Identifying Risk Factors and Preferred Hospitals for Hip/Knee Replacements: An Analysis of 2019-2022 Hospital Readmission Reduction Program Data

Adeline Casali* and Scott Eugley*

* Each author contributed equally to the design, coding & development, analysis, and writing of this project

August 15, 2024

Abstract

This study utilized machine learning to analyze data from the Hospital Readmission Reduction Program (HRRP) to identify significant risk factors associated with hospital readmission rates for hip/knee replacement surgery patients and classify hospitals as either preferred or non-preferred for this procedure. Machine learning models were trained using HRRP data from 2019-2022, and tested on 2024 HRRP data.

The key question of the analysis is: What risk factors are most significantly linked to readmission rates after hip and knee replacement surgery? The analysis found that the most important predictors were complications like accidental cuts during treatment, wounds reopening after surgery, and kidney problems following the procedure. Overall hospital satisfaction ratings from Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) surveys were also a top factor. This supports the hypothesis that higher patient satisfaction scores often correlates with better patient outcomes.

A Random Forest model was able to accurately classify hospitals as preferred or non-preferred with 98% accuracy. This suggests it could be a useful tool for health insurance companies to guide patients to hospitals with better outcomes. However, there is also concern about potential overfitting and data leakage within the model that warrants further investigation.

The findings from the analysis highlight areas in which hospitals could focus on to reduce readmissions, such as reducing surgical complications and improving the overall patient experience. With further investigation and refinement, this approach may help improve patient outcomes and also reduce healthcare costs associated with hospital readmissions after hip/knee replacement surgeries.

Background and Question

Since 2012, the Centers for Medicare and Medicaid Services have implemented the Hospital Readmission Reduction Program (HRRP). This program tracks hospital readmission rates and incentivizes hospitals to reduce unnecessary readmissions through financial penalties. Using 2019-2022 readmission data from the HRRP, this analysis aims to identify the preferred and non-preferred hospitals for hip and knee replacements for a health insurance company. Furthermore, it will examine the risk factors associated with higher readmission rates for these procedures. The insights from this analysis can be used to improve hospital performance, enhance patient care, and reduce costs. As of 2019, the average cost of readmission after hip/knee surgery was $8,588, and avoiding that cost would be highly beneficial for health insurance companies and consumers alike (Phillips et al., 2019). Understanding these risk factors can help health insurance companies guide patients towards hospitals with better outcomes, thereby improving patient outcomes and reducing costs associated with readmissions.

In this analysis, we seek to answer the following question: What risk factors are associated with hospital readmission rates for hip/knee replacements? We hypothesize that hospitals with better Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) scores will have lower readmission rates for hip/knee replacements because higher patient satisfaction often correlates with better overall care quality and patient outcomes, including reduced complications and better post-discharge support (Edwards et al., 2015). It is anticipated that patient satisfaction, clinical treatment quality, and operational efficiency will be the most important risk factor categories associated with hospital readmission rates among patients who have had hip or knee replacement surgery. The most important variables influencing readmission rates for patients after hip or knee replacement surgery will most likely be a combination of high HCAHPS scores and reduced rates of clinical complications.

Data

We use datasets from the Centers for Medicare and Medicaid Services, which allow for the comparison of the quality of care provided in Medicare-certified hospitals, Veterans Administration (VA) medical centers, and Department of Defense (DoD) hospitals nationwide (Centers for Medicare & Medicaid Services, 2024). We utilize predictors from the HCAHPS (Hospital Consumer Assessment of Healthcare Providers and Systems) dataset as well as Timely and Effective Care, containing information on average wait times and vaccination compliance, Complications and Deaths, containing information about the frequency of deaths and complications for procedures, and Payment and Spending metrics, which includes the costs associated with procedures. Our target variable is the predicted readmission rate after hip/knee surgery, which is a measure of the number of readmissions within 30 days predicted based on the hospital's performance with its observed case mix. The predicted number of readmissions is estimated using a hospital-specific intercept, and is intended to reflect the annual expected performance of the hospital given its historical case and patient mix and performance. Utilizing this metric instead of the crude observed readmission rate reduces the number of missing values and increases comparability. For our training dataset, we include HRRP data from 2019-2022. We then test the model using the most recent snapshot of the data from 04-24-2024. Although the original dataset had many missing values (129177 in initial dataset), our final dataset contained 1833 observations across 21 variables. See Appendix 2 for a list of all variables. Consequently, the biggest caveat to our dataset choice is that many variables had to be removed due to missing values, so there were many risk factors that we could not take into consideration for this analysis.

Preprocessing and Feature Engineering

To ensure the dataset's suitability for analysis, several preprocessing steps were employed. First, categorical variables were identified, and appropriate encoding methods were applied. The EDV column, which contains ordinal data, was dummy encoded based on predefined levels, transforming it into a numeric format that reflects its ordinal nature. This approach preserves the inherent order of the data while making it suitable for statistical modeling. Additionally, categorical states were encoded using a numeric mapping to eliminate potential ordinal relationships and standardize the representation of states. This choice was made to facilitate model compatibility and avoid any unintended ordinal implications in the analysis. Unnecessary or highly correlated variables were then removed to streamline the dataset. Variables with significant missingness or redundancy, such as certain patient survey ratings and summary statistics, were excluded. The decision to remove these variables was guided by their high correlation with other variables, high missingness, or redundancy with already included metrics. For instance, combining various patient survey metrics into a single comprehensive score helped reduce collinearity and simplify the feature set. See table 1 for a list of all variables that were removed and justification for doing so.

Table 1: Justifications for Removal of Variables

Variable	Justification
OP_18c	Removed due to high correlation and low relevance.
OP_22	Removed due to high correlation and low relevance.
ED_2_Strata_1	Removed due to high percentage of missingness.
OP_23	Removed due to high percentage of missingness.
VTE_2	Removed due to high percentage of missingness.
STK_02	Removed as stroke data is not relevant to Hip/Knee Surgery.
STK_05	Removed as stroke data is not relevant to Hip/Knee Surgery.
STK_06	Removed as stroke data is not relevant to Hip/Knee Surgery.
HcahpsLinearMeanValue_H_RECMND_LINEAR_SCORE	Removed due to strong correlation with overall hospital rating.
ExcessReadmissionRatio_HIP-KNEE	Removed due to high correlation with target variable.
ExpectedReadmissionRate_HIP-KNEE	Removed due to high correlation with target variable.
NumberOfReadmissions_HIP-KNEE	Removed as it is influenced by hospital size, which is not available.
Sepsis Variables	Removed due to unclear definition in dataset dictionary.
Score_PSI_90	Removed because it is a summary of other PSI variables, making it redundant.
Patient Survey Data	Removed due to collinearity and redundancy with other patient survey metrics.

