Understanding the Roles of Age, Ejection Fraction, Serum Creatinine, and Follow-up Time in Death after Heart Failure

Abstract

This study investigates the factors contributing to the risk of death in heart failure (HF) patients, focusing on age, ejection fraction (how much blood the heart efficiently pumps out with each beat), serum creatinine (a waste product in the blood produced by the body's muscles that assesses how well the kidneys' are filtering blood) levels, and follow-up time after the patient is discharged. Given the growing prevalence of HF and its associated risks, including sudden death and organ failure, understanding these factors is crucial for improving treatment and prevention strategies. We performed logistic regression analysis on data from 299 HF patients to explore how our chosen variables influence mortality risk. Our findings reveal that increasing age and serum creatinine levels is associated with higher odds of death, while a higher ejection fraction and longer follow-up time are linked to decreased mortality risk. These results support existing research and highlight the importance of monitoring kidney function and follow-up care in HF patients.

Introduction

Heart failure (HF) is a growing global health concern, responsible for millions of hospitalizations and deaths each year. Understanding the complex interplay of risk factors that contribute to HF is crucial for improving prevention and treatment strategies. HF occurs when the heart is unable to pump blood effectively, often causing damage to one's organs (NHLBI, 2025). This could lead to progressive or sudden death. We are interested in discovering which factors influence the odds of death after HF.

The prevalence of HF increases significantly with advancing age. According to the National Heart, Lung, and Blood Institute, as individuals age, the likelihood of developing conditions and structural cardiac changes such as coronary artery disease, hypertension, and ventricular stiffness increases, all of which can contribute to the development of HF (NHLBI, 2025). Ejection fraction (EF), a measure of the percentage of blood pumped out of the left ventricle with each heartbeat, is another process associated with HF (Golla et al., 2024). Lower EF has been found to lead to worse outcomes in HF patients, including increased mortality (Perea-Armijo et al., 2023). Serum creatinine, a waste product of muscle metabolism circulated in the blood, is another predictor of HF as it is a marker of kidney function (Metra et al., 2012). The kidneys play a critical role in maintaining homeostasis, and kidney dysfunction can exacerbate heart failure (Metra et al., 2012). Elevated creatinine levels often reflect impaired kidney function because the kidneys are not able to filter out waste products, leading to a buildup of creatinine. Additionally, it has been found that elevated creatinine levels are associated with higher mortality rates (Fried et al., 2003).

For our analysis, we want to understand how age, EF, and creatinine are related to the risk of death after HF by performing logistic regression analyses. We are also interested in two additional factors: how age changes the effects of creatinine level on risk of death and whether the follow-up time length after discharge has an effect. As an individual ages, they are more likely to have less functioning kidneys, leading to a possible interaction between age and creatinine level. Follow-up times can be made earlier by a doctor if the patient is not recovering well after HF, therefore, adding it to the model could cover attributes not otherwise included.

Data

We analyzed a dataset consisting of 299 HF patients between the ages of 40 and 95, admitted to the Institute of Cardiology and Allied Hospital Faisalabad-Pakistan, between April and December of 2015 (Ahmad et al., 2017). The patients were in the New York Heart Association HF classes 3 or 4 (the highest possible HF classification), with moderate or severe symptoms and limitations of normal activities (Russell et al., 2009). From Ahmad et al., each patient is a case in the dataset, with variables recorded during the hospital stay. A variable time was also recorded, representing how many days after leaving the hospital the patient was checked on again, ranging from 4 to 285 days. (Ahmad et al., 2017).

