further impacting quality of life. Clinicians could utilize KCCQ-OS scores to identify high-risk patients who may benefit from more intensive guideline-directed medical therapy and to educate patients about elevated risk, potentially motivating improvements in self-care behaviors such as medication adherence. The optimal use of the KCCQ-12 in routine clinical practice is nuanced, however. The PRO-HF (Patient-Reported Outcome Measurement in Heart Failure Clinic) trial demonstrated that using the KCCQ-12 in outpatient clinics improved patient-clinician communication,[53] but not clinical outcomes such as hospitalizations and mortality. However, the much higher baseline KCCQ-OS scores in the PRO-HF trial indicated that these patients were in a lower risk category than patients in the current study.[6] Notably, 1 study of patients with heart failure seen in the emergency room demonstrated that self-care interventions may improve KCCQ-OS scores.[54] The degree to which the KCCQ-12 can be used both as a predictive tool and a target for interventions warrants further investigation.

STUDY LIMITATIONS. Our study has several limitations beyond those already mentioned. First, our outcome data were limited to information extracted from the VUMC EHR, potentially missing some deaths and hospitalizations at outside facilities, though close follow-up likely minimized this issue. Second, some general cardiology patients were seen in the VUMC Heart Failure Clinics in which the KCCQ-12 was assigned, potentially affecting HFpEF classification. However, the final cohort was composed of 59% HFrEF and 41% HFpEF, consistent with the overall prevalence of these heart failure phenotypes,[55] and both groups had KCCQ-12 scores consistent with moderate heart failure-related impairment. Third, several variables had substantial rates of missingness, particularly NT-proBNP. This missingness may not be random, as clinicians likely prioritize biomarker testing in more symptomatic patients, potentially underestimating the apparent importance of biomarkers in our machine learning models. However, this pattern of selective biomarker testing reflects real-world clinical decision-making, wherein comprehensive laboratory assessment is not uniformly applied to all patients. Fourth, our analyses used EHR variables closest in temporal proximity to

the KCCQ-12 completion. This approach reflects real-world clinical practice where data elements are collected at different intervals based on clinical necessity rather than simultaneously with patient-reported outcomes. The KCCQ-OS score's emergence as the dominant predictor of hospitalizations and mortality highlights its robust utility in routine clinical settings where perfectly contemporaneous measurements are rarely available. Finally, clinicians had full access to KCCQ-12 scores during clinical encounters, potentially leading to intensified management for patients with poor health status. This access would likely bias our findings toward the null hypothesis, however, suggesting that the observed associations between KCCQ-12 scores and outcomes may underestimate the instrument's true prognostic value.

CONCLUSIONS

Our study suggests that the KCCQ-12 has significant predictive utility for hospitalizations and mortality in patients with heart failure when administered during routine outpatient clinical care. The strong associations of KCCQ-OS scores with clinical outcomes combined with the high feature importance of the questionnaire in machine learning models predicting these outcomes underscores the potential of this tool to inform clinical decisions. Further research is necessary to optimize the use of the KCCQ-12 in routine heart failure care and evaluate interventions based on questionnaire scores to enhance quality of life and reduce mortality.

DATA SHARING STATEMENT

The data utilized in this study were derived from patient medical records at Vanderbilt University Medical Center. Given the longitudinal, granular, and sensitive nature of these data, comprehensive deidentification would be impractical without significantly compromising scientific utility. Investigators interested in further information about the methods or analytical approaches are welcome to contact the corresponding author.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

