FINAL PROJECT REPORT

LHS 610

ABSTRACT

Impact on the mortality rate for a population with SAPS scores between 5-15, according to the age group, on patients with and without heart disease.

Ashruti Tuteja

Exploratory Data Analysis

Table of Contents

Dataset

Indwelling arterial catheters (IACs) are used extensively in the ICU for hemodynamic monitoring and for blood gas analysis. IAC use also poses potentially serious risks, including bloodstream infections and vascular complications. In 2015, *Hsu et al* published a study to assess whether IAC use was associated with mortality in patients who are mechanically ventilated and do not require vasopressor support. This dataset was created for the purpose of a case study in the book: Secondary Analysis of Electronic Health Records, published by Springer in 2016. The dataset in question was used throughout Chapter 16 (Data Analysis) by *Raffa J. et al.* to investigate the effectiveness of indwelling arterial catheters in hemodynamically stable patients with respiratory failure for mortality outcomes. The dataset is derived from MIMIC-II, the publicly-accessible critical care database. It contains a summary of clinical data and outcomes for 1,776 patients. The dataset (full_cohort_data.csv) is a comma-separated value file that includes a header with descriptive variable names.

To Access the dataset

Clinical data from the MIMIC-II database for a case study on indwelling arterial catheters. Accessed on: 1st February 2022. https://physionet.org/content/mimic2-iaccd/1.0/

Primary Usage of the Dataset

The primary use of this dataset is to carry out the case study in Chapter 16 of Secondary Analysis of Electronic Health Records. The case study data walks the reader through the process of examining the effect of indwelling arterial catheters (IAC) on 28-day mortality in the intensive care unit (ICU) in patients who were mechanically ventilated during the first day of ICU admission.

Strengths and Weaknesses of the Dataset

The dataset is of MIMIC thus, its trustability of it is assured since MIMIC is a reputed data source for medical data. It is also an open-source data set and can be accessed by everyone. It is an extensive dataset spanning numerous attributes like physiological parameters, body constituents, disease presence, and so on. The data dictionary is self-explanatory. Most importantly, the data does not contain any missing values or parameters in it; the completeness of the dataset is a major advantage for any analysis. Despite the cleanliness and completeness of the dataset, the fact that there are only 1776 instances is less for in-depth detailed analysis and model building. If there were more instances or patients recorded as part of the dataset, the subsequent study and its finding would be more inclusive and meaningful which can stand true in numerous cases. The data was collected using IAC and it is not known whether the readings obtained were cross-verified against other standard tests for each parameter.

Health-related question

"To find the mortality rate for a population with SAPS scores between 5-15, according to the age group, on patients with and without heart disease."

***P**opulation*: MIMIC data, population with SAPS score between 5-15
SAPS score is the risk of mortality of the patient in the ICU based on the severity of disease condition.
***G**rouping*: age-groups, below and above 60 years.

***I**ntervention or Exposure Variable*: Congestive heart failure (chf_flg) is a binary variable where 0 indicates the negative outcome and 1 indicates the positive outcome.

***C**omparison*: We aim to compare patients with congestive heart failure and without congestive heart failure. Congestive heart failure and chronic renal disease had a correlation of 0.2475 with mortality (relatively higher than the other variables in the dataset), which led us to choose congestive heart failure and chronic renal disease as the exposure variable and confounder respectively.

***O**utcome Variable*: The outcome variable is censored or death (censor_flg) which is a binary variable indicative of death when equal to 0 and indicative of censored when equal to 1. Also, because the SAPS score is an indication of mortality, hence it was more relevant to choose mortality as an outcome variable.

***C**onfounder*:
As a categorical confounder, chronic renal disease (renal_flg) was chosen. It is medically observed that having chronic kidney disease (CKD) implies a greater chance of having heart disease[1]. CKD can cause heart disease, and heart disease can cause CKD. In fact, heart disease is the most common cause of death among people on dialysis. Thus, choosing chronic renal disease as a categorical confounder seemed to be logical and as observed after analysis, it can be said that a person having both the chronic diseases - cardiac and renal has an almost equally likely chance of being alive or dead irrespective of the age group category. In other words, renal disease is a confounder that can affect or impact both the exposure variable of heart disease and the outcome variable of mortality.