Feature engineering focused on creating meaningful variables from existing data and removing irrelevant or redundant features. A notable example is the creation of the Score_Ovr_MORT variable, which aggregates multiple mortality-related scores into a single measure. This new feature consolidates related variables, reducing dimensionality while preserving critical information on mortality rates. For preprocessing, first, categorical variables were appropriately encoded, and redundant variables were removed, as described in the previous section. Collinearity was greatly reduced, as evident by the heat maps in Figure 1 and Figure 2. Additionally, all rows with NA values for the target variables were removed. We also decided to remove the one facility that had an NA value, which happened to be the same observation with a missing state value. Variables with less than 5% missing data were imputed by the median, as this has a small impact on the variation of the predictors. Any variables with greater than 5% missing values were imputed using kNN to allow for more accurate imputation and mitigate reduction in variation. Finally, all remaining features were scaled and normalized. All steps were repeated on the test dataset, which is the most recent snapshot of the data from April 24, 2024. Finally, descriptive statistics were computed for all variables. This included descriptive statistics for all numeric variables (Table 2), histograms for all numeric variables (Figure 3), the spread of the target variable (Figure 4), as well as geographic explorations into the number of facilities per state (Figure 5), and the average predicted readmission rate per state (Figure 6).

Figure 1. Correlation Heatmap of Numeric Variables Before Preprocessing

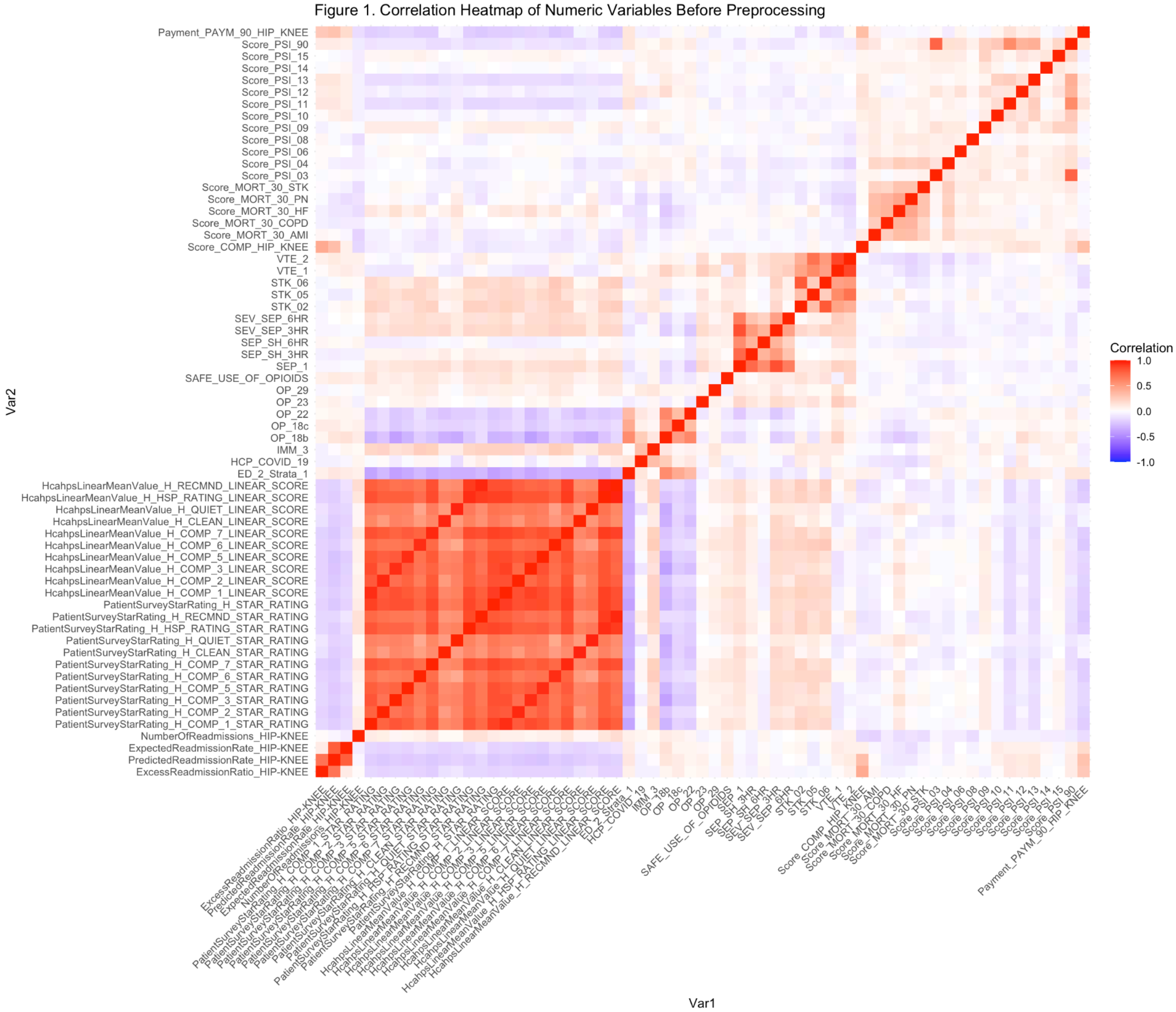


Figure 2. Correlation Heatmap of Numeric Variables After Preprocessing

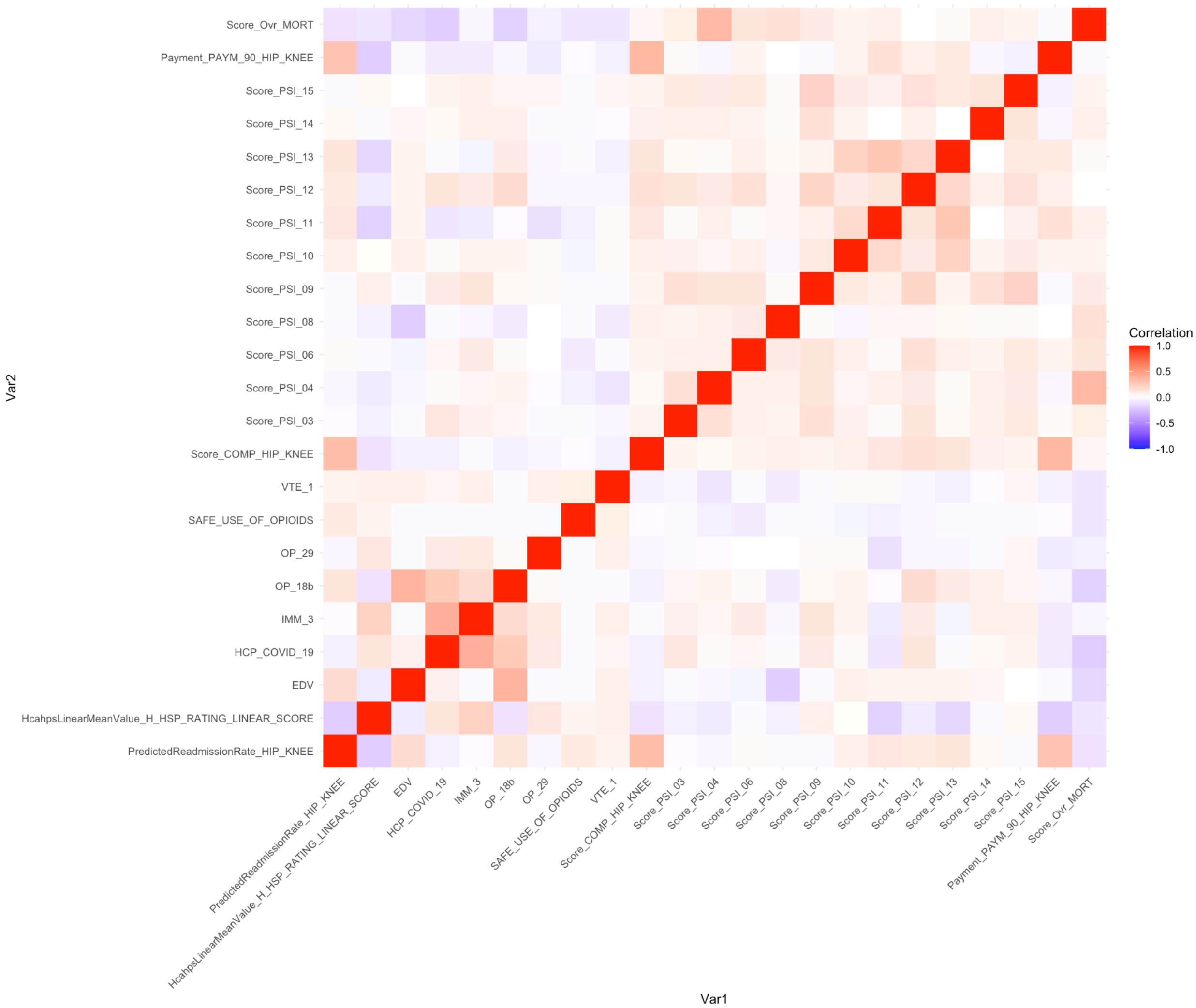


Table 2. Descriptive Statistics for All Numeric Variables in Final Dataset

	mean	sd	mad	min	max	range	skew	kurtosis	se
