Sepsis Insight and Management Platform (SIMPL): Utilizing Comorbidity Subgroups For Enhanced ICU Patient Care

Rihana Mohamed rmohamed@ucsd.edu
Jiyeon Song jis036@ucsd.edu
Vibha Sastry vsastry@ucsd.edu
Oakkar Aung oaaung@ucsd.edu

Kyle Shannon
kshannon@ucsd.edu

Abstract

The lack of incorporation of medical history in current sepsis risk prediction tools leaves a gap in the reality of sepsis development and the ability of a machine learning model to mimic this process. Inspired by the finding of Zador et al. (2019), which challenges the current diagnostic system by emphasizing the significance of interplay between pre-existing diseases in a patient's sepsis development, we decided to create a comprehensive sepsis risk assessment dashboard that unlike other prediction tools capitalizes on our new insights relating to comorbidity. Our dashboard aims to categorize ICU patients into risk categories based on comorbidity specific vitals and convey this categorization with relevant JIT information to facilitate sepsis risk assessment for medical professionals.

The prevalent approach in medical diagnostics is to treat patients based on singular, isolated symptoms observed during clinical evaluations, which may overlook the complex landscape of multimorbidity where multiple conditions coexist and interact within a patient. By analyzing and classifying ICU patients into seven distinct subgroups found using Latent Class Analysis, we can understand how diseases correlate with more serious diagnoses like sepsis. This relays us into different datasets to create seven machine learning models, one for each significant subgroup, allowing us to predict sepsis risk categories for those with different comorbidities. Using these predictions, a workable dashboard with patient information and their sepsis risk category is created. In addition their comorbidity subgroup is also assigned to better understand why their risk is categorized.

Website: https://oakkaraung.github.io/website_dsc180/

Code: https://github.com/jiyeonsongg/dsc180b-wi24-quarter2.git

1 Introduction

With the understanding that sepsis may differ based on a patients' biological disease makeup and having access to immense patient data, a data-driven sepsis analysis adds information to the current vital analysis protocol. Current protocol utilizes risk assessment scoring systems, such as the Sepsis-related Organ Failure Score (SOFA) and the Oxford Acute Severity Illness Score (OASIS), which often only focus on measured physiological data at admission. With the addition of patient history, this current practice has space to evolve. Current medical education, treatment protocols, and clinical trial designs continue to maintain a singular disease focus, which may contribute to the high mortality rate in clinical trials and sepsis patients. Based on Zador's paper regarding sepsis comorbidity groups, we created a sepsis dashboard showcasing vitals that medical providers usually use to predict sepsis, and additionally showcase a basic sepsis model that outputs sepsis risk categories. This dashboard broadens the typical vital focused sepsis diagnosis and gives doctors more information when diagnosing patients.

1.1 Discussion of Prior Work

Incorporating insights taken from numerous studies on sepsis and its protocol, as well as gaining expertise from industry professionals, allowed us to create an evolved dashboard, incorporating machine learning and just in time data.

In the paper "Multimorbidity states associated with higher mortality rates in organ dysfunction and sepsis: a data-driven analysis in critical care" (Zador et al. 2019) the idea that sepsis protocol needs to be adjusted to include patient disease composition is discussed and how this takes the step in trying to reduce the current mortality rate of sepsis patients. In order to do this they discussed the idea of parsing critical care patient populations into distinct comorbidity groups using Latent Class Analysis. Similarly to our dashboard, Zador created six distinct statistically significant subgroups, where our dashboard utilizes seven. This showcases that patients with different disease compositions have different rates of getting sepsis. Our dashboard utilizes this idea and includes current sepsis protocol practices but with the addition of Zador's research idea on disease composition.

When talking to industry experts, we recognized UCSD Health utilizes SBAR (Situation, Background, Assessment, Recommendation) protocol for sepsis. (Muarry 2024) The protocol is as follows:

- Situation: patient's situation and concluding whether a patient has previously been screened positive for sepsis or severe sepsis.
- Background: admission date, diagnosis for admission, recent surgery
- Assessment: blood pressure, heart rate, respiratory rate, oxygen saturation
- Recommendation: how to continue with the sepsis diagnosis

Our model briefly folloewd this protocol, but showed better results on certain changes.

Furthermore, the article "Validation of a Screening Tool for the Early Identification of Sepsis" Moore et al. 2009 discusses procedure in which patients sepsis prone are screened three times for sepsis, proving how deadly sepsis is and how screening for it has its nuances. The screening looks at heart rate, temperature, white blood cell count, and respiratory rate.

2 Methods

The development of an effective sepsis identification demands data-driven insights and user-friendly visualization tools. To create this, we embarked on an exploratory journey using Exploratory Data Analysis (EDA), Latent Class Analysis (LCA), machine learning methodologies, and the creation of an interactive web dashboard. Leveraging patient vital sign data from the MIMIC III Johnson et al. 2016 database, our EDA involved data cleaning and preprocessing to identify anomalies and ensure data integrity. Subsequently, we did Latent Class Analysis to identify distinct subgroups within the patient population based on their morbidity profiles. Furthermore, our utilization of machine learning, specifically the Random Forest Classifier, enabled us to predict specific sepsis risk classifications based on our understanding of patient subgroups and their associated risk profiles. To show our findings we created our final product, an interactive web dashboard, providing healthcare professionals with intuitive access to subgroup-based sepsis risk information.