As the continuous confounder, we chose the first hemoglobin count (hgb_first) which is taken at the time of admission of a patient to ICU. Reduced hemoglobin in patients with congestive heart failure (CHF) has been shown to be independently associated with an increased risk of hospitalization and all-cause mortality. Findings suggest a linear association between reduced hemoglobin and increased mortality risk. In studies that analyzed hemoglobin as a continuous variable, a 1-g/dL decrease in hemoglobin was independently associated with significantly increased mortality risk (Tang, Y. D., & Katz, S. D. (2006). Anemia in chronic heart failure: prevalence, etiology, clinical correlates, and treatment options. Circulation, 113(20), 2454-2461.)

Bradford Hill Criteria

It is a set of 9 criteria to provide evidence of causation within your data, between exposure or other variables with the outcome of interest [3]. It helps to evaluate the hypothesized relationship between the exposure and outcome. For the application of Bradford's criteria on MIMIC data, we were able to find relevancy for 7 out of 9 criteria.

Strength
The greater the association between our exposure variable and the outcome, more likely it is to be causal.
It could be of higher importance if we are able to entail which variable would lay a stronger association with our outcome of interest.
For example, in our study, a high SAPS score (5-15) is more likely to cause mortality as compared with low SAPS score.

Consistency
A study might have various variables of which we limit the study to only one exposure (considering others constant or confounders). The variety of variables pointing to support our null hypotheses is when we call the data to have consistency.
"Consistent findings observed by different people in different places with different samples strengthens the likelihood of an effect."
For instance, if we would have divided the data on the basis of gender also, it would have directed in the same direction that on the grounds of sex of a person if there is found a high SAPS score than the person is likely to face mortality irrespective of the gender of the patient.
Hence, as per Bradford, the "consistency criteria reveal a consistent story across multiple disciplines or practices" [4] within our data and also helps diminish internal validity.

Specificity
Associations are causal when we are aware that the association is specific to one outcome.
For an instance, in MIMIC data though there might be other factors associated with mortality but the high SAPS score as the exposure is at the nearest proximities to the mortality as an outcome.

Biological Plausibility

This factor relates fundamental concepts in explaining how biological concepts support the outcome. It states the plausible relation between cause and effect, by establishing the cause to interact with the outcome.

It is limited to the knowledge discovered during the time [5].

For MIMIC data, people with chronic diseases have their vital parameters in a critical or weakened state and hence it influences their SAPS score which in turn increases their likelihood to mortality.

Coherence

This criteria states how well our cause and effect are defined and justify the relationship irrespective of other factors present in the data.

For MIMIC data, irrespective of the age group or gender the chronic heart disease condition will remain to have the same continued impact on the mortality as outcome, justifying the chronic condition in relation with the outcome.

Experiment

If the study is conducted to remove the cause from the environment, the effect or the probability of outcome to happen will also diminish.

In our case, if chronic heart disease is removed as a condition from a patient, the chances of patient survival will increase.

Present Evidences about the Question

The Cardiovascular Disease (CVD) is a major source of mortality leading up to 17.9 million deaths in the year of 2019, before the start of the Covid-19 Pandemic. It is 32% of the total deaths recorded in that year. The most common type of CVD is heart attack and stroke which form 85% of these deaths.[6]

There are numerous scoring systems developed to measure the severity of the patient's disease condition out of which SAPS is a commonly used one which measures the illness score of the patient in the ICU with the values from 24 hours of observation. In [7], 274 patients were studied, where the SAPS score proved to be a predictor of in-hospital mortality. The outcomes were compared with the post-cardiac survival patients. Along with the SAPS score, the patient's age, neurologic status, and other factors also helped in determining the mortality and its discrimination of outcomes in patients having the illness. The SAPS score was not sufficient for predicting the neurological status of the patient.[7]

