ANALYSIS OF RACIAL DIFFERENCES IN ICD SHOCK BURDEN USING COX PROPORTIONAL HAZARDS REGRESSION

by

Matthew P. Johnson

B.A. Mathematics, St. John's University, 2010

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This thesis was presented

by

Matthew P. Johnson

It was defended on

June 2, 2016

and approved by

Thesis Advisor:

Jeanine M. Buchanich, M.Ed., Ph.D, Research Assistant Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Committee Members:

Ada O. Youk, Ph.D., Associate Professor of Biostatistics, Epidemiology and Clinical & Translational Science, Department of Biostatistics, Graduate School of Public Health, School of Medicine, University of Pittsburgh

Indrani Halder, Ph.D., Assistant Professor of Medicine, Division of Cardiology, Heart and Vascular Institute, University of Pittsburgh Medical Center

Jeanine M. Buchanich, M.Ed., Ph.D

ANALYSIS OF RACIAL DIFFERENCES IN ICD SHOCK BURDEN USING COX PROPORTIONAL HAZARDS REGRESSION

Matthew P. Johnson, M.S.

University of Pittsburgh, 2016

ABSTRACT

Background/Research Aims: Sudden cardiac arrest (SCA) and sudden cardiac death (SCD) are major public health issues that most often result in death and affect a significant number of adults. Implantable cardiac defibrillators (ICDs) are an important tool used in SCA/SCD prevention. Racial differences with respect to ICD firings have not been very well explored with only limited previous research on the topic. This study seeks to analyze a large dataset of patients with defibrillators by race and also build a risk model for ICD shock using Cox proportional hazards regression.

Methods: Using data from the GRADE study, N= 1770 patients (1449 Whites and 321 African Americans) were initially compared by race for almost 80 baseline variables. N= 1524 patients (1275 Whites and 249 African Americans) had information on ICD shocks and were compared by race via Kaplan-Meier survival curves. A Cox proportional hazard regression analysis was performed to produce hazard ratios and evaluate each baseline characteristic with respect to defibrillator shock over time individually, followed by a multivariable model building process to examine race in the context of other significant covariates. Missing data was also examined in the variables comprising the final model and multiple imputations were performed for any variables deemed to have excessive missingness.

Results: Overall, African Americans were younger, had more nonischemic cardiomyopathy, had a higher prevalence of hypertension, smoked more, and had a lower ejection fraction than whites in the study. In comparing Kaplan-Meier survival curves, African Americans had a higher burden of shocks over time than whites by the end of the 5 year time period ($p<0.01$). Over 20 variables were individually related to shock at 60 months in a Cox regression setting (including race). The final multivariable model consisted of seven variables: race (not statistically significant but forced into the model), ejection fraction, history of NSVT, antiarrhythmic medication, diagnosis, age, and BUN. One variable was found to have excessive missingness (BUN) and after performing multiple imputations, results were overall similar in the second analysis except race became statistically significant ($p= 0.04$, HR= 1.385) while BUN went from statistically significant to not.

Conclusions: The public health impact of this study is in trying to build on the limited previous research and paint a more complete portrait of the relationship between race and appropriate defibrillator firings. After accounting for data missingness, race was found to be statistically significant in its relationship to ICD shock over time when adjusting for several additional covariates in the final multivariable model.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would first like to sincerely thank my thesis and academic adviser Dr. Jeanine Buchanich. She has provided me with invaluable guidance navigating the Biostatistics program since I first arrived at Pitt. I very much appreciate all of the time she has dedicated to meeting with me and helping me work through everything from choosing classes my first semester to completing this project. She put me in a position to succeed from day one and for that I am forever grateful.

I would like to sincerely thank Dr. Indrani Halder for introducing me to the topic of my thesis and the GRADE dataset. She provided the foundation for this project in asking the fundamental research question of the study. She has helped to guide and review my research process from the beginning for which I am very grateful. Dr. Halder additionally facilitated meetings with Dr. Samir Saba of UPMC Heart and Vascular Institute, who very kindly provided a clinical perspective for my research and I also would like to thank.

Thank you to Dr. Ada Youk for serving on my thesis committee and giving guidance on my project.

Lastly, thank you to all of the wonderful professors I have had for classes or heard give talks in the Biostatistics department.

1.0 INTRODUCTION

1.1 CARDIAC ARREST AND RACIAL DIFFERENCES

Sudden cardiac arrest (SCA) and sudden cardiac death (SCD) are major public health problems with significant associated mortality. SCA is the abrupt loss of heart function in a person. This results in blood not flowing to the brain and other vital organs and is a medical emergency that can result in SCD within minutes if not treated.

According to the American Heart Association in their 2015 update[1], over 300,000 adults experienced an EMS-assisted, out-of-hospital cardiac arrest in the United States that year with an estimated risk-adjusted incidence rate of 76 per 100,000 people and a median age of 66 years old. For the patients lucky enough to be treated by EMS, the survival rate to hospital discharge was still only 10.6%. As these statistics illustrate, SCA/SCD are very serious issues that can end a person's life very suddenly at a time when they still may have a significant amount of time left to live.

SCA is normally preceded by an abnormal heart rhythm called an arrhythmia, which is a disturbance or abnormality in the heart's electrical conduction system. Underlying causes and risk factors for a ventricular arrhythmia (the cause of most cardiac arrests) leading to SCA can include acute myocardial infarction (AMI), myocardial ischemia (lack of blood flow to the heart muscle),

severe electrolyte imbalances, drug toxicities, weakened left ventricular function, some heart rhythm issues like long Q-T syndrome, and congenital heart disease, among others.

Previous studies have shown that there is an increased risk of SCA/SCD in the African American population[1; 2; 3; 4; 5; 6] and also that African-Americans are less likely to survive an out of hospital cardiac arrest[1; 2; 3; 4; 5; 6]. These differences are most likely multifactorial[3] and have traditionally been attributed to common cardiac risk factors in the African American population such as hypertension, left ventricular hypertrophy (or LVH, an enlargement/thickening of the left ventricle), heart failure, obesity, and diabetes mellitus. Socioeconomic status (SES) and genetic factors also very likely play a role in the differing burdens between populations.

One example of a recent study illustrating these findings is the Oregon Sudden Unexpected Death Study (SUDS)[2]. This was a prospective observational study that took place in Portland, OR. Briefly, approximately 2000 patients (just under 200 African Americans) over a 10-year period who experienced a cardiac arrest were identified through paramedics, the state medical examiner, and state hospitals and compared by race for incidence rates, patient demographics, cardiac arrest circumstances, and clinical history. The study found that the burden of SCA was significantly higher among blacks compared to whites, with the age-adjusted incidence rates about two-fold as high for black men and women compared to whites. Blacks were also found to be younger at the time of their cardiac arrest and with a higher prearrest prevalence of hypertension, diabetes mellitus, and chronic renal insufficiency along with a higher prevalence of congestive heart failure (CHF) and LVH.

1.2 ICD THERAPY AND RACIAL DIFFERENCES

One important and common method of treatment for preventing cardiac arrest and sudden death in those patients at risk is the implantable cardioverter defibrillator (ICD). In broad terms, the ICD is an implanted device that works by tracking a patient's heart rate and rhythm and if something dangerous is detected, it delivers an electric shock to restore a normal heartbeat via the implanted wires that connect the device to the heart. ICDs are used as a form of both primary and secondary prevention for patients[7; 8]. Multiple studies have shown ICDs to be effective at decreasing patient mortality related to SCA in multiple situations[7; 8] and the devices are a normal course of treatment for appropriate patients.

The large majority of prior research studies examining and showing the effectiveness of ICDs in reducing the risk of SCD for patients have taken place in the context of mostly white cohorts, even though African Americans are at a higher risk for SCD than the general population. There is evidence to support a utilization disparity among African Americans who are eligible for an ICD[9; 10], however, outcomes as they relate to race have not been as well explored.