2.1 Exploratory Data Analysis

We began our Exploratory Data Analysis by systematically extracting patient vital sign data from the MIMIC III Johnson et al. 2016 database, where each vital sign was distributed across 17 distinct "chartevents" tables. After extracting all available data for temperature, respiratory rate, heart rate, blood pressure, white blood cell count, and oxygen saturation, we assessed the viability of the data. During data cleaning and preprocessing, we addressed several issues, such as mismatched values and unreasonable outliers. Many entries were skewed due to errors, either human or machine-generated, and required careful correction. Further examination of these outliers suggested that they were likely measurement errors, as some values appeared to be inflated by factors of 1,000 or 100,000. For example, a respiratory rate of "29.390770 insp/min" may be mistakenly represented as "29,390,770 insp/min". Given the expected human range for vital signs, we adjusted these outliers to a more reasonable range, reducing noise in the dataset. After processing, the data quality improved, making it more suitable for analysis. Some vitals such as blood pressure and white blood cell count had cutoffs applied to them to keep the values in a humanly possible range. The proportion of unreasonable values in our datasets was approximately 0.05% or less, which was deemed manageable for our analysis. In exploring the distributions of vital signs between subgroups and the original dataset, we observed many discrepancies. Visual analyses using boxplots and bar charts revealed that patients with shared comorbidity profiles differed from the overall data. This indicates that splitting the vital sign data

by subgroup could impact the model. Summary statistics were conducted for each vital sign in both the original cohort and each subgroup. Finally, we conducted a comparison of the vital signs between different comorbidity subgroups and the entire cohort using bootstrapping. Bootstrapping is a resampling technique that allows us to estimate the sampling distribution of a statistic and calculate its variability. We conducted bootstrapping for each vital sign to compare the means between the subgroups and the entire cohort. This analysis involved generating multiple bootstrap samples from the cohort data and calculating the mean for each sample. By comparing the means of the subgroup samples to the bootstrap distribution of the entire cohort, we were able to determine if the differences in means were statistically significant. All assumptions for bootstrap analysis were met. The results of the bootstrapping analysis provided insights into how vital signs varied between different comorbidity subgroups, helping us understand the impact of comorbidities on patient vitals.

2.2 Statistical Analysis

In alignment with our first-quarter project, we targeted patients aged 16 and above, consistent with the age group treated with the same sepsis regimen. We expanded our cohort to include patients with readmissions, enriching our dataset with extensive comorbidity information.

2.2.1 Latent Class Analysis (LCA)

We employed Latent Class Analysis (LCA) to categorize patients into distinct subgroups based on their comorbidity profiles. We examined models with 5 to 12 classes, ultimately selecting the model that presented the lowest Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values, indicating the most suitable balance between model fit and simplicity. This process identified seven optimal subgroups characterized by distinct health profiles, namely Cardiopulmonary, Young, Hepatic/Addiction, Complicated Diabetes, Uncomplicated Diabetes, and an 'Unknown' category for individuals not fitting into the other defined groups.

2.2.2 Logistic Regression Analysis

We conducted logistic regression to investigate the relationships and characteristics of comorbidities within each subgroup. This analysis revealed significant associations(p-value $<= 0.1$), such as a strong correlation between congestive heart failure and other conditions like cardiac arrhythmias, chronic pulmonary disorders, and hypertension within certain subgroups.

2.2.3 Kruskal-Wallis Test

To assess differences among the identified subgroups, we applied the Kruskal-Wallis test to compare median Sequential Organ Failure Assessment (SOFA) scores. The SOFA scores, indicative of organ failure severity, showed significant disparities among subgroups. The test results led us to reject the null hypothesis, with a p-value < 0.05, confirming that the median SOFA scores differ significantly across subgroups.

2.2.4 Post-Hoc Analysis

Following the Kruskal-Wallis test, we executed a post-hoc analysis for pairwise subgroup comparisons. We employed the Z test for each comparison and adjusted the p-values to account for multiple comparisons, controlling the family-wise error rate. This adjustment highlighted statistically significant differences between the median SOFA scores across subgroups, further underscoring the distinct nature of each identified patient subgroup.

2.3 Machine Learning

The machine learning method has been used in two parts of our project. One model was used to predict the likelihood of a patient being in their assigned subgroup and the other was the main machine learning we utilized in predicitng the patients sepsis risk level.

2.3.1 Gradient Boosting Machine Model for Subgroup Likelihood

Following the patient group assignment via Latent Class Analysis (LCA), our objective was to assess the precision of these classifications and explore the potential for alternate group affiliations. To this end, we applied eXtreme Gradient Boosting (XGBoost) for the model development phase.

In the process of parameter optimization, we selected 'multi:softprob' as the learning objective. This specific setting is designed to output the predicted probability distribution of each patient across the various subgroups initially identified by the LCA. This probabilistic approach facilitates a nuanced evaluation, offering insights into not only the primary group assignment but also the likelihood of membership in alternative subgroups.

Upon obtaining these probability distributions, we systematically organized and presented the data according to each patient's LCA-derived group. This structured presentation serves a dual purpose: firstly, it aids healthcare providers in understanding the model's confidence in each group assignment, and secondly, it substantiates the initial grouping with quantitative evidence. This methodology not only enhances transparency but also provides a comprehensive view of the patient categorization, thereby supporting informed decision-making by medical professionals and other stakeholders.

