Table 1: **Table: Summary Statistics by 30-Day Complication Status**

Variable	**No Complication** N = 927[1]	**Complication** N = 73[1]	**p-value**[2]
Age	59.22 (14.06)	65.99 (12.11)	<0.001
BMI	27.00 (4.69)	27.49 (5.11)	0.6
Surgery_Duration_Minutes	95.95 (39.79)	113.55 (47.64)	0.001
Estimated_Blood_Loss	126.47 (98.13)	143.85 (107.14)	0.2

[1] Mean (SD)
[2] Wilcoxon rank sum test

Complication Risk Model: SDSC Analytic Pitch

Walker Blackston, MSPH

2025-04-04

Background

To demonstrate how I think through surgical outcome prediction and statistical problems, I built a prototype risk model for 30-day complications based on simulated surgical data.

In this hypothetical dataset of 1,000 patients, we sought to compare predictors of complications at 30-days follow up from surgery. We simulated a variety of demographic and clinical variables based on seeds of real-world data and expected distributions in a random sample of patients. The following brief will highlight steps in the analysis, modeling and interpretation of findings. A full printout of code used to generate data and this document are available in the Github provided in my pitch.

Data Summary

These data suggest relatively balanced groups for those with and without complications within 30 days post-op, except that those experiencing complications are more likely to be older ($p < .001$), experience a longer surgery ($p < .01$) and experience intra-operative events ($p < .01$).

Visualizing these data can help us understand these relationships more clearly.

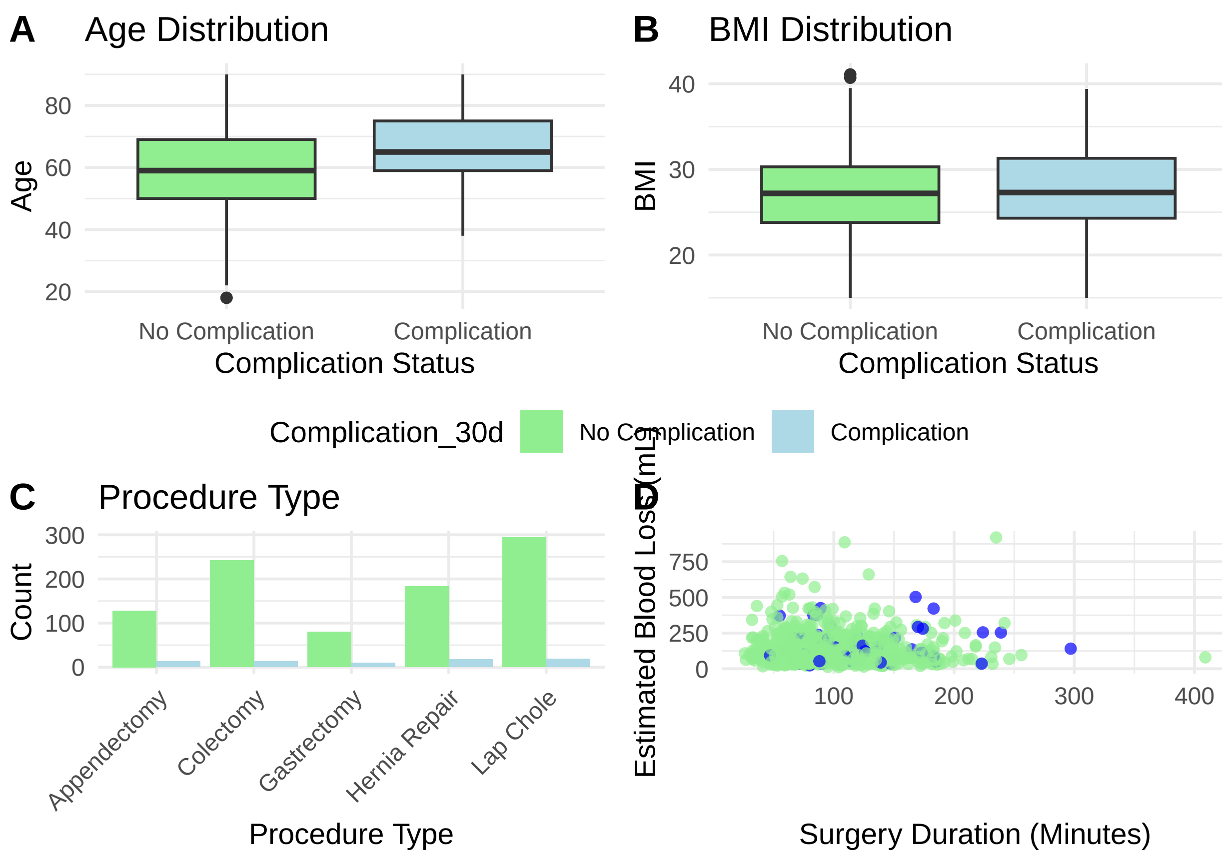
Table 2: **Table: Procedure, Events and ASA class by 30-Day Complication Status**