There are a few previous studies of note to review pertaining to racial differences and ICDs. The Sudden Cardiac Death in Heart Failure (SCD-HeFT) study[11] was a National Heart, Lung, and Blood institute sponsored, multi-center, randomized trial that demonstrated ICD therapy significantly improved survival compared with medical therapy alone in stable moderately symptomatic heart failure patients with a left ventricular ejection fraction (or LVEF, a measurement of the percentage of blood leaving the heart during each contraction) of <35%. A follow up study examined the outcomes in African Americans. Of the approximately 2500 patients enrolled in the study, 17% were African American. Baseline demographics, clinical variables, SES, and long-term outcomes were compared according to race. African Americans

were found to be younger and had more non-ischemic heart failure (heart failure caused by something other than coronary artery disease), lower EF's, worse NYHA functional class (a classification of the extent/severity of a patient's heart failure based on their symptoms and limitations), and higher prevalence of a history of non-sustained ventricular tachycardia (NSVT, a ventricular arrhythmia). No significant differences were found in the rate of ICD discharge by race. Adjusted mortality risk was significantly higher in African Americans compared with whites but mortality was equally reduced in both race groups receiving ICD.

The Multicenter Automatic Defibrillator Implantation Trial (MADIT-II) trial[12] involved approximately 1200 subjects (8% African American) and examined the effectiveness of ICD therapy. The study found that ICDs were associated with a significant reduction in mortality and sudden cardiac death in whites but not blacks. Even after adjusting for relevant covariates, the ICD therapy/conventional therapy hazard ratios for total mortality and sudden cardiac death showed a trend toward benefit in whites but not African Americans.

The Prospective Observational Study of Implantable Cardioverter-Defibrillators (PROSE-ICD) study[10] was a multicenter prospective observational study of patients with systolic heart failure who underwent ICD implantation for primary prevention of SCD conducted from 2003 to 2013. The study was conducted to determine the association between race and outcomes in this cohort of patients and the primary endpoint was appropriate ICD shock (a shock for a ventricular arrhythmia). There were just under 1200 patients in the study (~450 African Americans). African Americans were found to be at an increased risk of all-cause mortality without receiving an appropriate ICD shock, but not to be at increased risk for appropriate ICD shock. Ejection fraction appeared to explain over 20% of the excess risk of mortality in African Americans and diabetes and hypertension each were responsible for over 10%.

1.3 RESEARCH OBJECTIVES

This study seeks to build off of the limited previous research to examine the racial differences in patients with ICDs with respect to defibrillator shocks received. Many prior studies examining racial differences in SCA/SCD have only been descriptive in nature, and the prior research related specifically to racial differences and ICDs is also limited. This study looks to build a risk model relating race and appropriate ICD shock (defined as a shock for a ventricular arrhythmia) using Cox proportional hazards regression in examining a large cohort of patients with ICDs implanted. The main objective of the study is to build and use this model to further examine differences in shocks received by race.

2.0 METHODS

2.1 STUDY SAMPLE

The current study utilized data from the Genetic Risk Assessment of Defibrillator Events (GRADE) study. The GRADE study was a NHLBI sponsored prospective multi-center study that aimed to identify genetic modifiers of arrhythmic risk. The coordinating center was at the University of Pittsburgh Medical Center. Subjects with left ventricular systolic dysfunction (EF≤30%) and an ICD placed within the prior 6 months were enrolled. Recruitment took place from 2002 to 2010 and patients were followed annually for up to 5 years.

2.1.1 Sample size

There were initially 1823 patients included in the dataset. Thirty-three patients were missing a recorded race and were excluded from the analysis. An additional four patients had race of American Indian, ten patients had race of Asian, and four had race of "other". Due to their extremely small sample sizes and the desire to compare whites and African Americans directly, these patients was also excluded. Of the remaining 1772 patients, two had ages under 18 and were excluded. The sample size of 1770 unique patients with race of either white or African American was used for all descriptive analyses (321 African Americans, ~18.1%). For all additional analyses, a total of 246 patients were missing the outcome of interest (appropriate defibrillator shock at 60 months) and thus not included in the survival or model building process for a total N= 1524 patients (249 African Americans, ~16.3%).

2.2 PRIMARY OUTCOME

The primary outcome for this study was appropriate ICD shock at 60 months. Some patients were followed for longer than 60 months, however 60 months was the last planned measurement for the study and thus was set as the maximum time frame used for the outcome variable. Measurements were recorded both as a continuous variable noting the time to shock in months, as well as a binary yes/no shock value at 60 months.

2.2.1 Independent variables

There were almost 80 independent variables that were measured at baseline included in the dataset that were initially analyzed in this study. The variables included patient demographics (ex. age, gender), disease status (ex. cardiac diagnosis, NYHA class), past medical histories (ex. history of hypertension, history of diabetes), baseline clinical values (ex. systolic blood pressure, weight), medication profile (ex. beta blocker, diuretics), baseline lab values (ex. sodium, creatinine), baseline EKG values (ex. QRS interval, LVH), baseline echocardiogram values (ex. left ventricular ejection fraction, valvular diagnoses), and baseline heart catheterization values (ex. CAD, various heart pressures).

Comparisons between whites and African Americans were performed using Pearson's chi-squared test for all categorical variables and either t-test or Wilcoxon rank-sum test depending on normality for all continuous variables. Each variable is measured at baseline only and is not time dependent. Routine data maintenance was performed prior to the analysis with obvious data entry errors, such as incorrectly coded variables, and highly improbable medical/biological values, such

as negative numbers for a variable that cannot be negative or values several orders of magnitude away from all other values, dropped.

2.3 STATISTICAL ANALYSIS

2.3.1 Survival analysis

In order to examine the data for any potential overall trends in shock differences between races, a standard survival analysis was conducted for the outcome variable with the event/censoring times. The survival function, *S(t)*, was defined as the probability of an individual surviving, or not having an event, beyond time *t* (Mathematically, $S(t) = P(T>t)$). Survival curves were created with Kaplan-Meier estimates using the conditional survival probability formula:

$$\hat{S}(t) = \prod_{i=t_i \leq t} \left(1 - \frac{d_i}{n_i}\right)$$

where n_i represents the number of subjects at risk at event time $t_{(i)}$ and d_i represents the number of events at event time $t_{(i)}$. Survival curves were compared by race using a standard log-rank test, where the null hypothesis was that both curves were equal (H_0: $S_1(t)=S_2(t)$ for all t). The test statistic was calculated as follows:

$$Q = \frac{[\sum_{i=1}^{m}(d_{i1} - \widehat{e_{1i}})]^2}{\sum_{i=1}^{m} \widehat{v_{1i}}}$$

where d_{i1} represents the total number of deaths at event time $t_{(i)}$ in group 1, which follows a hypergeometric distribution with mean $\widehat{e_{1i}} = d_i(\frac{n_{i1}}{n_i})$ and variance $\widehat{v_{1i}} = d_i(\frac{n_i - d_i}{n_i - 1})(\frac{n_{1i}}{n_i})(1 - \frac{n_{1i}}{n_i})$. Q follows a chi-squared distribution with 1 degree of freedom.

2.3.2 Cox proportional hazards regression modeling

Cox proportional hazards regression models were used to model hazard functions and produce hazard ratios with regard to a patient receiving a defibrillator shock. The hazard function, as the conditional instantaneous failure rate at time *t*, is defined as $h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t+\Delta t | T \geq t)}{\Delta t}$. The Cox proportional hazard model is defined as

$$h(t; X) = h_0(t)\exp(X\beta)$$

where $h_0(t)$ is the baseline hazard function with covariate matrix X and parameter vector β. The hazard ratio (HR) for defibrillator shock between units of a specified covariate is easily calculated from the model as $HR = \exp(\beta)$. Of note, continuous covariates were centered at their mean to improve interpretability prior to model building.

Covariates were first tested univariably using Wald's test to assess individual significance as related to defibrillator shock at 60 months. From there, two-variable models were built to examine all univariably significant variables plus race. This was done to explore the relationship between race and the outcome adjusting for each covariate individually. Again, the Wald test was used to assess both individual covariate and overall model significance. A p-value <0.10 was considered significant for both the univariable and two-variable models.