Results

We started with exploratory data analysis of the dataset, plotting the empirical log odds of death against our predictor variables: ejection fraction percentage, age in years, serum creatinine level in mg/dL, and time in days. We chose to take the reciprocal of ejection fraction and of serum creatinine in order to satisfy our model assumptions. Plotting the empirical log odds of death versus age demonstrates a moderately weak, positive linear association (Figure A1), meaning as

an individual gets older, the odds of death increase. Figure A2 shows a moderately weak, negative linear association of the empirical log odds of death versus the reciprocal creatinine level. Because higher levels of creatinine indicate worse kidney function, it makes sense that as the reciprocal increases (meaning creatinine is decreasing), the log odds of death decrease. There is a potential interaction in Figure A2 with younger ages having lower empirical log odds of death at similar levels of creatinine. Figure A3 shows a moderately strong, positive correlation between empirical log odds of death and the reciprocal EF percentage. Because lower EF levels indicate the heart is not pumping enough blood, it makes sense that as the reciprocal increases (meaning EF percentage decreases), the log odds of death increase. Figure A4 shows a moderately weak, negative linear relationship between empirical log odds of death and time. This follows our initial hypothesis that the follow-up time would be shorter if the patient is not recovering well, which would lead to higher log odds of death.

Models	ANOVA	AIC
Model 1: $Y \overset{\text{indep}}{\sim} \text{Bernoulli}(p)$ $\log(\frac{p}{1-p}) = \beta_0 + \beta_1(\frac{1}{\text{EF}}) + \beta_2(\text{Age}) + \beta_3(\frac{1}{\text{Creatinine}})$	-	297.4
Model 2: $Y \overset{\text{indep}}{\sim} \text{Bernoulli}(p)$ $\log(\frac{p}{1-p}) = \beta_0 + \beta_1(\frac{1}{\text{EF}}) + \beta_2(\text{Age}) + \beta_3(\frac{1}{\text{Creatinine}}) + \beta_4(\frac{1}{\text{Creatinine}} \times \text{Age})$	0.8352	299.4
Model 3: $Y \overset{\text{indep}}{\sim} \text{Bernoulli}(p)$ $\log(\frac{p}{1-p}) = \beta_0 + \beta_1(\frac{1}{\text{EF}}) + \beta_2(\text{Age}) + \beta_3(\frac{1}{\text{Creatinine}}) + \beta_4(\text{Time})$	$< 2.2 \times 10^{-16}$	225.1

Table 1. Three compared models, ANOVA F-test p-value results of model 1 vs. 2 and 1 vs. 3, and AIC values

We compared three models (Table 1), the first with no interactions, the second with an interaction between creatinine and age, and the third with time added. We performed ANOVA F-tests to determine the explanatory abilities of the models. First, comparing model 1 and model 2, we found the p-value was bigger than 0.05, which means that model 2 did not statistically significantly better explain the data, so we continued with model 1. We then compared model 1 and model 3, with the added time variable. The p-value was smaller than 0.05, which suggests the addition of time makes the model explain significantly more. We verified this choice by calculating the Akaike Information Criterion (AIC) for each model. Model 3 has the lowest AIC value compared to the other two models, finalizing our choice to proceed with model 3 (Table 1).

All of our models satisfied the assumptions of linearity and independence. We incorporated the reciprocal of EF and creatinine because it made our initial empirical log odds plots more linear. Residual plots for model 3 (Figure A5) show linearity because of the random scatter. Residual plots for models 1 and 2 also showed linearity, but are omitted from this report for space. In model 3, we found no dependencies in gender, smoking status, or blood pressure as our residual plots (Figure A6) have random scatter. We have some concerns about additional sources of dependence that could exist in these patients. Lifestyle and health factors, such as stress and diet, could cause unaccounted for dependence because they have been found to be associated with HF (CDC, 2024).

There was no collinearity in model 3 because the Variance Inflation Factor (VIF) was below 4 for all variables. We removed one potentially influential point that had a studentized residual greater than 3 (Figure A7). The patient died, but had a low age, a high EF percentage, and low amounts of creatinine, causing a larger residual as the model predicted this patient would survive. The fact that this patient died despite these favorable indicators suggests that their outcome may have been influenced by factors not captured in the dataset.

Because EF and creatinine are transformed to be reciprocals, we will use partial residual plots to interpret their effect. Figure A8 shows that after accounting for the other predictors, the effect of EF is a statistically significant ($p < 0.001$) downward trend on the proportion of death. This downward trend is segmented, with a sharp decrease from below 35% and then a gradual decrease after 40%. This general trend is expected, as a higher EF means that a higher percentage of blood is successfully leaving the heart, meaning higher chances of survival. An EF under 40% indicates reduced EF, while anything above 40% is normal, which can explain why the slope changes at around 40% (Bhatia et al., 2006).