This material is based upon work supported by Career Development Award 1IK2HX003021 (Dr Bachmann) from the United States Department of Veterans Affairs, Health Services Research Service; Rapid-Cycle Research Award RI-RCIP-1812-002 (Dr Bachmann) from the Patient-Centered Outcomes Research Institute; and T32 HG008341 (Dr El-Sabawi) from the National Institutes of Health. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States government. Dr Lindenfeld is a consultant for Abbott, Abiomed, Adona, Alleviant, Boston Scientific, CVRx, Edwards Lifesciences, Fire 1, Intershunt, Medtronic, Merck, Orchestra Biomed, VWave, Whiteswell, and Vascular Dynamics. Dr Gupta receives research support from the NIH; and is a consultant for Novo Nordisk. Dr Spertus has provided consultative services on patient-reported outcomes and evidence evaluation to Alnylam, AstraZeneca, Bayer, Janssen, Bristol Myers Squibb, Terumo, Cytokinetics, BridgeBio, VentricHealth, and Imbria; holds research grants from the National Institutes of Health, the Patient-Centered Outcomes Research Institute, the American College of Cardiology Foundation, Lexicon, Imbria, and Janssen; owns the copyright to the Seattle Angina Questionnaire, Kansas City Cardiomyopathy Questionnaire, and Peripheral Artery Questionnaire; and serves on the Board of Directors for Blue Cross Blue Shield of Kansas City. Dr Collins receives research support from NIH, PCORI, and DOD and is a consultant for Reprieve Cardiovascular, Prenosis, Abbott, Boehringer Ingelheim and Tosoh. Dr Bachmann's spouse is employed by Amgen, Inc. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr Justin M. Bachmann, Cardiology Section, VA Tennessee Valley Healthcare System, 2525 West End Avenue, Suite 300-A, Nashville, Tennessee 37203, USA. E-mail: justin.m.bachmann@vanderbilt.edu. X handle: @JMBachmann.

REFERENCES

1. Rahimi K, Malhotra A, Banning AP, Jenkinson C. Outcome selection and role of patient reported outcomes in contemporary cardiovascular trials: systematic review. *BMJ*. 2010;341:c5707.

2. Bachmann JM, Posch DR, Hickson GB, et al. Developing an implementation strategy for systematic measurement of patient-reported outcomes at an academic health center. *J Healthc Manag*. 2020;65:15–28.

3. Gibbons C, Porter I, Goncalves-Bradley DC, et al. Routine provision of feedback from patient-reported outcome measurements to healthcare providers and patients in clinical practice. *Cochrane Database Syst Rev*. 2021;10(10):CD011589. https://doi.org/10.1002/14651858.CD011589.pub2

4. Spertus JA, Jones PG. Development and validation of a short version of the Kansas City Cardiomyopathy Questionnaire. *Circ Cardiovasc Qual Outcomes*. 2015;8:469–476.

5. Green CP, Porter CB, Bresnahan DR, Spertus JA. Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure. *J Am Coll Cardiol*. 2000;35:1245–1255.

6. Sandhu AT, Calma J, Skye M, et al. Clinical impact of routine assessment of patient-reported health status in heart failure clinic: the PRO-HF trial. *Circulation*. 2024;149:1717–1728.

7. Kosiborod MN, Bhatt AS, Claggett BL, et al. Effect of dapagliflozin on health status in patients with preserved or mildly reduced ejection fraction. *J Am Coll Cardiol*. 2023;81:460–473.

8. Butler J, Filippatos G, Jamal Siddiqi T, et al. Empagliflozin, health status, and quality of life in patients with heart failure and preserved ejection fraction: the EMPEROR-Preserved trial. *Circulation*. 2022;145:184–193.

9. Butler J, Anker SD, Filippatos G, et al. Empagliflozin and health-related quality of life outcomes in patients with heart failure with reduced ejection fraction: the EMPEROR-Reduced trial. *Eur Heart J*. 2021;42:1203–1212.

10. Kosiborod MN, Jhund PS, Docherty KF, et al. Effects of dapagliflozin on symptoms, function, and quality of life in patients with heart failure and

reduced ejection fraction: results from the DAPA-HF trial. *Circulation*. 2020;141:90–99.

11. Chandra A, Vaduganathan M, Lewis EF, et al. Health-related quality of life in heart failure with preserved ejection fraction: the PARAGON-HF trial. *JACC Heart Fail*. 2019;7:862–874.

