An integrative predictive model for hospital acquired complications

by

Jenny Liu

Master of Engineering in Electrical Engineering and Computer Science

SUBMITTED TO THE
DEPARTMENT OF ELECTRICAL ENGINEERING
AND COMPUTER SCIENCE
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF ENGINEERING IN ELECTRICAL ENGINEERING
AND COMPUTER SCIENCE
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2015

Signature of Author: ______________________________
Department of Electrical Engineering and Computer Science
May 1, 2015

Certified by: ______________________________
Gil Alterovitz
Assistant Professor of Harvard Medical School; Research Affiliate CSAIL
Thesis Supervisor

Accepted by: ______________________________
Prof. Albert R Meyer
Chairman
Masters of Engineering Thesis Committee

An integrative predictive model for hospital acquired complications
by Jenny Liu
Submitted to the Department of Electrical Engineering and Computer Science
May 1, 2015
In Partial Fulfillment of the Requirements for the Degree of Master of Engineering in Electrical Engineering and Computer Science

ABSTRACT

Hospital Acquired Complications (HACs) are a serious problem affecting modern day healthcare institutions. It is estimated that in US hospitals, HACs cause an approximately 10% increase in total inpatient hospital costs. With US hospital spending totaling nearly $900 billion per year, the tremendous damages caused by HACs is no small matter. Early detection and prevention of HACs could greatly reduce strains on the US healthcare system and improve mortality rates. Here we show a machine-learning model for predicting the occurrence of HACs using clinical data limited to short periods following Intensive Care Unit (ICU) admission. In addition, we also identify several keystone features that demonstrate high predictive power HACs during certain time periods following patient admission. Based on our research, we can reduce excessive hospital costs due to HAC by at least $10 billion annually. We can also reduce the number of excessive hospital stay days by 4.6 million days, and potentially reduce patient mortality by at least 10k patients. The classifiers and features analyzed in this study show high promise of being able to be used for accurate prediction of HACs in clinical settings long before the complication symptoms are manifested. These findings could provide a great aid to doctors and other healthcare professionals in containing the damages caused by HACs in healthcare institutions nationwide.

1 Introduction

Hospital Acquired Complications (HACs) are complications affecting patients after a hospital admission. Most commonly, these issues are caused by infection and other issues caused as a side effect of primary treatments. However, many HACs are not identified in a timely manner, and often lead to further medical issues causing increased hospitalization time or even death. It is estimated that hospital acquired complications alone cause nearly 99,000 deaths annually in the United States[3]. Early identification of HACs will improve their treatment and mortality rates, ultimately increasing patient care quality and reducing unnecessary inpatient costs.

The development of large-scale Electronic Medical Records (EMR) databases containing diverse clinical data have led to a surge in medical informatics over the past several years[4-6]. Previous findings have demonstrated that patient phenomic information can be seen as a function of their temporal clinical features, such as laboratory values and medication information[7]. Since the manifestation of a HAC is also a patient phenotype, it follows that there must be a way to identify and predict HACs using such temporal features.

The Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) database is a collection of EMRs for a cohort of critically ill patients admitted to the Beth Israel Deaconess Medical Center's (BIDMC) Intensive Care Unit (ICU) between 2001 and 2007[8]. The MIMIC database contains the full medical records of 24,580 adults admitted to BIDMC's ICU during this eight-year period. Analysis of this rich medical cohort can lead to new discoveries for the prediction and early prevention of problematic HACs for clinical use.

2 Goal

To build a model that predicts whether patients will have hospital acquired complications using machine learning methods.

Table 1. State of the art vs. M.Eng Thesis

Author	Prediction Topic	Dataset	Classifier	AUC
Escobar[9]	Unplanned transfer from the medical-surgical ward to ICU	Kaiser Permanente Medical Care Program. n= 102,422	Logistic Regression	0.78
Tabak[10]	Morbidity of inpatients	Care Fusion. n=770,523	Logistic Regression	0.91
Dai[11]	Hospitalization due to heart disease	Boston Medical Center. n= 45,579	Supervised Machine Learning	0.82
Tran[12]	Suicide risks	Barwon Health. n = 7,399	Logistic Regression	0.79
Roubinian[13]	Red blood cell transfusion	Kaiser Permanente Northern California. n = 275,874	Logistic Regression	0.87
Rana[14]	Unplanned readmission after myocardial infarction	Barwon Health. n = 1660	Logistic Regression	0.78

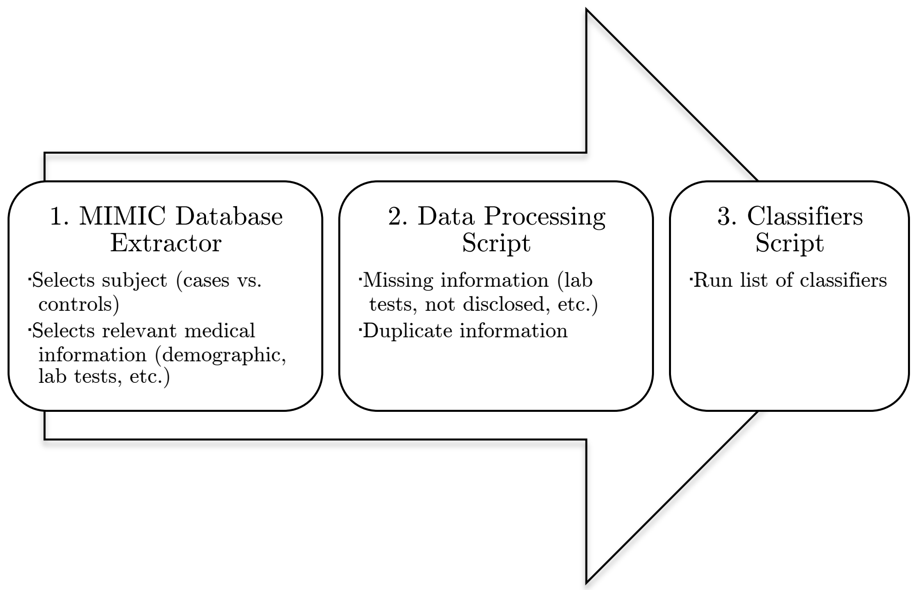
M.Eng Thesis	Hospital Acquired Complications	MIMIC Database. n= 10,386	Decision Tree	0.46-0.91

3 Methods

Dr. Jeremy Warner and Peijin Zhang established the following pipeline (Figure 1), and I collaborated with them for my thesis.