Figure A9 shows after accounting for other predictors in the model, the effect of creatinine shows a statistically significant ($p < 0.001$) upward trend on the estimated probability of death. This is expected following initial research, as higher creatinine can indicate an impaired kidney due to poor cardiac output (Metra et al., 2012).

As seen in Table A1, holding all other predictors constant, a ten year increase in age is associated with a 52.1% increase in the estimated odds of death (95% CI: 12.4%, 109%). This is unsurprising given our initial research because as people age, the body starts to function at a less effective level. Holding all other predictors constant, a seven day increase in follow-up time is associated with an estimated 14.0% decrease in the odds of death (95% CI: 10.5%, 17.8%). This is congruent with our initial hypothesis, because the follow-up time for a patient could be chosen based on how healthy the patient was while in the hospital. Both age ($p < 0.01$) and time ($p < 0.001$) were found to be significant effects (Table A1).

Discussion

We wanted to understand the effects of age, EF, creatinine, and follow-up time on the odds of death for heart failure patients. Our final model was found to be valid, given valid assumptions of linearity and independence and no evidence of collinearity. We found that as age and creatinine levels increase, the estimated odds of death also increase, and as EF and time after follow-up appointment increase, the estimated odds of death decrease. These results match what we were expecting to see, based on our initial research.

Some limitations of this analysis include an uneven distribution of gender, with about 200 male patients, but only 100 female patients, meaning the results may not be representative for all people. Similarly, this data was collected in Pakistan in 2015, meaning the results may not be applicable to populations in other regions of the world, or to past and future populations. Besides smoking, this study did not look into how lifestyle choices affect death after HF. Future studies could examine whether other choices increase the odds of death, such as average drinks, hours of physical activity, and hours of sleep per week.

Ethical statistical research considers whether technology used or created could discriminate against certain groups of people. This study did not record the race or ethnicity of the patients, meaning the data could disproportionately represent certain groups. This would mean an inaccurate model for underrepresented groups, leading to misguided judgments of heart disease risk. This is worrying given that certain communities are "more likely to have conditions that increase their risk for cardiovascular disease" (CDC, 2024).

It is important to remember that every person is biologically unique, and that the results of this study show general trends that do not necessarily apply to every individual. Also, many other factors not reported on here can influence heart disease. Overall, maintaining a healthy diet, regular physical activity, and periodically checking in with health care providers can help minimize the risks of heart disease and heart failure (CDC, 2024).

References

Ahmad, T., Munir, A., Bhatti, S. H., Aftab, M., & Raza, M. A. (2017). Survival analysis of heart failure patients: A case study. *PLoS ONE, 12*(7). https://doi.org/10.1371/journal.pone.0181001.

Bhatia, R. S., Tu, J. V., Lee, D. S., Austin, P. C., Fang, J., Haouzi, A., Gong, Y., & Liu, P. P. (2006). Outcome of Heart Failure with Preserved Ejection Fraction in a Population-Based Study. *The New England Journal of Medicine, 355*(3). https://doi.org/10.1056/NEJMoa051530.

Centers for Disease Control and Prevention. (2024, December 2). Heart disease risk factors. *Centers for Disease Control and Prevention.* https://www.cdc.gov/heart-disease/risk-factors/index.html.

Fried, L. F., Shlipak, M. G., Crump, C., Kronmal, R. A., Bleyer, A. J., Gottdiener, J. S., Kuller, L. H., & Newman, A. B. (2003). Renal insufficiency as a predictor of cardiovascular outcomes and mortality in elderly individuals. *Journal of the American College of Cardiology, 41*(8), 1364–1372. https://doi.org/10.1016/s0735-1097(03)00163-3.

Golla, M. S. G., Hajouli, S., & Ludhwani, D. (2024, May 5). Heart Failure and Ejection Fraction. *National Library of Medicine*; StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK553115/.

Metra M., Cotter G., Gheorghiade M., Cas L. D., Voors A. A. (2012, September) The role of the kidney in heart failure. European Heart Journal, 33(17). 2135–2142, https://doi.org/10.1093/eurheartj/ehs205.