Procedure Type	**No Complication** N = 927[1]	**Complication** N = 73[1]	**p-value**[2]
Procedure_Type			0.2
Appendectomy	128 (14%)	13 (18%)	
Colectomy	242 (26%)	13 (18%)	
Gastrectomy	80 (8.6%)	10 (14%)	
Hernia Repair	183 (20%)	18 (25%)	
Lap Chole	294 (32%)	19 (26%)	
Intraoperative_Events	138 (15%)	21 (29%)	0.002
ASA_Class			0.039
1	83 (9.0%)	8 (11%)	
2	290 (31%)	18 (25%)	
3	352 (38%)	22 (30%)	
4	159 (17%)	16 (22%)	
5	43 (4.6%)	9 (12%)	

[1] n (%)

[2] Pearson's Chi-squared test; Fisher's exact test

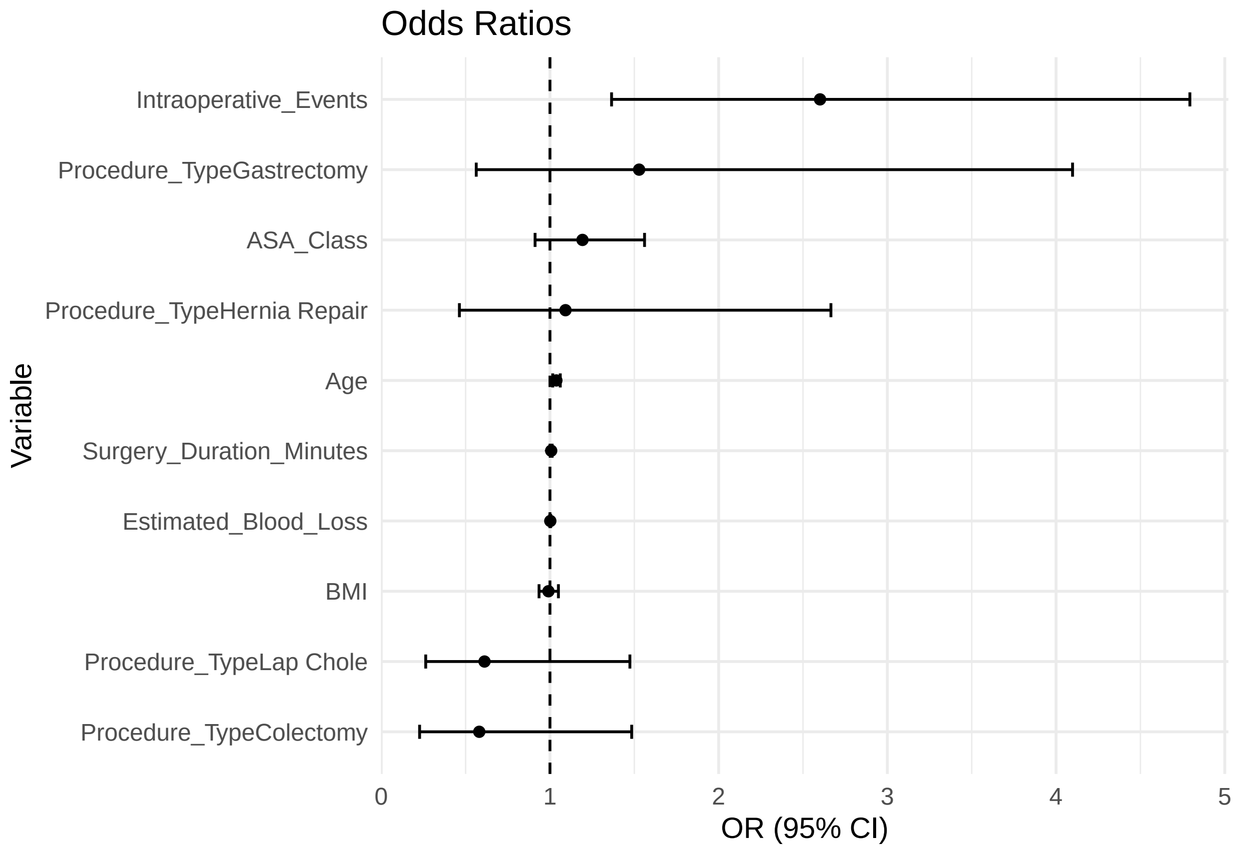
Exploring the data



Predictive Modeling

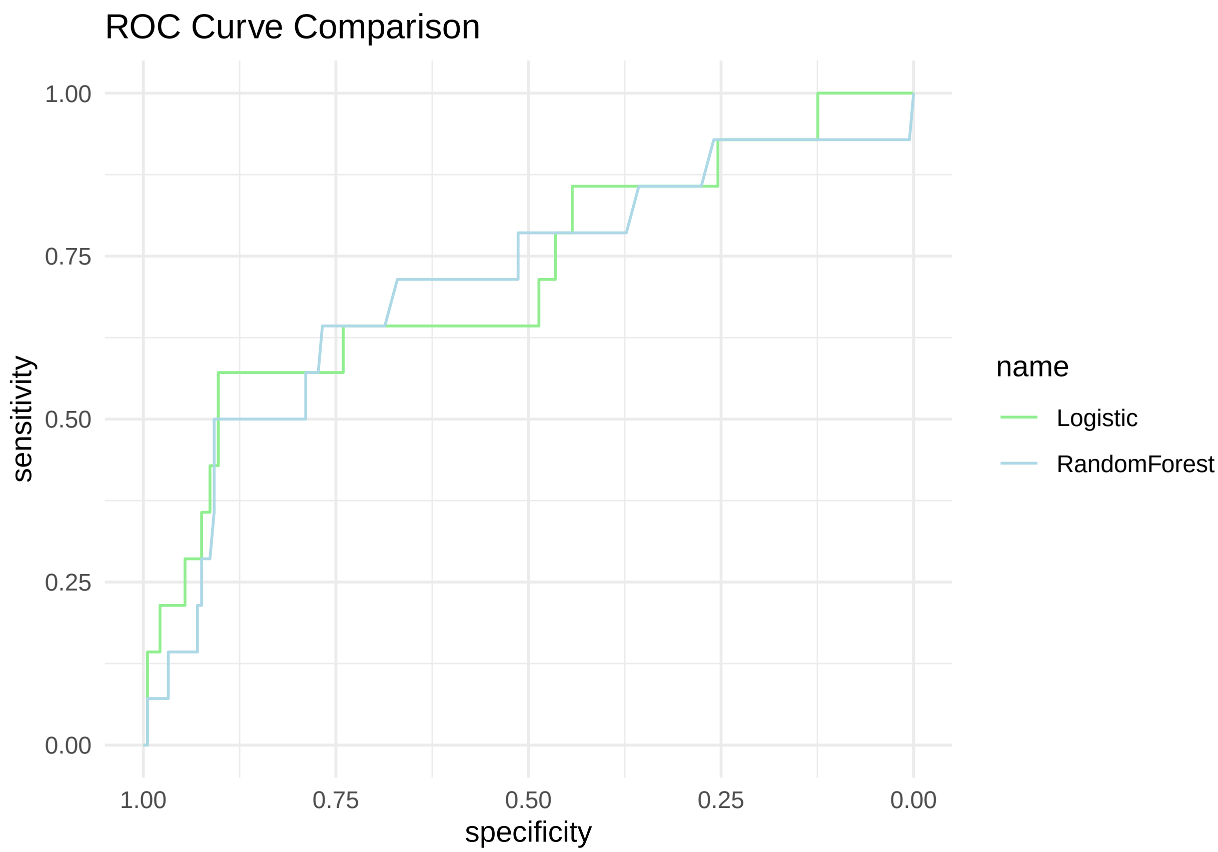
Logistic Model

In order to best predict post-op complications, we want to build a series of robust, interpretable models. Often prediction can come at the expense of interpretability, so in this case we first want to build an explainable logistic model and evaluate what variables most contribute to increased risk of complication.



While important to analyze output and diagnostics of the models (below), visual inspection of Odds Ratios (ORs) can provide useful information for the most important variables driving our outcome of interest. Here, as only age and the presence of an intra-operative event were significant positive predictors of complications. We can determine this simply and visually by assessing which plots and their standard error *do not* cross the reference line where OR = 1.

Machine Learning Model (Random Forest)



Results

