Data-Driven Medicine: Using Post-Operative Vital Signs as Biomarkers for Pancreatic Cancer Surgical Complications

Sneha Iyer and Preeti Iyer

Princeton University, Department of Computer Science, Princeton, NJ | Thomas Jefferson Hospital Department of Surgery, Philadelphia, PA

ABSTRACT

This study aims to formulate a more intelligent and proactive way to detect post operative complications in patients undergoing pancreaticduodenectomy surgeries by understanding post operative bio behavior from a real patient set sources from Thomas Jefferson hospital. Using vital signs in the days following surgery and patient medical history, its hopeful that we'll be able to create a predictive model to help health systems have a more comprehensive way to monitor patients and provide appropriate medical attention those who are at high risk for certain complications.

BACKGROUND AND RELEVANT WORK

Data-driven medicine is a burgeoning and exciting field, only just now rapidly growing and starting to tap into the power of analytics. to drive smarter decisions. Hospital departments, especially emergency medicine care, have started to use different markers already present in medical data collection to create live time triage systems and risk prediction monitoring.

A pancreaticduodenectomy ("Whipple surgery") is a high-risk surgical procedure performed for tumors of the pancreatic head and other periampullary structures; while mortality has drastically decreased over the years for this high risk procedure, still the rate of post-operative complications - many of which are medically severe - remains high and there is a need to better understand the nature of common and unusual post-operative complications to hopefully inform predictive analytics which can provide more proactive care during post operative recovery. We hope to get specific insights into ways different vital signs and patient history data can be used to more intelligently monitor patients. Drawing on the

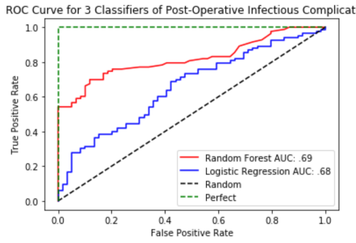
More broadly we hope to truly find a way to make post operative complications more manageable for hospitals by translating our findings into tangible deliverable for the Thomas Jefferson hospital's post operative team (using checklist integration with EHR systems, changing existing medical protocol, etc).

DATA / CLASSIFICATION

POST-OPERATIVE TIME LINE – HOURS AFTER SURGERGY

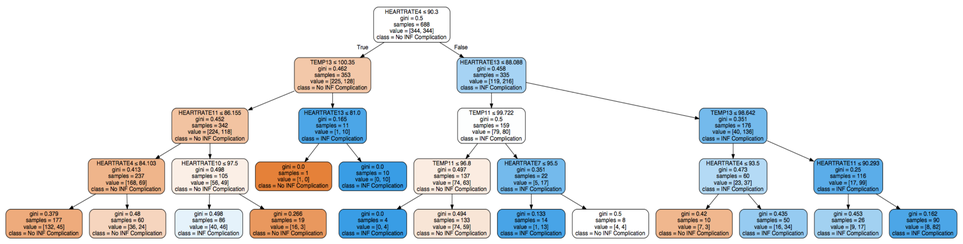
MOST IMPRTANT TIMES AND VITAL SIGNS FOR
INFECTIOUS COMPLICATIONS AND COMPLICATIONS PREDICTION

ROC Curve for 3 Classifiers of Post-Operative Infectious Complications

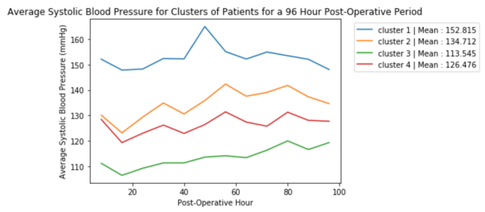


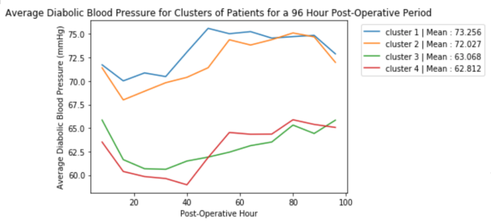
The most important timed predictors of infection and overall complications are shown above. It's interesting that there seems to be a cyclic pattern for infectious and it might offer some earlier signs of importance compared with the features at the end of the 4th day of post operative care being most important for general complications. In general infectious complications also had a better accuracy (**69% versus 61%).**

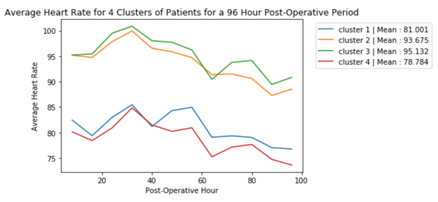
In order to have more understanding of the behavior of the features which were important a decision tree classifier was built with the subset of most important features. To implement checks for high risks patients in a clinal setting it will be useful to have thresholds; we can use the decisions tree node splits as and informative basis:

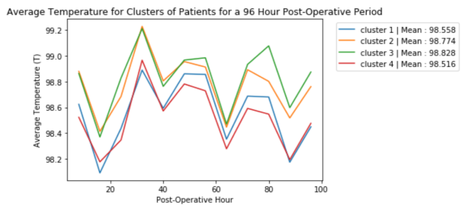


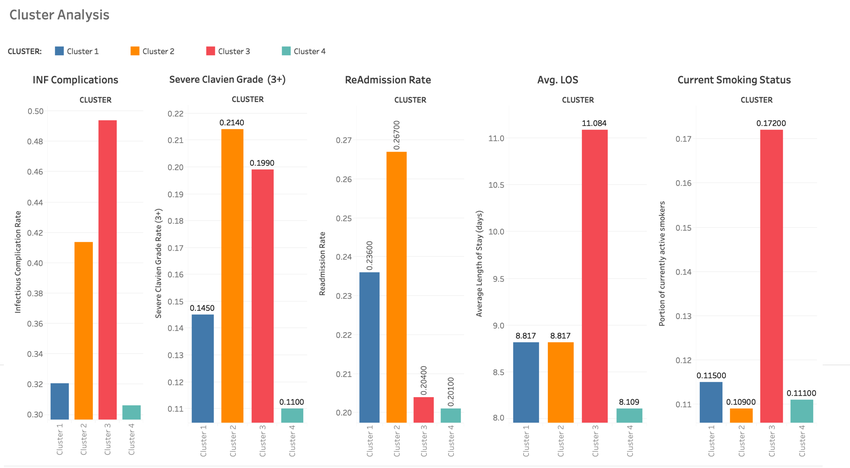
RESULTS











Unsupervised Machine Learning : Clustering Patient Profiles
Using Gaussian Mixture Models, 4 distinct patient clusters were found. These clusters were analyzed with respect to vital signs and other diagnostic metrics like Length of Stay , Clavien Grade (how severe a complication is), and patient history. There very visibly was a cluster (*Cluster 3)* with the highest infection and complication rates, highest LOS, highest average heart rates, who were also prevalent smokers however these patients had less severe complications. There was another cluster (*Cluster 2)* which seemed to also have extremely high heart rates and complication rates which also had patients who markedly had more severe complication, leading to a very high readmission rate by these patients down the line.

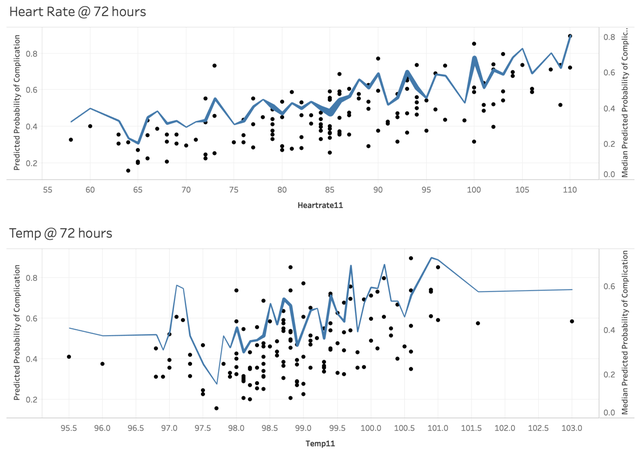
It was seen that temperature follows uniform trajectory between all classes of patients (while those most prone to complications did have higher temperatures), while heart rate offered a more divergent descriptor between classes that had higher complication rates and those that did not.

Heart Rate as a predictor:

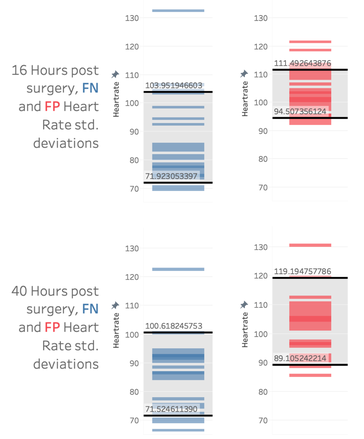
Heart rate was a strong indicator for infectious complication rates. It was particularly interesting that the most important features in a Random Forest Classifier predicting for infectious complications were heart rates at increments of 24 hours after surgery (see figure to the left). To the right it can be seen that there is a slight positive trend between heart rate and temperature at 72 hours (which both proved to be indicators) and the predictive probability of an infectious complication.

It could be seen that False Negatives were more likely to have lower heart rates and false positives were most likely to have higher heart rates.

Predicted Probabilities of Infectious Complications with Vital Signs at 72 Hours
aggreagated by median and with width = No. samples



False Negative and False Positive
Heart Rate standard deviations



ACKNOWLEGMENTS

We'd like to thank Thomas Jefferson hospital for collaboration in this project, we hope to share the results with them and together work towards a checklist based system that can be implemented by nurses to make sure high risk patients can be given appropriate care during the post-operative treatment process.

daguiar@cs.princeton.edu