

1 **TITLE:** The Colorado Heart Failure Acuity Risk Model (CHARM) Score: A Mortality Risk  
2 Model for Waitlisted Cardiac Transplant Patients

3

4 **AUTHORS:**

5 Rachel D. Murphy, MD<sup>1</sup>

6 Sarah Y. Park, MD<sup>2</sup>

7 Larry A. Allen, MD, MHS<sup>2</sup>

8 Amrut V. Ambardekar, MD<sup>2</sup>

9 Joseph C. Cleveland, Jr., MD<sup>3</sup>

10 Michael T. Cain, MD<sup>3</sup>

11 Bruce Kaplan, MD<sup>4,5</sup>

12 Jordan R.H. Hoffman, MPH, MD<sup>3</sup>

13 John S. Malamon, PhD<sup>4,5</sup>

14

15 **AFFILIATIONS:**

16 <sup>1</sup>University of Colorado Anschutz Medical Campus, Department of Medicine

17 <sup>2</sup>University of Colorado Anschutz Medical Campus, Department of Medicine, Division of  
18 Cardiology

19 <sup>3</sup>University of Colorado Anschutz Medical Campus, Department of Surgery, Division of  
20 Cardiothoracic Surgery

21 <sup>4</sup>University of Colorado Anschutz Medical Campus, Department of Surgery, Division of  
22 Transplant Surgery

23 <sup>5</sup>Colorado Center for Transplantation Care, Research and Education (CCTCARE)

24

25 **WORD COUNT:** 2,987

26 **LAST REVISION:** 08/10/2023

27 **CORRESPONDING AUTHOR:**

28 John Stephen Malamon, PhD

29 University of Colorado, Anschutz Medical Campus

30 1635 Aurora Ct., Aurora CO 80045

31 Phone: 1-720-848-9063

32 Email: john.malamon@cuanschutz.edu

33 **KEY POINTS:**

34 **Question:** Can pre-operative patient characteristics be used to develop a formal system  
35 to accurately estimate, rank, and predict the relative short-term mortality of waitlisted  
36 heart transplant patients?

37

38 **Findings:** Using twelve patient attributes, we derived three linear regression equations  
39 to accurately predict the 30-day, 90-day, and 1-year mortality of waitlisted heart  
40 transplant patients. We developed and calibrated a seven-tiered risk index for each  
41 model that was 99% correlated to the observed mortality rate. Using several  
42 independent validation methods, we achieved extreme sensitivity (>98%) in ordinal  
43 ranking patient groups who were more likely to survive 30 days on the waitlist. Model  
44 performance was measured using the area under the receiver operating characteristic  
45 (ROC) curve. Using six interaction terms, the area under the ROC curve was 96.4%  
46 (30-day), 90.4% (90-day), and 78% (1-year).

47

48 **Meaning:** Our models accurately discriminate among patient subgroups who are more  
49 likely to die while waitlisted. Because our tiered ranking system is simple, extremely  
50 sensitive, and well calibrated, it is ideal for prioritizing waitlisted heart transplant patients  
51 based on a well-defined medical urgency score. These models are generalized and  
52 therefore extensible to defining medical urgency in larger patient populations  
53 experiencing end-stage heart failure.

54

55

56 **ABSTRACT:**

57 **Importance:** Although the Organ Procurement and Transplantation network provides  
58 structured policies and guidance for waitlisted cardiac transplant patients, the heart  
59 transplantation community lacks a mathematical model that can accurately estimate the  
60 short-term risk of death associated with being waitlisted. Importantly, the CHARM score  
61 provides a risk management and ranking system for patients based on a well-defined  
62 and sensitive medical urgency metric.

63 **Objective:** We had three primary objectives in completing this study. First, to increase  
64 relevance and applicability, we selected patient attributes that were clinically justified  
65 and readily available. Second, we designed and implemented an intuitive, formal  
66 system that accurately defined the relative risk of death while being waitlisted at 30-day,  
67 90-day, and 1-year censoring periods. Third, we developed and validated a medical  
68 urgency metric that is intuitive, practical, and can be implemented nationally.

69 **Design:** We present a multivariable, prognostic model and risk management strategy  
70 for adult waitlisted heart transplant patients (N=1,965) from the Scientific Registry of  
71 Transplant Recipients (SRTR) database that were waitlisted from January 1, 2008, to  
72 September 2, 2022. To independently validate each model, we randomly split this  
73 cohort into a discovery set (N=1,174) and validation set (N=784). Twelve independent  
74 patient attributes were selected, and three linear regression formulas were derived to  
75 estimate and rank the relative risk of dying while waitlisted. Four independent validation  
76 methods were used to measure each model's performance as a classifier and ranking  
77 system.

78 **Setting:** The United States

79 **Participants:** This cohort (N=1,965) consisted of adult heart transplant candidates  
80 without missing laboratory data who were placed on the waitlist from January 1, 2008,  
81 to September 2, 2022. Patients listed for multi-organ transplantation were excluded.  
82 Patients with missing laboratory data were analyzed independently.

83 **Exposures:** The short-term risk of death remaining on the heart transplant waitlist.

84 **Main Outcomes and Measures:** The primary outcome of this study was the design,  
85 development, and validation of a formal risk management system for waitlisted heart  
86 transplant candidates experiencing end-stage heart failure. We derived three linear  
87 regression formulas and calibrated a seven-tiered risk index to accurately rank patients  
88 who were more likely to die on the waitlist at 30-day (30D), 90-day (90D), and 1-year  
89 (1Y) censoring periods. Four independent validation methods were used to measure  
90 each model's classification and ranking performance.

91 **Results:** Using six interaction terms, we applied the 5-fold cross-validation procedure to  
92 the CHARM to discover an area under the ROC curve of 96.4%, 90.4.%, and 78% for  
93 the 30D, 90D, and 1Y models, respectively. The mean positive predictive values of the  
94 tiered risk system were 99.2% (30D), 94.1% (90D), and 88% (1Y). Risk indices for all  
95 three models were >99% correlated to the observed mortality rate across the seven  
96 tiers for the 30D, 90D, and 1Y models.

97 **Conclusions and Relevance:** We designed, implemented, and validated an intuitive  
98 and formal risk scoring and ranking system which is ideal for prioritizing waitlisted heart  
99 failure patients based on a well-defined medical urgency metric. The CHARM score  
100 provides extreme sensitivity in predicting short-term mortality outcomes. The CHARM  
101 score is extensible to larger patient populations experiencing end-stage heart failure.

## 102 1. INTRODUCTION

103 The Organ Procurement and Transplantation Network (OPTN) has worked to develop a  
104 cardiac transplant allocation system that distributes donor hearts to the most critically ill  
105 patients<sup>1</sup>. The 2018 policy revision created a 6-tiered “status”-based stratification  
106 system that attempted to encompass the increasing complexity of managing critical  
107 cardiac illness. However, this policy has been critiqued for its subjectivity and  
108 heterogeneity in accurately discriminating patient risk while on the waitlist. For example,  
109 a large percentage of transplant candidates are now stratified to Status 1 or Status 2 by  
110 exception rather than by the standard criteria, which can be modified by a physician’s  
111 practice<sup>2</sup>. Importantly, studies have demonstrated that the 2018 policy revision is  
112 associated with a significant reduction in post-transplant survival<sup>3,4</sup>. A formal prognostic  
113 model that accurately stratifies waitlisted end-stage heart failure (HF) patients based on  
114 medical urgency is nonexistent. The need for such a model was recently highlighted by  
115 Pelzer *et al.*, who determined the current allocation system had only a moderate ability  
116 to successfully rank transplant candidates according to medical urgency<sup>5</sup>.