2.3.2 Random Forest Classifier Subgroup Population

In order to tailor sepsis diagnosis strategies to the unique characteristics of patient subgroups identified through Latent Class Analysis (LCA), we created machine learning models for each subgroup. Recognizing that the primary users of our dashboard are medical professionals, we prioritized the interpretability of these models over accuracy. Consequently, we narrowed down our selection to three understandable classification algorithms: logistic regression, decision tree classification, and random forest classification. Logistic regression was employed as a baseline model. Logistic regression is a classification model that utilizes data to predict the probability of that classification happening. Given the output variables 0, 1, 2 matching to the risk categories "low risk", "moderate risk", and "high risk", logistic regression outputs the category value with the highest probability of occurring based on the input features. However it has limited accuracy, yielding less than 40%. Decision trees are the baselines for a random forest classification which creates splits based on different features and continues splitting until a return variable is calculated. The splits are calculated by maximizing how many leaf nodes (return variables) are correctly identified. Decision tree classification runs the risk of overfitting and exhibits suboptimal performance on the testing set. This is because the decision tree tries to be completely right by creating a split that perfectly predicts every single point in the training set.

The model that we ended up using to predict sepsis risk categorization is the Random Forest Classifier. (See Figure 2) A random forest classifier is a tree model that is similar to a decision tree but solves the problem of overfitting by creating multiple decision trees. A random forest classifier utilizes root nodes, internal nodes, and leaf nodes. In order to create a single tree for a random forest classifier the algorithm must decide which feature to use as the root node for the initial split, which features to use as a splitting variable, and when a leaf node (or the return variable, sepsis risk category) should be returned. The splits for each non-leaf node is decided using different indices. This split is decided differently compared to a singular decision tree model since the split for a random forest classifier uses a random subset of the data to train the model. In addition, the random forest classifier creates multiple trees on different subsets of the data. By creating these different trees the return variable is decided by finding the sepsis risk that is returned in the greatest frequency by all the trees in the random forest.

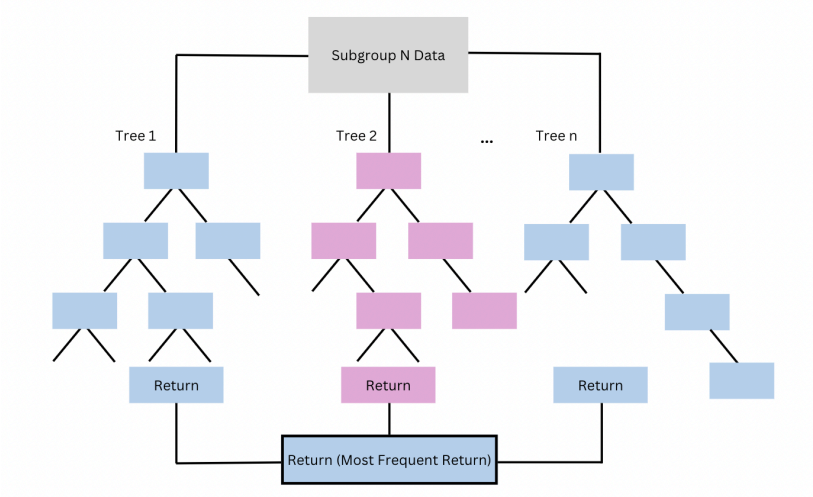

Figure 1: Random Forest Illustration

To tune the model to get the best accuracy, we utilized grid search with cross-validation to tune hyperparameters. Grid search takes a list of hyperparameters to tune and figures out which one of each hyperparameter to configure the random forest with. The hyperparameters we decided to tune were bootstrap, the max depth of the decision trees, the max features used to split between nodes, the minimum number of samples needed to split a node, the minimum number of data points needed to set a node, and number of trees built. The grid search calculated all possibilities and returned the best combination of hyperparameters by highest accuracy. Given the vast amount of data available in the MIMIC III database, we opted to validate hyperparameters on a subset of each subgroup's data. This allowed us to create a balance between efficiency and model performance. Each model itself was trained on a portion of the available data per subgroup. The return variable we utilized was breaking up the range (0-24) of the SOFA (Sequential Organ Failure) into different categories. We know that SOFA scores range from 0-24 and a score of 11 or higher means the patient has a high chance of mortality. The SOFA score "was designed as a research tool so that groups of patients (e.g., those with sepsis, and infection in the bloodstream which can lead to shock and death) could be categorized based on their risk of death" HHS ASPR TRACIE 2199. By creating the return variable, the range of when a patient has high mortality in relation to sepsis, we can properly alert medical professionals.

2.3.3 Random Forest Classification Overall Population

The goal of this dashboard is to incorporate a model in predicting sepsis and creating a random forest classifier per subgroup, as we believe patients within each subgroup share similar comorbidities and thus will have similar patterns in being diagnosed for sepsis. To create a comparable model to ensure our subgroup models are helping in predicting sepsis risk categories, we built a random forest classifier that utilized the entire population as one and used the same features in predicting the sepsis risk categories.

2.4 Dashboard

We developed an interactive web dashboard to present our findings from our analyses. This web application was constructed using HTML, TailwindCSS, Tableau, and Javascript and aims to present subgroup-based sepsis risk information in an easily digestible manner for all levels of medical staff.

Upon entering the webpage, the user will be presented with three colored cards that represent the three categories we decided on categorizing patients into: green, yellow, and red. These cards are affiliated with a risk level and contain select patients' data. Clicking on a patient's card takes the user to the dashboard that displays specific information for that patient that a healthcare worker may need to see when it comes to dealing with possibe septic patients.

The dashboard consists of four main parts: the Just In Time (J.I.T.) Information section, visualization section, Subgroup analysis, and Sepsis Risk. The J.I.T. information is presented in the top left and contains the most important vital information that a doctor may need to see about the patient such as age, identifier, and vital signs.