After univariable and two-variable models were built, a multivariable model was constructed using a backwards stepwise selection process. Criterion for inclusion in the initial multivariable model was univariable significance at the p=0.10 level and also significance with race in the two-variable models. Variable removal was based on individual covariate Wald tests. A p-value>0.05 was considered not significant and eligible for removal. Covariates with the highest p-values were removed two at a time and the model was then rerun with two additional

variables removed. This process was repeated until all variables were statistically significant at the 0.05 level. The partial likelihood ratio test was performed between each model to confirm the lack of statistical significance for the removed variables. As race was the main covariate of interest in this study, it was forced into the model regardless of statistical significance.

Model diagnostics were performed on the final model. Overall model significance was assessed and the proportional hazards assumption was also tested. Additionally, multicollinearity between variables in the final model was assessed by VIF criteria.

Briefly, VIF reflects the degree to which the variances of other coefficients are increased due to the inclusion of covariate x_j. It is calculated as $VIF_j = \frac{1}{1-R_j^2}, j = 1,2,\dots,k$ where R_j^2 is the squared multiple correlation coefficient based on regressing x_j on the remaining k-1 predictors. A VIF of around 1 corresponds to no effect and a VIF larger than 10 was deemed to be of concern.

2.3.3 Missing data

Data missingness was examined for the variables that comprised the final multivariable model. Missing data are commonly categorized in three ways with terminology first conceived by Rubin[13]: Missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR).

For missing data to be MCAR, it must be totally random and independent of all observed and unobserved values. Formally (following notation from Little and Rubin)[14], Let $Y = (y_{ij})$ denote an (n x K) rectangular dataset without missing values, with the ith row $y_i = (y_{i1},\dots,y_{iK})$ where y_{ij} is the value of variable Y_j for subject i. With missing data, define the missing-data indicator matrix $M = (m_{ij})$, such that $m_{ij} = 1$ if y_{ij} is missing and $m_{ij} = 0$ if y_{ij} is present. The matrix M then defines

the pattern of missing data. The missing data mechanism is characterized by the conditional distribution of M given Y, say $f(M|Y,\phi)$, where ϕ denotes unknown parameters. If missingness does not depend on the values of the data Y, missing or observed, that is, if $f(M|Y,\phi) = f(M|\phi)$ for all Y, ϕ, the data are MCAR.

Missing data are MAR if there is a relationship between these data and observed values, but no relationship with unobserved values. Formally, now let Y_{obs} denote the observed components or entries of Y and Y_{miss} the missing components. Then if missingness depends only on the components Y_{obs} of Y that are observed, and not on the components that are missing, that is, $f(M|Y,\phi) = f(M|Y_{obs},\phi)$ for all Y_{miss}, ϕ, the data are MAR.

Finally, missing data are MNAR if they are neither MCAR or MAR. In this case, missingness is not independent and depends on more than the observed components. It depends on the unobserved values, such as the missing observations themselves. Valid estimation then requires that the missing-data mechanism be modeled as part of the estimation process.

Missingness that is truly MCAR can effectively be ignored because the missing data are simply a random sample of the original sample and analyses of the complete cases will be valid and unbiased (although will have lower statistical power due to the decreased sample size)[14; 15; 16]. MCAR was informally examined in this study by looking at the association between missingness for a certain variable in terms of the other variables in the model and was also formally tested with Little's test[17]. There is no test to distinguish between MAR and MNAR as this would depend on knowing the missing values themselves.

Both statistical packages used for this analysis, R and Stata, treat missing data the same- by using complete case analysis or listwise deletion. These are two terms that mean that any observations with any missing values for the variables used in the model are automatically

dropped. Excessive missingness in any variable therefore can become a problem by limiting the amount of information being used in the analysis. This will decrease the statistical power and precision and possibly introduce bias if the missing data are not MCAR, as well as lead to larger standard errors[14; 15; 16]. All of these issues can lead to incorrect results and inferences.

2.3.4 Multiple imputation

There are several common ways to deal with missing data besides complete case analysis. Another common approach is to impute, or fill in, missing values. Single imputation techniques, such as filling in the mean, "hot deck", or regression prediction, involve using the observed values to try to fill in a best guess for the missing values and then the analysis is carried forward in the usual way with both the observed and imputed observations. The major issue with single imputation, however, is that it does not take into account the uncertainty of the estimates which can lead to underestimates of standard errors[14; 15; 16]. This in turn affects p-values and confidence intervals. Variances for the variables with missing data tend to be underestimated as well and can lead to biased estimates of parameters[15; 16].

The solution to this problem is multiple imputation, which is appropriate for many situations including this study- an observational study that has covariates with missing observations[18]. Multiple imputation does account for the uncertainty in the estimates and produces parameter estimates that are approximately unbiased with correct standard errors[15; 16]. Multiple imputation follows a similar strategy to single imputation, but repeats the imputation process to simulate the missing data more than once. Data are assumed to be MAR under this approach. In general, there are three steps that are completed in this process:

1. *M* complete datasets are generated under the chosen imputation model.

2. The Cox proportional hazard regression model is run separately on each imputed dataset m=1,2,…,M.
3. The results from each analysis are pooled together to give a single model result.

This process was first developed by Rubin (1987).

Briefly, overall parameter estimates are created by taking the mean of all parameter estimates from the individual analyses of each imputed dataset. Overall standard error estimates are created in a few steps. First, the 'within' variance is created by taking the average of the squared standard errors in each analysis. The 'between' variance is the sample variance of the parameter estimates across the several analyses. One expression of the formula for combining the 'between' and 'within' variances given by Allison is as follows[15]:

$$\sqrt{\frac{1}{M}\sum_{m=1}^{M} s_m^2 + (1+\frac{1}{M})(\frac{1}{M-1})\sum_{m=1}^{M}(b_m - \bar{b})^2}$$

where M is the number of datasets, s_m is the standard error in the m^{th} dataset, and b_m is the parameter estimated in the m^{th} dataset.

For this study, the multiple imputation process was carried out in Stata using the mi package for any variable that needed to be imputed and any imputation was only carried out one variable at a time. Most literature, for example Rubin (1987), suggests that M=5 is sufficient to obtain valid inference (corresponding to asymptotic relative efficiency of 95% for 50% information missing). Increasing number of imputations produces increasing asymptotic efficiency[15] and M=20 imputations was chosen for this study.

The specific imputation method used was based on the type of variable that needed to be imputed, for example linear regression was used for any continuous variables with excessive missingness or logistic regression for any binary variables. The other variables comprising the

final model were chosen as the predictor variables for any imputation models. Briefly, each imputation in step 1 uses a slightly different regression model by modeling the coefficients from their own distributions. So the assigned values for the missing data are slightly different in each imputation to account for the inherent uncertainty in the estimates, and when the imputations are combined this uncertainty is taken into account.

3.0 RESULTS

3.1 DESCRIPTIVE ANALYSIS

3.1.1 Primary outcome

A total of 314 patients (20.6%) experienced a defibrillator shock by 60 months. Table 1 shows the events broken down by race. Just under 25% of African Americans received a shock and almost 20% of whites were shocked.

Table 1. Proportion of patients experiencing shock at 60 months by race

Shock at 60 months	**Total (N=1,524)**	**White (n=1,275)**	**AA (n=249)**
No	1,210 (79.4%)	1,021 (80.1%)	189 (75.9%)
Yes	314 (20.6%)	254 (19.9%)	60 (24.1%)

Among all individuals who received shocks, the average time to shock was 17.5 months with a standard deviation of 15.1 months. Whites were shocked at an average of 17.5 months and African Americans at 17.7 months. Table 2 shows time to shock broken down by race and Figure 1 illustrates these results as a histogram.