National Heart, Lung and Blood Institute. (2022, March 24). Heart Failure - Causes and Risk Factors. National Heart, Lung, and Blood Institute. https://www.nhlbi.nih.gov/health/heart-failure/causes.

Packer M. (2020). What causes sudden death in patients with chronic heart failure and a reduced ejection fraction?. *European heart journal, 41*(18), 1757–1763. https://doi.org/10.1093/eurheartj/ehz553.

Perea-Armijo, J., López-Aguilera, J., González-Manzanares, R., Pericet-Rodriguez, C., Castillo-Domínguez, J. C., Heredia-Campos, G., Roldán-Guerra, Á., Urbano-Sánchez, C., Barreiro-Mesa, L., Aguayo-Caño, N., Delgado-Ortega, M., Crespín-Crespín, M., Ruiz-Ortiz, M., Mesa-Rubio, D., Osorio, M. P., & Anguita-Sánchez, M. (2023). The Worsening of Heart Failure with Reduced Ejection Fraction: The Impact of the Number of Hospital Admissions in a Cohort of Patients. *Journal of clinical medicine, 12*(18), 6082. https://doi.org/10.3390/jcm12186082.

Russell, S. D., Saval, M. A., Robbins, J. L., Ellestad, M. H., Gottlieb, S. S., Handberg, E. M., Zhou, & Y., Chandler, B. (2009). New York Heart Association functional class predicts exercise parameters in the current era. *American heart journal, 158*(4), 24–30. https://doi.org/10.1016/j.ahj.2009.07.017.

Appendix

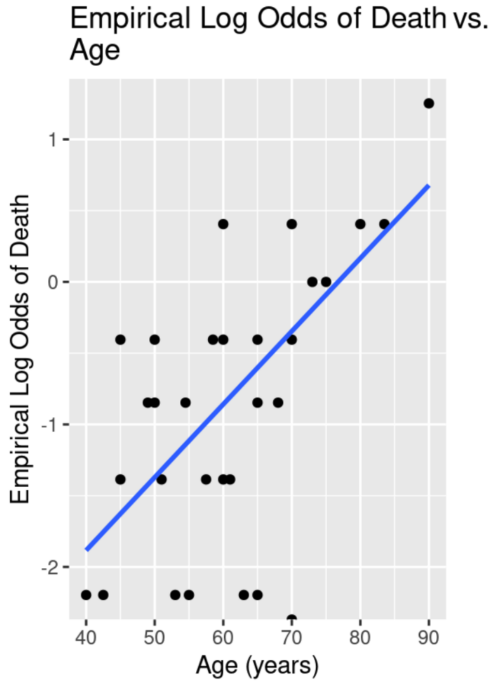

Figure A1. Empirical log odds of death vs. age.

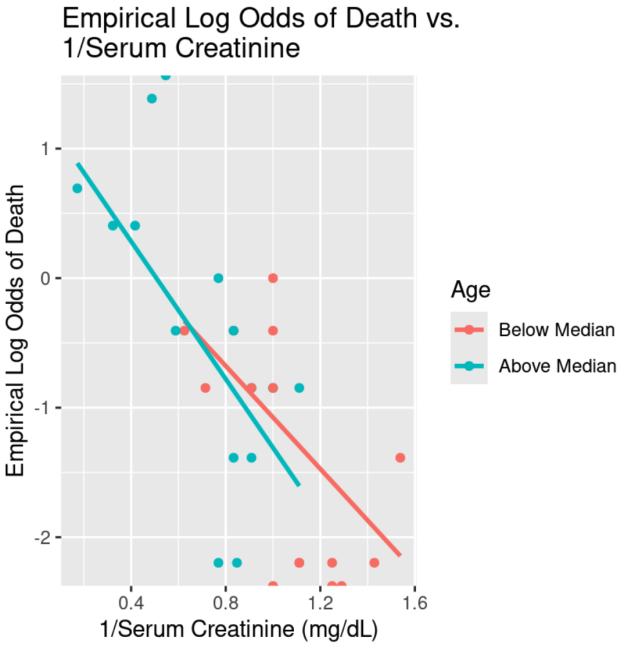

Figure A2. Empirical log odds of death vs. reciprocal creatinine.