117

118 While useful prognostic tools for patients with HF have been introduced, such as the  
119 Heart Failure Survival Score<sup>6</sup> and the Seattle Heart Failure Model<sup>7</sup>, they have not  
120 accurately predicted waitlist mortality<sup>8</sup>. The Ottawa Heart Failure Risk Scale provides a  
121 risk stratification tool designed for acute heart failure patients in emergency  
122 departments<sup>9</sup>. However, this score has not been applied to or validated against a  
123 waitlisted heart transplant patient population. Similarly, the Meta-Analysis Global Group  
124 in Chronic Heart Failure (MAGGIC) risk score performs well for HF patients with

125 preserved ejection fraction, but it has not been externally validated for reduced ejection  
126 fraction, which would exclude most of the population on a transplant list<sup>10</sup>.

127

128 Predictive models have been successfully implemented in other organ transplant  
129 systems. The Model for End-stage Liver Disease including Sodium or MELD-Na score  
130 for liver transplantation accurately predicts 90-day patient waitlist mortality and is the  
131 most significant metric in liver allocation<sup>11</sup>. A similar model is desperately needed in  
132 cardiac transplantation to accurately prioritize the most critically ill patients. We used the  
133 Scientific Registry of Transplant Recipients (SRTR) database to develop and validate  
134 three predictive mortality models for waitlisted patients with end-stage HF. Our models  
135 utilize objective physiological data to determine the most urgent heart transplant  
136 candidates and stratify their relative risk of waitlist mortality at 30 days, 90 days, and 1  
137 year.

138

## 139 **2. METHODS**

### 140 **2.1 Data sources**

141 This study utilized retrospective data from the SRTR. The SRTR system includes data  
142 on all donors, waitlisted candidates, and transplant recipients in the US, submitted by  
143 the members of OPTN<sup>12</sup>. The Health Resources and Services Administration, U.S.  
144 Department of Health and Human Services provides oversight to the activities of the  
145 OPTN and SRTR contractors. The data reported here have been supplied by the  
146 Hennepin Healthcare Research Institute as the contractor for SRTR. The interpretation

147 and reporting of these data are the responsibility of the author(s) and in no way should  
148 be seen as an official policy of or interpretation by the SRTR or the U.S. Government.

149

## 150 **2.2 Study population**

151 The study population included adult waitlisted cardiac transplant patients with complete  
152 laboratory results who registered for a single-organ heart transplant (N=1,965) between  
153 January 1, 2008, and September 2, 2022. **Supplemental Figure 1** provides a  
154 participant workflow diagram with the number of participants retained for each exclusion  
155 criterion. All steps in this analysis were conducted for 30-day (30D), 90-day (90D), and  
156 1-year (1Y) all-cause mortality. Survival times for waitlisted candidates started at the  
157 date of listing and were censored at the date of death and upon removal from the  
158 waitlist. The population was randomly split (60%/40%) into two cohorts, a discovery set  
159 (N=1,179) and a validation set (N=786). The discovery set was used for variable  
160 selection, the generation of three linear regression equations, and the creation of a  
161 tiered risk index. The validation set was used to independently evaluate each model's  
162 performance in predicting mortality outcomes and ranking patients' relative risk of death.  
163

## 164 **2.3 Study Ethics**

165 Informed consent was obtained for all study participants. This study was reviewed by an  
166 ethical committee (Colorado Multiple Institutions Review Board) and was determined to  
167 be non-human subjects research. This study was not grant funded. The authors do not  
168 have any conflicts of interest to disclose. All authors reviewed the results and approved  
169 the final manuscript.

170

## 171 **2.4 Statistical approach**

### 172 **2.4.1 Missingness and Sensitivity Analysis**

173 We compared the patient populations with some missing laboratory data (N=20,991) to  
174 those without missing data (N=1,965). For continuous variables, two-way ANOVA was  
175 used to test the observed differences in patient characteristics and the twelve  
176 independent variables used in this study. The chi-squared test was used to measure the  
177 significance of categorical and indicator (binary) variables. Missing data were significant  
178 in the laboratory variables; therefore, we analyzed missingness patterns and performed  
179 three related sensitivity analyses using: 1) participants without missing data (N=1,965);  
180 2) all participants with missing laboratory data (N=22,949); and 3) all participants with  
181 missing laboratory values imputed (N=22,949). The Multivariate Imputation by Chain  
182 Variables or MICE algorithm was used to impute missing values.

183

### 184 **2.4.2 Clinical Variable Selection and Importance**

185 For clinical applicability, we selected twelve patient variables that were readily available  
186 and clinically justified. These twelve variables should also exhibit discriminating power  
187 and low collinearity. To measure collinearity, we calculated the Pearson's cross-  
188 correlation coefficient (R) and variance inflation factor (VIF) for all twelve independent  
189 variables. Low collinearity was defined as two independent variables with an R value  
190 less than 0.7. The VIF is used to determine the correlation between independent  
191 variables in a logistic regression model. A VIF of 1 provides no correlation, whereas  
192 values above 2.5 indicate considerable multicollinearity<sup>13</sup>. Using the discovery set, we



193 measured and ranked the relative importance of each patient variable via the  
194 unsupervised Random Forest (RF) algorithm<sup>14</sup> with the number of days survived as the  
195 outcome. The RF algorithm provided an orthogonal method for comparing the relative  
196 effect of each independent variable used in these models.

197

### 198 **2.4.3. Independent Variable Definitions**

199 Life Support refers to any waitlisted heart transplant recipient who received any cardiac  
200 support prior to transplant including intravenous (IV) inotropic infusion, or any circulatory  
201 support device such as right ventricular assist devices (RVAD), left ventricular assist  
202 devices (LVAD), total artificial hearts (TAH), extracorporeal membrane oxygenation  
203 (ECMO), or intra-aortic balloon pumps (IABP)<sup>15</sup>. The Ventilator variable refers to the  
204 need for any mechanical ventilation prior to transplant. The use of mechanical  
205 ventilators has been shown to significantly increase patients' risk of death post-  
206 transplant<sup>16</sup>. The Prior Cardiac Surgery indicator variable includes patients with any  
207 history of cardiac surgery.

208

### 209 **2.4.4 Regression Formulas for Calculating Patient Risk Scores**

210 We set out to develop a short-term prognostic model to formally define a patient risk  
211 score also known as the CHARM score based on a patient's estimated likelihood of  
212 death while waitlisted. Thus, we constructed three linear regression equations using the  
213 logistic link function for 30D, 90D, and 1Y mortality outcomes (**Equations 1-3**). All  
214 laboratory values were transformed to the logarithmic scale prior to calculation. We  
215 calculated CHARM scores for all patients in the discovery set to develop these

216 equations. For consistency across the models, we used **Equation 4** to normalize  
217 CHARM scores to range from 0 to 1.

218

#### 219 ***2.4.5 Model Calibration and the Tiered Risk Index***

220 A well calibrated ranking system is required to accurately estimate the relative medical  
221 urgency of waitlisted heart transplant patients. Thus, we developed a seven-tiered risk  
222 index system based on the CHARM score which ranges from one to seven. We then fit  
223 the risk tiers to the observed mortality rate to maximize the statistical discrimination in  
224 ranking the relative mortality likelihood observed in these patient subgroups or indices.  
225 We defined tiered risk thresholds and calibrated each regression model by maximizing  
226 the goodness-of-fit between the observed mortality and the risk indices, which are  
227 positively correlated. Here, we calculated the goodness-of-fit (R) using linear  
228 regression.

229

#### 230 ***2.4.6 Validating Dichotomous Outcomes using Logistic Regression***

231 We measured each model's performance in predicting mortality events at three  
232 censoring periods using 5-fold cross-validation (in-sample) and supervised sample hold-  
233 out validation (out-sample) using the validation set. We evaluated both additive models  
234 and those using interaction terms. Age, Albumin, Creatinine, Circulatory Support,  
235 Previous Transplant, and Prior Cardiac Surgery were the interaction terms. For sample  
236 hold-out validation, we used the supervised Random Forest classification method<sup>14</sup> with  
237 70 trees. Priors were calculated using the discovery set and predictions were

238 independently validated using the validation set. Finally, we calculated the area under  
239 the ROC curve<sup>17</sup> for all three models.