Table 2. Mean time to shock (in months) by race

	Total (N=314) **Mean (SD)**	**White (n=254)** **Mean (SD)**	**AA (n=60)** **Mean (SD)**
Time to shock (mo.)	17.5 (15.1)	17.5 (15.1)	17.7 (14.7)

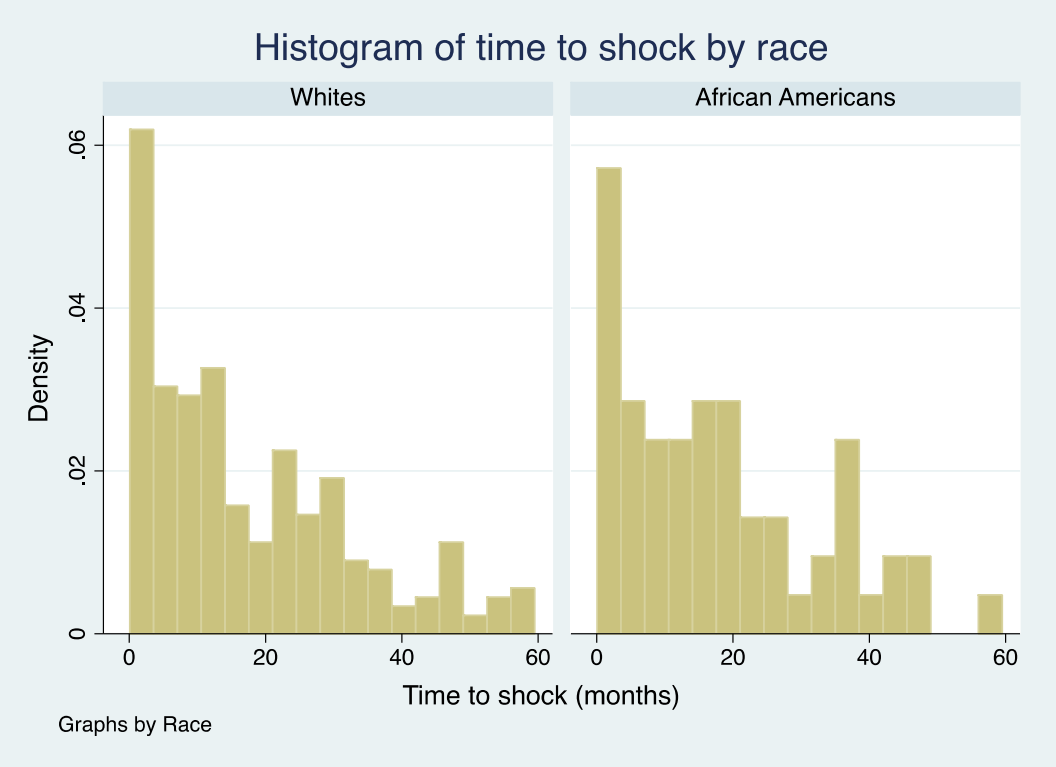

Figure 1. Histogram of time to shock by race

3.1.2 Independent variables

Each independent variable analyzed is presented in the tables below, grouped by variable category. Table 3 shows baseline patient demographics and disease status broken down by race. African Americans had a higher proportion of females and were younger overall. Whites had a higher proportion of patients with ischemic cardiomyopathy. NYHA class was also different by race.

Table 3. Patient Baseline Demographics and Disease Status

Covariate	Total (N=1770)	White (n=1449)	AA (n=321)	P-value
	N (%)	n (%)	n (%)	
Gender				<0.001
Male	1405 (79.4%)	1182 (81.6%)	223 (69.5%)	
Female	365 (20.6%)	267 (18.4%)	98 (30.5%)	
Dx				<0.001
Ischemic Cardiomyopathy	1240 (70.1%)	1084 (74.9%)	156 (48.4%)	
Idiopathic Cardiomyopathy	396 (22.4%)	259 (17.9%)	137 (42.5%)	
Myocarditis/Other	133 (7.5%)	105 (7.2%)	28 (8.7%)	
NYHA Class				0.04
1	230 (13.3%)	181 (12.8%)	49 (15.4%)	
2	950 (55.0%)	789 (55.9%)	161 (50.9%)	
3	529 (30.6%)	422 (29.9%)	107 (33.6%)	
4	19 (1.1%)	19 (1.3%)	0 (0%)	
Device Type				<0.001
ICD	73 (4.1%)	62 (4.3%)	11 (3.4%)	
Single chamber	396 (22.5%)	288 (20.0%)	108 (33.8%)	
Dual chamber	502 (28.5%)	443 (30.7%)	59 (18.4%)	
ICD/BiV	790 (44.9%)	648 (45.0%)	142 (44.4%)	
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	
Age (Years)	62.6 (12.2)	63.8 (11.6)	57.1 (13.4)	<0.001

Table 4 shows patient past medical histories by race. African Americans had a higher prevalence of hypertension and current smokers, but diabetes showed no statistically significant difference by race. Whites had a higher prevalence of prior heart attacks (MI).

Table 4. Patient Past Medical Histories

Covariate	Total (N=1770)	White (n=1449)	AA (n=321)	P-value
	N (%)	n (%)	n (%)	
Hx Myocardial Infarction				<0.001
No	812 (46.0%)	614 (42.5%)	198 (61.7%)	
Yes	954 (54.0%)	831 (57.5%)	123 (38.3%)	
Hx Diabetes Mellitus				0.32
No	1134 (64.1%)	936 (64.6%)	198 (61.7%)	
Yes	635 (35.9%)	512 (35.4%)	123 (38.3%)	
Hx Hypertension				<0.001
No	615 (34.8%)	570 (39.4%)	45 (14.0%)	
Yes	1154 (65.2%)	878 (60.6%)	276 (86.0%)	
Hx Hypercholesterolemia				<0.001
No	609 (34.5%)	468 (32.4%)	141 (43.8%)	
Yes	1156 (65.5%)	976 (67.6%)	180 (56.2%)	
Hx Smoking				0.81
No	832 (47.1%)	680 (47.0%)	152 (47.5%)	
Yes	936 (52.9%)	768 (53.0%)	168 (52.5%)	
Current Smoker				<0.01
No	789 (80.0%)	660 (81.8%)	129 (72.1%)	
Yes	197 (20.0%)	147 (18.2%)	50 (27.9%)	
Hx PVD				0.16
No	1607 (91.1%)	1309 (90.7%)	298 (93.1%)	
Yes	157 (8.9%)	135 (9.3%)	22 (6.9%)	
Hx Atrial Tachycardia				<0.01
No	1098 (62.4%)	875 (60.7%)	223 (70.1%)	
Yes	661 (37.6%)	566 (39.3%)	95 (29.9%)	
FHx Sudden Death				0.90
No	1313 (74.4%)	1075 (74.5%)	238 (74.1%)	
Yes	451 (25.6%)	368 (25.5%)	83 (25.9%)	

Table 4 Continued				
Hx Syncope				<0.001
No	1269 (71.7%)	1073 (74.1%)	196 (61.1%)	
Yes	500 (28.3%)	375 (25.9%)	125 (38.9%)	
Hx SVT				0.79
No	1362 (77.1%)	1117 (77.2%)	245 (76.6%)	
Yes	404 (22.9%)	329 (22.8%)	75 (23.4%)	
Hx NSVT				0.31
No	1054 (59.7%)	871 (60.3%)	183 (57.2%)	
Yes	711 (40.3%)	574 (39.7%)	137 (42.8%)	
Hx Ablation				0.08
No	1596 (90.2%)	1298 (89.6%)	298 (92.8%)	
Yes	173 (9.8%)	150 (10.4%)	23 (7.2%)	
Hx CABG				<0.001
No	1141 (64.5%)	875 (60.4%)	266 (82.9%)	
Yes	627 (35.5%)	572 (39.6%)	55 (17.1%)	
Hx Valve				0.23
No	1619 (91.7%)	1321 (91.4%)	298 (93.4%)	
Yes	146 (8.3%)	125 (8.6%)	21 (6.6%)	
Prior Shock				0.02
No	1414 (80.4%)	1173 (81.5%)	241 (75.5%)	
Yes	345 (19.6%)	267 (18.5%)	78 (24.5%)	
Mean (SD)	**Mean (SD)**	**Mean (SD)**		
Pack Years	45.4 (47.8)	47.1 (43.9)	38.9 (60.2)	<0.001
# of Episodes	5.0 (10.7)	5.2 (11.4)	4.6 (8.8)	0.14
# of Grafts	3.4 (1.4)	3.5 (1.4)	3.1 (1.3)	0.08

Table 5 shows patient vital signs by race. African Americans had higher blood pressures and baseline heart rate but weight was not different by race.