Medical records including lab and medication information were extracted from the MIMIC database. Patient ICD-9-CM discharge codes were used to identify patients suffering from HACs. A list of ICD-9-CM codes indicative of HACs is shown later in this essay. A total of 5193 hospital admissions in MIMIC were identified to have HACs by these definitions. An equal number of 5193 hospital admissions were randomly sampled from the remaining population excluding patients whose hospital stay durations were beyond the 1^{st} or 99^{th} percentiles to serve as controls.

Figure 1. Pipeline



3.1 MIMIC Database Extractor

3.1.1 Iteration 1

Initial contact time of a patient were algorithmically derived times of admission to the hospital. While some patients were directly admitted to the ICU, other patients were in a normal hospital bed for some period of time before admitted to the ICU. The last contact time was defined as the time of ICU discharge for same day discharge patients, or the time of the last laboratory test or POE stop order for patients with longer hospitalizations. Clinical information was then extracted for each admission based on time from first contact. For intervals of 1, 2, 3, 6, 12, 18, 24, 48, 72, 96, and 120 hours from first contact, information about patients were extracted by recording their maximum, minimum, and median values for all lab results, as well as all POEs assigned to the patient. Information about patient demographics such as age, gender, and ethnicity were included as well. Patients whose last contact times were outside of the specified interval were excluded for the analysis.

3.1.2 Iteration 2: Expand ICD-9-CM complications list

We expanded the list of complications by appending possible HAC ICD-9-CM codes to our complications.expanded.csv file. See https://github.com/GilAlterovitz/phewas/blob/master/complications.candidates.expanded.csv

By expanding the list, we increased our sample size of HAC patients. We hypothesized that increasing our sample size will help increase our AUC. See https://github.com/GilAlterovitz/phewas/blob/master/t.phenome.extractor.14.1.R

3.1.3 Iteration 3: Append categories to patient data.

Next, we decided to categorize the HACs into five different categories and append categories to patient information. Categories were assigned depending on the type of ICD9 complication. See https://github.com/GilAlterovitz/phewas/blob/master/complication.candidates.csv

The five categories are defined in Table 2.

Table 2. Category number and labels

Category Number	Category Label
1	Infectious Complications
2	Hemostatic or hemolytic complications
3	Surgical/procedural complications
4	Medical complications
5	Other complications

We hypothesized that subdividing HACs into five different categories would increase AUC because we would be predicting for a subset of complications that would have more similarities within one another. See https://github.com/GilAlterovitz/phewas/blob/master/t.phenome.extractor.14.2.R

3.1.4 Iteration 4: Append ICD-9-CM codes to patient data

Finally, we decided to append ICD-9-CM codes to patient information. We hypothesized that subdividing the HACs to their specific ICD-9-CM codes, would help increase our AUC because we would be predicting for a finer granular specific ICD-9-CM codes.

I modified the script to extract the ICD-9-CM codes to append to hadm˙id and to merge with the cohort. There can be at least one ICD-9CM code per patient so they were appended as a string. See https://github.com/GilAlterovitz/phewas/blob/master/t.phenome.extractor.14.3.1.R

3.2 Data Processing Script

3.2.1 Iteration 1

To prevent excessive noise and cluttering of our data, lab features missing in greater than 50% of our selected cohort and POE features missing in greater than 95% of our cohort were removed prior to analysis. Continuous features were then discretized into three equal frequency bins (low, medium, high) prior to classifier training. All missing values were treated as its own, distinct 4th bin. De-duplication was also performed on the data set due to synonymous naming of categorical variables such as “Amber” vs. “AMB”.

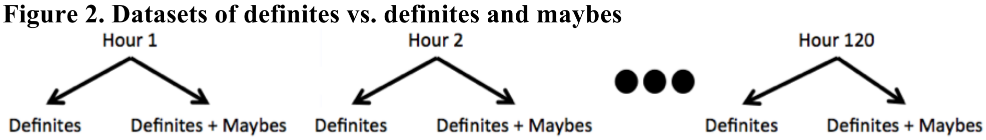
3.2.2 Iteration 2: Definites vs. definites and maybes

We split data into definites vs. definites and maybes.

The “maybes” in this iteration are defined as the expanded list of candidate ICD-9-CM codes.

As a result, we have a total of 11 (different set of hours) * 2 (definites vs. definites and maybes) = 22 data sets.