240

#### 241 ***2.4.7 Validating the Seven-Tiered Risk Index System***

242 To validate our risk classification system using a univariate test, we leveraged the Cox  
243 proportional-hazard regression (CPHR) method<sup>18</sup> to measure tier performance in terms  
244 of survival time. Here the unit of measurement was the concordance index<sup>19</sup>. To further  
245 evaluate the performance of the tiered risk system, we used **Equation 5** to calculate the  
246 rank precision or the positive predictive value (PPV) for each risk tier as compared to all  
247 others. A true positive (TP) occurred when a patient with a lower risk index outlived a  
248 patient in any higher tier. A false positive (FP) occurred when a patient with a higher risk  
249 index outlived a patient in any lower tier<sup>20</sup>. All analyses were performed using the R  
250 statistical language version 4.1.2<sup>21</sup>.

251

### 252 **3. RESULTS**

#### 253 ***3.1 Population Characteristics***

254 The discovery, validation, and total patient population characteristics are depicted in  
255 **Table 1**. The mean (SD) age of the 1,965 study participants with no missing laboratory  
256 data was 48.5 (14.5) years; 1,432 (73.1%) participants were male, 1,402 (71.6%) were  
257 White/Caucasian, 374 (19.1%) were Black or African American, 104 (5.3%) were  
258 Hispanic or Latino, 56 (2.9%) were Asian, 9 (0.5%) were Native Hawaiian or Other  
259 Pacific Islander, and 6 (0.3%) were American Indian or Alaska Native. In summary,

260 there were no statistically significant differences in the patient characteristics or  
261 independent variables when comparing the discovery set to the validation set.

262

### 263 **3.2. Missingness and Sensitivity Analysis**

264 **Supplemental Table 1** summarizes the missing data for the twelve independent  
265 variables used in this model. More than 90% of total waitlisted patients were missing  
266 laboratory data. **Supplemental Table 2** provides the patient population characteristics  
267 for those with some missing data (N=20,991) and those without missing data (N=1,965).  
268 On average, patients without missing data were younger and more likely to be female or  
269 Caucasian. Statistically significant differences between the two cohorts were observed  
270 in five of the twelve independent variables. **Supplemental Figure 2** provides a missing  
271 value map for all independent variables. In summary, Age, Circulatory Support,  
272 Previous Transplant, Sodium, and Ventilator were missing at random. The other seven  
273 variables were missing completely at random<sup>22</sup>. **Supplemental Table 3** provides the  
274 AUC of the ROC for participants without missing data, with missing data, and with  
275 missing values imputed.

276

### 277 **3.3 Variable Selection and Importance**

278 **Supplemental Table 4** provides the relative variable importance in the RF model,  
279 measured by the Gini Index, a statistic dispersion value notated as the Increase in Node  
280 Purity (IncNodePurity). These data were normalized using **Equation 4** and are useful  
281 for ranking the relative importance of each independent variable. All logistic regression  
282 coefficients, p-values, and VIF values are provided in **Supplemental Table 5**. The

283 correlation coefficients of the independent variables were below the absolute value of  
284 0.21, and VIF values ranged from 1 to 1.532 (**Figure 1**). No significant co-linearity was  
285 observed in the variables or models.

286

### 287 **3.4 Mathematical Formulas**

288 Equation 1: The 30-day CHARM score

$$\begin{aligned} \mathbf{CHARM}_{30D} = & \sum_{i=1}^P 10.5 + (0.0129 * Age_i) - (0.116 * Cardiac\ Surgery_i) \\ & + (0.6512 * Life\ Support_i) - (0.7136 * \log(Albumin_i)) \\ & + (0.4314 * \log(Creatinine_i)) - (15.83 * Ventilator_i) \\ & + (1.088 * Previous\ Transplant_i) + (0.0001 * \log(BNP_i)) \\ & + (0.6119 * \log(AST_i)) - (2.992 * \log(Sodium_i)) \\ & + (0.5185 * \log(Bilirubin_i)) - (0.0711 * Waitlist\ Days_i) \end{aligned}$$

289

290 Equation 2: The 90-day CHARM score

$$\begin{aligned} \mathbf{CHARM}_{90D} = & \sum_{i=1}^P -6.61 + (0.076 * Age_i) + (0.123 * Cardiac\ Surgery_i) \\ & + (1.111 * Life\ Support_i) + (0.76 * \log(Albumin_i)) \\ & + (0.423 * \log(Creatinine_i)) + (0.336 * Ventilator_i) \\ & + (0.029 * Previous\ Transplant_i) - (0.205 * \log(BNP_i)) \\ & + (0.12 * \log(AST_i)) - (0.124 * \log(Sodium_i)) \\ & + (0.089 * \log(Bilirubin_i)) - (7.433 * Waitlist\ Days_i) \end{aligned}$$

291

292 Equation 3: The 1-year CHARM score

$$\begin{aligned} \mathbf{CHARM}_{1Y} = & \sum_{i=1}^P -2.485 + (0.1 * Age_i) + (0.164 * Cardiac\ Surgery_i) \\ & + (0.814 * Life\ Support_i) + (0.285 * \log(Albumin_i)) \\ & + (0.554 * \log(Creatinine_i)) + (0.134 * Ventilator_i) \\ & + (0.013 * Previous\ Transplant_i) - (0.148 * \log(BNP_i)) \\ & + (0.255 * \log(AST_i)) - (0.175 * \log(Sodium_i)) \\ & + (0.136 * \log(Bilirubin_i)) - (1.382 * Waitlist\ Days_i) \end{aligned}$$

293

294 Equation 4: Normalization by scaling between 0 and 1

$$NPRS = \sum_{i=1}^P \frac{(PRS_i) - \min(PRS_{i:P})}{\max(PRS_{i:P}) - \min(PRS_{i:P})}$$

295

296 Equation 5: Positive Predictive Value of the Tiered Risk System

$$PPV = TP / (TP + FP)$$

297

### 298 **3.4 Model Calibration and the Tiered Risk Index**

299 The observed mortality rate is presented in **Figure 2** per risk index as a function of the  
300 CHARM score. **Supplemental Figures 3-5** provide the CHARM score distributions for  
301 each model. For the 90-day model, Risk Index (RI) 1 had an observed mortality rate of  
302 1.52%, RI 2 had a mortality rate of 3.54%, RI 3 had a mortality rate of 7%, RI 4 had a  
303 mortality rate of 11.25%, RI 5 had a mortality rate of 13.5%, RI 6 had a mortality rate of  
304 20.8%, and RI 7 had a mortality rate of 26.5%. The risk indices were >99% correlated to  
305 the observed mortality rate across the seven tiers for all three models.

306

### 307 **3.5 Logistic Regression for Predicting Short-term Mortality Outcomes**

308 The 5-fold cross-validation procedure produced an AUC of 94.8%, 86.7%, and 74.2%  
309 for the 30D, 90D, and 1Y additive models, respectively (**Supplemental Figures 6-8**).  
310 Using interaction terms, the AUC was 96.4%, 90.4%, and 78%, respectively (**Figure 3**).  
311 Sample hold-out validation produced an AUC of 93.8%, 92.5%, and 74.7% for the 30D,  
312 90D, and 1Y additive models, respectively (**Supplemental Figures 9-11**). In summary,  
313 these models and the tiered risk system provide a reliable and highly accurate  
314 methodology for ranking the short-term survival of waitlisted heart transplant patients.