Table 5. Patient Baseline Vital Signs

Covariate	Total (N=1770)	White (n=1449)	AA (n=321)	P-value
	Mean (SD)	Mean (SD)	Mean (SD)	
Heart Rate	73.0 (12.4)	72.5 (12.6)	74.9 (11.3)	<0.01
SBP	119.8 (19.3)	118.9 (18.5)	123.6 (22.3)	<0.01
DBP	71.5 (11.8)	70.6 (11.4)	75.4 (13.0)	<0.001
Weight	197.2 (48.6)	196.1 (46.0)	202.2 (58.7)	0.08
Height	68.7 (3.8)	68.8 (3.7)	68.2 (4.3)	<0.01

Table 6 highlights patient's medication profiles. A higher proportion of African Americans were on beta blockers and diuretics while there was no difference by race for antiarrhythmic medications.

Table 6. Patient Baseline Medications

Covariate	Total (N=1770)	White (n=1449)	AA (n=321)	P-value
	N (%)	n (%)	n (%)	
ACE				0.02
No	595 (33.8%)	469 (32.5%)	126 (39.4%)	
Yes	1167 (66.2%)	973 (67.5%)	194 (60.6%)	
A2 Blocker				0.48
No	1498 (85.0%)	1230 (85.3%)	268 (83.8%)	
Yes	264 (15.0%)	212 (14.7%)	52 (16.2%)	
Beta blocker				0.01
No	261 (14.8%)	228 (15.8%)	33 (10.3%)	
Yes	1504 (85.2%)	1217 (84.2%)	287 (89.7%)	

Table 6 Continued				
Aldactone				<0.001
No	1289 (73.0%)	1098 (76.0%)	191 (59.5%)	
Yes	476 (27.0%)	346 (24.0%)	130 (40.5%)	
Antiarrhythmic				0.11
No	1381 (78.2%)	1119 (77.5%)	262 (81.6%)	
Yes	384 (21.8%)	325 (22.5%)	59 (18.4%)	
Diuretic				<0.01
No	478 (27.1%)	412 (28.6%)	66 (20.6%)	
Yes	1285 (72.9%)	1030 (71.4%)	255 (79.4%)	
Digoxin				0.01
No	995 (56.6%)	794 (55.2%)	201 (62.8%)	
Yes	763 (43.4%)	644 (44.8%)	119 (37.2%)	
Nitrates				0.13
No	901 (74.5%)	724 (75.5%)	177 (70.8%)	
Yes	308 (25.5%)	235 (24.5%)	73 (29.2%)	
Hydralazine				<0.001
No	1090 (92.7%)	898 (97.1%)	192 (76.5%)	
Yes	86 (7.3%)	27 (2.9%)	59 (23.5%)	

Table 7 shows patient laboratory values by race. African Americans had higher creatinine but lower BUN which seems to paint a mixed picture about any differences in kidney function.

Table 7. Patient Baseline Laboratory Values

Covariate	**Total (N=1770)**	**White (n=1449)**	**AA (n=321)**	**P-value**
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	
Ca	9.1 (0.8)	9.0 (0.8)	9.2 (0.6)	0.01
Mg	1.9 (0.3)	1.9 (0.3)	1.8 (0.3)	<0.01
Na	138.2 (3.2)	138.2 (3.2)	138.1 (3.2)	0.56
Creat	1.4 (1.0)	1.4 (1.0)	1.5 (1.3)	0.01
BUN	23.6 (13.2)	24.0 (13.0)	22.1 (13.8)	<0.001

Baseline EKG values by race are shown in Table 8. Whites had a higher proportion of atrial fibrillation, and African Americans had a higher proportion of LVH.

Table 8. Patient Baseline EKG Values

Covariate	Total (N=1770)	White (n=1449)	AA (n=321)	P-value
	N (%)	n (%)	n (%)	
Atrial Fibrillation				0.02
No	1473 (90.3%)	1195 (89.4%)	278 (93.9%)	
Yes	159 (9.7%)	141 (10.6%)	18 (6.1%)	
LVH				<0.001
No	1483 (91.2%)	1245 (93.6%)	238 (80.4%)	
Yes	143 (8.8%)	85 (6.4%)	58 (19.6%)	
RVH				0.31
No	1615 (99.1%)	1323 (99.2%)	292 (98.6%)	
Yes	14 (0.9%)	10 (0.8%)	4 (1.4%)	
AV Block				0.67
No	1451 (89.1%)	1186 (89.0%)	265 (89.8%)	
Yes	177 (10.9%)	147 (11.0%)	30 (10.2%)	
RBBB				0.45
No	1517 (93.2%)	1239 (93.0%)	278 (94.2%)	
Yes	110 (6.8%)	93 (7.0%)	17 (5.8%)	
LBBB				0.10
No	1407 (86.5%)	1143 (85.9%)	264 (89.5%)	
Yes	219 (13.5%)	188 (14.1%)	31 (10.5%)	
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	
PR	169.3 (44.7)	170.0 (45.6)	166.6 (41.3)	0.25
QRS	132.2 (44.1)	134.8 (42.9)	120.5 (47.6)	<0.001
QTC	473.0 (53.4)	472.6 (53.4)	474.6 (53.5)	0.57

Table 9 shows baseline echocardiogram values by race. African Americans had a lower left ventricular ejection fraction and higher proportion of patients with a valvular diagnosis.

Table 9. Patient Baseline Echocardiogram Values

Covariate	**Total (N=1770)**	**White (n=1449)**	**AA (n=321)**	**P-value**
	N (%)	**n (%)**	**n (%)**	
RV Enlarged				<0.001
No	1008 (79.3%)	891 (81.1%)	117 (68.0%)	
Yes	263 (20.7%)	208 (18.9%)	55 (32.0%)	
RV Hypokinetic				0.04
No	1011 (79.2%)	885 (80.2%)	126 (73.3%)	
Yes	265 (20.8%)	219 (19.8%)	46 (26.7%)	
Pleural Effusion				0.01
No	1166 (91.0%)	1019 (91.8%)	147 (85.5%)	
Yes	116 (9.0%)	91 (8.2%)	25 (14.5%)	
Valvular Dx				0.02
No	375 (28.7%)	339 (29.9%)	36 (20.9%)	
Yes	931 (71.3%)	795 (70.1%)	136 (79.1%)	
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	
LVEF	0.21 (0.06)	0.21 (0.06)	0.19 (0.07)	<0.001
LV Diastole	6.3 (1.0)	6.3 (1.0)	6.6 (1.0)	<0.01
LV Systole	5.4 (1.1)	5.4 (1.1)	5.6 (1.2)	0.04
Septum	2.1 (2.9)	2.2 (3.0)	1.4 (2.0)	0.08
Post Wall	1.9 (2.6)	2.0 (2.7)	1.3 (1.4)	0.03
LA Diameter	4.7 (1.0)	4.7 (1.0)	4.6 (1.0)	0.08

Finally, Table 10 shows heart catheterization values by race. Whites had a significantly higher proportion of coronary artery disease.

Table 10. Patient Baseline Heart Catheterization Values

Covariate	**Total (N=1770)**	**White (n=1449)**	**AA (n=321)**	**P-value**
	N (%)	**n (%)**	**n (%)**	
CAD				<0.001
No	124 (16.4%)	84 (13.0%)	40 (35.1%)	
Yes	637 (83.6%)	563 (87.0%)	74 (64.9%)	
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	
RA	9.2 (5.3)	9.0 (5.3)	10.3 (5.7)	0.29
RV Systolic	46.5 (16.9)	46.1 (17.0)	50.0 (15.1)	0.26
RV Diastolic	9.7 (6.4)	9.8 (6.6)	9.3 (4.9)	0.74
PA Systolic	46.7 (17.5)	46.4 (17.5)	49.4 (17.3)	0.40
PA Diastolic	21.8 (9.0)	21.6 (9.1)	23.8 (8.2)	0.18
PA Mean	31.0 (10.7)	30.8 (10.7)	32.2 (11.0)	0.59
PCWP	20.6 (8.6)	20.4 (8.6)	22.8 (8.0)	0.23
CO	4.6 (1.6)	4.6 (1.6)	4.2 (1.5)	0.17
CI	2.3 (0.8)	2.3 (0.8)	2.2 (0.8)	0.31
PA Sat	61.0 (10.1)	61.5 (10.0)	56.6 (9.8)	0.04

3.2 SURVIVAL ANALYSIS

Figure 2 shows the Kaplan-Meier survival curves for appropriate defibrillator shock by race. African-Americans had a higher burden of shocks over time. A log rank test shows a statistically significant difference in the curves by race ($p<0.01$). The curves seem to follow a

similar pattern out to approximately 10-15 months with African Americans suffering a slightly higher burden of shocks, however after that point the curves begin to show increasing separation all the way to the study endpoint at 60 months.