PredictedReadmissionRate_HIP_KNEE	0.00000	1.00000	0.9431408	-2.879575	4.4232211	7.302796	0.4373061	0.4908058	0.0233571
HcahpsLinearMeanValue_H_HSP_RATING_LINEAR_SCORE	0.00000	1.00000	0.7964386	-5.566501	3.0285467	8.595047	-0.5851721	1.5050011	0.0233571
EDV	0.00000	1.00000	1.4463754	-1.653088	1.2736124	2.926700	-0.1385316	-1.1578643	0.0233571
HCP_COVID_19	0.00000	1.00000	0.8684108	-5.885694	1.1640463	7.049740	-1.3148000	2.1219028	0.0233571
IMM_3	0.00000	1.00000	0.9189058	-4.244216	1.1648908	5.409107	-1.0958217	0.7241943	0.0233571
OP_18b	0.00000	1.00000	0.8953193	-2.567998	7.9999828	10.567980	0.9817063	3.3080181	0.0233571
OP_29	0.00000	1.00000	0.3787292	-7.847297	0.6676801	8.514977	-3.3057142	14.4641144	0.0233571
SAFE_USE_OF_OPIOIDS	0.00000	1.00000	0.6976501	-3.663836	6.9237316	10.587568	0.5890847	3.1474800	0.0233571
VTE_1	0.00000	1.00000	0.4857731	-9.400663	0.9749047	10.375567	-4.3773341	29.2741519	0.0233571
Score_PSI_03	0.00000	1.00000	0.5434188	-1.037499	11.0387303	12.076230	3.5027779	21.8351982	0.0233571
Score_PSI_04	0.00000	1.00000	0.8206364	-4.317401	3.8370394	8.154441	-0.0656641	1.2542628	0.0233571
Score_PSI_06	0.00000	1.00000	0.6552102	-2.801201	5.8164972	8.617698	0.9085914	1.9551905	0.0233571
Score_PSI_08	0.00000	1.00000	0.0000000	-3.651722	4.8844865	8.536209	0.5502254	1.7580966	0.0233571
Score_PSI_09	0.00000	1.00000	0.6884533	-2.855072	7.2396114	10.094683	1.2648156	4.8211377	0.0233571
Score_PSI_10	0.00000	1.00000	0.3972884	-2.955550	7.9775177	10.933067	1.8287437	7.4056616	0.0233571
Score_PSI_11	0.00000	1.00000	0.7145473	-2.004891	12.8618308	14.866722	2.7479837	22.3393252	0.0233571
Score_PSI_12	0.00000	1.00000	0.8397775	-2.483138	4.9432855	7.426424	0.9512312	1.6446910	0.0233571
Score_PSI_13	0.00000	1.00000	0.7188248	-3.022597	5.3360566	8.358654	1.0438233	2.7498825	0.0233571
Score_PSI_14	0.00000	1.00000	0.5399830	-2.635057	6.6944124	9.329469	1.9129489	6.4857308	0.0233571
Score_PSI_15	0.00000	1.00000	0.6797006	-2.300412	7.1131306	9.413543	1.7127063	5.4101230	0.0233571
Score_Ovr_MORT	0.00000	1.00000	0.8935151	-4.216908	3.7555248	7.972433	-0.0667113	0.5511229	0.0233571