### 315 **3.6 Validation of the Seven-Tiered Risk Index System**

316 Using the validation cohort, we found that the PPV of the seven-tiered risk index system  
317 ranged from 98.3% to 100% (30D), 83.7% to 100% (90D), and 0.668% to 100% (**Table**  
318 **2**). The mean PPVs were 99.2% (30D), 94.1% (90D), and 88% (1Y). Using the Cox  
319 proportional-hazard regression methodology as a univariate test where the unit of  
320 measurement was the risk index, we found a significant difference in survival times by  
321 risk tier (**Supplemental Figure 12-14**). For example, CPHR revealed that a patient with  
322 a RI of 7 had a 26% chance of death after 90 days on the waitlist, while a patient with a  
323 RI of 1 had about a 2% risk of death. The mortality rates for the risk indices provided in  
324 **Figure 2** are nearly identical to the inverse survival rates provided by the CPHR  
325 analysis.

## 326 **4. DISCUSSION**

327  
328 There is currently no model that accurately stratifies waitlisted cardiac transplant  
329 patients based on medical urgency. Formal pre-transplant predictive models have been  
330 successfully developed in liver transplantation. Evans *et al.* demonstrated an overall 1-  
331 year survival rate increase of 18% in high-acuity patients in the 15 years following the  
332 national implementation of the MELD-Na score<sup>23</sup>. A similar metric is needed for cardiac  
333 transplant to prioritize the most critically ill waitlisted patients.

334  
335 To this end, we created the Colorado Heart failure Acuity Risk Model or CHARM score.  
336 The models provided herein precisely rank patient subgroups based on waitlist mortality  
337 with mean PPVs of 99.2% (30D), 94.1% (90D), and 88% (1Y). They demonstrate  
338 excellent accuracy with AUC values of 96.4%, 90.4%, and 78%, respectively. While

339 previous scores or models have aimed to assess heart failure illness severity, the  
340 CHARM score is the first to present this level of accuracy and predictive performance  
341 among waitlisted HF patients.

342  
343 The CHARM score is intended to be used in cardiac transplant candidates at the time of  
344 listing. The 90D time frame will allow for frequent reevaluation while maintaining a high  
345 degree of prediction accuracy. The CHARM score can substantially improve heart  
346 allocation within the current “status”-based system by providing an objective prognostic  
347 measurement of medical urgency. The CHARM score will inform the continuous  
348 distribution for CD system, as it utilizes a framework that is point-based rather than  
349 “status”-based, in which candidates are prioritized for transplant through designation of  
350 a composite score from a variety of attributes. Staged implementation of the CD system  
351 is currently underway with anticipated completion of the heart allocation system within  
352 the next few years. A substantial portion (25%) of the lung transplant composite score is  
353 comprised of estimated waitlist mortality, further emphasizing the importance of pre-  
354 transplant mortality estimates for assessing patient acuity. The CHARM score provides  
355 a simple, accurate measure of pre-transplant mortality that can easily be incorporated  
356 into a heart transplant composite score.

357  
358 The twelve independent patient variables incorporated into the CHARM score were  
359 chosen for their objectivity, clinical availability, and relevance to cardiac illness. Each  
360 variable was determined to significantly contribute to the predictive value of waitlist  
361 mortality. Laboratory values were selected for their evaluation of crucial organ function



362 in the setting of severe HF. A large meta-analysis of sixty-four models that predicted  
363 death or hospitalization from HF determined renal function is one of the most significant  
364 factors in these outcomes<sup>24</sup>. Renal function was included in the CHARM score through  
365 serum sodium and creatinine. Multiple studies have demonstrated worsened short-term  
366 mortality for HF patients related to low serum sodium<sup>25,26</sup>. Creatinine is a standard  
367 measure of renal function, often used as a surrogate for eGFR. While eGFR was  
368 considered as a measure of renal function, it is not a value directly recorded in the  
369 SRTR database and can be calculated differently by institutions. Both age and BNP are  
370 also known prognostic factors in severe HF patients<sup>17</sup>. While age influences a variety of  
371 physiological processes, advanced age in HF (>65 years) affects vascular resistance  
372 and heart rate responsiveness, likely due to increased circulating norepinephrine  
373 levels<sup>27</sup>. BNP is a biomarker exclusively produced by cardiac tissue. It serves as an  
374 objective marker of cardiac stretch since it is influenced by level of end-diastolic volume,  
375 and it can also be an indicator of responsiveness to diuretic management<sup>28</sup>.

376  
377 In addition to physiologic data, indicator variables such as previous cardiac surgery,  
378 previous cardiac transplant, ventilator support, and life support proved to be large  
379 contributors to pre-transplant mortality. These interventions often serve as a “bridge” to  
380 transplant, reserved for the most critically ill patients. Previous cardiac surgeries or  
381 transplants are also indicative of a more extensive history of cardiac illness. We  
382 specifically chose not to include variables that depend upon a physician’s practice or  
383 measurement, such as right heart catheterization data or the level of inotropic support.

384 These variables were excluded with the intent of reducing bias based on treatment  
385 variation.

386  
387 Though we present accurate predictive models, they must be viewed through the  
388 confines of the study limitations, including sample size and missing data. Sample size  
389 was dependent on the availability of data within the SRTR, reduced to 1,965  
390 participants to account for missing laboratory variables. A thorough missingness  
391 analysis was performed to account for selection biases; however, a larger population  
392 with complete laboratory data would increase the power of these models. Additionally,  
393 variables chosen for the model were limited by the type of data recorded within SRTR.  
394 As such, certain serum markers were not readily available within the database and  
395 therefore could not be used. Similarly, the specificity of some variables was limited. For  
396 example, the term “Life Support” is broad and included anyone on inotropic support  
397 alone or varying degrees of temporary or durable mechanical support. Therefore, “Life  
398 Support” currently serves as a binary variable within the model, understanding that  
399 different modalities of mechanical support may contribute differently to patient risk. The  
400 next steps for further utilization of the CHARM score will include simulation modeling  
401 and a prospective, multi-center validation study in which these additional variables are  
402 collected and analyzed.

403

#### 404 **Conclusion**

405 The Colorado Heart failure Acuity Risk Model (CHARM) score provides a novel,  
406 validated model with strong positive predictive value for short-term mortality among

407 patients waitlisted for cardiac transplantation. We anticipate the CHARM score will be  
408 useful in the era of continuous distribution to standardize organ allocation by providing  
409 an objective and intuitive system for stratifying waitlisted heart failure patients based on  
410 medical urgency.

411

412 **Author Contributions:** John Stephen Malamon (JSM) had full access to all the data in  
413 the study and takes responsibility for the integrity of the data and the accuracy of the  
414 data analysis.

415 *Concept and design: JSM*

416 *Acquisition and analysis of data: JSM*

417 *Interpretation of data: All authors.*

418 *Drafting of the manuscript: RDM, SYP, JRHH, JSM*

419 *Critical revision of the manuscript: All authors.*

420 *Statistical analysis: JSM*

421 *Obtained funding: BK, JRHH*

422 *Administrative, technical, or material support: RDM, JSM*

423 *Supervision: BK, JRHH, JSM*

424

425 **CONFLICTS OF INTEREST:**

426 The authors do not have any conflicts of interest to disclose.

427

428 **ACKNOWLEDGMENT:** The authors have not acknowledgment to make.