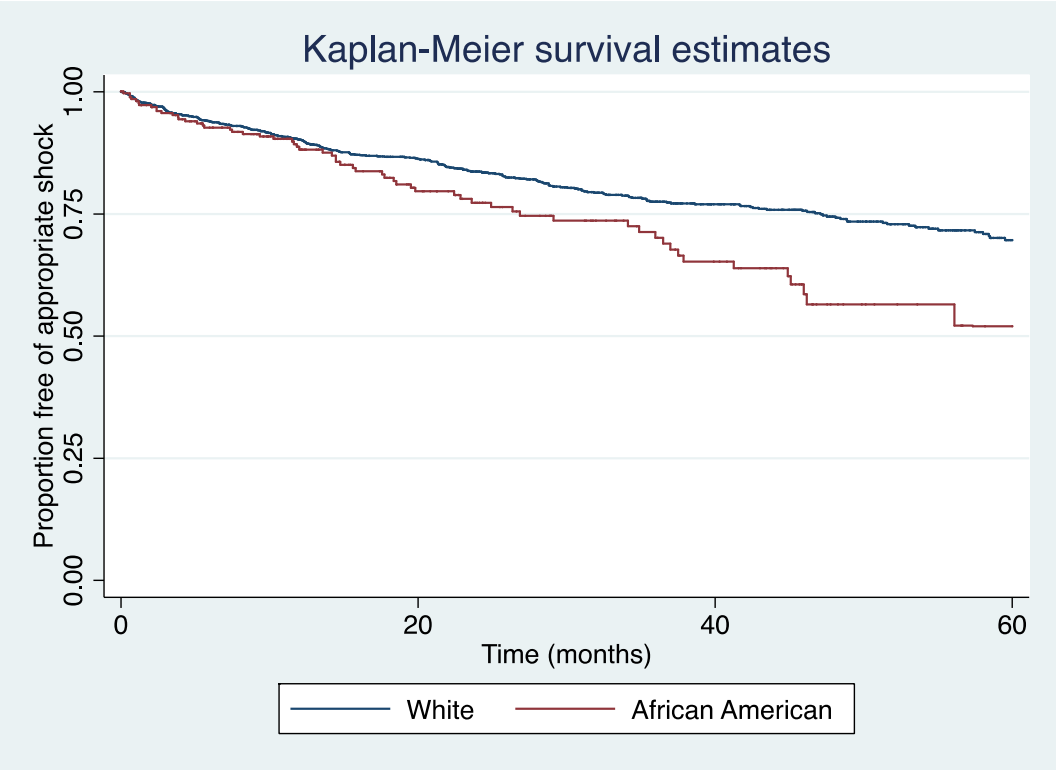

Figure 2. Shock burden over time by race. Event is appropriate shock up to 60 months. P<0.01.

3.3 COX PROPORTIONAL HAZARD REGRESSION MODELING

3.3.1 Univariable analysis and two-variable models

All 77 independent variables were tested in univariable Cox regressions to examine their individual statistical significance related to appropriate shock. Of these, 23 variables were statistically significant at the 0.10 level. Table 11 displays each statistically significant variable.

Of note, race was found to be statistically significant in the univariable models with a hazard ratio of ~1.5 (higher risk in African Americans).

Table 11. Significant variables (p<0.10) for univariable Cox regressions

Covariate	Univariable p-value	Hazard ratio	95% CI
LVEF	<0.001	0.010	(0.001,0.064)
Hx SVT	<0.001	1.725	(1.361,2.187)
Hx NSVT	<0.001	1.560	(1.250,1.947)
Antiarrhythmics	<0.001	1.753	(1.374,2.236)
Prior shock	<0.001	2.228	(1.748,2.840)
Diagnosis	<0.001		
LV diastole	<0.001	1.605	(1.395,1.847)
LV systole	<0.001	1.500	(1.317,1.708)
SBP	<0.01	0.990	(0.984,0.996)
Race	<0.01	1.522	(1.147,2.019)
Gender	0.01	0.631	(0.459,0.867)
Age	0.01	0.987	(0.978,0.996)
BUN	0.01	1.012	(1.003,1.021)
ACE	0.02	0.755	(0.598,0.951)
Aldactone	0.03	1.319	(1.036,1.679)
DBP	0.03	0.989	(0.979,0.999)
Digoxin	0.03	1.282	(1.027,1.6)
RV enlarged	0.03	1.374	(1.031,1.833)
Hx of syncope	0.03	1.307	(1.020,1.675)
Nitrates	0.04	1.399	(1.018,1.923)
Na	0.04	0.958	(0.920,0.998)
Diuretics	0.04	1.308	(1.007,1.698)
Height	0.10	1.026	(0.996,1.057)

Table 12 gives the results of the two-variable model building process sorted by covariate p-value. For this step, the height variable was combined with weight to form BMI as BMI is a more clinically meaningful measure than height alone. In each of the 23 two-variable models built except three, both race and the covariate were statistically significant. Two covariates (LV diastole, LV systole) removed the race effect and one covariate (BMI) was itself not significant. These three variables were dropped from consideration for multivariable model building.

Table 12. Two-variable model building process

Covariate	Overall model p-value (Wald)	Covariate p-value	Race p-value
LVEF	<0.001	<0.001	0.04
Hx SVT	<0.001	<0.001	<0.01
Hx NSVT	<0.001	<0.001	0.01
Antiarrhythmics	<0.001	<0.001	<0.01
Prior shock	<0.001	<0.001	0.01
LV diastole	<0.001	<0.001	0.23
LV systole	<0.001	<0.001	0.13
SBP	<0.001	<0.01	<0.01
Gender	<0.001	<0.01	<0.01
DBP	<0.001	<0.01	<0.01
BUN	<0.01	<0.01	0.02
Age	<0.01	0.02	0.01
Digoxin	<0.01	0.02	<0.01
ACE	<0.01	0.03	<0.01
Na	0.02	0.04	0.04
Nitrates	0.01	0.05	0.02
Hx of syncope	<0.01	0.05	<0.01
RV enlarged	<0.01	0.05	0.03
Diuretics	<0.01	0.06	<0.01
Aldactone	<0.01	0.08	0.01

Table 12 Continued			
Diagnosis	<0.001	0.08	<0.01
BMI	0.02	0.96	<0.01

3.3.2 Multivariable model building

After completing the model building process, the final model consisted of seven variables. These are displayed in Table 13 along with associated hazard ratios and p-values in the final model. Race was not statistically significant when adjusting for the other variables in the model. The hazard ratio of ~1.3 showed a higher adjusted risk of shock for African Americans. The other variables in the final model that were found to be significantly associated with appropriate shock over time were EF, personal history of NSVT, antiarrhythmic medication, cardiac diagnosis, age, and blood urea nitrogen (BUN, a measure of kidney function).

Table 13. Final multivariable model with p-values

Covariate	**HR**	**95% CI**	**p-value**
Race	1.282	(0.903,1.820)	0.17
LVEF	0.010	(0.001,0.106)	<0.001
Hx NSVT	1.435	(1.091,1.888)	0.01
Antiarrhythmics	1.721	(1.274,2.327)	<0.001
Diagnosis(2)	0.709	(0.488,1.029)	0.07*
Diagnosis(3)	0.493	(0.256,0.949)	0.03*
Age	0.981	(0.969,0.994)	<0.01
BUN	1.014	(1.004,1.024)	<0.01

*Overall diagnosis p-value is 0.04

The final model was overall statistically significant- confirmed with $p<0.001$ for the likelihood ratio test. The proportional hazards assumption was met ($p=0.77$). There were no

statistical issues with multicollinearity between variables in the final model. Table 14 shows individual VIF values for each variable in the final model. The mean VIF was 1.10 for all variables.