Figure 4. Histogram of Predicted Readmission Rate for Hip/Knee Replacement

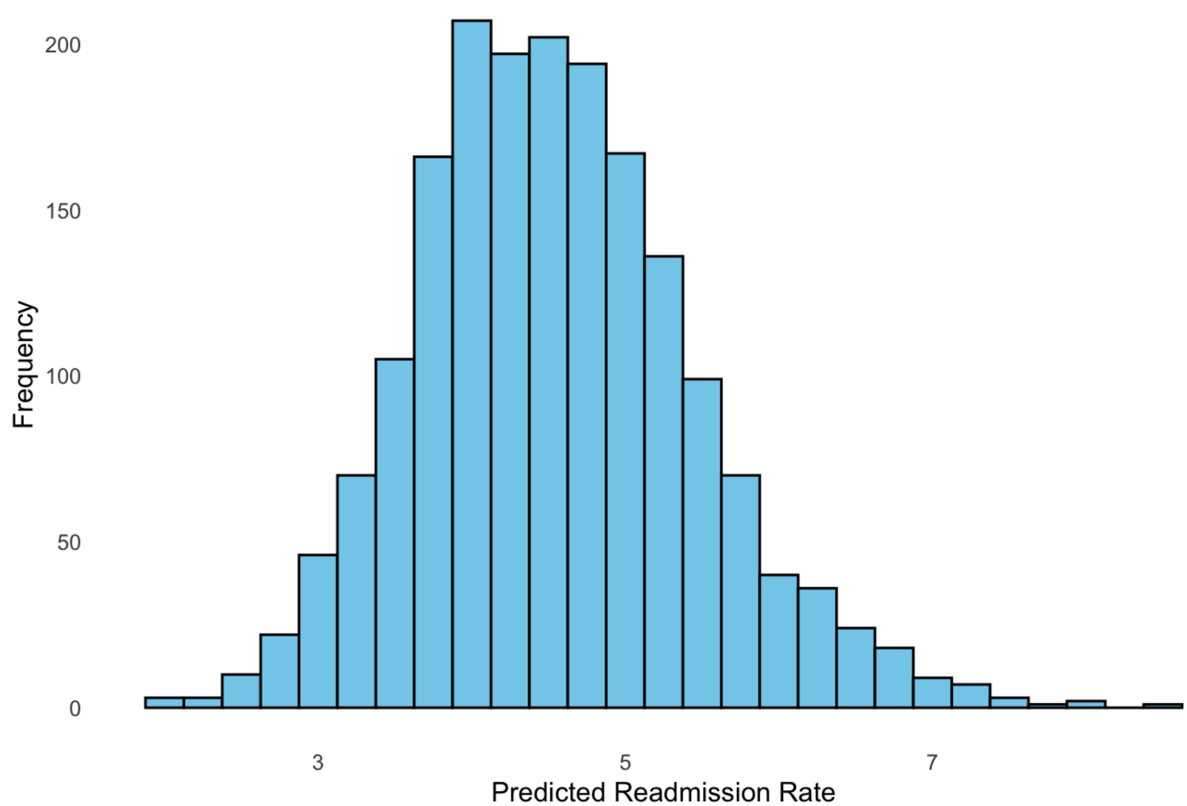


Figure 5. Number of Facilities per State

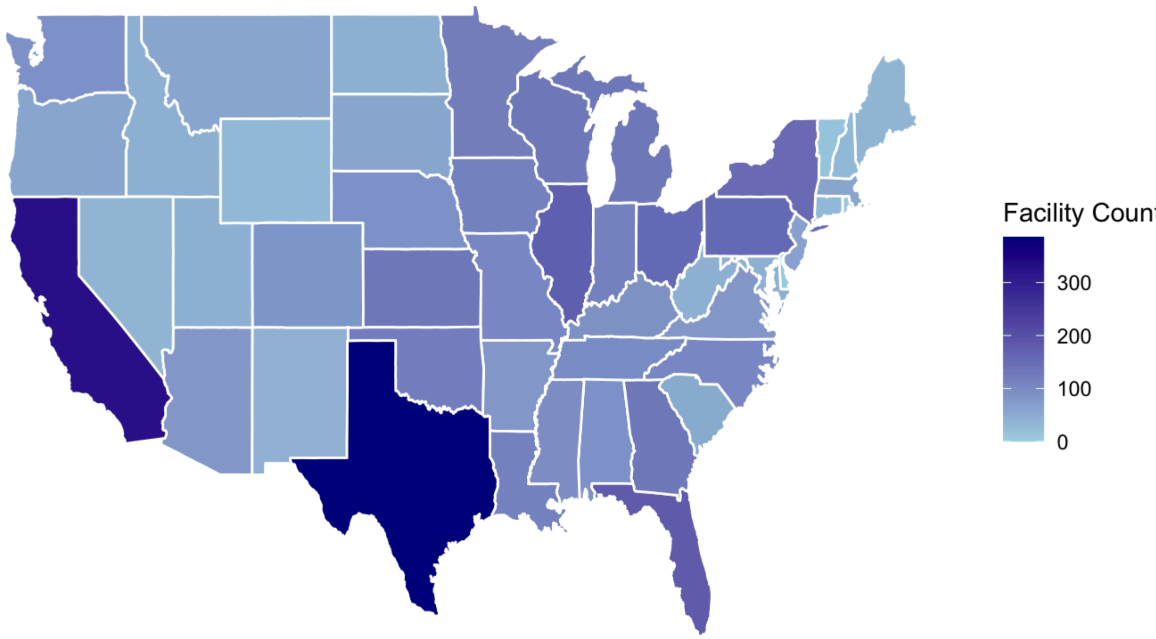
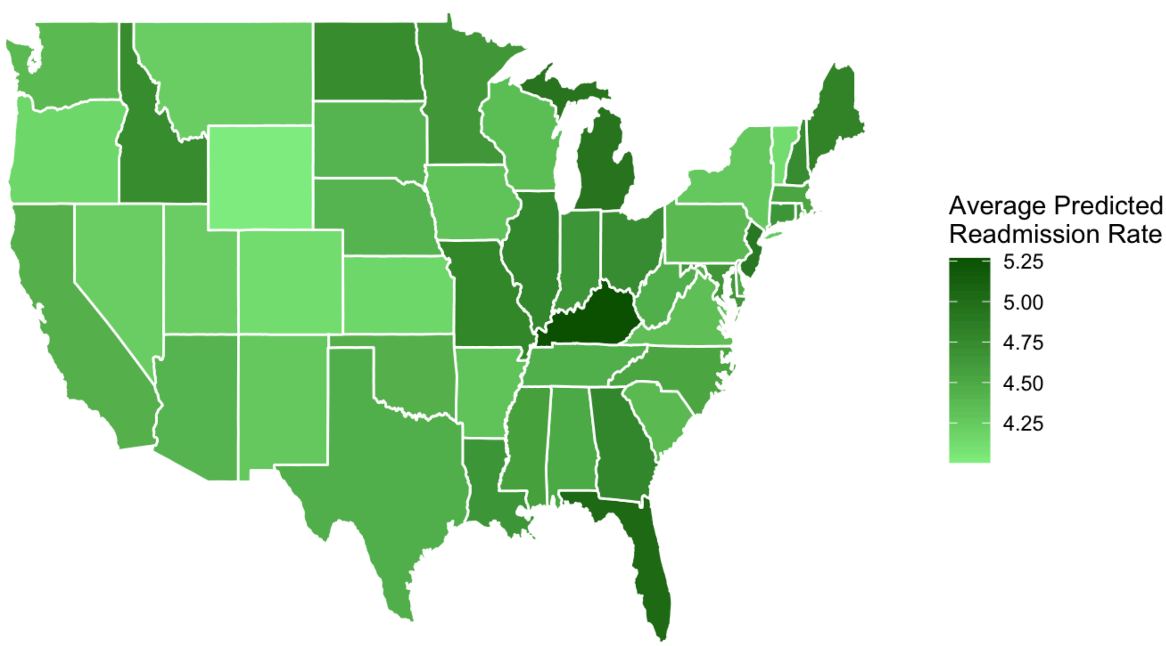


Figure 6. Average Predicted Readmission Rate for Hip/Knee Replacement per State



Unsupervised Learning Models

For unsupervised machine learning, we employed both k-means clustering and hierarchical clustering to explore the underlying structure of the dataset. For k-means clustering, we first standardized the numeric features and used the elbow method to determine the optimal number of clusters, which was found to be three. The k-means algorithm was then applied, and cluster characteristics were analyzed by computing the mean of numeric features within each cluster. Visualizations of feature distributions across clusters and the overall cluster structure were created to gain insights into the clustering results (Figure 7). Hierarchical clustering was also performed to validate the k-means results. We computed a distance matrix and applied Ward's method to obtain cluster solutions. An analysis of within-cluster sum of squares (WCSS) for different cluster counts was conducted to determine the optimal number of clusters, which again suggested three clusters. The hierarchical clustering results were compared with those from k-means by visualizing clusters in the PCA-reduced feature space and analyzing cluster characteristics (Figure 8). Both clustering methods revealed distinct patterns, but the preliminary findings suggest that the clusters may not be highly meaningful due to potential issues with the data's inherent structure. Overall, the preliminary clustering analysis indicates that while the k-means and hierarchical methods can segment the data, the results may not be sufficiently informative or actionable at this stage.