429

430 **5. REFERENCES:**

- 431 1. Stevenson LW. Crisis Awaiting Heart Transplantation: Sinking the Lifeboat. *JAMA Intern*  
432 *Med* 2015;175(8):1406-9. DOI: 10.1001/jamainternmed.2015.2203.
- 433 2. Maitra NS, Dugger SJ, Balachandran IC, Civitello AB, Khazanie P, Rogers JG. Impact of  
434 the 2018 UNOS Heart Transplant Policy Changes on Patient Outcomes. *JACC Heart*  
435 *Fail* 2023;11(5):491-503. DOI: 10.1016/j.jchf.2023.01.009.
- 436 3. Cogswell R, John R, Estep JD, et al. An early investigation of outcomes with the new  
437 2018 donor heart allocation system in the United States. *J Heart Lung Transplant*  
438 2020;39(1):1-4. DOI: 10.1016/j.healun.2019.11.002.
- 439 4. Kilic A, Hickey G, Mathier MA, et al. Outcomes of the First 1300 Adult Heart Transplants  
440 in the United States After the Allocation Policy Change. *Circulation* 2020;141(20):1662-  
441 1664. DOI: 10.1161/CIRCULATIONAHA.119.045354.
- 442 5. Pelzer KM, Zhang KC, Lazenby KA, et al. The Accuracy of Initial U.S. Heart Transplant  
443 Candidate Rankings. *JACC Heart Fail* 2023;11(5):504-512. DOI:  
444 10.1016/j.jchf.2023.02.005.
- 445 6. Lin EY, Cohen HW, Bhatt AB, et al. Predicting Outcomes Using the Heart Failure  
446 Survival Score in Adults with Moderate or Complex Congenital Heart Disease. *Congenit*  
447 *Heart Dis* 2015;10(5):387-95. DOI: 10.1111/chd.12229.
- 448 7. Levy WC, Mozaffarian D, Linker DT, et al. The Seattle Heart Failure Model: prediction of  
449 survival in heart failure. *Circulation* 2006;113(11):1424-33. DOI:  
450 10.1161/CIRCULATIONAHA.105.584102.
- 451 8. Goda A, Williams P, Mancini D, Lund LH. Selecting patients for heart transplantation:  
452 comparison of the Heart Failure Survival Score (HFSS) and the Seattle heart failure  
453 model (SHFM). *J Heart Lung Transplant* 2011;30(11):1236-43. DOI:  
454 10.1016/j.healun.2011.05.012.

- 455 9. Lustig DB, Rodriguez R, Wells PS. Implementation and validation of a risk stratification  
456 method at The Ottawa Hospital to guide thromboprophylaxis in ambulatory cancer  
457 patients at intermediate-high risk for venous thrombosis. *Thromb Res* 2015;136(6):1099-  
458 102. DOI: 10.1016/j.thromres.2015.08.002.
- 459 10. Pocock SJ, Ariti CA, McMurray JJ, et al. Predicting survival in heart failure: a risk score  
460 based on 39 372 patients from 30 studies. *Eur Heart J* 2013;34(19):1404-13. DOI:  
461 10.1093/eurheartj/ehs337.
- 462 11. Austin MT, Poulouse BK, Ray WA, Arbogast PG, Feurer ID, Pinson CW. Model for end-  
463 stage liver disease: did the new liver allocation policy affect waiting list mortality? *Arch*  
464 *Surg* 2007;142(11):1079-85. DOI: 10.1001/archsurg.142.11.1079.
- 465 12. OPTN. A Guide to Calculating and Interpreting the Estimated Post-Transplant Survival  
466 (EPTS) Score Used in the Kidney Allocation System (KAS). April 21, 2020  
467 ([https://optn.transplant.hrsa.gov/media/1511/guide\\_to\\_calculating\\_interpreting\\_epts.pdf](https://optn.transplant.hrsa.gov/media/1511/guide_to_calculating_interpreting_epts.pdf)).
- 468 13. Johnston R, Jones K, Manley D. Confounding and collinearity in regression analysis: a  
469 cautionary tale and an alternative procedure, illustrated by studies of British voting  
470 behaviour. *Qual Quant* 2018;52(4):1957-1976. DOI: 10.1007/s11135-017-0584-6.
- 471 14. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP. Random forest: a  
472 classification and regression tool for compound classification and QSAR modeling. *J*  
473 *Chem Inf Comput Sci* 2003;43(6):1947-58. DOI: 10.1021/ci034160g.
- 474 15. Colvin M, Smith JM, Ahn Y, et al. OPTN/SRTR 2020 Annual Data Report: Heart. *Am J*  
475 *Transplant* 2022;22 Suppl 2:350-437. DOI: 10.1111/ajt.16977.
- 476 16. Miller PE, Mullan CW, Chouairi F, et al. Mechanical ventilation at the time of heart  
477 transplantation and associations with clinical outcomes. *Eur Heart J Acute Cardiovasc*  
478 *Care* 2021;10(8):843-851. DOI: 10.1093/ehjacc/zuab063.

- 479 17. Santaguida PL, Don-Wauchope AC, Oremus M, et al. BNP and NT-proBNP as  
480 prognostic markers in persons with acute decompensated heart failure: a systematic  
481 review. *Heart Fail Rev* 2014;19(4):453-70. DOI: 10.1007/s10741-014-9442-y.
- 482 18. Cox DR. *Regression Models and Life-Tables*. Breakthroughs in Statistics Methodology  
483 and Distribution. New York: Springer; 1992:527-541.
- 484 19. Harrell FE, Jr., Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical  
485 tests. *JAMA* 1982;247(18):2543-6. (<https://www.ncbi.nlm.nih.gov/pubmed/7069920>).
- 486 20. Malamon JS, Jackson WE, Saben JL, et al. A model for calculating the long-term  
487 estimated post-transplant survival of deceased donor liver transplant patients.  
488 *EBioMedicine* 2023;90:104505. DOI: 10.1016/j.ebiom.2023.104505.
- 489 21. Team RC. *R: A language and environment for statistical computing (Version 4.0. 2)*. R  
490 Foundation for Statistical Computing 2020.
- 491 22. Pedersen AB, Mikkelsen EM, Cronin-Fenton D, et al. Missing data and multiple  
492 imputation in clinical epidemiological research. *Clin Epidemiol* 2017;9:157-166. DOI:  
493 10.2147/CLEP.S129785.
- 494 23. Evans MD, Diaz J, Adamusiak AM, et al. Predictors of Survival After Liver  
495 Transplantation in Patients With the Highest Acuity (MELD  $\geq$ 40). *Ann Surg*  
496 2020;272(3):458-466. DOI: 10.1097/SLA.0000000000004211.
- 497 24. Rahimi K, Bennett D, Conrad N, et al. Risk prediction in patients with heart failure: a  
498 systematic review and analysis. *JACC Heart Fail* 2014;2(5):440-6. DOI:  
499 10.1016/j.jchf.2014.04.008.
- 500 25. Gheorghiade M, Abraham WT, Albert NM, et al. Relationship between admission serum  
501 sodium concentration and clinical outcomes in patients hospitalized for heart failure: an  
502 analysis from the OPTIMIZE-HF registry. *Eur Heart J* 2007;28(8):980-8. DOI:  
503 10.1093/eurheartj/ehl542.

- 504 26. Klein L, O'Connor CM, Leimberger JD, et al. Lower serum sodium is associated with  
505 increased short-term mortality in hospitalized patients with worsening heart failure:  
506 results from the Outcomes of a Prospective Trial of Intravenous Milrinone for  
507 Exacerbations of Chronic Heart Failure (OPTIME-CHF) study. *Circulation*  
508 2005;111(19):2454-60. DOI: 10.1161/01.CIR.0000165065.82609.3D.
- 509 27. Cody RJ, Torre S, Clark M, Pondolfino K. Age-related hemodynamic, renal, and  
510 hormonal differences among patients with congestive heart failure. *Arch Intern Med*  
511 1989;149(5):1023-8. (<https://www.ncbi.nlm.nih.gov/pubmed/2655542>).
- 512 28. Mueller C, McDonald K, de Boer RA, et al. Heart Failure Association of the European  
513 Society of Cardiology practical guidance on the use of natriuretic peptide concentrations.  
514 *Eur J Heart Fail* 2019;21(6):715-731. DOI: 10.1002/ejhf.1494.
- 515

516 **FIGURE LEGENDS:**

517

518 **Figure 1. Correlation Heatmap of Independent Variables.** Pearson's correlation  
519 coefficients were calculated using all patients (N=1,965) and independent variables  
520 (N=12) used to construct the three models (30D, 90D, and 1Y). This is used to provide a  
521 measure of collinearity. Blue indicates a positive correlation, and red indicates a  
522 negative correlation. The color saturation and circle area increase as the correlation  
523 coefficients increase in magnitude.