Below and to the right of the J.I.T. section we have our visualization section which consists of a dropdown menu and containers for Tableau visualizations. The dropdown menus contain some options when it comes to graphs to be displayed, such as various vitals to explore or groups of patients. This allows the visual that is chosen to be as visible and legible as possible as only one visualization appears at a time.

Our final two sections are where the results of our analyses on this patient are shown. These dictate the subgroup they were found in, and their sepsis risk category level as found by our analyses. The subgroup section also has a dropdown with various diseases that shows other diseases in that subgroup that are highly correlated with the selected condition. These sections are built not only to show our findings for this patient, but to also help explain the reasoning behind our automated analyses.

3 Results

3.1 Exploratory Data Analysis

Visual analyses and summary statistics indicate differences in the distributions of vitals between subgroups as well as between subgroups and the original cohort. The vital sign to show the most variance between subgroups as well as between subgroups and the original cohort was blood pressure, while temperature showed the least. Bootstrapping found significant differences for the following vitals: blood pressure, heart rate, oxygen saturation, white blood cell count; with the exception of temperature. Subgroup 4 did not have a statistically significant bootstrap result for respiratory rate and blood pressure.

3.2 Authenticating distinct comorbidity subgroups

In the first quarter project, we implemented Latent Class Analysis (LCA) to identify distinct patient subgroups, ultimately determining seven unique classes based on minimal values in the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). This decision was made at the juncture where both AIC and BIC values ceased to decline.

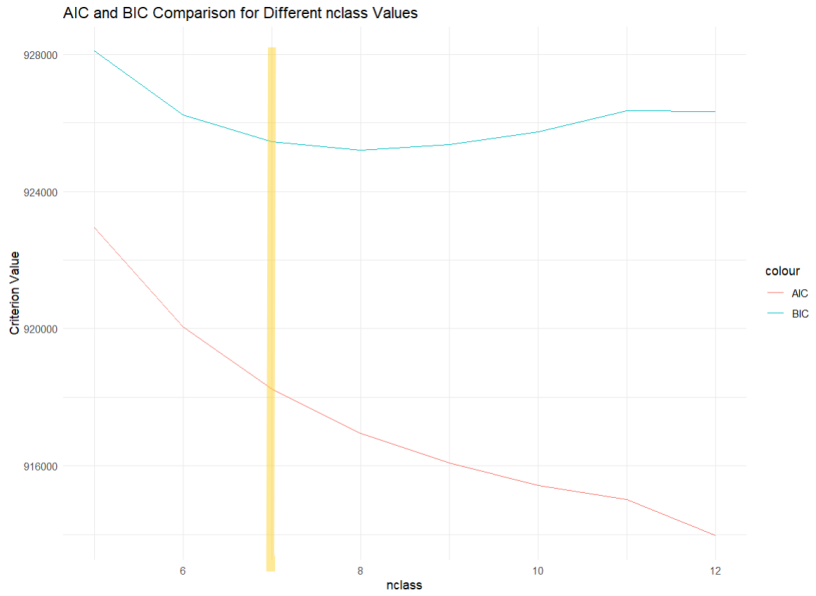

Figure 2: AIC and BIC Comparison Across Different Number of Classes in Latent Class Analysis

Subsequent exploratory data analysis revealed specific characteristics of these seven subgroups, with the five most prevalent morbidity indices being Cardiac, Cardiopulmonary, Diabetes (both Complicated and Uncomplicated), Young, Hepatic, Addiction, and an ”Unknown” category for those not fitting into the aforementioned classifications. To ensure the accuracy of these subgroup allocations, we applied a Gradient Boosting Machine (GBM) model as a secondary verification step. This approach allowed us to refine the initial subgroup assignments by identifying and correcting misclassifications. The model was trained using 80% of the dataset and tested with the remaining 20%, resulting in a training accuracy of 99.34% and a test accuracy of 96.44%. Further examination of specific patients revealed a majority correctly placed within their respective groups with complete certainty, though a minority exhibited only a 74% likelihood of correct classification.

3.3 Assessing distinct characteristics and implications for morbidity

Upon validating the patient subgroups, our next objective was to assess the interdependence of diseases within each subgroup to pinpoint prevalent comorbidities. We employed Logistic Regression to map out potential comorbid conditions among patients. For instance,

within the Cardiac group, individuals with cardiac arrhythmias were found to have a heightened likelihood of also suffering from additional ailments such as Congestive Heart Failure, Valvular Disease, Hypertension, Paralysis, Uncomplicated Diabetes, Renal Failure, Obesity, Lymphoma, and Metastatic Cancer. These associations were deemed significant, with p-values less than or equal to 0.1.

Further analysis was conducted to identify disparities among subgroups by comparing their median Sequential Organ Failure Assessment (SOFA) scores, reflecting medical severity. The Kruskal-Wallis test was utilized for this comparison, under the null hypothesis that the median severity scores across the subgroups were equivalent. The resultant p-value was less than 2.2e-16, leading to the rejection of the null hypothesis and confirming significant variances in the severity scores between the groups.