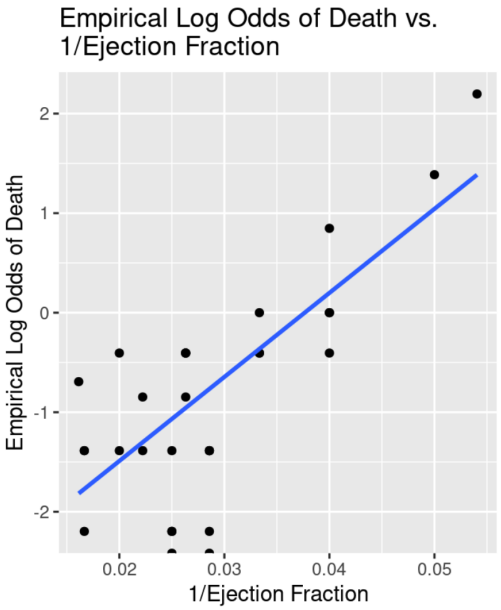

Figure A3. Empirical log odds of death vs. reciprocal EF.

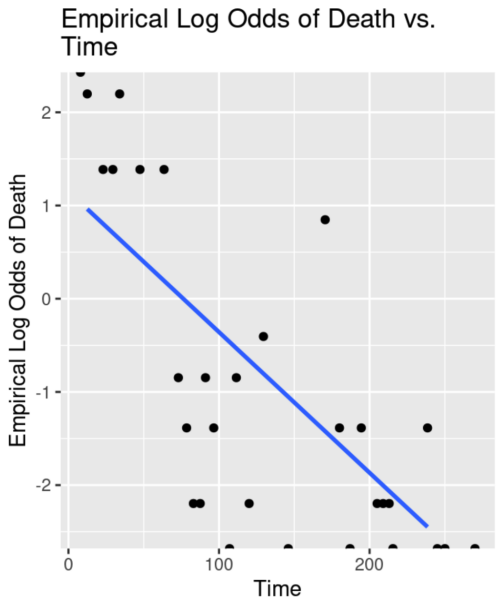

Figure A4. Empirical log odds of death vs. follow up time.