524

525 **Figure 2. Calibration Plots for 30-day, 90-day, and 1-year Models.** Patient risk  
526 scores (PRS) were calculated for all patients in the discovery cohort (N=1,179) and are  
527 provided as a function of observed patient mortality rate per tier for the 30-day, 90-day,  
528 and 1-year models. 30-day, 90-day, and 1-year models are labeled A, B, and C,  
529 respectively. The goodness-of-fit (R) was calculated using linear regression and was  
530 greater than 0.99 for all three models.

531

532 **Figure 3. Area Under the Receiver Operating Characteristic Curve using**  
533 **Interaction Terms.** Logistic regression was performed using the generalized linear  
534 model with interaction terms for the 30-day (A), 90-day (B), and 1-year (C) models. The  
535 area under the receiver operating characteristic curve or AUC was reported at 0.964,  
536 0.904, and 0.78 for the 30-day, 90-day, and 1-year models, respectively.

537



538 TABLES:

<b>Table 1. Patient Characteristics and Laboratory Values for Waitlisted Heart Transplant Patients, 2008 to 2022</b>					
		<b>Discovery (N=1,179)</b>	<b>Validation (N=786)</b>	<b>All (N=1,965)</b>	<b>p-value</b>
<b>Age</b>					
	Mean (SD)	48.9 (14.6)	47.8 (14.2)	48.5 (14.5)	0.243
	Median [Min, Max]	53.0 [18.0, 74.0]	50.0 [18.0, 73.0]	52.0 [18.0, 74.0]	
<b>Sex</b>					
	Female	312 (26.6%)	214 (27.3%)	526 (26.9%)	0.94
	Male	862 (73.4%)	570 (72.7%)	1432 (73.1%)	
<b>Race</b>					
	Caucasian	851 (72.5%)	551 (70.3%)	1402 (71.6%)	0.995
	Hispanic/Latino	57 (4.9%)	47 (6.0%)	104 (5.3%)	
	Black or African American	222 (18.9%)	152 (19.4%)	374 (19.1%)	
	Asian	31 (2.6%)	25 (3.2%)	56 (2.9%)	
	American Indian or Alaska Native	4 (0.3%)	2 (0.3%)	6 (0.3%)	
	Native Hawaiian or Pacific Islander	5 (0.4%)	4 (0.5%)	9 (0.5%)	
<b>Ethnicity</b>					
	Latino	59 (5.0%)	47 (6.0%)	106 (5.4%)	0.65
	Non-Latino or unknown	1115 (95.0%)	737 (94.0%)	1852 (94.6%)	
<b>Education</b>					
	High School (9-12)	394 (33.6%)	252 (32.1%)	646 (33.0%)	0.994
	Attended College/Technical School	324 (27.6%)	225 (28.7%)	549 (28.0%)	
	Associate/Bachelor Degree	251 (21.4%)	164 (20.9%)	415 (21.2%)	
	Post-college Graduate Degree	114 (9.7%)	88 (11.2%)	202 (10.3%)	
	Grade School (0-8)	21 (1.8%)	15 (1.9%)	36 (1.8%)	
	None	70 (6.0%)	40 (5.1%)	110 (5.6%)	

<b>Independent Variables</b>					
<b>Albumin</b>					
	Mean (SD)	3.59 (0.578)	3.59 (0.611)	3.59 (0.591)	0.991
	Median [Min, Max]	3.60 [1.70, 6.40]	3.60 [0.7, 6.50]	3.60 [0.7, 6.50]	
<b>AST</b>					
	Mean (SD)	39.4 (74.0)	39.8 (83.1)	39.6 (77.7)	0.993
	Median [Min, Max]	27.0 [0.1, 1640]	26.0 [0.1, 1810]	26.5 [0.1, 1810]	
<b>Bilirubin</b>					
	Mean (SD)	0.975 (1.83)	1.09 (1.77)	1.02 (1.80)	0.377
	Median [Min, Max]	0.7 [0.1, 40.6]	0.7 [0.1, 34.7]	0.7 [0.1, 40.6]	
<b>BNP</b>					
	Mean (SD)	1690 (2280)	1560 (2180)	1640 (2240)	0.478
	Median [Min, Max]	776 [0, 10000]	696 [5, 10000]	741 [0, 10000]	
<b>Cardiac Surgery</b>					
	Yes	506 (43.1%)	332 (42.3%)	838 (42.8%)	0.947
	No	668 (56.9%)	452 (57.7%)	1120 (57.2%)	
<b>Creatinine</b>					
	Mean (SD)	1.49 (1.16)	1.53 (1.14)	1.51 (1.15)	0.725
	Median [Min, Max]	1.24 [0.2, 16.1]	1.25 [0.14, 10.8]	1.25 [0.14, 16.1]	
<b>Circulatory Support</b>					
	Yes	436 (37.1%)	296 (37.8%)	732 (37.4%)	0.962
	No	738 (62.9%)	488 (62.2%)	1226 (62.6%)	
<b>Previous Transplant</b>					
	Yes	97 (8.3%)	75 (9.6%)	172 (8.8%)	0.607
	No	1077 (91.7%)	709 (90.4%)	1786 (91.2%)	
<b>Sodium</b>					
	Mean (SD)	135 (4.39)	135 (4.63)	135 (4.50)	0.13
	Median [Min, Max]	136 [117, 150]	136 [109, 149]	136 [109, 150]	
<b>Ventilator</b>					
	Yes	18 (1.5%)	5 (0.6%)	23 (1.2%)	0.197
	No	1156 (98.5%)	779 (99.4%)	1935 (98.8%)	
<b>Waitlist Days</b>					
	Mean (SD)	260 (388)	268 (412)	263 (398)	0.899
	Median [Min, Max]	110 [1.00, 4090]	119 [1.00, 3370]	113 [1.00, 4090]	
<i>AST=aspartate aminotransferase; BNP=Brain natriuretic peptide</i>					