The model achieved an AUC (Logistic): 0.72 and AUC(RandomForest): 0.71, respectively with higher odds of complication associated only with:

- Advanced age
- Intra-operative events

Accuracy and Interpretation

Table 3: Confusion Matrix (Logistic Regression)

Prediction \ Reference	No Complication	Complication
No Complication	185	14
Complication	0	0

Table 4: Overall Performance Metrics (Logistic Regression)

Accuracy	Kappa	AccuracyLower	AccuracyUpper	AccuracyNull	AccuracyPValue	McnemarPValue

0.93	0	0.88	0.96	0.93	0.57	0

Table 5: Class-Specific Performance Metrics (Logistic Regression)

Sensitivity	Specificity	Pos Pred Value	Neg Pred Value	Precision	Recall	F1	Prevalence	Detection Rate	D
1	0	0.93	NaN	0.93	1	0.96	0.93	0.93	

Table 6: Confusion Matrix (Random Forest)

Prediction \ Reference	No Complication	Complication
No Complication	185	14
Complication	0	0

Table 7: Overall Performance Metrics (Random Forest)

Accuracy	Kappa	AccuracyLower	AccuracyUpper	AccuracyNull	AccuracyPValue	McnemarPValue
0.93	0	0.88	0.96	0.93	0.57	0

Table 8: Class-Specific Performance Metrics (Random Forest)

Sensitivity	Specificity	Pos Pred Value	Neg Pred Value	Precision	Recall	F1	Prevalence	Detection Rate	D
1	0	0.93	NaN	0.93	1	0.96	0.93	0.93	

Next Steps

This model demonstrates early potential to extract explainable risk signals from surgical outcomes data. If advanced with real-world video and procedural metadata, the insights could power a live surgeon feedback loop or benchmark dashboard.