Figure 7. K-Means Clustering Results

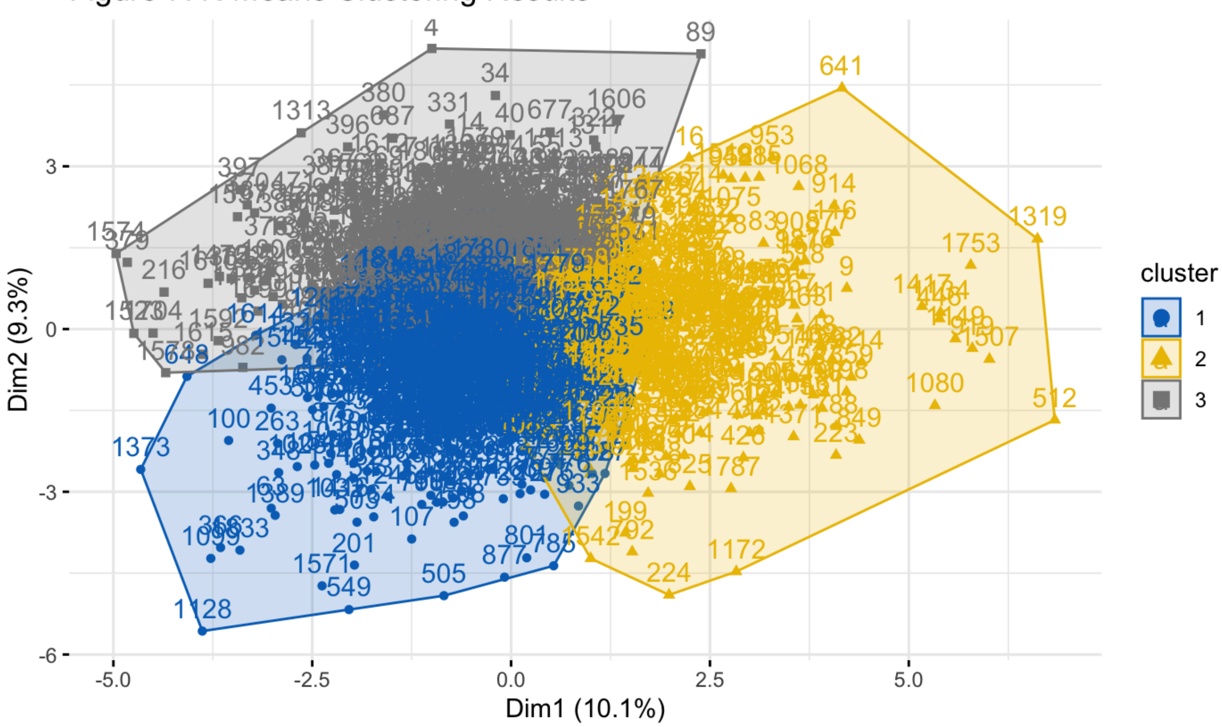
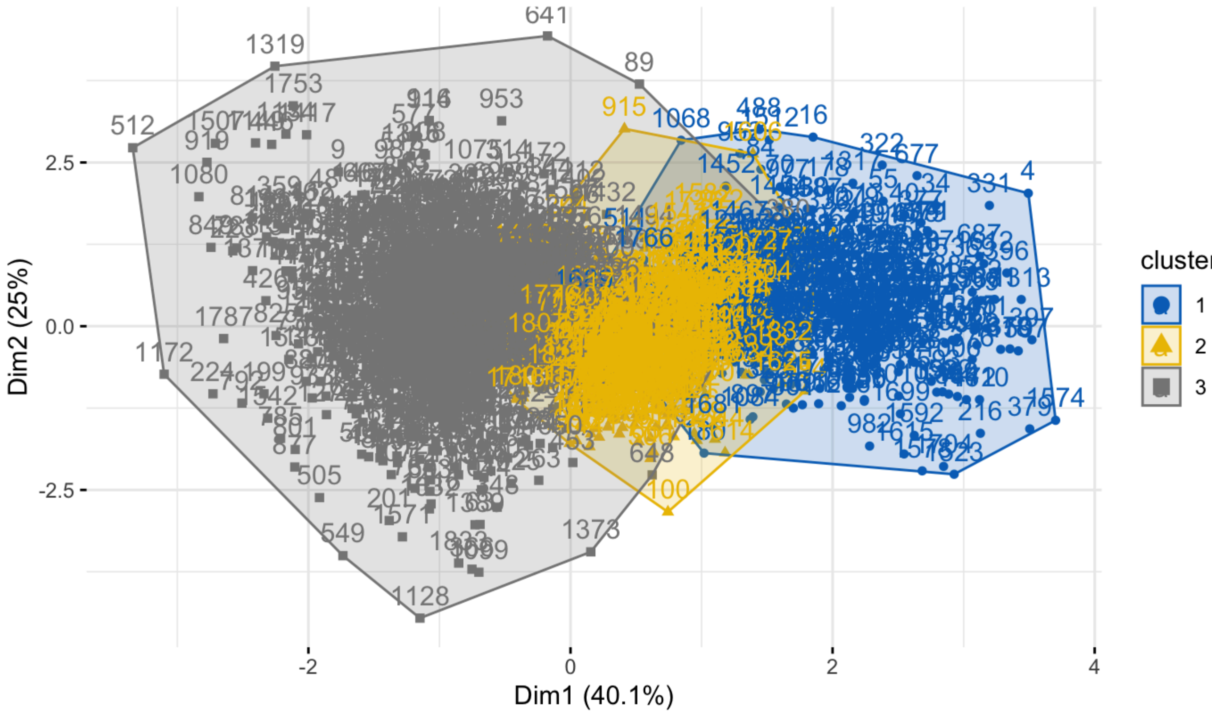


Figure 8. Hierarchical Clustering Results (PCA)



Supervised Learning Models

We implemented three supervised learning models on the final dataset to determine important risk factors associated with hip/knee readmission rates and to predict preferred and non-preferred hospitals, based on the national median readmission rate. For our initial model, we chose the Random Forest algorithm due to its robustness and ability to handle a large number of input variables without overfitting. Random Forests are ensemble learning methods that construct multiple decision trees during training and output the mode of the classes (classification) or mean prediction (regression) of the individual trees. This approach helps in managing the complexities and interactions between numerous features in our dataset. We used a grid search with 7-fold cross-validation to tune the mtry parameter, which determines the number of variables considered at each split in the trees. Next, we chose to create a Logistic Regression model with 5-fold repeated cross-validation to create a baseline comparison with which to compare the Random Forest model. This model is straightforward, easily interpretable, and lists significant predictors. Finally, we created a Support Vector Machine (SVM) model, chosen for its ability to handle complex, non-linear relationships. To train this model, we conducted a 10-fold cross-validation grid search for the most accurate kernel, cost, and gamma values.

Random Forests make minimal assumptions about the underlying data distribution, which makes them versatile. However, they assume that the data is independent and identically distributed, and the model benefits from a large number of features and observations. To validate these assumptions, we conducted residuals analysis, ensuring no patterns indicating violations of independence or homoscedasticity. The Durbin-Watson test was used to check for autocorrelation in the residuals, and residual plots were examined for any non-random patterns. The results showed that the residuals are relatively randomly distributed, with a slight right skew (Figure 9). However, the Q-Q plot indicates that there are more extreme values than should typically be expected in a normal distribution (Figure 10). To address and mitigate overfitting, we utilized training and testing with completely separate datasets and also utilized cross-validation ensure each model's performance was consistent across different subsets of the data.