This paper [8] deals with utilizing SAPS scores to determine the severity of the patients and their vital status in the ICU. The prevalence of major disease affects the condition of the patient and developed prognostic models to determine the status of the patient. Predicting the probability of mortality of the patient is the next step in the process which can help the physicians determine suitable care methods for the patient and also inform the family members of the condition of the patient.[8]

Among the several scoring systems for predicting mortality in intensive care patients, a research study explores the predictive value of three scoring systems such as LODS, OASIS, and SAPS II.[9]

In this study, the MIMIC-IV database was used comprising 76, 540 ICU admissions, and 522 patients were enrolled in this study out of which 104 died and 418 survived in the hospital. Age, heart rate, and SAPS II score, were significantly higher in the death group than in the survival group. It has been confirmed that SAPS II have a good predictive value, with an area under the curve (AUC) of greater than 0.85. Thus, choosing a patient population with a higher SAPS score for further analysis seemed to be more valuable than choosing otherwise.[10]

In this research study, decision curve analysis (DCA) was used to assess the benefits of SAPS II and SOFA scores and compare their predictive power. As a result, it was seen that the SAPS II scoring system had more net benefits in assessing the long-term mortality compared with the SOFA scoring system. In addition, another research study [6] which included more than 3,600 post-cardiac surgery patients has demonstrated that the SAPS II scoring system had improved capabilities of predicting hospital mortality when compared to the SOFA score. This motivated us to choose the SAPS score over the SOFA score for analysis in chronic cardiac disease patients.[11]

Triple Aim

Healthcare organizations that attain the Triple Aim will have healthier populations, in part because of new designs that better identify problems and plausible solutions. Elderly patients represent nearly 50% of all ICU admissions and account for 60% of ICU days[12].

SAPS Score is an ICU scoring system that has the ability to predict survivorship and evaluate the predicted mortality against the observed mortality[13] Thus, keeping the SAPS Scoring system in mind and its usefulness in reducing mortality, it was ideal to choose the population aged over 60 with a high SAPS Score. This population also has chronic heart disease as stated in the health-related question, which limits the population to observe those individuals with this chronic disease.

Improving the experience of care

In this dataset and our analysis of the heart disease with respect to mortality, healthcare organizations might consider utilizing a greater portion of the facilities for patients aged above 60 having a high SAPS score and chronic heart disease. The experience of care for such a population can be improved where patients can expect less complex and more coordinated care, thus reducing the burden of illness. An effective health system provides a better experience of care for patients aged above 60 with a higher SAPS score and heart disease by including ways to track this experience in ambulatory settings, including patient engagement and other clinical preventive strategies.

Improving the Health of Populations

An analysis of high SAPS score population with chronic heart disease can help diminish the chances of deteriorating conditions of such patients and help the caregivers ensure that the SAPS score remains within the threshold level. Eliminating overuse or misuse of therapies or diagnostic tests can lead to both reduced costs and improved outcomes, especially for patients in the ICU.

Reducing per capita Costs of Healthcare

Decision-related information derived from such a SAPS scoring system will likely play an important role in guiding physicians in decision-making and may facilitate evidence-based allocation of limited healthcare resources in the future. Providing medical attention and care earlier to such a population aged over 60 with high SAPS scores and having chronic heart disease might reduce the risk of mortality and cost associated with it such as medications, length of stay, and other overheads.

Importantly, reducing the per capita cost of care for populations will give businesses the opportunity to be more competitive, lessen the pressure on publicly funded health care budgets, and provide communities with more flexibility to invest in activities, such as schools and the lived environment, that increase the vitality and economic wellbeing of their inhabitants [same as above citation, please paraphrase this sentence].

Visual Evaluation

Hypothesis

Null hypothesis
There is no significant relationship between the mortality of the patients and the presence of congestive heart failure.