Figure 2. Datasets of definites vs. definites and maybes



3.2.3 Iteration 3: Categories 1-5

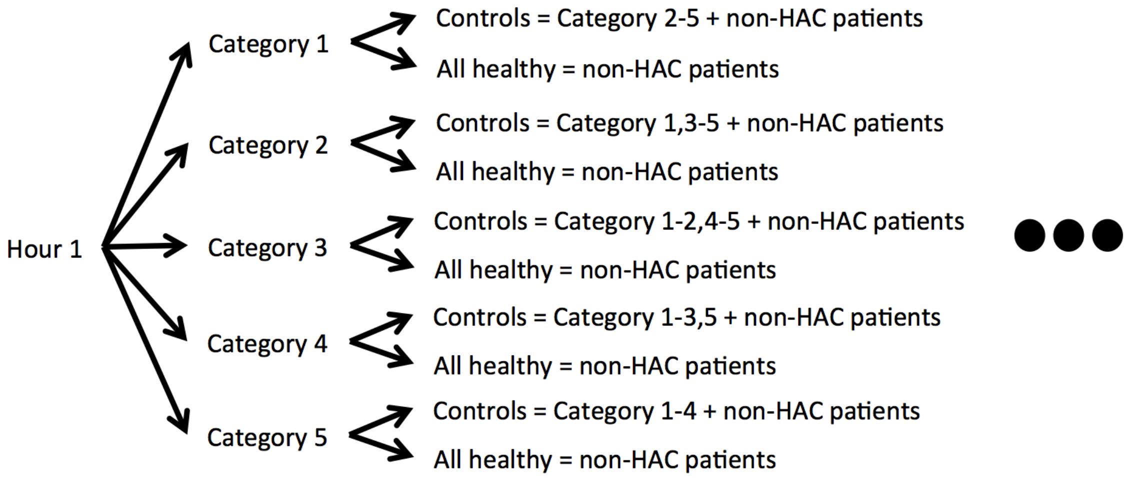
We split data into five categories with controls vs. all healthy.

In total, we have 11 (different set of hours) * 5 (categories) * 2 ("controls" category N-5 with non-HAC patients vs. "all healthy" non-HAC patients only) = 110 data sets.

Controls mean Category N is marked as Case Yes, and all other Categories N-5 and all non-HAC patients are marked as Case No. For example, we have a set of data where Category 1 is marked as Case Yes, and Category 2-5 and non-HAC patients are marked as Case No.

All healthy means that Category N is marked as Case Yes, and all non-HAC patients are marked as Case No. Category N-5 are excluded from this data set. For example, we have a set of data where Category 1 is marked as Case Yes, and non-HAC patients are marked as Case No.

Figure 3. Datasets of categories with controls vs. all healthy



3.2.4 Iteration 4: Categories 1-5 with discretization

We Split data into five categories discretized vs. partial discretized vs. none discretized.

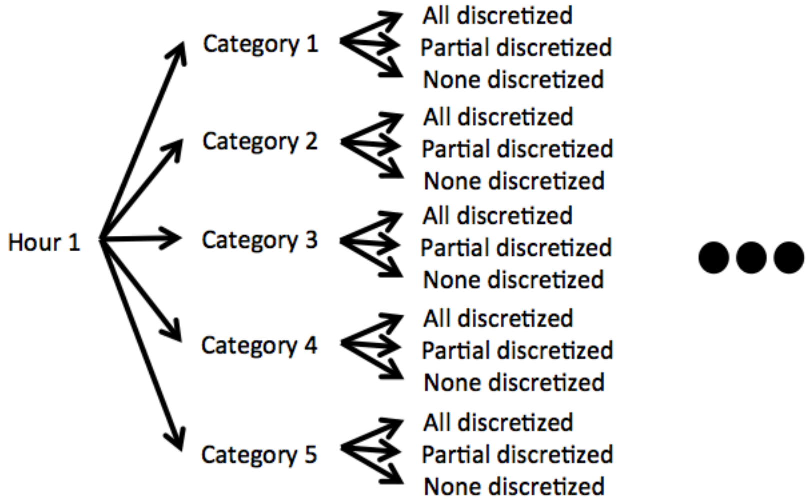
In total, we have: 11 (different set of hours) * 5 (categories) * 3 (all discretized, partial discretization vs. none discretized) = 225 data sets.

All discretized means that all continuous features discretized.

Partial discretized means that features with less than 25% missing were discretized, whereas the rest were imputed using mean of the missing feature.
None discretized meant that all continuous features were imputed with the mean. This was achieved via R using aregImpute from the HMisc package. See https://github.com/GilAlterovitz/phewas/blob/master/NewExtractFix.R

I wrote one of the data processing scripts here: https://github.com/GilAlterovitz/phewas/blob/master/dataprocessor.R

Figure 4. Datasets of categories with all discretized vs. partial vs. none

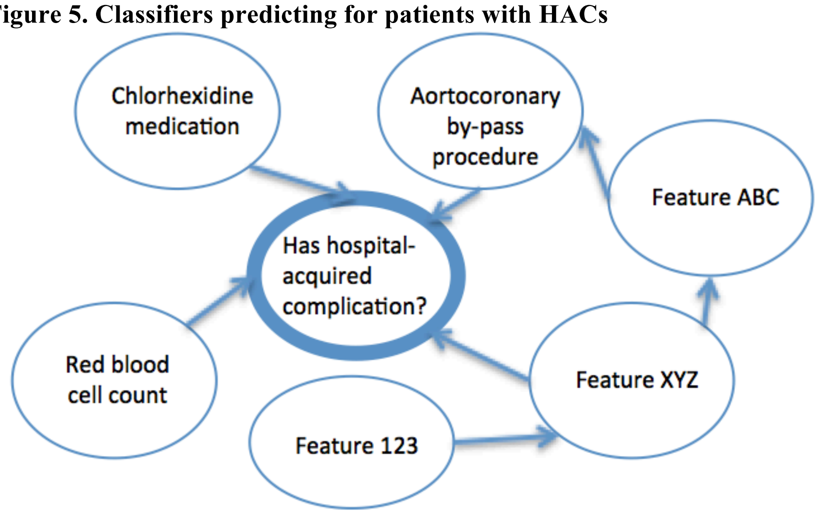


3.3 Classifiers Script

3.3.1 Iteration 1

Classifiers were then trained on the processed dataset to predict the occurrence of HACs using the temporal information extracted across each time interval. Ten fold cross validation was performed to assess the performance of our generated network on additional external data[15-17]. We ran the following list of classifiers seen in Table 3.