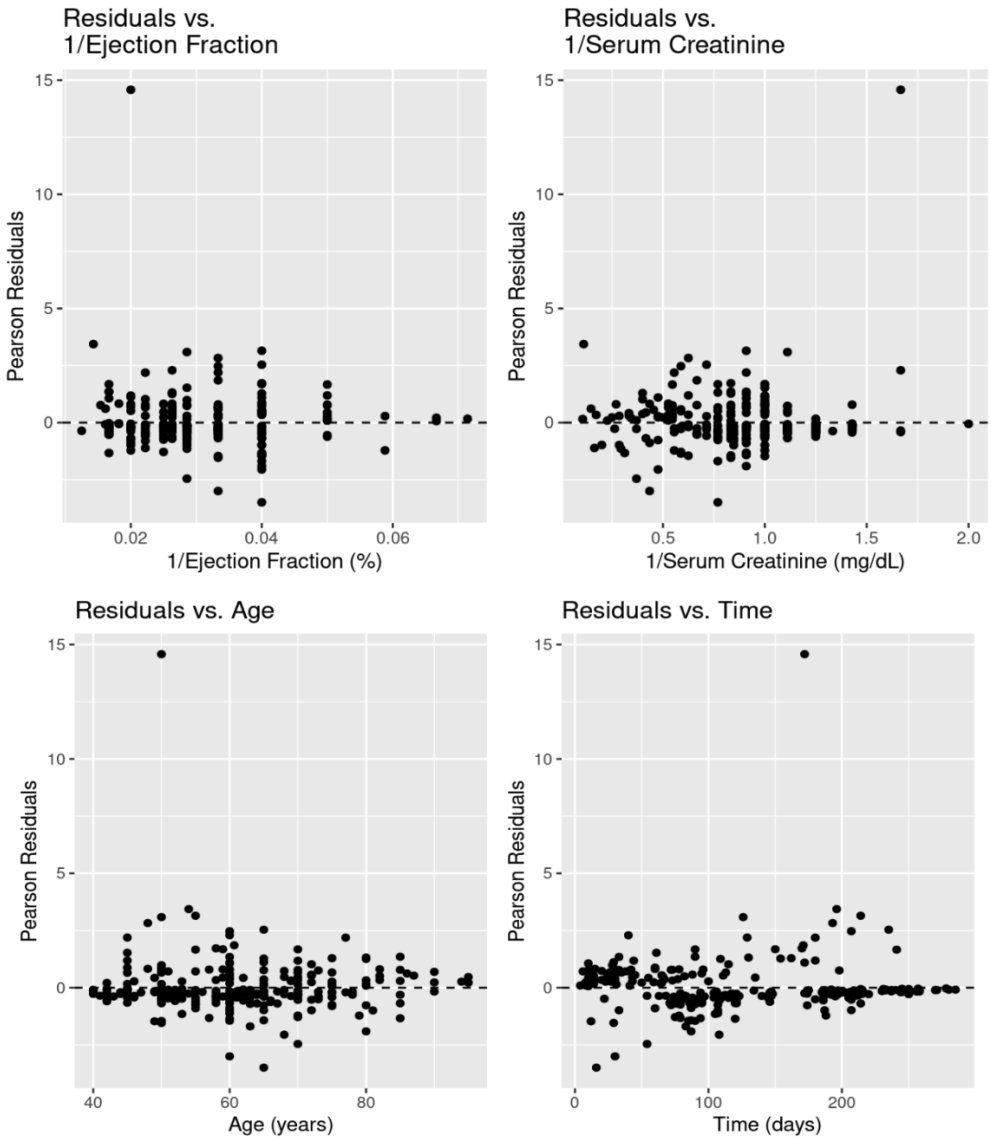

Figure A5. Pearson residual plots for the explanatory variables in model 3, showing no linearity concerns.

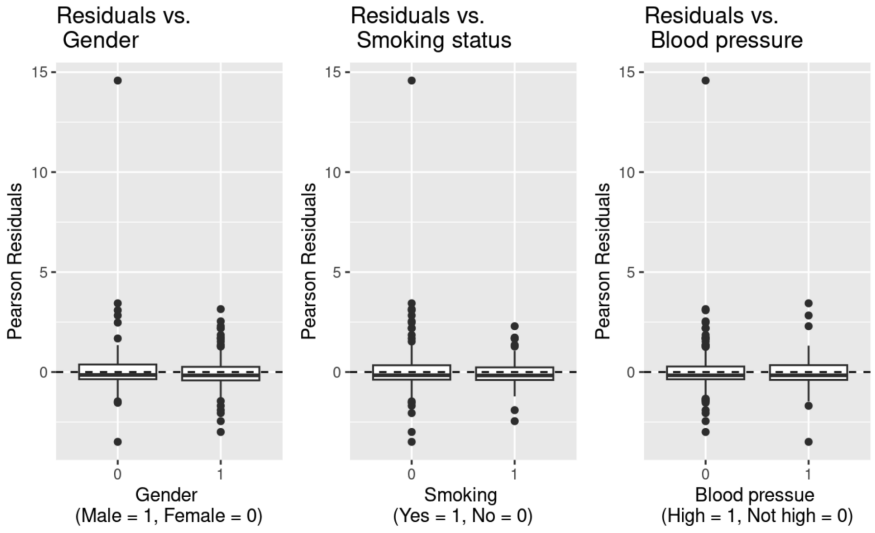

Figure A6. Pearson residual plots for possible dependencies in model 3, showing no dependence concerns.