12. Lewis EF, Claggett BL, McMurray JJV, et al. Health-related quality of life outcomes in PARADIGM-HF. *Circ Heart Fail*. 2017;10(8):e003430. https://doi.org/10.1161/CIRCHEARTFAILURE.116.003430

13. Lewis EF, Kim HY, Claggett B, et al. Impact of spironolactone on longitudinal changes in health-related quality of life in the Treatment of Preserved Cardiac Function Heart Failure With an Aldosterone Antagonist trial. *Circ Heart Fail*. 2016;9:e001937.

14. Spertus J, Peterson E, Conard MW, et al. Monitoring clinical changes in patients with heart failure: a comparison of methods. *Am Heart J*. 2005;150:707–715.

15. Pokharel Y, Khariton Y, Tang Y, et al. Association of serial Kansas City Cardiomyopathy Questionnaire assessments with death and hospitalization in patients with heart failure with preserved and reduced ejection fraction: a secondary analysis of 2 randomized clinical trials. *JAMA Cardiol*. 2017;2:1315–1321.

16. Heidenreich PA, Spertus JA, Jones PG, et al. Health status identifies heart failure outpatients at risk for hospitalization or death. *J Am Coll Cardiol*. 2006;47:752–756.

17. Angraal S, Mortazavi BJ, Gupta A, et al. Machine learning prediction of mortality and hospitalization in heart failure with preserved ejection fraction. *JACC Heart Fail*. 2020;8:12–21.

18. Ford I, Norrie J. Pragmatic trials. *N Engl J Med*. 2016;375:454–463.

19. Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-world evidence - what is it and what can it tell us? *N Engl J Med*. 2016;375:2293–2297.

20. Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *Eur Heart J*. 2017;38:1805–1814.

21. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One*. 2017;12:e0174944.

22. Quer G, Arnaout R, Henne M, Arnaout R. Machine learning and the future of cardiovascular care: JACC state-of-the-art review. *J Am Coll Cardiol*. 2021;77:300–313.

23. Liu T, Krentz A, Lu L, Curcin V. Machine learning based prediction models for cardiovascular disease risk using electronic health records data: systematic review and meta-analysis. *Eur Heart J Digit Health*. 2025;6:7–22.

24. Liu W, Laranjo L, Klimis H, et al. Machine-learning versus traditional approaches for atherosclerotic cardiovascular risk prognostication in primary prevention cohorts: a systematic review and meta-analysis. *Eur Heart J Qual Care Clin Outcomes*. 2023;9:310–322.

25. Batra P, Khera AV. Machine learning to assess coronary artery disease status-is it helpful? *Lancet*. 2023;401:173–175.

26. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007;370:1453–1457.

27. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015;350:g7594.

28. Steitz BD, Wong JIS, Cobb JG, Carlson B, Smith G, Rosenbloom ST. Policies and procedures governing patient portal use at an academic medical center. *JAMIA Open*. 2019;2:479–488.

29. Microsoft. Azure Data Studio (Version 1.49.1): Microsoft; 2024.

30. Sandhu AT, Zheng J, Skye M, Heidenreich PA. Updating the accuracy of administrative claims for identifying left ventricular ejection fraction among patients with heart failure. *Circ Cardiovasc Qual Outcomes*. 2023;16:e008919.

31. Pocock SJ, Ariti CA, McMurray JJ, et al. Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies. *Eur Heart J*. 2013;34:1404–1413.

32. Simpson J, Jhund PS, Lund LH, et al. Prognostic models derived in PARADIGM-HF and validated in ATMOSPHERE and the Swedish Heart Failure Registry to predict mortality and morbidity in chronic heart failure. *JAMA Cardiol*. 2020;5:432–441.

33. Levy WC, Mozaffarian D, Linker DT, et al. The Seattle Heart Failure Model: prediction of survival in heart failure. *Circulation*. 2006;113:1424–1433.

34. Myhre PL, Vaduganathan M, Claggett BL, et al. Association of natriuretic peptides with cardiovascular prognosis in heart failure with preserved ejection fraction: secondary analysis of the TOP-CAT randomized clinical trial. *JAMA Cardiol*. 2018;3:1000–1005.

35. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care*. 2005;43:1130–1139.