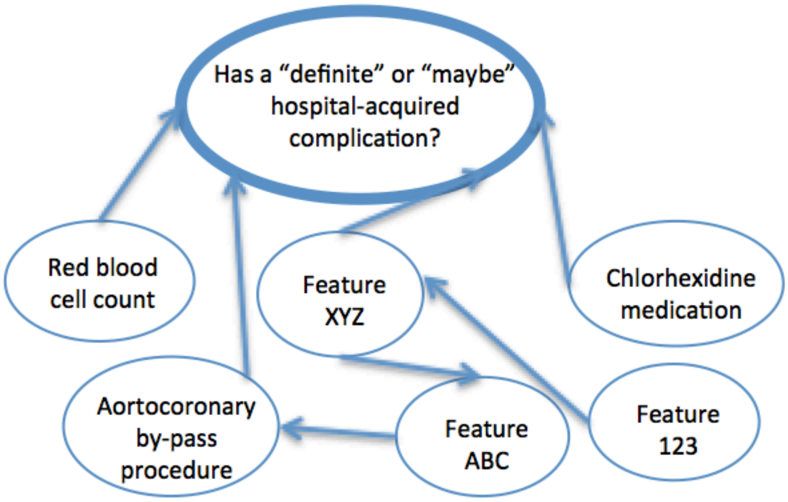
Figure 5. Classifiers predicting for patients with HACs



3.3.2 Iteration 2: Predicting for Case Yes = maybes or definites

Classifiers were then trained to predict Case Yes, which included both definite complications and maybe complications. Case No were non-HAC patients.

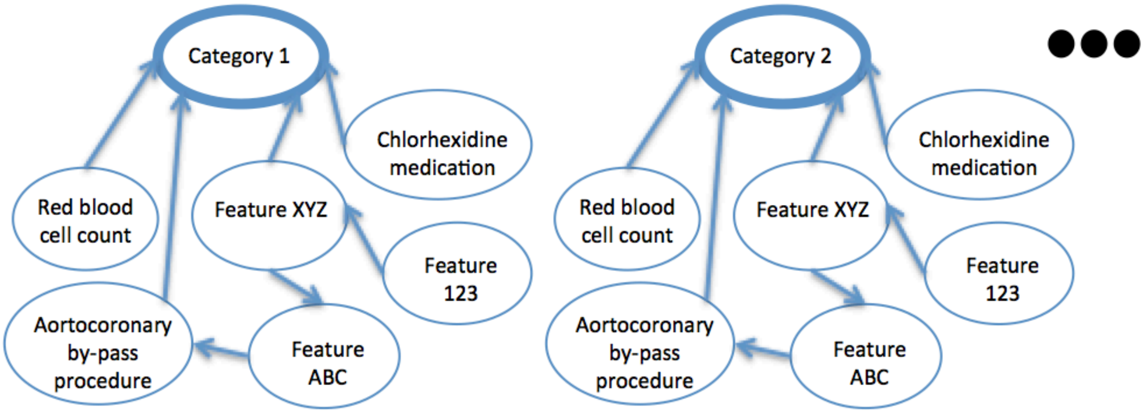
Figure 6. Classifiers predicting for patients with definite HACs or maybe HACs



3.3.3 Iteration 3: Predicting for Case Yes for each category

Classifiers were then trained to predict Case Yes for each Category 1 through 5. Case No depended on which dataset you used. If you used the Controls dataset, Case No included the other categories, e.g. 2-5, and also included non-HAC patients. If you used the All healthy dataset, then Case No only included non-HAC patients.

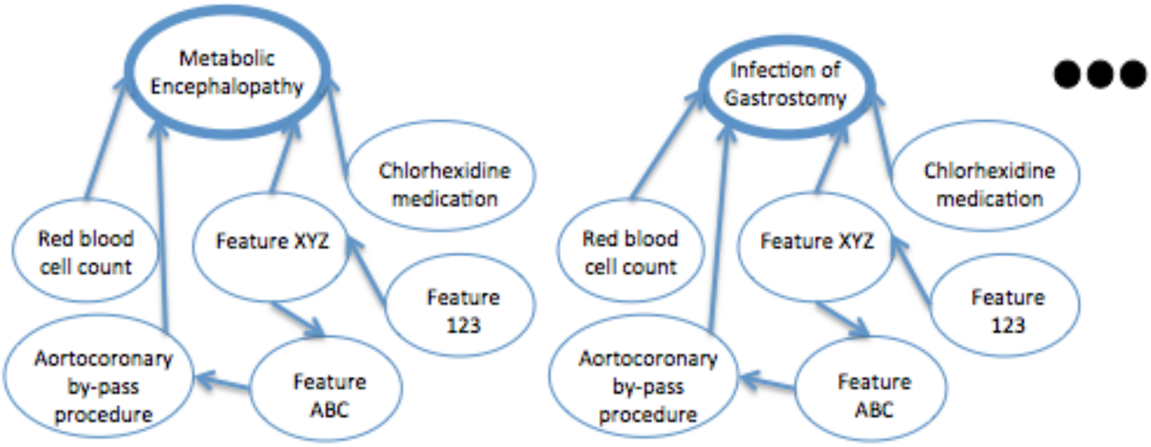
Figure 7. Classifiers predicting for patients with different HAC categories



3.3.4 Iteration 4: Predicting for ICD-9-CM codes

Classifiers were trained to predict each of the 43 ICD-9-CM codes.