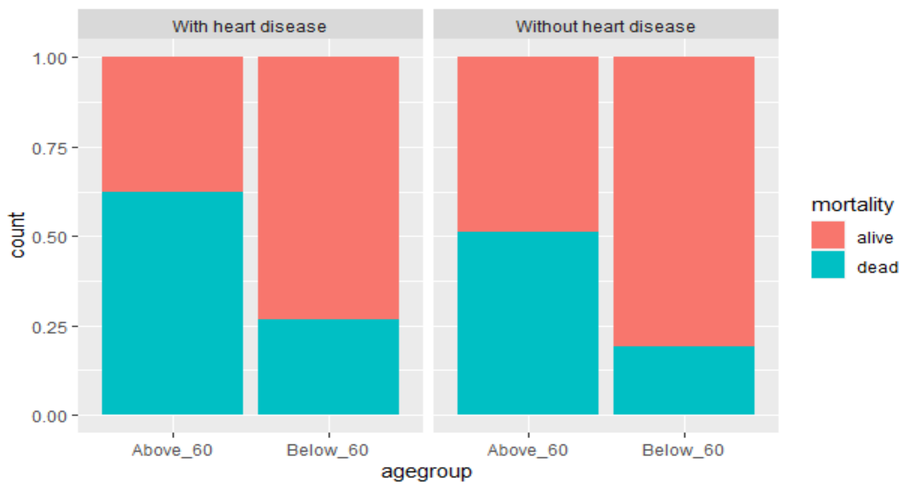
Alternative Hypothesis
There is a significant relationship between the above variables where the presence of heart disease affects the mortality of the patients.

Expected
It is expected that the null hypothesis will be rejected, therefore proving the alternative hypothesis as valuable. It is common knowledge that the chances of survival or leading a healthy life for a cardiac patient is less when compared to people without any heart ailments.

Graphical Representation

Without Confounder (Unstratified Analysis)

Y is the mortality rate; X is age groups – with and without Heart disease amongst the population with high SAPS scores of 5 to 15.



Interpretation
The above 60 years age group population with high SAPS score have a higher proportion of death rate when suffering from heart disease.

Statistical Analysis

The null hypothesis is that congestive heart failure (chf_flg, a binary variable) is not related to mortality (alive or dead). We see here that we have two categories (congestive heart failure and without congestive heart failure) for the population and hence, this is a two-sample test. Thus, we perform a Chi-square test to test our actual hypothesis against the null hypothesis. There is no need to check distribution in the case of two categorical variables. As a result, we see that the p-value is very less than 0.05 ($p = 1.647e-05$) or we can say that there is a <1% probability congestive heart failure is not related to mortality. Thus, we can possibly reject the null hypothesis on the notion that there is <1% probability that we would observe this data based on random choice. The p-value is statistically insignificant.

```
chisq.test(without_conf$mortality,without_conf$chf_flg)
```

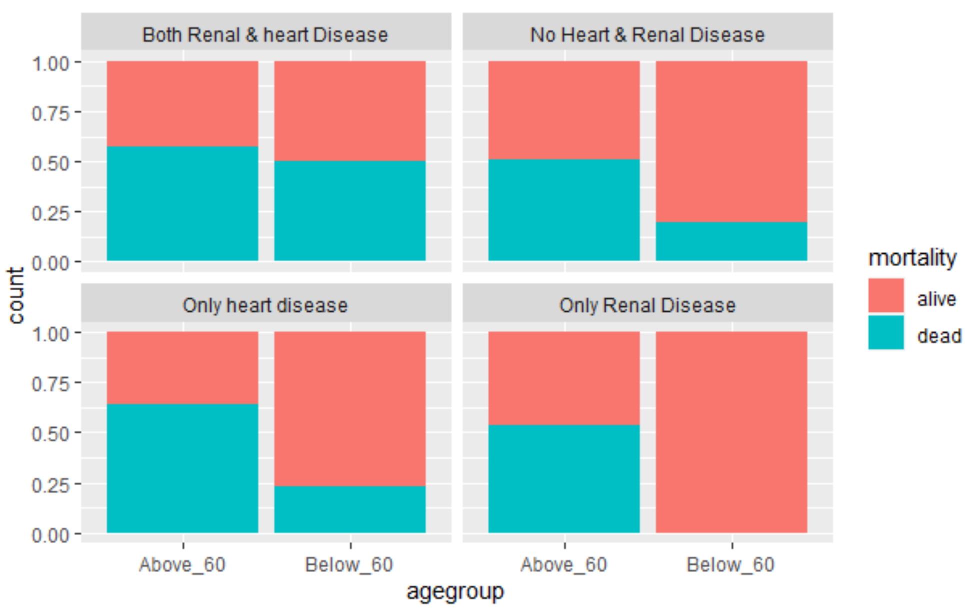
```
    Pearson's Chi-squared test with Yates' continuity
    correction

data:  without_conf$mortality and without_conf$chf_flg
X-squared = 18.56, df = 1, p-value = 1.647e-05
```