Figure 9. Histogram of Residuals

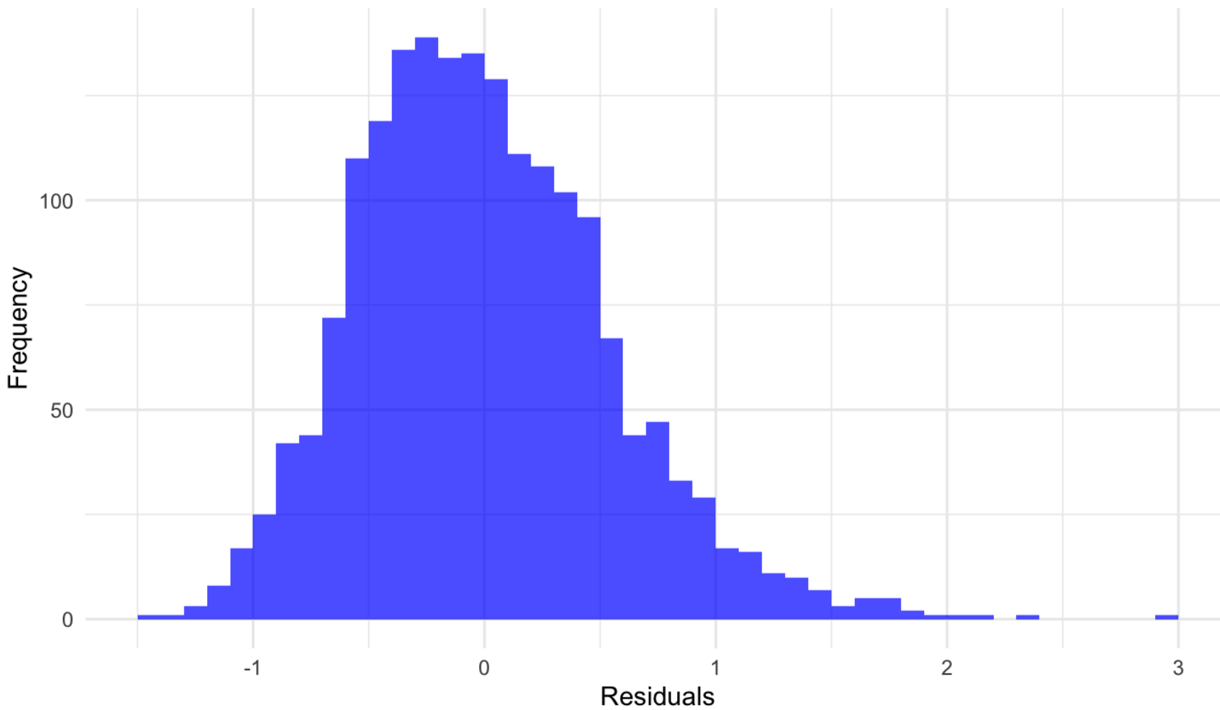
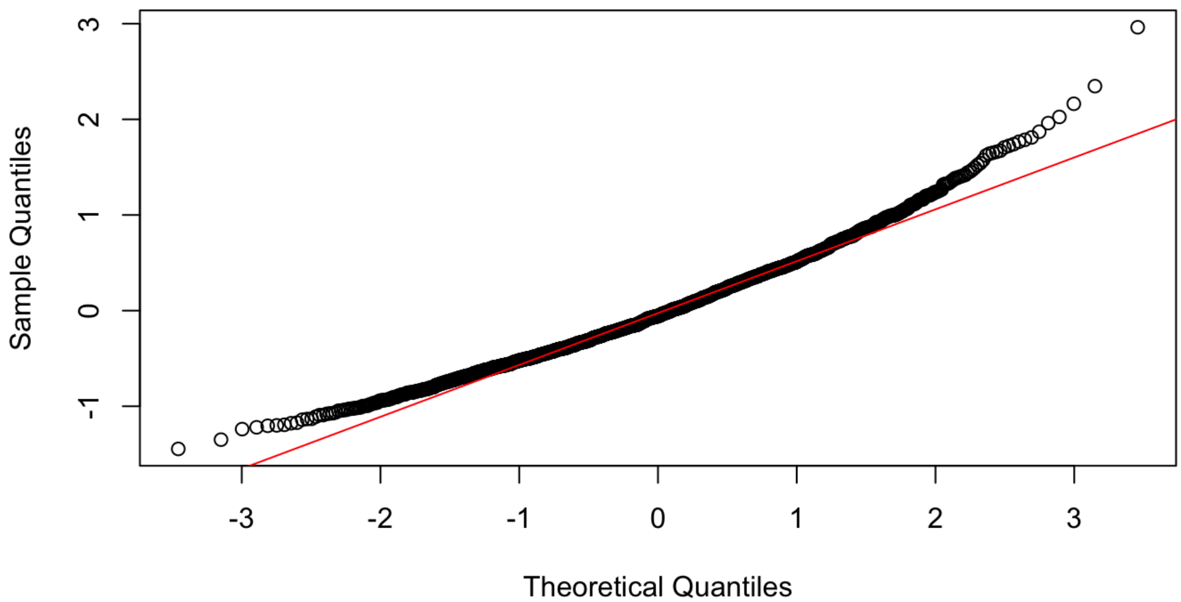


Figure 10. QQ Plot of Residuals



Final Model

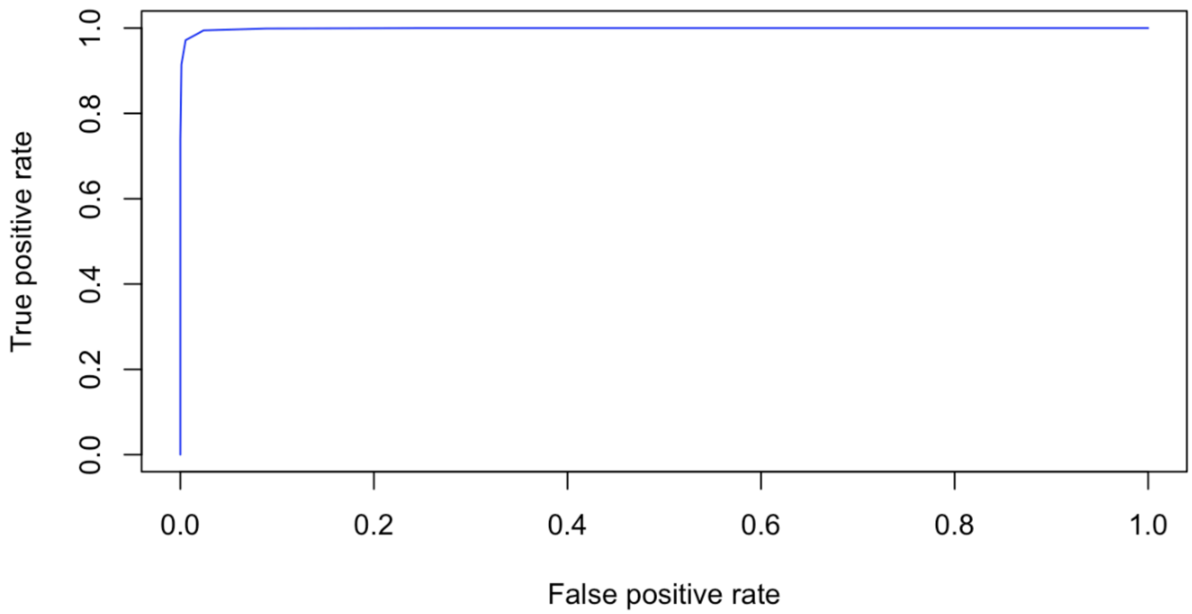
The final model is the Random Forest qualitative model. This model was chosen due to its impressive accuracy and its ability to handle the complex relationships between features.