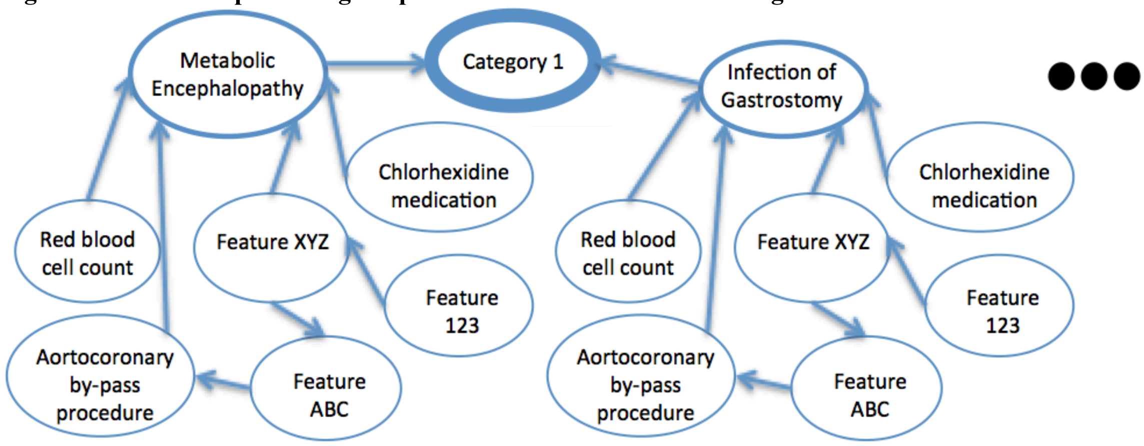
Figure 8. Classifiers predicting for patients with different ICD-9-CM codes



3.3.5 Iteration 5: Dual-layer prediction

Classifiers were trained to predict each of the 43 ICD-9-CM codes, and then taking that prediction to predict for either Category 1-5 complication.

Figure 9. Classifiers predicting for patients with different HAC categories



4 Results

4.1.1 Iteration 1

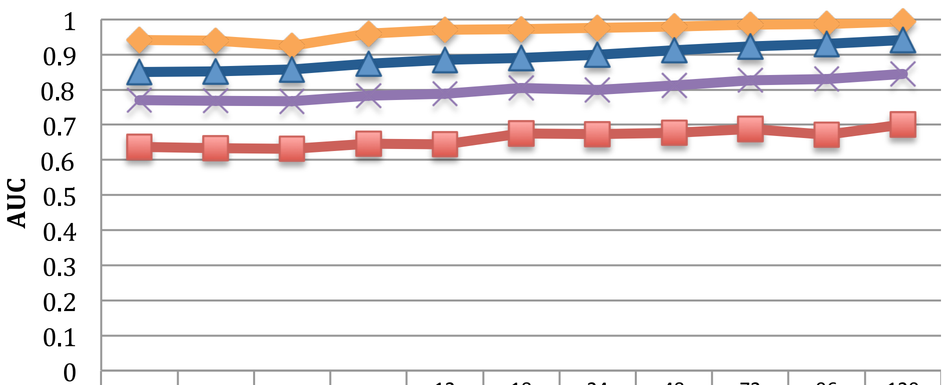
N/A I wasn't part of the project when they performed the first iteration.

4.1.2 Iteration 2: Definites vs. maybes and definites

Below were my observations.

Figure 10. Comparison of Validation AUC 0.6363 – 0.8457 between Maybe and Definites vs. Definites only

AUC increases over time, Maybe is more stable than Definites only model.



	1 hours	2 hours	3 hours	6 hours	12 hours	18 hours	24 hours	48 hours	72 hours	96 hours	120 hours
AUC-Definite-Training	0.9417	0.9403	0.9247	0.9591	0.9698	0.973	0.9768	0.9793	0.985	0.9862	0.9943
AUC-Definite-Validation	0.6363	0.6328	0.6311	0.6456	0.6438	0.6752	0.673	0.6778	0.6884	0.6711	0.7003
AUC- Maybes- Training	0.85	0.8517	0.857	0.8731	0.8857	0.8909	0.8991	0.9122	0.9241	0.93	0.9414
AUC- Maybes- Validation	0.7692	0.7678	0.7671	0.783	0.7878	0.8043	0.7984	0.8118	0.8273	0.83	0.8457

Figure 11. Rules for each hour generated by classifier

HOURS	RULES
1	50444.median>=14.8&50408.max<8.1&50442.min<4.48
2	50444.median>=14.9
3	50386.median<13.35&50444.median>=14.9&50408.median<8.5&vancomycin.IV.rx>=1
6	50444.median>=14.15&vancomycin.IV.rx<1&50386.min<13.5&50090.max<3.7&50140.median>=0.7&phenytoin.POIV.rx<1&aspirin.POPR.rx>=1&age<51.59803
12	50444.median>=13.95&vancomycin.IV.rx>=1&50412.median<31.85
18	50444.median>=14.2&vancomycin.IV.rx>=1&50412.median<31.85
24	50444.median>=13.85&vancomycin.IV.rx<1&50333.median<0.05&PACKED CELL TRANSFUSION.cpt>=1
48	50444.median>=13.9&vancomycin.IV.rx>=1&50412.min<31.6
72	50444.median>=13.95&REOPEN RECENT LAP SITE.cpt>=0&50408.min<7.3&INCISION W/REMOVAL FOREI .cpt>=1
72	50444.median>=13.95&REOPEN RECENT LAP SITE.cpt<0&50386.min>=10.2&50468.max>=33.5
72	50444.median>=13.95&REOPEN RECENT LAP SITE.cpt>=0&50408.min>=7.3&SM BOWEL ENDOSCOPY NEC.cpt>=1&k-phos.IV.rx>=1
96	50444.median>=14.25&PARENTERAL INFUS CONC NU .cpt>=0&50010.median>=0.95&vancomycin.IV.rx<1&50383.min<30.6&PERITONEAL ADHESIOLYSIS.cpt>=1
120	50444.median>=14.3&50386.min<9.7&PARENTERAL INFUS CONC NU .cpt>=0&SM BOWEL ENDOSCOPY NEC.cpt<1&EXTRACORPOREAL CIRCULAT.cpt>=1&REOPEN THORACOTOMY SITE.cpt>=1