With Confounder (Stratified Analysis)

~**Categorical variable**

Y is the mortality rate, X is age-groups – with and without Heart disease & renal disease amongst the population with high SAPS score (5-15).



Interpretation

The confounding variable which was selected is the presence of other chronic diseases along with heart disease. The confounding disease in question is renal disease which is a categorical variable. Since there is a presence of two categorical variables, a total of 4 classes can be visualized - both Heart and Renal disease, Only Heart disease, Only Renal Disease, and neither Heart nor renal disease.

The first graph shows the difference in the mortality rate for both age groups having both disease conditions. The difference in the mortality rate in this cohort is minimal.

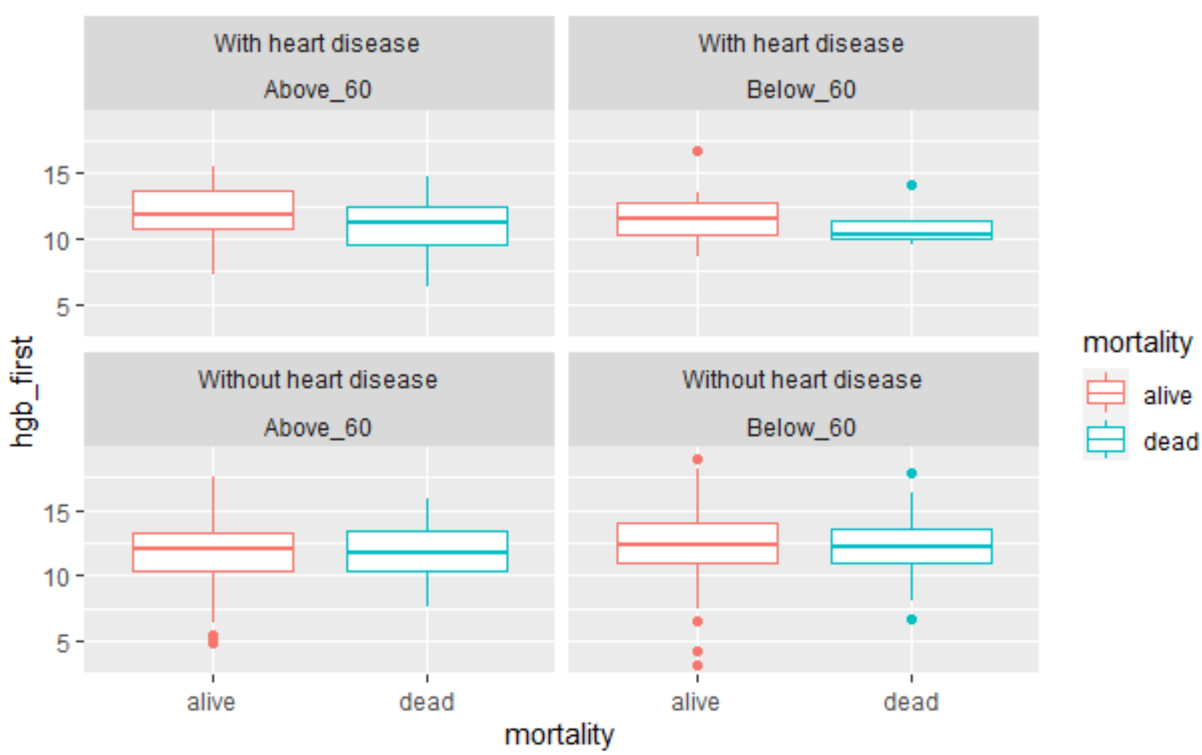
The second graph marks the difference in patients without any disease. There is more fatality for people above the age limit.