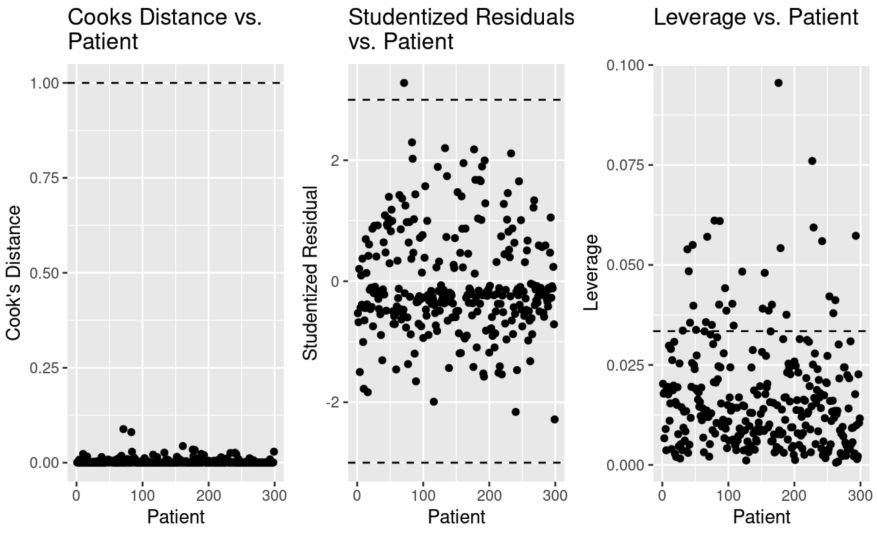

Figure A7. Influential point testing plots of model 3.

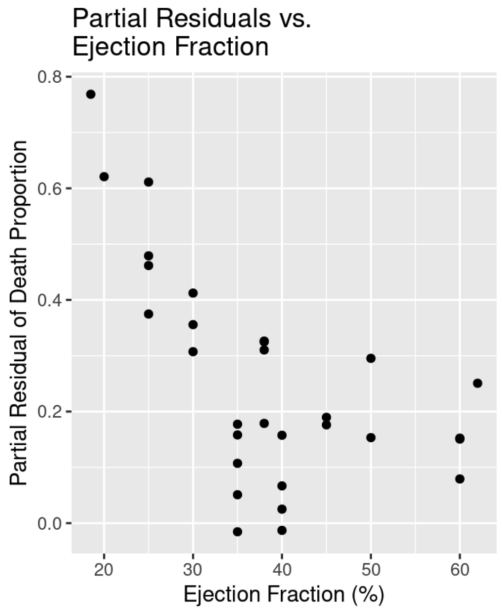

Figure A8. Partial residual plot of death proportion to interpret ejection fraction in model 3.

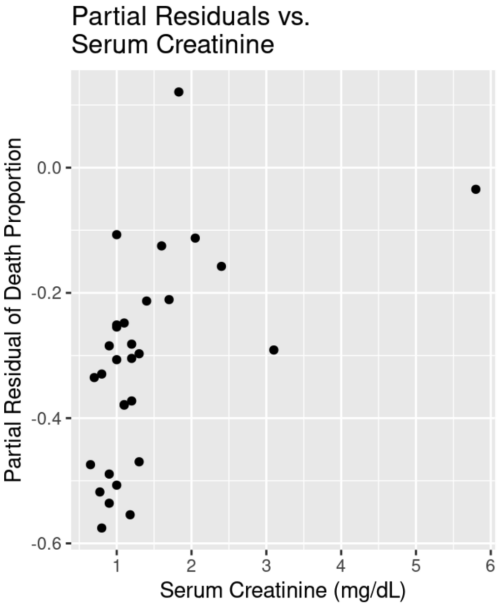

Figure A9. Partial residual plot of death proportion to interpret creatinine in model 3.

Coefficient	**Estimate**	**CI Lower Bound**	**CI Upper Bound**	**P-value**
Intercept ($\widehat{\beta_0}$)	-1.645	-4.898	0.727	0.151
Reciprocal EF ($\widehat{\beta_1}$)	100.756	63.591	143.575	5.91×10^{-7}
Age ($\widehat{\beta_2}$)	0.0419	0.0117	0.073	0.008
Reciprocal Creatinine ($\widehat{\beta_3}$)	-2.752	-3.596	-1.220	9.04×10^{-5}
Time ($\widehat{\beta_4}$)	-0.0216	-0.026	-0.0159	4.69×10^{-12}

Table A1. Estimates and CIs on the transformed scale, and p-values for the coefficients of our final model, model 3.