36. Centers for Disease Control and Prevention. Agency for Toxic Substances Disease Registry, and Geospatial Research, Analysis, and Services Program. *CDC/ATSDR Social Vulnerability Index*. 2024. Accessed December 12, 2024. https://www.atsdr.cdc.gov/place-health/php/svi/svi-data-documentation-download.html

37. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2011;12:2825–2830.

38. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr*. 1974;19:716–723.

39. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2024.

40. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Krishnapuram B, Shah M, Smola AJ, Aggarwal C, Shen D, Rastogi R, eds. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM; 2016:785–794. https://doi.org/10.1145/2939672.293978

41. Lundberg S, Lee S-I. A unified approach to interpreting model predictions. In: Guyon I, Von Luxburg U, Bengio S, et al., eds. *Advances in Neural Information Processing Systems 30: 31st Annual Conference on Neural Information Processing Systems (NIPS 2017). Neural Information Processing Systems Foundation, Inc.* NeurIPS); 2017:4765–4774.

42. Hsich E, Gorodeski EZ, Blackstone EH, Ishwaran H, Lauer MS. Identifying important risk factors for survival in patients with systolic heart failure using random survival forests. *Circ Cardiovasc Qual Outcomes*. 2011;4:39–45.

43. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *Ann Appl Stat*. 2008;2:841–860, 20.

44. Pölsterl S, Navab N, Katouzian A. Fast training of support vector machines for survival analysis. In: Appice A, Rodrigues P, Santos Costa V, Gama J, Jorge A, Soares C, eds. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2015*. Lecture Notes in Computer Science; 9285.

45. Python Software Foundation. *Python 3.11.9*. Python Software Foundation; 2024.

46. Hu D, Liu J, Zhang L, et al. Health status predicts short- and long-term risk of composite clinical outcomes in acute heart failure. *JACC Heart Fail*. 2021;9:861–873.

47. Bernstein DN, Karhade AV, Bono CM, Schwab JH, Harris MB, Tobert DG. Sociodemographic factors are associated with patient-reported outcome measure completion in orthopaedic surgery: an analysis of completion rates and determinants among new patients. *JB JS Open Access*. 2022;7(3):e22.00026. https://doi.org/10.2106/JBJS.OA.22.00026

48. Unni E, Coles T, Lavallee DC, Freel J, Roberts N, Absolom K. Patient adherence to patient-reported outcome measure (PROM) completion in clinical care: current understanding and future recommendations. *Qual Life Res*. 2024;33:281–290.

49. Barbera L, Sutradhar R, Seow H, et al. The impact of routine Edmonton Symptom Assessment System (ESAS) use on overall survival in cancer patients: results of a population-based retrospective matched cohort analysis. *Cancer Med*. 2020;9:7107–7115.

50. Kim J, Lonner JH, Nelson CL, Lotke PA. Response bias: effect on outcomes evaluation by mail surveys after total knee arthroplasty. *J Bone Joint Surg Am*. 2004;86:15–21.

51. Ludemann R, Watson DI, Jamieson GG. Influence of follow-up methodology and completeness on apparent clinical outcome of fundoplication. *Am J Surg*. 2003;186:143-147.

52. Porter ME, Teisberg EO. *Redefining Health Care: Creating Value-Based Competition on Results*. Harvard Business School Press; 2006.

53. Sandhu AT, Zheng J, Kalwani NM, et al. Impact of patient-reported outcome measurement in heart failure clinic on clinician health status assessment and patient experience: a substudy of the PRO-HF trial. *Circ Heart Fail*. 2023;16:e010280.

54. Stubblefield WB, Jenkins CA, Liu D, et al. Improvement in Kansas City Cardiomyopathy Questionnaire scores after a self-care intervention in patients with acute heart failure discharged from the emergency department. *Circ Cardiovasc Qual Outcomes*. 2021;14:e007956.

55. Dunlay SM, Roger VL, Redfield MM. Epidemiology of heart failure with preserved ejection fraction. *Nat Rev Cardiol*. 2017;14:591-602.

KEY WORDS health informatics, heart failure, machine learning, patient-reported outcomes, prognosis

APPENDIX For supplemental figures and tables, please see the online version of this paper.