The third graph and the fourth graph depict the presence of either one of the diseases that have a stark difference in mortality rate between the age groups. Above the threshold age, the percentage of deaths from heart disease is more when compared with the patients below the threshold. On the renal side, for patients below, there is no death, while patients above the age bar have a more than 50% mortality rate.

With Confounder (Stratified Analysis)

~ Continuous variable

Y is the mortality rate, X is age-groups – with and without Heart disease & first WBC count amongst the population with high SAPS score



Interpretation
Graphs 1&2: If at the time of admission to ICU, the hemoglobin count (hgb_first) is high for a patient of any age (whether below or above the age of 60) having heart disease, the mortality rate is lower for the population.

Graphs 3&4: If at the time of admission to ICU, the hemoglobin count is high for a patient of any age (whether below or above the age of 60) not having heart disease, the mortality rate is lower for the population. However, here the hemoglobin count taken at the time of admission (hgb_first) has minimal difference in the mortality rate in the population (alive or dead).

Prediction Modeling

Presence of "Perfect" Model

A "perfect" machine learning model would predict the mortality of the patient based on the numerous health conditions as the input. Based on the variables taken into consideration in the initial parts, mortality is the predicted outcome variable or target while factors like SAPS score, heart disease condition, renal disease, age, etc. are the independent explanatory variables. The model which will be trained with the dataset containing these variables' instances will be able to predict the output when the input parameters of the new testing patients are entered into the model. For the ideal model to work perfectly, each parameter must be handled robustly to determine the mortality.

The model could be implemented as a software extension in the computer with the existing EHR system as a CDSS allowing the user to enter the values of the input attribute from the patient record and determine the mortality. Another aspect of implementation will be in a patient monitoring device or mobile monitor of the user utilizing it on the go or at the location of the patient. The potential users will be healthcare professionals like physicians most probably in the ICU or ED. Other care-providing assistants like nurses also could handle this model.

Knowing the possibility of mortality of the patient in crisis could assist the physician team in deciding what the next course of action should be based on the severity of the patient's condition. Shifting the plan of care based on the model's prediction can save the patient and/or decrease the cost value of the doctors.

Preprocessing of the Data
The dataset was vast including numerous health parameters of the patient from weight, BMI, sodium, potassium, and other human body constituents. But the main criteria chosen for this analysis is the chronic disease condition in the heart and renal. Thus, SAPS score, heart and renal

disease element, age, hemoglobin, and WBC count. Days in hospital or ICU also were included. These parameters are common and critical attributes that factor in the target variable of mortality. The patient's age contributes to their average health condition. The HB, WB, and instances of the chronic disease show the extent to which the disease condition and its impact on their health are at. The days in the ICU also add insight into how serious their health condition is.

There was no missingness in the dataset. The data obtained from the PhysioNet database was clean without any missing values in them.

Machine Learning Algorithms

Models utilized are Decision Tree, Support Vector Machine, Neural Network - Perceptron, Naive Bayes, Bagged Tree Models, K- Nearest Neighbours, Logistic Regression, and Linear Discriminant Analysis.

Performance Analysis

The different performance metrics of all the implemented models are compared against each other to see which model can be selected for use. As seen in the table below, the SVM model with a polynomial kernel has the best performance in accuracy, positive predictive value, ROC (area under the curve), and PR curve area. The negative predictive value is also the second-highest of all the given models with a 0.03 difference from the maximum observed value of NPV. The accuracy of 71% refers to the fact that the model is able to correctly predict both classes. The PPV and NPV are the ratios of patients having their mortality rated correctly predicted to the number of patients predicted with the respective class. ROC shows the ability of the model to distinguish between the two classes. Thus, considering these elements, SVM is considered the most suitable model out of all the given models.

The question about whether this model is good enough to be used in the clinical practice still exists. The SVM works better when compared to the other models. Instead of relatively speaking, considering the absolute nature of the model, the accuracy of 71% is less for introducing the model in real-time. The model's performance can be improved by increasing the dataset instances; having a larger number of entries in the training dataset will improve the robustness of the model and allow it to predict the outcome more efficiently.