Table 14. VIF for each variable in final model

Covariate	VIF
Race	1.10
LVEF	1.08
Hx NSVT	1.02
Antiarrhythmics	1.02
Diagnosis(2)	1.15
Diagnosis(3)	1.09
Age	1.26
BUN	1.09

3.4 ASSESSMENT OF MISSING DATA

One variable in the final model (BUN) had an issue with excessive missingness. Of the 1524 observations used in the model building process, 492 were missing this covariate (~32.3%). No other variable in the final model was missing more than 3% of its values.

3.4.1 Testing for MCAR

All variables in the final model were tested, and both LVEF and race were found to be significantly different by missingness in BUN. Both of these variables were also found to significantly predict missingness of BUN in a logistic regression setting. Finally, Little's MCAR test confirmed that BUN was indeed not MCAR.

3.4.2 Multiple imputation

Table 15 shows summary statistics for BUN under the first, tenth, and last imputation as compared to the original variable. Linear regression was used for each imputation as BUN is a continuous variable. The means and standard deviations of each of these imputations do not show any concerning departures from the original values. The max values are all very similar, however there is a wider range in the minimum values among these imputations.

Table 15. BUN summary statistics under original, first, tenth, and last imputation

Imputation	N	Mean	Std. Dev.	Min	Max
Original	1240	0.000	13.218	-22.270	66.431
1	1742	-0.031	13.260	-32.369	66.431
10	1742	0.021	13.581	-47.196	66.431
20	1742	-0.182	13.341	-50.647	66.431

Table 16 compares coefficients, standard errors, and p-values for each variable in the model between the original model and the imputed model. The most significant finding when comparing the two models is that race becomes statistically significant while BUN goes from being significant to not. The p-value for race decreased by 0.13 and the p-value for BUN increased by 0.07. Coefficients and standard errors are relatively stable between models with no large changes in any variable. Each standard error from the imputed model is slightly smaller than from the original model. The largest change in coefficient is in LVEF which changes by ~0.65 between models. No other coefficient changes by more than 0.09. The largest change in standard error is also in LVEF which changes by 0.21, while no other standard error changes by more than 0.04.

Table 16. Comparison of original model (O) vs multiple imputation model (MI)

Coviariate	Coeff (O)	Coeff (MI)	Std. Error (O)	Std. Error (MI)	p-value (O)	p-value (MI)
Race	0.248	0.326	0.179	0.155	0.17	0.04
LVEF	-4.615	-3.963	1.210	1.000	<0.001	<0.001
Hx NSVT	0.361	0.369	0.140	0.117	0.01	<0.01
Antiarrhythmics	0.543	0.527	0.154	0.129	<0.001	<0.001
Diagnosis(2)	-0.345	-0.382	0.190	0.155	0.07*	0.01
Diagnosis(3)	-0.708	-0.618	0.335	0.283	0.03*	0.03
Age	-0.019	-0.017	0.007	0.005	<0.01	<0.01
BUN	0.014	0.009	0.005	0.005	0.01	0.08

*Diagnosis was overall significant in the original model (p=0.04)

Table 17 shows the hazard ratios derived from the imputed model compared to those from the original model. Again, the values of the point estimates are relatively consistent between models with no large changes. The hazard ratio associated with race changes by only ~0.1. No other hazard ratio changes by more than ~0.05. Most confidence intervals are slightly more narrow for the imputed model than the original model.

Table 17. Comparison of Hazard Ratios for original (O) and imputed (MI) models

Covariate	HR (O)	95% CI (O)	HR (MI)	95% CI (MI)
Race	1.282	(0.903,1.820)	1.385	(1.023,1.876)
LVEF	0.010	(0.001,0.106)	0.019	(0.003,0.135)
Hx NSVT	1.435	(1.091,1.888)	1.447	(1.149,1.821)
Antiarrhythmics	1.721	(1.274,2.327)	1.694	(1.315,2.182)
Diagnosis(2)	0.709	(0.488,1.029)	0.683	(0.503,0.926)
Diagnosis(3)	0.493	(0.256,0.949)	0.539	(0.310,0.938)
Age	0.981	(0.969,0.994)	0.983	(0.972,0.993)
BUN	1.014	(1.004,1.024)	1.009	(0.999,1.020)

4.0 DISCUSSION

In this cohort of patients with ICDs from the University of Pittsburgh GRADE study, African Americans received a statistically significant higher burden of appropriate shocks over a 5-year period than whites. Because appropriate ICD shock occurs in the setting of a dangerous ventricular heart rhythm that would otherwise lead to SCA if not terminated, and it is known that African Americans have higher rates of SCA[1; 2; 3; 4; 5; 6], the finding does seem plausible.

Of 77 independent variables tested in relation to the outcome, 23 were found to be individually statistically significant in a Cox proportional hazards regression setting. Using these variables as the basis for a backwards stepwise multivariable model building process, the final model consisted of the following covariates: left ventricular ejection fraction, history of NSVT, antiarrhythmic medication, diagnosis, age, and BUN. Race was not statistically significant but was forced into the model.

Missing data was found to be excessive in BUN and a multiple imputation procedure was performed to develop completed datasets. After analyzing and combining the simulated datasets, findings were similar for most covariates in terms of coefficients, hazard ratios and standard errors. One covariate (LVEF) did have slightly larger changes than the rest. Standard errors were slightly smaller and confidence intervals for hazard ratios in the imputed model were narrower than for the original model. Overall these findings seem to imply that the missing data did not introduce substantial bias into the original analysis (with the possible exception of LVEF), however the imputation of the missing data did increase the precision of the estimates.

Two additional notable changes occurred as well in the imputed model. Race became statistically significant and BUN went from being statistically significant to not. While the changes

in p-value were not extremely large, they were enough to change the statistical significance. This is mostly likely due to a combination of the increased statistical power from having more observations, in addition to the (limited) previous bias in the original analysis being corrected. Even though the coefficients and standard errors for race and BUN did not change much between the two models, the somewhat larger potential bias in LVEF may be playing a role in the relationship between all of these variables and shock. Together these two factors decreased the variability of the estimates and increased the statistical precision enough to change the p-values. Overall, even though most changes were subtle, the multiple imputation process seems to be a worthwhile procedure in the analysis.

In examining the hazard ratios, it can be seen that African American race, lower LVEF, personal history of NSVT, being on an antiarrhythmic medication, lower age, and higher BUN were all associated with an increased risk of defibrillator shock over the time period studied. Additionally, having a diagnosis of ischemic cardiomyopathy was associated with a higher hazard rate compared to both idiopathic cardiomyopathy and myocarditis/other.

This study has both similarities and differences when compared to the results of prior research. Compared to whites, African Americans in this cohort were younger, had more non-ischemic cardiomyopathy, lower EF, worse NYHA functional class, and more hypertension. These findings are similar to other cohorts studied with respect to both ICDs and SCA/SCD in general[2; 3; 4; 6; 10; 11; 12]. One major difference and important finding from this study is that while neither the SCD-HeFT study nor the PROSE-ICD study found a differing burden of ICD shocks between whites and African Americans, this study did. This appears to be a new finding.

Another important aspect of this study and one of its major strengths is that it examined a much larger set of independent variables than prior studies examining race and ICDs. Using such

a large number of variables has allowed for the relationship between appropriate shock and many more patient characteristics to be explored. The greater than 20 variables that were found to be individually statistically significant include variables that are commonly explored in the literature with respect to race and ICDs[10; 11; 12], such as ejection fraction, age, gender, ischemic vs nonischemic cardiomyopathy diagnosis, blood pressure, and kidney function, but also variables that are much less commonly examined, if at all. Examples include antiarrhythmic medications, RV enlargement, serum sodium levels, and having a history of syncope or SVT. The two variable models allowed the race effect to be examined in the context of each variable individually. Multivariable model building with this many variables also allows for race to be adjusted for many more variables than prior studies, which provided greater insight into the relationship between race and appropriate shock. The seven variable final model, like the individual variable results, also includes a mix of covariates previously seen in the literature (LVEF, diagnosis, age, BUN) and two that are not as frequently seen (Hx NSVT and antiarrhythmic medication).