To deepen our analysis, we conducted a Post-hoc test for pairwise comparisons among the subgroups to understand their severity score disparities more comprehensively. The analysis revealed significant differences between most group pairs, with p-values of 0. However, the Cardio and Cardiopulmonary groups showed a p-value of 1 thus supporting the null hypothesis for these two groups. This nuanced understanding of comorbidity patterns and severity score disparities among the subgroups serves to enhance the precision of patient categorization and inform targeted medical interventions.

```
  Kruskal-Wallis rank sum test

data: x and group
Kruskal-Wallis chi-squared = 6257.1459, df = 6, p-value = 0


                         Comparison of x by group
                               (Bonferroni)
Col Mean-|
Row Mean |          1          2          3          4          5          6
---------+------------------------------------------------------------------
       2 |   9.701216
         |     0.0000*
         |
       3 |  -9.482843  -24.65916
         |     0.0000*    0.0000*
         |
       4 |   34.37022   37.59820   55.09062
         |     0.0000*    0.0000*    0.0000*
         |
       5 |   43.90841   49.17757   63.94338   16.95055
         |     0.0000*    0.0000*    0.0000*    0.0000*
         |
       6 |   22.53922   19.08350   39.19455  -15.20452  -29.06203
         |     0.0000*    0.0000*    0.0000*    0.0000*    0.0000*
         |
       7 |   8.212430   0.475530   18.84885  -24.45903  -34.75100  -12.89901
         |     0.0000*     1.0000    0.0000*    0.0000*    0.0000*    0.0000*

alpha = 0.05
Reject Ho if p <= alpha/2
There is a significan difference in the groups according to the median values of the SOFA score.
It rejects the Null Hypothesis.$chi2
```

Figure 3: Pairwise Test

3.4 Machine Learning for Sepsis Risk Categorization

3.4.1 Random Forest for Risk Prediction

We trained seven separate random forest classification trees for each subgroup. Each subgroup utilized a different set of hyperparameters, each having similar values. Although most of them are in the similar range Figure 3 shows the different accuracy rates per sub-

group and what specific comorbidities made up the majority of the subgroup. The accuracies range from 69-79%, with the highest accuracy for Subgroup 5 (Young Subgroup) with the lowest accuracy being Subgroup 2 (Cardiac).

Patient Subgroup	Accuracy
Subgroup 1 (Comp. Diabetes)	75.63%
Subgroup 2 (Cardiac)	68.60%
Subgroup 3 (Hepatic/Addiction)	71.64%
Subgroup 4 (Uncomp. Diabetes)	71.83%
Subgroup 5 (Young)	79.44%
Subgroup 6 (Unknown)	72.88%
Subgroup 7 (Cardiopulmonary)	68.73%

Figure 4: Percent Accurate Per Subgroup

Within each subgroup we can also see the different features are important utilizing the random forest classifications' feature importance attribute. Using Figure 4 as an example, the rest of the figure being outlined in the appendix, we can see that different subgroups weight features differently. Subgroup 1 (the Complicated Diabetes subgroup) uses age as a significant feature when predicting the risk category for the patient. However, this specific subgroup does not utilize oxygen saturation as much as a significant feature. Furthermore, Subgroup 2 (Young subgroup) has great weightage on the hours spent in the hospital, compared to other subgroups. Another important feature we might recognize is the importance of age for Subgroup 7 (Cardiopulmonary subgroup), which is higher than other subgroups.

Patient Subgroup	Vital	Feature Importance
Subgroup 1 (Comp. Diabetes)	Temperature	0.119
	Heart Rate	0.126
	Respiratory Rate	0.0954
	Oxygen Saturation	0.059
	Blood Pressure	0.117
	Number of Admission	0.067
	Hours in the Hospital	0.174
	Age	0.242

Figure 5: Feature Importance Subgroup 1 Example

3.4.2 Random Forest for Overall Population

The overall population random forest utilized the hyperparameters bootstrap which was True , the max depth of the decision trees was 90, the max features used to split between nodes is 2, the minimum number of samples needed to split a node is 2, the minimum number of data points needed to set a node 5, and number of trees built was 600. The model fitted on the whole population resulted in an accuracy of roughly 61%.

	Accuracy
All Patients	61.95%

Figure 6: Percent Accuracy Full Population

3.5 Dashboard

The finished dashboard displays the most relevant information for our users dealing with sepsis. Vital signs include heart rate, temperature, respiratory rate, blood pressure, blood oxygen level, and white blood count. These vitals were chosen after consulting with many healthcare stakeholders.

Our visualizations consist of a subgroup visualization to the right of the Just In Time information section and a patient specific visualization section towards the bottom. Both sections use Tableau to make interactive dashboard visualizations. The Subgroup visualization section consists of bar charts representing the breakdown of age and disease composition within the subgroup that the patient was found in. These help the user understand our subgroup based approach to sepsis treatment. In our patient specific visualization section, there are scatter plot graphs for patient vitals and a procedure timeline so the user could understand the patient's specific medical history.

To the right of these visuals, the Patient Group Analysis section helps the stakeholder understand the patient's role in the subgroup. At the top, a card tells you the likelihood our analyses found for our patient to belong to the subgroup while the patient's elixhauser comorbidity categories are found below that. Using those categories, a user can select a condition from a dropdown list that would display how significantly associated these different diseases are.

To the far right of the screen is a simple sepsis risk alert that shows red, green, or yellow based on the patient's predicted risk score. This calculation is made using our machine learning methods and shows information and directions for the risk level and what to do next.