The model can predict the mortality of the patient given their disease inputs. Upon implementation and utilization in the field by the medical professionals, they can tune their level of care based on the model's prediction. For instance, if the model determines the patient has a higher probability of being alive, and if other vitals and parameters show improvement, then the patient can be shifted from the ICU but still be placed under complete observation so that in case of future complications they can be treated in short notice. On the other hand, if the patient's prediction is critical and chances of mortality are high, then the physicians can increase their efforts and offer state-of-the-art treatment to the patient to bring them to a stable position. It is important to set a threshold as there are to determine which probability of mortality can be set a threshold to determine when the patient can be considered to have a high chance of living or dying.

If possible, tuning the data and the model to include a multiclass classification instead of a binary class, will be more helpful. The first class would be a high chance of living, the second would be chances of death but hope to save the patient, while the last would be an unfortunate and hopeless situation for the patient.

The threshold can be determined by FP: TP which here is 43: 112. The threshold will then be 112 / (112 + 43) = 112 / 155 = 0.7225 ~ 72.2 %. Thus, setting the threshold at this limit would benefit the model by bringing out the full cost-benefit potential.

	Accuracy	PPV	NPV	ROC_AUC	PR_AUC
Decision tree	0.65	0.65	0.63	0.69	0.78
SVM	0.71	0.72	0.7	0.76	0.81
Neural Network	0.63	0.62	0.67	0.68	0.78
Naïve Bayes	0.65	0.63	0.73	0.73	0.76
Log Regression	0.68	0.68	0.68	0.76	0.81
Bagged trees	0.65	0.65	0.65	0.71	0.77
KNN	0.61	0.63	0.56	0.65	0.76
LDA	0.68	0.68	0.68	0.76	0.81

Confusion Matrix - SVM		TRUTH	
		Alive	Dead
PREDICTED	Alive	112	43
	Dead	29	68

Conclusion

Based on the initial overviewing of the dataset, mapping the Bradford's causal criteria, sketching the triple aim of US healthcare, and by performing statistical and modelling analysis, illustrated the importance of health conditions in determining the mortality of a patient in the hospital. Understanding the danger and risk of mortality a patient is in, guides the physicians in deciding the level of care given to them. Also, the family or relatives can be notified and care options can be explored by them. Our analysis, provides basis to such decisions and scope for focusing on what is best for healthcare delivery.

The data and the subsequent analysis supported how the variables of having a heart disease, the age factor, SAPS score which describes the severity of the patient, impact mortality. The different models were compared and amongst those, SVM was chosen.

The metric display for the functionality of each model, the confusion matrix and the threshold were calculated in an attempt to answer the health-related question. Changing the threshold from the default value can improve the experience of using the model in real time. Despite these, there is scope for improvement when additional data is available as the current population isn't very large.

References

[1] (https://www.kidneyfund.org/all-about-kidneys/risk-factors/heart-disease-and-chronic-kidney-disease-ckd).

[2] https://sph.unc.edu/wp-content/uploads/sites/112/2015/07/nciph_ERIC11.pdf

[3] https://en.wikipedia.org/wiki/Bradford_Hill_criteria

[4] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4589117/
[5]
https://www.tandfonline.com/doi/full/10.1080/10408444.2018.1518404?casa_token=RVeMSLdSSZ8AAAAA%3ADRCVh3shqK6SkEczgp-7q1SHyxLEEkAXp-TgA7MUZWwCP3Ag9aajmfF9-DRns82AtZa_gAg-RCvM

[6] https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[7] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3760015/

[8] https://link.springer.com/article/10.1007/s00134-005-2763-5

[9] https://www.nature.com/articles/s41598-021-03397-3.pdf?proof=t+target%3D

[10] https://www.frontiersin.org/articles/10.3389/fcvm.2021.774935/full

[11] Schoe A, Bakhshi-Raiez F, de Keizer N, van Dissel JT, de Jonge E. Mortality prediction by SOFA score in ICU-patients after cardiac surgery; comparison with traditional prognostic-models. *BMC Anesthesiol.* (2020) 20:65. doi: 10.1186/s12871-020-00975-2