4.1 PUBLIC HEALTH SIGNIFICANCE

The public health significance of the findings of this study is in trying to build on the limited previous research and paint a more complete portrait of the relationship between race and appropriate defibrillator firings. An increased burden of appropriate defibrillator firings for African Americans is an important new finding because it has not been seen before in the literature, and also because this measurement is a surrogate for SCA/SCD. Patients who have a defibrillator firing are patients who would have otherwise gone into cardiac arrest and most likely died. Recurrent ventricular arrhythmias could affect the heart and rest of the body in negative ways, in

addition to the physical discomfort a patient receives from a shock. These patients may have other underlying medical issues related to their ventricular arrhythmias that need to be addressed. And defibrillators are not perfect devices so there is a risk of a patient dying without receiving the necessary shock.

The race effect with respect to defibrillator firings over a five-year period has been adjusted for a large number of covariates to construct the final model. In practice this model could be used to assess those who may be at higher risk of receiving a shock, specifically with respect to a patient's race, and these patients could have changes made to their care to account for this fact. This model also provides new insight into several patient characteristics that possibly should be given further consideration in a clinical setting when examining risk factors for a patient receiving a defibrillator shock.

4.2 FUTURE CONSIDERATIONS

Confounding and interactions were not examined in this analysis but would be very logical next steps to undertake. A dataset with this many covariates almost certainly is affected by certain variables confounding or modifying the effects of others. The multiple comparisons problem is another statistical issue to consider. Testing this many independent variables presents the issue of some of the findings of variable significance occurring simply by chance. A commonly used solution to this problem would be to introduce a Bonferroni correction to control the familywise error rate in the univariable models prior to multivariable model building.

Some variables may be more clinically relevant than others or known to clinically affect others in meaningful ways. This analysis was carried out with a standard statistical model building

algorithm where all independent variables were equally considered in both the univariable and multivariable process. Future approaches could be more clinically or biologically inclined. The variables used could be limited to those most applicable to the "real world" of patient care or from a more biological point of view. Additionally, the large number of covariates as they relate to ICD shocks could be analyzed in specific subgroups, such as medical history alone or test/laboratory values alone.

Another future consideration would be to examine missing data from the start of the analysis as opposed to only in the final model. While imputing a significant amount of variables (or another missing data method) may be very time consuming, it may yield different results or affect the analysis in other ways. The outcome variable could also be imputed.

Examining patient mortality by race would be another logical next step, as almost every other previous study that has examined appropriate defibrillator firings has also examined patient mortality in some way. The GRADE dataset does include mortality data.

4.3 CONCLUSION

In conclusion, this study examined a very large set of covariates with respect to race and appropriate defibrillator firings in a cohort of patients with ICDs. The study has provided new insights that can be applied to patient care in hopes of identifying those at increased risk for ICD shock. This study has also provided a stepping stone for future research in an area with many opportunities for further exploration. SCA/SCD, and by extension appropriate ICD firing, are very serious public health issues and continuing to explore its underlying risk factors should be of great importance.

BIBLIOGRAPHY

[1] MOZAFFARIAN, D. et al. Heart disease and stroke statistics--2015 update: a report from the American Heart Association. **Circulation,** v. 131, n. 4, p. e29-322, Jan 2015. ISSN 1524-4539. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/25520374 >.

[2] REINIER, K. et al. Distinctive Clinical Profile of Blacks Versus Whites Presenting With Sudden Cardiac Arrest. **Circulation,** v. 132, n. 5, p. 380-7, Aug 2015. ISSN 1524-4539. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/26240262 >.

[3] FENDER, E. A.; HENRIKSON, C. A.; TERESHCHENKO, L. Racial differences in sudden cardiac death. **J Electrocardiol,** v. 47, n. 6, p. 815-8, 2014 Nov-Dec 2014. ISSN 1532-8430. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/25155390 >.

[4] COWIE, M. R. et al. Out-of-hospital cardiac arrest: racial differences in outcome in Seattle. **Am J Public Health,** v. 83, n. 7, p. 955-9, Jul 1993. ISSN 0090-0036. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/8328616 >.

[5] BECKER, L. B. et al. Racial differences in the incidence of cardiac arrest and subsequent survival. The CPR Chicago Project. **N Engl J Med,** v. 329, n. 9, p. 600-6, Aug 1993. ISSN 0028-4793. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/8341333 >.

[6] WILDE, E. T.; ROBBINS, L. S.; PRESSLEY, J. C. Racial differences in out-of-hospital cardiac arrest survival and treatment. **Emerg Med J,** v. 29, n. 5, p. 415-9, May 2012. ISSN 1472-0213. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/21546508 >.

[7] DIMARCO , J. P. Implantable Cardioverter–Defibrillators. **New England Journal of Medicine,** v. 349, n. 19, p. 1836-1847, 2003. Disponível em: < http://www.nejm.org/doi/full/10.1056/NEJMra035432 >.

[8] KADISH, A.; MEHRA, M. Heart failure devices: implantable cardioverter-defibrillators and biventricular pacing therapy. **Circulation,** v. 111, n. 24, p. 3327-35, Jun 2005. ISSN 1524-4539. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/15967863 >.

[9] THOMAS, K. L. et al. Racial disparity in the utilization of implantable-cardioverter defibrillators among patients with prior myocardial infarction and an ejection fraction of <or=35%. **Am J Cardiol,** v. 100, n. 6, p. 924-9, Sep 2007. ISSN 0002-9149. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/17826371 >.

[10] ZHANG, Y. et al. Outcomes in African Americans undergoing cardioverter-defibrillator implantation for primary prevention of sudden cardiac death: findings from the Prospective Observational Study of Implantable Cardioverter-Defibrillators (PROSE-ICD). **Heart**

Rhythm, v. 11, n. 8, p. 1377-83, Aug 2014. ISSN 1556-3871. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/24793459 >.

[11] MITCHELL, J. E. et al. Outcome in African Americans and other minorities in the Sudden Cardiac Death in Heart Failure Trial (SCD-HeFT). **Am Heart J,** v. 155, n. 3, p. 501-6, Mar 2008. ISSN 1097-6744. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/18294487 >.

[12] VOROBIOF, G. et al. Effectiveness of the implantable cardioverter defibrillator in blacks versus whites (from MADIT-II). **Am J Cardiol,** v. 98, n. 10, p. 1383-6, Nov 2006. ISSN 0002-9149. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/17134634 >.

[13] RUBIN, D. Inference and Missing Data. **Biometrika,** v. 63, n. 3, p. 581-592, 1976.

[14] LITTLE, R.; RUBIN, D. **Statistical Analysis with Missing Data**. 2nd Edition. Wiley, 2002.

[15] ALLISON, P. Missing Data. In: MILLSAP, R. e MAYDEU-OLIVARES, A. (Ed.). **The SAGE Handbook of Quantitative Methods in Psychology**: SAGE Publications, 2009. cap. 4, p.72-89.

[16] DONDERS, A. R. et al. Review: a gentle introduction to imputation of missing values. **J Clin Epidemiol,** v. 59, n. 10, p. 1087-91, Oct 2006. ISSN 0895-4356. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/16980149 >.

[17] LITTLE, R. A Test of Missing Completely at Random for Multivariate Data with Missing Values. **Journal of the American Statistical Association,** v. 83, n. 404, p. 1198-1202, 1988.

[18] KENWARD, M. G.; CARPENTER, J. Multiple imputation: current perspectives. **Stat Methods Med Res,** v. 16, n. 3, p. 199-218, Jun 2007. ISSN 0962-2802. Disponível em: < http://www.ncbi.nlm.nih.gov/pubmed/17621468 >.