4 Discussion

4.1 Exploratory Data Analysis

Our hypothesis that vital patterns would differ between subgroups and from the original held true. Almost all the vitals displayed statistical significance from the original subgroup, with the exception of subgroup 4 (Comp. Diabetes) which had a less powerful bootstrap result due to its small size. However, the p-values for blood pressure and respiratory rate in subgroup 4 were on the verge of being statistically significant, suggesting that filtering vitals by subgroup for prediction could still be beneficial. Furthermore, subgroup 4's bootstrap was still statistically significant for heart rate, oxygen saturation, and white blood cell

count, despite its small size further solidifying that vitals differ significantly by subgroup. Some limitations with the data and analysis are that we are treating vitals as non time-series data, which could potentially misrepresent the data. This does not undermine the significant differences between the overall distributions of vitals amongst patient groups, but rather is something that could be added to increase the accuracy of the analysis. It should be noted that the composition of comorbidities within patient populations in the ICU can significantly impact vital sign measurements, reflecting the interwovenness of underlying health conditions and physiological responses. This knowledge can greatly aid in sepsis care in many ways. Understanding how specific comorbidity compositions affect vital signs can aid in more accurate risk stratification for various conditions, allowing healthcare providers to better identify high-risk patients who may require more intensive monitoring or intervention. Using knowledge of comorbidity profiles can also allow healthcare professionals to further tailor treatment to patients needs, potentially improving outcomes and reducing the risk of adverse events, such as death. Furthermore, understanding comorbidity profiles can allow hospitals to allocate healthcare resources more efficiently, saving money and prioritizing resources for patients most in need of them. Our findings related to the impact of comorbidities on vital signs has potential to improve many aspects of sepsis care, as well as contributing to future research and insights into new disease mechanisms or treatment targets. A.1

4.2 Enhanced Patient Categorization through Statistical Analysis

In this research, we have employed statistical analysis to enhance clinical decision-making. By leveraging Latent Class Analysis (LCA) and verification processes, we have identified six subgroups within our patient cohort, consistent with the classifications found in referenced literature, alongside an additional "Unknown" category, which underscores the complexities and limitations encountered in the absence of extensive domain-specific knowledge.

Significantly, each identified subgroup is characterized by distinct medical histories and conditions, underlining the nuanced nature of patient categorization. This segmentation enables medical providers to approach patient care with a more personalized strategy.

This subgroup clarity is important for medical practitioners, as it boosts their confidence in the models but also explains the patients' comorbidity subgroups within their care. Furthermore, the application of basic statistical analysis for assessing comorbidities and severity scores, illustrates the integration of statistical tests utilized in medical research. This validates the relevance of our findings.

4.3 Machine Learning Sepsis Risk Categorization

The Random Forest classification model for each subgroup has a range of accuracy between 69-79%. When looking at each subgroup's accuracy we can see that certain subgroups such as Subgroup 5 (Young Subgroup) have a high accuracy rate since the individuals within this subgroup hold a similar age range which would have similar health breakdown since age

related health problems would not impact the vitals and disease compositions of the patients. Patients that are younger are easier to diagnose with a sepsis risk category since their fluctuation of vitals would indicate a clear outlier Subgroup 2 (Cardiac) and Subgroup 7 (Cardiopulmonary) tend to have a lower accuracy. Subgroup 2 (Cardiac) include patients that have comorbidities related to a heart condition which may be difficult to predict sepsis for since vitals such as heart rate may be elevated for reasons other than high sepsis risk. The model, however, will understand it as being an indication of a high sepsis risk flag. Similarly, Subgroup 7 (Cardiopulmonary) includes patients that have comorbidities related to heart and lung conditions. These conditions create high heart rate and low blood oxygen level due to underlying health conditions rather than sepsis. However, this model will relate these vital signs to having a risk of sepsis, resulting in a low accuracy rate for these groups.

Looking at the feature importance within each subgroup some subgroups give higher importance to certain variables compared to other subgroups. Looking at Subgroup 3 (Hepatic/Addiction) A.2 we can notice that the highest feature importance is the heart rate. Individuals with substance addiction will have abnormal heart rates, and give heart rate a high feature importance. Subgroup 7 (Cardiopulmonary) A.2 has a high weightage on age which could mean that the age of the patient highly relates to what sepsis risk categorization they fall into.

4.4 Dashboard

The implemented dashboard is currently hosted as a basic static website using github pages. This was due to our lack of knowledge implemented databases into a full stack web application, but basic functionality for the sepsis risk has been achieved using just a static website and preprocessed CSV and JSON data. Future improvements for the website would include live database querying using AWS services, integration with Epic Systems Electronic Health Record services, and live-tracking of patient vitals.

5 Conclusion

The development of our dashboard to incorporate sepsis risk assessment hopes to further diagnose sepsis by utilizing and showcasing comorbidity composition for patients. By incorporating insights from Zador et al. (2019) Zador et al. 2019 regarding the relationship between pre-existing diseases and sepsis development, we have created a useable dashboard that both include predictions for sepsis risk categorization as well as show doctors patient vitals for professional opinions. Our dashboard goes beyond traditional approaches by providing a more holistic assessment of sepsis risk.

Our methodology explored patient vital sign data, utilization of Latent Class Analysis (LCA) to identify distinct comorbidity subgroups, and implementation of machine learning mod-

els tailored to each subgroup for predicting sepsis risk categories. The machine learning component, particularly the Random Forest Classifier, proved effective in predicting sepsis risk based on subgroup-specific characteristics compared to the overall population.

Furthermore, our dashboard provides healthcare professionals with access to subgroup-based sepsis risk information through an interactive web interface. It presents just-in-time patient information with visualizations of vital sign data overtime, subgroup analysis results, and sepsis risk categorization per patient. Our dashboard equips medical professionals with insights to support sepsis diagnosis .