539

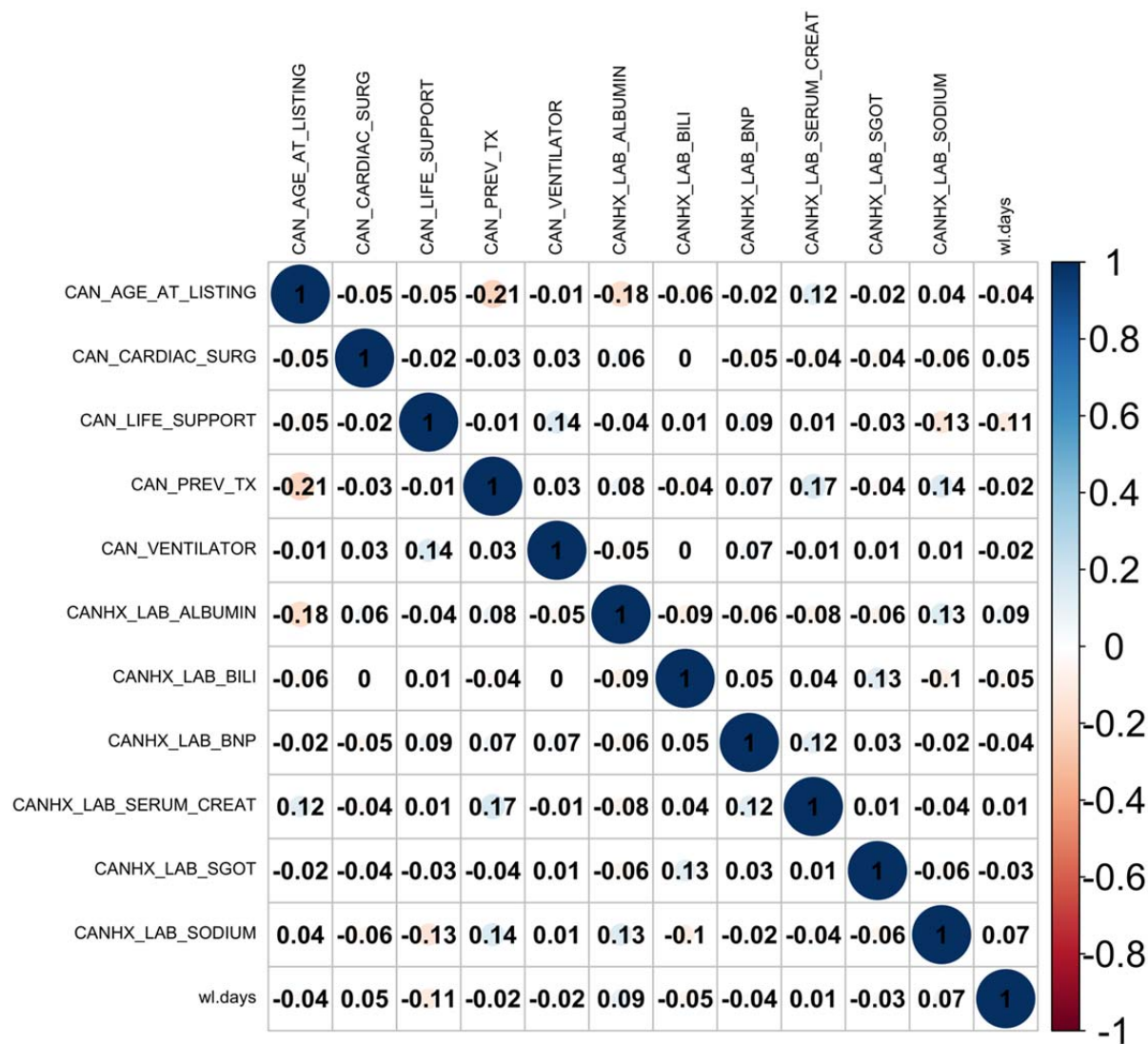
**Table 2. Tier Risk Index Precision for 30-day, 90-day, and 1-year Mortality Models**

<b>Tier</b>	<b>30-day Precision (PPV)</b>	<b>90-day Precision</b>	<b>1-year Precision</b>
7	0.983	0.942	0.857
6	0.987	0.927	0.862
5	0.989	0.924	0.825
4	0.993	0.837	0.668
3	0.986	0.969	0.966
2	1.0	0.985	0.985
1	1.0	1.0	1.0