Figure 12. Example of ROC curve for Hour 1 of Maybes and Definites

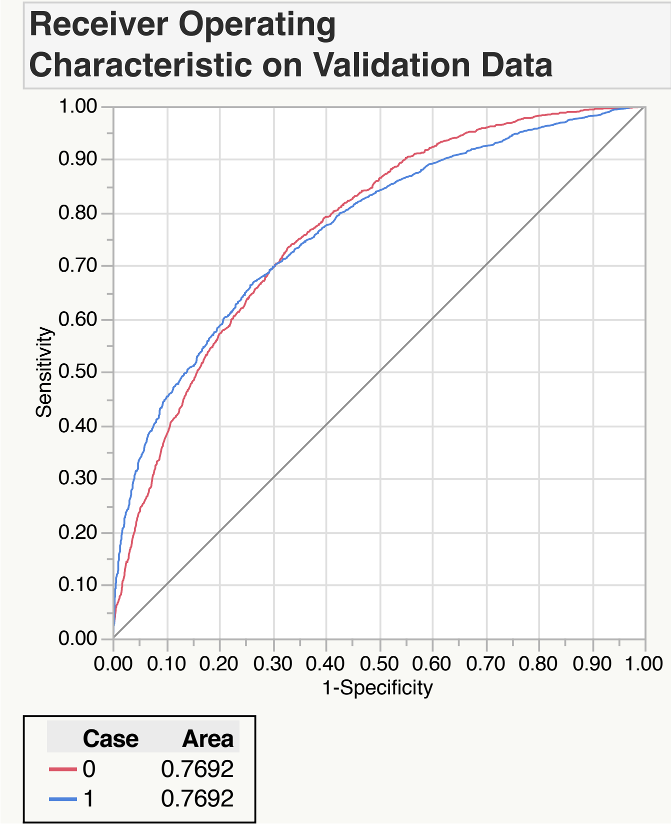


Table 3. Validation Values Breakdown

	Hour 1	2	3	6	12	18
AUC-Definite	0.6363	0.6328	0.6311	0.6456	0.6438	0.6752
True Pos	0.63	0.61	0.64	0.63	0.61	0.64
True Neg	0.57	0.56	0.56	0.58	0.61	0.6
AUC- Maybes	0.7692	0.7678	0.7671	0.783	0.7878	0.8043
True Pos	0.94	0.94	0.94	0.93	0.92	0.93
True Neg	0.25	0.26	0.25	0.33	0.36	0.37

	Hour 24	48	72	96	120
AUC-Definite	0.673	0.6778	0.6884	0.6711	0.7003
True Pos	0.62	0.65	0.71	0.72	0.82
True Neg	0.62	0.6	0.55	0.53	0.41
AUC- Maybes	0.7984	0.8118	0.8273	0.83	0.8457
True Pos	0.91	0.92	0.94	0.95	0.97
True Neg	0.38	0.38	0.34	0.31	0.28

Table 4. Top Ranked Features for Each Hour

Rank	Hour 1	2	3	6	12

1	Median RDW	Median RDW	Median RDW	Median RDW	Median Urea N
2	Minimum RDW	Median Urea N	Median Urea N	Median Urea N	Median RDW
3	Median Urea N	Maximum RDW	Minimum Urea N	Maximum Urea N	Maximum RDW
4	Maximum RDW	Minimum Urea N	Maximum RDW	Minimum Urea N	Maximum Urea N
5	Minimum Urea N	Minimum RDW	Maximum Urea N	Maximum RDW	Minimum RDW
6	age	Minimum Creat	Minimum RDW	Minimum RDW	Minimum Urea N
7	Minimum Creat	Maximum Urea N	Maximum Creat	Median Creat	Median Creat
8	Maximum Creat	Median Creat	Minimum Creat	Minimum Creat	Maximum Creat
9	Maximum Urea N	Maximum Creat	Median Creat	Median MCHC	Median MCHC
10	Median Creat	Minimum RBC	Median PT	Maximum MCHC	Minimum Creat
11	Minimum MCHC	Median MCHC	Maximum RBC	Maximum Creat	Minimum MCHC
12	Median RBC	Median NEUTS	Minimum RBC	Minimum MCHC	Minimum PT

Table 5. Definites- Top Predictors (Median)

1. Median RDW
2. Maximum RDW
3. Minimum RDW
4. Minimum HGB
5. Median HGB
6. Age
7. Maximum HGB
8. Minimum MCHC
9. Medicine: Vancomycin IV
10. Minimum WBC

Table 6. Definites- Biggest increase

1. Medicine: Vancomycin IV
2. Procedure: Parenteral Infusion
3. Procedure: Peritoneal Adhesiolysis
4. Medicine: Neutra-lactulose POPR
5. Minimum Calcium

Table 7. Top ranked features for each hour (continued)