References

HHS ASPR TRACIE ., "SOFA Score Fact Sheet." Online document. [Link]

Johnson, A. E. W., T. J. Pollard, L. Shen, L. H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark. 2016. "MIMIC-III, a freely accessible critical care database." https://mimic.physionet.org

Moore, LJ, SL Jones, LA Kreiner, B McKinley, JF Sucher, SR Todd, KL Turner, A Valdivia, and FA Moore. 2009. "Validation of a screening tool for the early identification of sepsis." *J Trauma* 66(6): 1539–1546. [Link]

Muarry, Scott. 2024. "Sepsis Protocol at UCSD Health." Personal interview

Zador, Zsolt, André Landry, Michael D. Cusimano et al. 2019. "Multimorbidity states associated with higher mortality rates in organ dysfunction and sepsis: a data-driven analysis in critical care." *Crit Care* 23, p. 247. [Link]

Appendices

A.1 Exploratory Data Analysis Patient Composition

	Entire Patient Cohort	Subgroup 1 (Cardiopulmonary)	Subgroup 2 (Young)	Subgroup 3 (Hepatic/Addiction)	Subgroup 4 (Comp. Diabetes)	Subgroup 5 (Uncomp. Diabetes)	Subgroup 6 (Cardiac)	Subgroup 7 (Unknown)
Patient Count	37519	2390	6688	3448	1998	11663	8333	6998
Gender (%)								
Male	43.30%	47.80%	59.40%	67.20%	56.40%	60.50%	52.10%	49.70%
Female	56.70%	52.20%	40.60%	32.80%	43.60%	39.50%	47.90%	50.30%
Age (% of patients per age group)								
16-24	2.53%	0.25%	15.75%	0.30%	0.58%	0%	0.08%	0.54%
25-44	11.12%	3.52%	46.96%	16.55%	12.58%	3.79%	0.29%	3.33%
45-64	33.64%	27.60%	36.99%	68.20%	42.05%	43.25%	5.27%	31.88%
65—84	41.26%	54.41%	0.25%	14.93%	41.67%	51.27%	59.68%	49.02%
85>	11.45%	14.22%	0.04%	0.02%	3.12%	1.69%	34.66%	15.29%
Vitals (mean)								
Temperature (°F)								
Mean	98.58	98.27	98.91	98.4	98.38	98.7	98.29	98.56
Standard Deviation	1.52	1.4	1.56	1.51	1.45	1.45	1.44	1.51
Median	98.6	98.2	98.8	98.3	98.3	98.7	98.24	98.5
Max	109	106.9	108.14	107.9	107	109	109	109
Respiratory Rate (insp/min)								
Mean	20.01	19.55	19.98	18.81	19.77	20.48	20.29	20.01
Standard Deviation	6.21	6.46	6.31	5.85	6.19	6.05	6.45	6.21
Median	19	19	19	18	19	20	20	19
Max	90	89	90	88	90	90	90	90
Blood Pressure (mm Hg)								
Mean	120.55	118.08	121.35	119.43	126.35	121.22	118.03	123.12
Standard Deviation	24.51	24.36	24.32	29.14	24.19	24.1	25.84	24.57
Median	110	119	117	125	119	115	121	119
Max	200	200	200	200	200	200	200	200
Heart Rate (bpm)								
Mean	87.33	85.57	91.19	92.04	84.59	85.23	86	87.28
Standard Deviation	18.32	18.24	19.9	19.34	16.84	16.78	18.13	17.85
Median	86	84	91	91	84	84	84	86
Max	300	256	223	245	240	300	300	300
Oxygen Saturation (SpO2)								
Mean	87.33	85.57	91.19	92.04	84.59	85.23	86	87.28
Standard Deviation	18.32	18.24	19.9	19.34	16.84	16.78	18.13	17.85
Median	86	84	91	91	84	84	84	86
Max	300	256	223	245	240	300	300	300
White Blood Cell Count (4-11,000)								
Mean	4.63	4.53	4.66	4.77	4.6	4.67	4.58	4.64
Standard Deviation	3.43	3.38	3.13	3.44	3.55	3.45	3.44	3.45
Median	2.89	2.77	3.9	3.49	2.37	2.82	2.72	2.85
Max	11	11	11	11	11	11	11	11

Figure A 1: Patient Composition

A.2 Random Forest Model Subgroup Feature Importance

Patient Subgroup	Vital	Feature Importance
Subgroup 1 (Comp. Diabetes)	Temperature	0.119
	Heart Rate	0.126
	Respiratory Rate	0.0954
	Oxygen Saturation	0.059
	Blood Pressure	0.117
	Number of Admission	0.067
	Hours in the Hospital	0.174
	Age	0.242
Subgroup 2 (Cardiac)	Temperature	0.149
	Heart Rate	0.14
	Respiratory Rate	0.11
	Oxygen Saturation	0.079
	Blood Pressure	0.144
	Number of Admission	0.031
	Hours in the Hospital	0.198
	Age	0.148
Subgroup 3 (Hepatic/Addiction)	Temperature	0.126
	Heart Rate	0.135
	Respiratory Rate	0.111
	Oxygen Saturation	0.07
	Blood Pressure	0.126
	Number of Admission	0.042
	Hours in the Hospital	0.18
	Age	0.211
Subgroup 4 (Uncomp. Diabetes)	Temperature	0.143
	Heart Rate	0.14
	Respiratory Rate	0.119
	Oxygen Saturation	0.071
	Blood Pressure	0.148
	Number of Admission	0.03
	Hours in the Hospital	0.18
	Age	0.168
Subgroup 5 (Young)	Temperature	0.117
	Heart Rate	0.121
	Respiratory Rate	0.092
	Oxygen Saturation	0.059
	Blood Pressure	0.129
	Number of Admission	0.03
	Hours in the Hospital	0.251
	Age	0.199
Subgroup 6 (Unknown)	Temperature	0.147
	Heart Rate	0.131
	Respiratory Rate	0.11
	Oxygen Saturation	0.071
	Blood Pressure	0.132
	Number of Admission	0.053
	Hours in the Hospital	0.183
	Age	0.182
Subgroup 7 (Cardiopulmonary)	Temperature	0.118
	Heart Rate	0.131
	Respiratory Rate	0.098
	Oxygen Saturation	0.066
	Blood Pressure	0.123
	Number of Admission	0.037
	Hours in the Hospital	0.172
	Age	0.256