A grid search with 7-fold cross-validation was implemented to optimize the mtry and ntrees parameters. The mtry parameter determines the number of features considered at each split in the trees. Values of 2, 4, 6, and 8 were considered and the optimal mtry value was found to be 4. Due to computing power, a separate grid search of number of trees (ntree) was implemented, testing values of 5, 10, and 12. The optimal number of trees was determined to be 10.

To validate the model's performance, current HRRP data from 2024 was used as the test set. This creates a real-world use-case for the model and ensures the model's ability to generalize to new, unseen data.

The qualitative Random Forest model achieved an impressive Area Under the Curve (AUC) of 0.99 on the test set, which indicates superior discriminative ability between preferred and non-preferred hospitals for hip/knee replacements (Figure 11). The model also had an impressive accuracy of 0.98, indicating the model correctly classified 98% of the instances. With a Kappa of 0.97, the model performs well beyond what would be expected by chance. The model also boasts impressive specificity and sensitivity, at 0.98 and 0.99, respectively, indicating balanced performance.

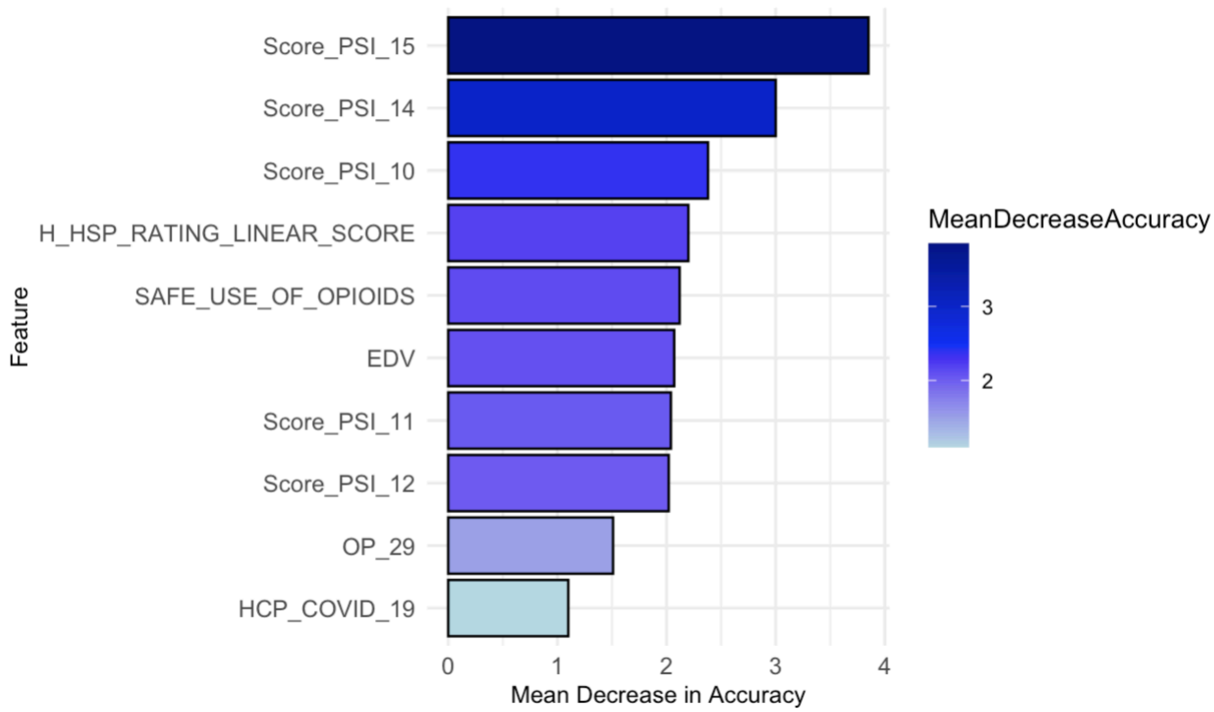
Figure 11. ROC Curve for Random Forest Model



ROC curve nearly touches the top-left corner of the plot, indicating that the True Positive Rate is very high while the False Positive Rate is very low.

A key component of the Random Forest model is that it allows us to assess the importance of each feature in making predictions. Figure 12 shows a list ranking features by their importance, according to Mean Decrease Accuracy (MDA). MDA indicates how much the model's accuracy decreases when the values of a particular feature are randomly permuted (in other words shuffled), while keeping all other features the same. A high MDA for a particular feature means that permuting its values results in a significant drop in model accuracy, which means the feature is highly important. A low MDA suggests the feature is less important, or possibly even harmful to model, introducing noise.

Figure 12. Top 10 Important Features from Random Forest



A ranked list of feature importances as determined by the Random Forest qualitative model. Ranking “Accidental cuts and tears from medical treatment”, “A wound that splits open after surgery on the abdomen or pelvis”, “Kidney and diabetic complications after surgery”, and “Overall hospital rating - linear mean score” as the most important features when determining hospital preferred status.

While the model's performance metrics are impressively high, with an AUC of 0.99 and accuracy of 0.98, this does raise concern regarding potential overfitting, and/or possible data leakage. To address this issue, the following steps would be recommended moving forward:

- Nested cross-validation: Implementation of nested cross-validation may help eliminate some overfitting. This involves an outer loop of cross-validation to assess performance and an inner loop for hyperparameter tuning. However, this is computationally intensive.
- Feature reduction: Experimentation with reducing the number of features may help, perhaps focusing on the top 10-15 most important features, to make a simpler model.
- Collecting more data: If possible, gathering data from a wider range of hospitals over a longer period of time may help the model generalize better across different contexts.

Conclusions

This analysis aimed to identify the risk factors associated with hospital readmission rates for hip and knee replacement patients and to determine preferred and non-preferred hospitals for these procedures. Using training data from the Hospital Readmission Reduction Program (HRRP) for 2019-2022, a qualitative analysis was performed using predictive modeling to determine which hospitals were preferred or non-preferred based on current 2024 data.

The primary hypothesis was that hospitals with better overall hospital satisfaction scores would have lower readmission rates. The Random Forest model, which emerged as our final and best-performing model, identified several key risk factors. Among the most important features were complications such as accidental cuts and tears from medical treatment, wounds splitting open after surgery, and kidney and diabetic complications after surgery. The overall hospital rating was also among the top predictors, agreeing with our initial hypothesis.

The Random Forest model exhibited superior performance, with an AUC of 0.99 and an accuracy of 0.98 on the test set. However, these exceptional performance metrics also raise concern with potential overfitting or data leakage, which warrants further investigation.

Our findings agree with previous research, for example, the study conducted by Edwards et al. (2015), also found that higher patient satisfaction often correlates with better overall patient outcomes. The results of this analysis supports this, as the overall hospital satisfaction rating was among the most significant features in the Random Forest model.

Discussion and Next Steps

<u>Key Takeaways</u>

- Referring back to the initial question of what risk factors are associated with readmission rates in hip/knee patients, complications such as accidental cuts and tears from medical treatment, wounds splitting open after surgery, and kidney and diabetic complications post-surgery emerged as the most important predictors of readmission rates (Figure 12).
- Overall hospital satisfaction rating from HCAHPS scores, was also among the top features, which supports our initial hypothesis about the relationship between patient satisfaction and readmission rates.
- The Random Forest model's high accuracy suggests that it can be a valuable tool for health insurance companies in guiding patients to hospitals with better outcomes for hip and knee replacements.