Rank	Hour 18	24	48	72	96	120
1	Median RDW	Median RDW	Median Urea N	Median RDW	Median RDW	Maximum Urea N
2	Median Urea N	Median Urea N	Median RDW	Median Urea N	Maximum Urea N	Median Urea N
3	Maximum RDW	Maximum RDW	Maximum RDW	Maximum Urea N	Median Urea N	Median RDW
4	Minimum RDW	Maximum Urea N	Maximum Urea N	Maximum RDW	Maximum RDW	Maximum RDW
5	Maximum Urea N	Minimum RDW	Minimum Urea N	Minimum Urea N	Median Lymphs	Median PO2
6	Minimum Urea N	Minimum Urea N	Minimum RDW	Minimum RDW	Maximum NEUTS	Maximum Creat
7	Median Creat	Median Creat	Maximum Creat	Maximum Creat	Median PO2	Med: Levofloxacin POIV
8	Maximum Creat	Maximum Creat	Median Lymphs	Maximum NEUTS	Med: Levofloxacin POIV	Minimum RDW
9	Maximum NEUTS	Med: Levofloxacin POIV	Med: Levofloxacin POIV	Med: Levofloxacin POIV	Maximum Creat	Med: Metronidazole POIV
10	Median HGB	Minimum Chloride	Maximum NEUTS	Minimum Lymphs	Med: Metronidazole POIV	Med: Oxycodone PO
11	Med: Vancomycin IV	Minimum MCHC	Median Creat	Med: Metronidazole POIV	Minimum Lymphs	Minimum MCHC
12	Med: Levofloxacin POIV	Minimum Creat	Minimum MCHC	Median PO2	Median NEUTS	Med: Vancomycin IV

Table 8. Definites- Biggest decrease

1. Medicine: Integrelin IV
2. Procedure: Inj or Inf Platelet
3. Procedure: Inhibit
4. Procedure: PTCA without Thrombolytic AG
5. Procedure: RT Heart Cardiac Catheter

Table 9. Definites- Biggest squared change:

1. Medicine: Integrelin IV
2. Minimum Lactate
3. Procedure: Infusion of Vasopressor
4. Procedure: Inj or Inf of Platelet Inhibit
5. Procedure: Parenteral Infus Conc Nu

Table 10. Maybes- Top Predictors (Median):

1. Median RDW
2. Median Urea Nitrogen
3. Maximum Urea Nitrogen
4. Maximum RDW
5. Minimum Urea Nitrogen
6. Minimum RDW
7. Maximum Creatinine
8. Median Creatinine
9. Minimum Creatinine
10. Minimum MCHC tied with Medicine: Levofloxacin POIV

Table 11. Maybes- Biggest increase

1. Medicine: Milk of Magnesia PO
2. Medicine: Oxycodone
3. Medicine: Vancomycin
4. Medicine: Furosemide POIV
5. Medicine: Metronidazole POIV

Table 12. Maybes- Biggest decrease

1. Age
2. Excision or destruction of Esophageal Lesion
3. Median RBC
4. Minimum Creatinine
5. Maximum MCHC

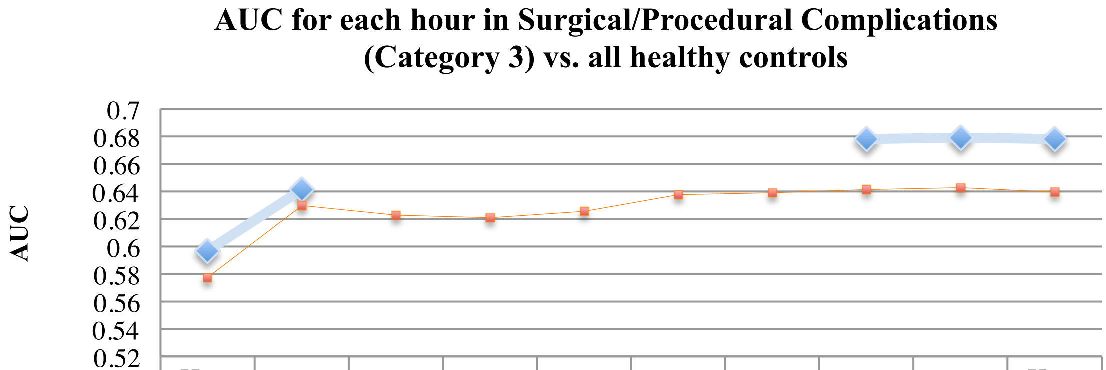
Table 13. Maybes- Biggest squared change

1. Age
2. Medicine: Oxycodone
3. Medicine: Vancomycin
4. Medicine: Furosemide
5. Medicine: Metronidazole POIV

4.1.3 Iteration 3: Categories

Below were my observations for Surgical/Procedural Complications (Category 3) vs. all healthy controls.

Figure 13. AUC for each hour in Category 3 with all healthy controls



	Hour 1	2	3	6	12	18	24	48	72	Hour 96
Training	0.5965	0.6412						0.678	0.6787	0.6781
Validation	0.5772	0.6294	0.6226	0.6208	0.6255	0.6373	0.6391	0.6411	0.6424	0.6394