*PPV = Positive Predictive Value*

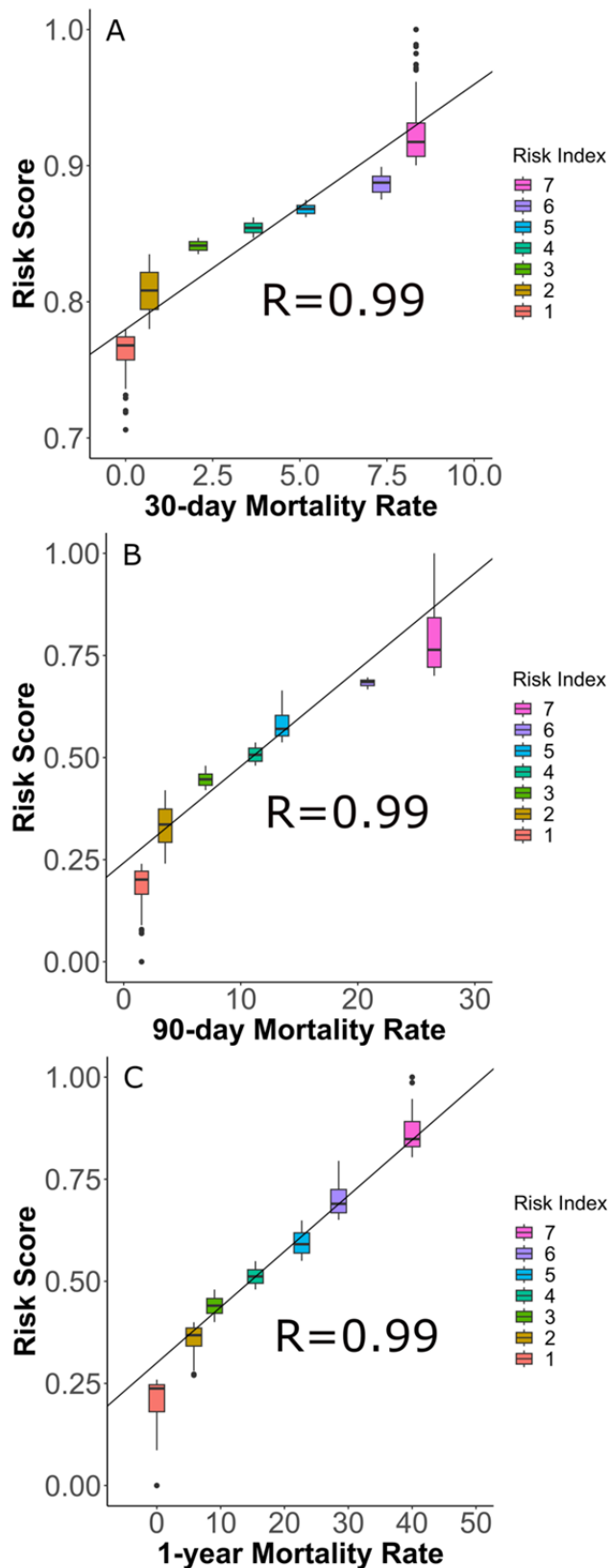
540

541 **Figure 1. Correlation Heatmap of Independent Variables.**



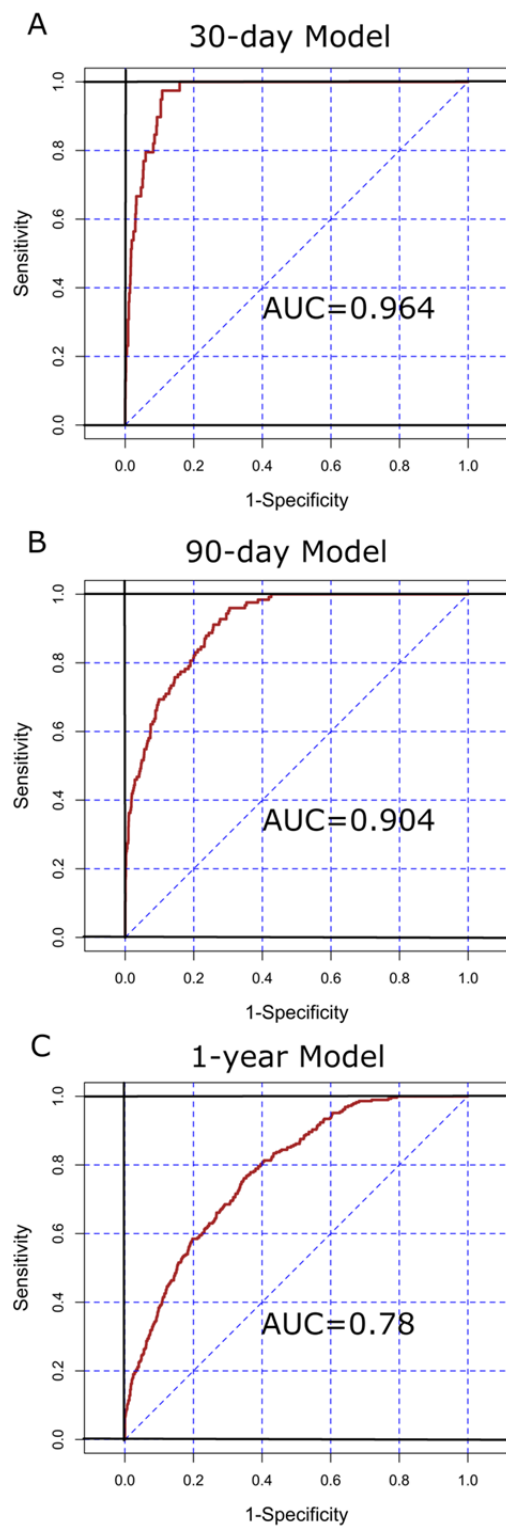
542

543 **Figure 2. Calibration Plots for 30-day, 90-day, and 1-year Models.**  
544



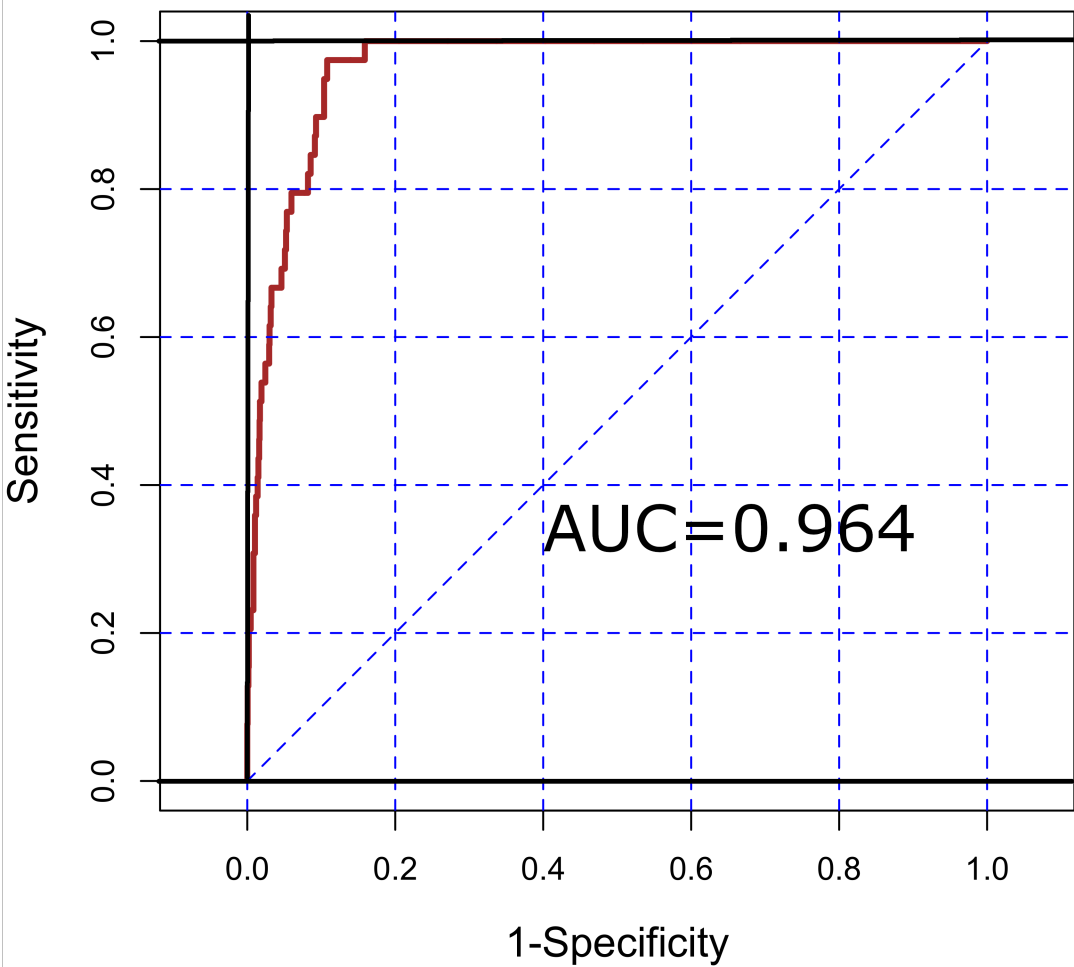
545

546 **Figure 3. Area Under the Receiver Operating Characteristic Curve using**  
547 **Interaction terms.**  
548



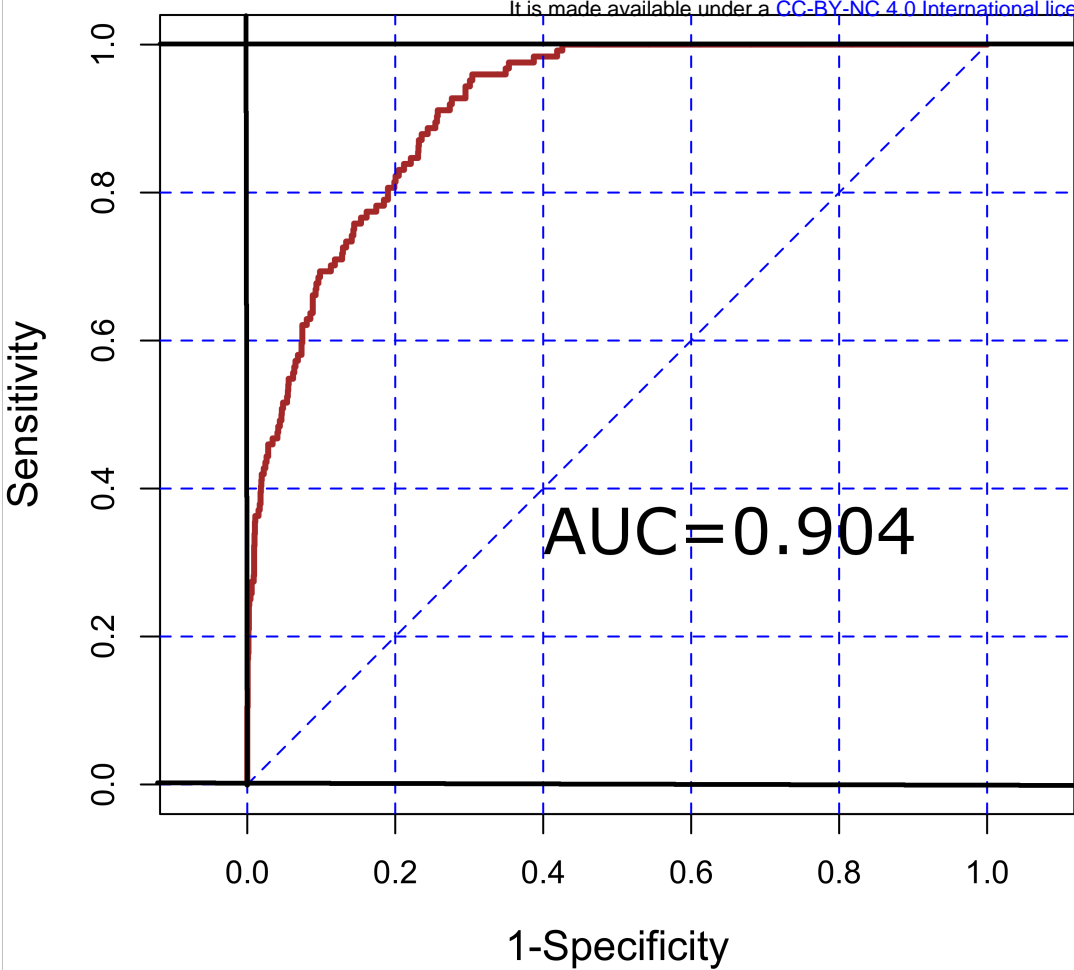
549

# A 30-day Model



# B 90-day Model

medRxiv preprint doi: <https://doi.org/10.1101/2023.08.30.23294870>; this version posted September 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-NC 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).



# C 1-year Model