While these results are promising, there are caveats and concerns that should be addressed:

- The exceptionally high performance metrics (0.99 AUC, 0.98 accuracy) raise concerns with potential overfitting and/or data leakage. This warrants further investigation to ensure the model's generalizability.
- The dataset had many missing values, across multiple features, resulting in the removal of several features (Table 1). This limitation may have excluded potentially important risk factors from the analysis.
- The clustering analysis (Figures 7 and 8) revealed that while the data could be segmented, the results may not be sufficiently informative or actionable at this stage.

Given these findings and limitations, this is a proposed list of recommendations and future steps:

- Implement nested cross-validation to help eliminate some possible overfitting and data leakage.
- Experiment with feature reduction, focusing on the top 10-15 most important features to create a simpler, more generalizable model.
- Consider another model, such as XGBoost, which is robust at handling missing values.
- Collect data from a wider range of hospitals over a longer period of time to improve the model's generalization ability.
- Develop a way for health insurance companies to utilize this model in their decision-making processes. This may involve creating a user interface for analyzing hospital performance and referring patients towards the preferred hospitals.
- Collaborate with healthcare professionals and hospitals to address the identified risk factors. For example, developing interventions that reduce complications and improve overall patient satisfaction.
- Conduct a cost-benefit analysis to quantify the potential savings from reduced readmission rates.

By addressing these issues and continuing with the next steps, we will be able to further our understanding of the risk factors associated with hip and knee replacement readmissions and develop effective strategies for improving patient outcomes and reducing healthcare costs.

Code Availability

All code is available in the following GitHub repository: https://github.com/adelinecasali4/hospital-readmission/.

Appendix 1: References

Centers for Medicare & Medicaid Services. (2024). *2024 Hospital Readmissions Reduction Program (HRRP) Data*. CMS. https://data.cms.gov/provider-data/topics/hospitals

Edwards, P. K., Levine, M., Cullinan, K., Newbern, G., & Barnes, C. L. (2015). Avoiding readmissions—support systems required after discharge to continue rapid recovery? *The Journal of Arthroplasty*, *30*(4), 527–530. https://doi.org/10.1016/j.arth.2014.12.029

Phillips, J. L. H., Rondon, A. J., Vannello, C., Fillingham, Y. A., Austin, M. S., & Courtney, P. M. (2019). How much does a readmission cost the bundle following primary hip and knee arthroplasty? *The Journal of Arthroplasty*, *34*(5), 819–823. https://doi.org/10.1016/j.arth.2019.01.029

Appendix 2: Data Dictionary

Measure ID ***Measure utitilized in final dataset**	**Description**
ED_2	Average (median) admit decision time to time of departure from the emergency department for emergency department patients admitted to inpatient status.
EDV*	Emergency department volume. Number based on the volume of patients submitted by a hospital used for the measure OP-22: Left without Being Seen.
ExcessReadmissionRatio_HIP-KNEE	The ratio of the predicted readmission rate to the expected readmission rate, based on an average hospital with similar patients. Performance is compared against a ratio of one, such that below one is better and above one is worse in terms of readmission rates.
ExpectedReadmissionRate_HIP-KNEE	The expected number of readmissions in each hospital is estimated using its patient mix and an average hospital-specific intercept. It is thus indirectly standardized to other hospitals with similar case and patient mixes.
FacilityId	Unique facility identifier.
FacilityName	Name of the facility.
H_HSP_RATING_LINEAR_SCORE*	Overall hospital rating - linear mean score. Employs all survey response items in each HCAHPS measure and are converted and combined into a 0-100 linear-scaled measure score.
H_RECMND_LINEAR_SCORE	Recommend hospital - linear mean score. From question: Would you recommend this hospital to your friends and family?

HCP_COVID_19*	COVID-19 vaccination coverage among healthcare providers.
IMM_3*	Healthcare workers given influenza vaccination.
NumberOfReadmissions_HIP-KNEE	Crude number of readmissions in each hospital within 30 days.
OP_18b*	Average (median) time patients spent in the emergency department before leaving from the visit.
OP_18c	Average time patients spent in the emergency department before being sent home (Median Time from ED Arrival to ED Departure for Discharged ED Patients – Psychiatric/Mental Health Patients).
OP_22	Percentage of patients who left the emergency department before being seen.
OP_23	Percentage of patients who came to the emergency department with stroke symptoms who received brain scan results within 45 minutes of arrival.
OP_29*	Percentage of patients receiving appropriate recommendation for follow-up screening colonoscopy.
Payment_PAYM_90_HIP_KNEE	Payment for hip/knee replacement - estimates of payments associated with a 90-day episode of care for hip/knee replacement.
PredictedReadmissionRate_HIP_KNEE*	The number of readmissions within 30 days predicted based on the hospital's performance with its observed case mix. The predicted number of readmissions is estimated using a hospital-specific intercept, and is intended to reflect the annual expected performance of the hospital given its historical case and patient mix and performance.

SAFE_USE_OF_OPIOIDS*	Percentage of patients who were prescribed 2 or more opioids or an opioid and benzodiazepine concurrently at discharge.
Score_COMP_HIP_KNEE	Rate of complications for hip/knee replacement patients.
Score_MORT_30_AMI	Death rate for heart attack patients.
Score_MORT_30_COPD	Death rate for chronic obstructive pulmonary disease (COPD) patients.
Score_MORT_30_HF	Death rate for heart failure patients.
Score_MORT_30_PN	Death rate for pneumonia patients.
Score_MORT_30_STK	Death rate for stroke patients.
Score_Ovr_MORT*	Summary measure (row-wise mean) of Score_MORT_30_AMI, Score_MORT_30_COPD, Score_MORT_30_HF, Score_MORT_30_PN, and Score_MORT_30_STK.
Score_PSI_03*	Rate of pressure sores.
Score_PSI_04*	Deaths among patients with serious treatable complications after surgery.
Score_PSI_06*	Collapsed lung due to medical treatment.
Score_PSI_08*	Broken hip from a fall after surgery.
Score_PSI_09*	Postoperative hemorrhage or hematoma rate.
Score_PSI_10*	Kidney and diabetic complications after surgery.
Score_PSI_11*	Respiratory failure after surgery.
Score_PSI_12*	Serious blood clots after surgery.
Score_PSI_13*	Blood stream infection after surgery.
Score_PSI_14*	A wound that splits open after surgery on the abdomen or pelvis.
Score_PSI_15*	Accidental cuts and tears from medical treatment.
Score_PSI_90	Serious complications (this is a composite or summary measure).
SEP_1	Severe sepsis and septic shock.
SEP_SH_3HR	Septic shock 3 hour.
SEP_SH_6HR	Septic shock 6 hour.

SEV_SEP_3HR	Severe sepsis 3 hour.
SEV_SEP_6HR	Severe sepsis 6 hour.
State	State where the facility is located.
STK_02	Percentage of ischemic stroke patients prescribed or continuing to take antithrombotic therapy at hospital discharge.
STK_05	Percentage of ischemic stroke patients administered antithrombotic therapy by the end of hospital day 2.
STK_06	Percentage of ischemic stroke patients who are prescribed or continuing to take statin medication at hospital discharge.
VTE_1*	Percentage of patients that received VTE prophylaxis after hospital admission or surgery.
VTE_2	Percentage of patients that received VTE prophylaxis after being admitted to the intensive care unit (ICU).