[12] P. E. Marik, "Management of the critically ill geriatric patient," *Critical Care Medicine*, vol. 34, no. 9, pp. S176–S182, 2006

[13] (Aftab Haq, Sachin Patil, Alexis Lanteri Parcells, Ronald S. Chamberlain, "The Simplified Acute Physiology Score III Is Superior to the Simplified Acute Physiology Score II and Acute Physiology and Chronic Health Evaluation II in Predicting Surgical and ICU Mortality in the "Oldest Old"", Current Gerontology and Geriatrics Research, vol. 2014, Article ID 934852, 9 pages, 2014. https://doi.org/10.1155/2014/934852).

Appendix

Data Dictionary

S.NO	Parameter Name	Meaning
1	aline_flg	IAC used (binary, 1 = year, 0 = no)
2	icu_los_day	length of stay in ICU (days, numeric)
3	hospital_los_day	length of stay in hospital (days, numeric)
4	age	age at baseline (years, numeric)
5	gender_num	patient gender (1 = male; 0=female)
6	weight_first	first weight, (kg, numeric)
7	bmi	patient BMI, (numeric)
8	sapsi_first	first SAPS I score (numeric)
9	sofa_first	first SOFA score (numeric)
10	service_unit	type of service unit (character: FICU, MICU, SICU)
11	service_num	service as a numeric (binary: 0 = MICU or FICU, 1 = SICU)
12	day_icu_intime	day of week of ICU admission (character)
13	day_icu_intime_num	day of week of ICU admission (numeric, corresponds with day_icu_intime)
14	hour_icu_intime	hour of ICU admission (numeric, hour of admission using 24hr clock)
15	hosp_exp_flg	death in hospital (binary: 1 = yes, 0 = no)
16	icu_exp_flg	death in ICU (binary: 1 = yes, 0 = no)
17	day_28_flg	death within 28 days (binary: 1 = yes, 0 = no)
18	mort_day_censored	day post ICU admission of censoring or death (days, numeric)

19	censor_flg	censored or death (binary: 0 = death, 1 = censored)
20	sepsis_flg	sepsis present (binary: 0 = no, 1 = yes -- absent (0) for all)
21	chf_flg	Congestive heart failure (binary: 0 = no, 1 = yes)
22	afib_flg	Atrial fibrillation (binary: 0 = no, 1 = yes)
23	renal_flg	Chronic renal disease (binary: 0 = no, 1 = yes)
24	liver_flg	Liver Disease (binary: 0 = no, 1 = yes)
25	copd_flg	Chronic obstructive pulmonary disease (binary: 0 = no, 1 = yes)
26	cad_flg	Coronary artery disease (binary: 0 = no, 1 = yes)
27	stroke_flg	Stroke (binary: 0 = no, 1 = yes)
28	mal_flg	Malignancy (binary: 0 = no, 1 = yes)
29	resp_flg	Respiratory disease (non-COPD) (binary: 0 = no, 1 = yes)
30	map_1st	Mean arterial pressure (mmHg, numeric)
31	hr_1st	Heart Rate (numeric)
32	temp_1st	Temperature (F, numeric)
33	spo2_1st	S_pO_2 (%, numeric)
34	abg_count	arterial blood gas count (number of tests, numeric)
35	wbc_first	first White blood cell count (K/uL, numeric)
36	hgb_first	first Hemoglobin (g/dL, numeric)
37	platelet_first	first Platelets (K/u, numericL)
38	sodium_first	first Sodium (mEq/L, numeric)
39	potassium_first	first Potassium (mEq/L, numeric)

40	tco2_first	first Bicarbonate (mEq/L, numeric)
41	chloride_first	first Chloride (mEq/L, numeric)
42	bun_first	first Blood urea nitrogen (mg/dL, numeric)
43	creatinine_first	first Creatinine (mg/dL, numeric)
44	po2_first	first PaO_2 (mmHg, numeric)
45	pco2_first	first PaCO_2 (mmHg, numeric)
46	iv_day_1	input fluids by IV on day 1 (mL, numeric)