Figure A 2: Feature Importance Per Subgroup

A.3 Dashboard

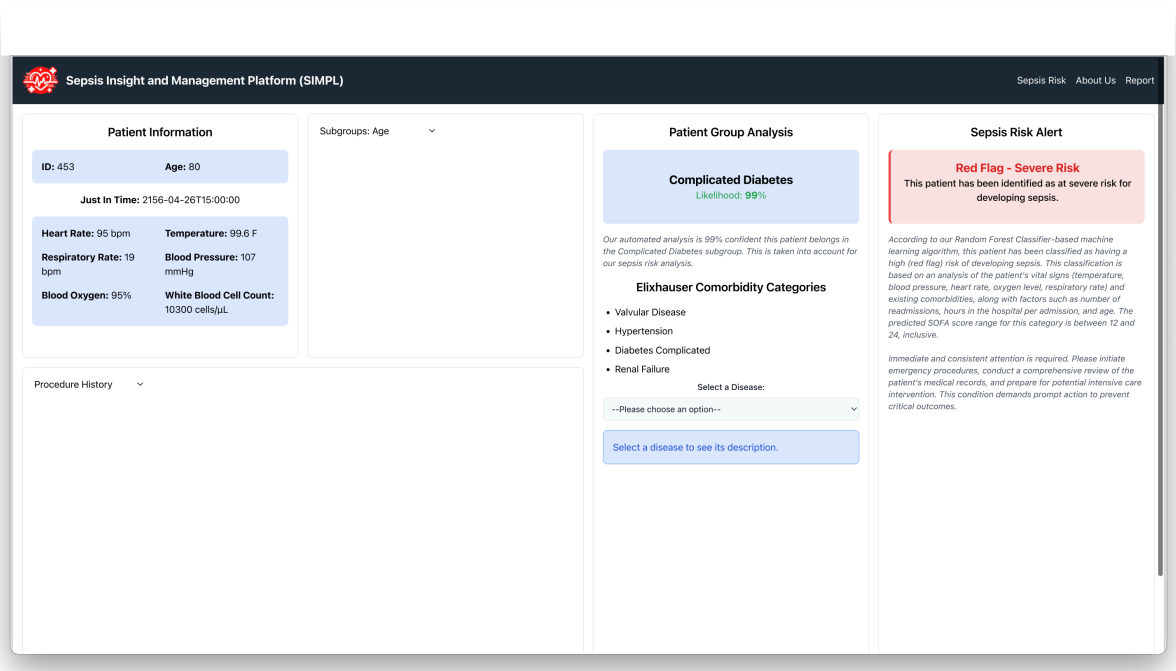

Figure A 3: Dashboard for Patient 453

A.4 Bootstrap

	p-value	Statistically Significant?
Temp		
Subgroup 1 (Cardiopulmonary)	0.5	No
Subgroup 2 (Young)	0.5	No
Subgroup 3 (Hepatic/Addiction)	0.5	No
Subgroup 4 (Comp. Diabetes)	0.5	No
Subgroup 5 (Uncomp. Diabetes)	0.5	No
Subgroup 6 (Cardiac)	0.5	No
Respiratory Rate		
Subgroup 1 (Cardiopulmonary)	0.001	Yes
Subgroup 2 (Young)	0.001	Yes
Subgroup 3 (Hepatic/Addiction)	0.001	Yes
Subgroup 4 (Comp. Diabetes)	0.073	No
Subgroup 5 (Uncomp. Diabetes)	0.001	Yes
Subgroup 6 (Cardiac)	0.001	Yes
Blood Pressure		
Subgroup 1 (Cardiopulmonary)	0.001	Yes
Subgroup 2 (Young)	0.001	Yes
Subgroup 3 (Hepatic/Addiction)	0.001	Yes
Subgroup 4 (Comp. Diabetes)	0.082	No
Subgroup 5 (Uncomp. Diabetes)	0.001	Yes
Subgroup 6 (Cardiac)	0.001	Yes
Heart Rate		
Subgroup 1 (Cardiopulmonary)	0.001	Yes
Subgroup 2 (Young)	0.001	Yes
Subgroup 3 (Hepatic/Addiction)	0.001	Yes
Subgroup 4 (Comp. Diabetes)	0.001	Yes
Subgroup 5 (Uncomp. Diabetes)	0.001	Yes
Subgroup 6 (Cardiac)	0.001	Yes
Oxygen Saturation		
Subgroup 1 (Cardiopulmonary)	0.001	Yes
Subgroup 2 (Young)	0.001	Yes
Subgroup 3 (Hepatic/Addiction)	0.001	Yes
Subgroup 4 (Comp. Diabetes)	0.001	Yes
Subgroup 5 (Uncomp. Diabetes)	0.001	Yes
Subgroup 6 (Cardiac)	0.001	Yes
White Blood Cell Count		
Subgroup 1 (Cardiopulmonary)	0.001	Yes
Subgroup 2 (Young)	0.001	Yes
Subgroup 3 (Hepatic/Addiction)	0.001	Yes
Subgroup 4 (Comp. Diabetes)	0.001	Yes
Subgroup 5 (Uncomp. Diabetes)	0.001	Yes
Subgroup 6 (Cardiac)	0.001	Yes

Figure A 4: Bootstrap Result