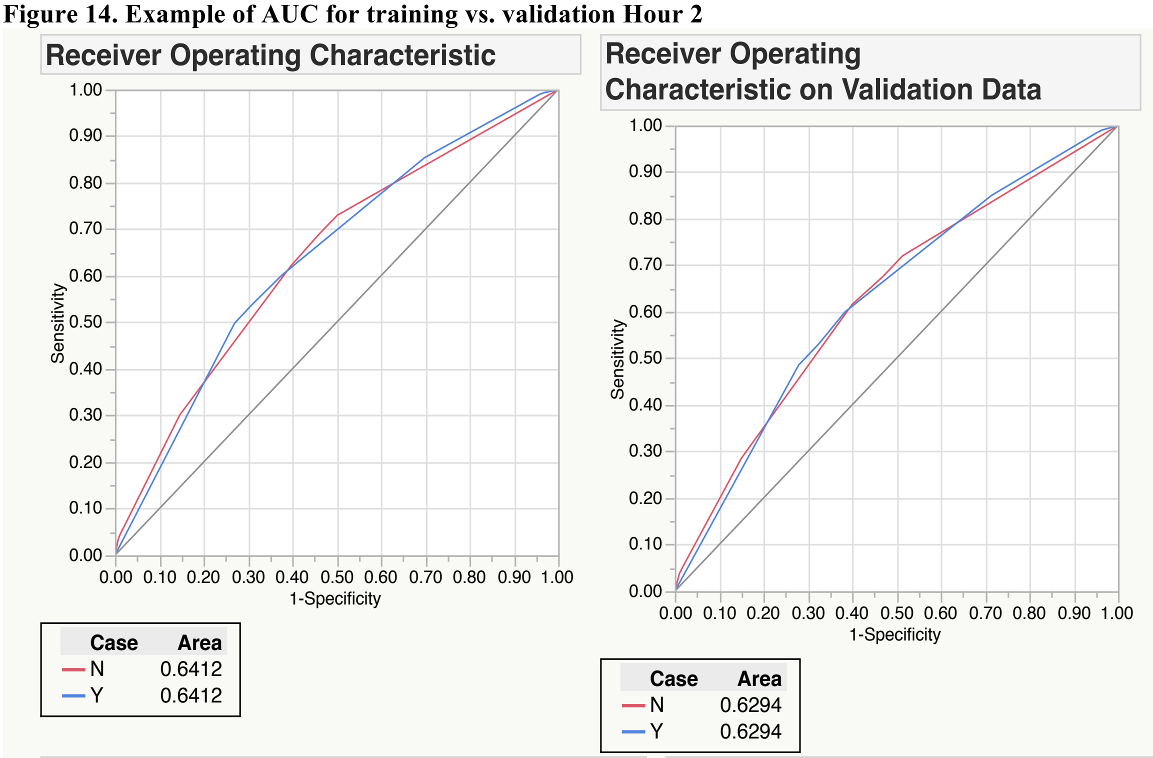
Table 14. Contributors for each hour

Rank	Hour 1	2	3	6	12	18
1	White Blood Cell Median (50468.median)	Lymphs Minimum (50408.min)	Neuts Median (50419.median)	Neuts Minimum (50419.min)	Lymphs Minimum (50408.min)	Lymphs Minimum (50408.min)
2	Potassium Median (50149.median)	HGB minimum (50386.min)	50386.median	50386.median	famotidine.POIV.rx	50386.max
3	50177.median	50149.max	50159.min	famotidine.POIV.rx	50386.max	50090.median
4		50412.max	vancomycin.IV.rx		50440.min	50653.min
5		50399.min	50444.max		age	famotidine.POIV.rx
6		famotidine.POIV.rx	fentanyl.IV.rx		metoprolol.POIV.rx	50399.median
7			50399.min		diphenhydramine.POIV.rx	age
8			50177.min		prednisone.PO.rx	vancomycin.IV.rx
9			famotidine.POIV.rx		phenytoin.POIV.rx	50444.max
10						ceftriaxone.IVIM.rx
11						CORONAR.ARTERIOGR.2.CATH..cpt
12						morphine.MULTI.rx
13						line.flush.IV.rx

Rank	Hour 24	48	72	96
1	Lymphs Minimum (50408.min)	Lymphs Median (50408.median)	Lymphs Maximum (50408.max)	Lymphs Maximum (50408.max)
2	famotidine.POIV.rx	NONOP.REMOVAL.HEART.ASSI..cpt	NONOP.REMOVAL.HEART.ASSI..cpt	amiodarone.POIV.rx
3	50386.median	50386.max	50444.median	NONOP.REMOVAL.HEART.ASSI..cpt
4	50386.max	50440.median	50149.median	50444.max

5	50444.max	mycophenolate.mofetil.POIV.rx	50383.min	mycophenolate.mofetil.POIV.rx
6	ceftriaxone.IVIM.rx	50383.min	50083.median	PARENTERAL.INFUS.CONC.NU..cpt
7	metoprolol.POIV.rx	50090.min	famotidine.POIV.rx	epinephrine.IV.rx
8		50090.max	metoclopramide.POIV.rx	50008.nn
9		50633.nn	PARENTERAL.INFUS.CONC.NU..cpt	senna.PO.rx
10		morphine.MULTI.rx	50468.min	acetaminophen.POPR.rx
11		X50068.max	X50170.median	SPINAL.TAP.cpt
12		ceftriaxone.IVIM.rx	amiodarone.POIV.rx	azithromycin.POIV.rx
13		PARENTERAL.INFUS.CONC.NU..cpt	CONTRAST.ARTERIOGRAMi.LEG..cpt	levophed.IV.rx
14		EXTRACORPOREAL.CIRCULAT.cpt		

Figure 14. Example of AUC for training vs. validation Hour 2



4.1.4 Iteration 4: Predicting for the 37 different ICD-9-CM codes

Below were my observations.

Table 12. Validation AUC for the 37 out of 43 different ICD-9-CM codes.
AUC 0.45911- 0.9066

Description	Hr 1	2	3	6	12	18	24	48	72	96

METABOLIC ENCEPHALOPATHY	0.7175	0.6931	0.6167	0.7319	0.7148	0.7206	0.6593	0.7865	0.7341	0.7064
TOXIC ENCEPHALOPATHY	0.7444	0.7559	0.7634	0.7326	0.7111	0.7592	0.7802	0.7407	0.7386	0.7495
IATROGENIC PULMONARY EMBOLISM AND INFARCTION	0.7326	0.7332	0.6819	0.7339	0.7676	0.7107	0.7797	0.7057	0.7753	0.7635
IATROGENIC PNEUMOTHORAX	0.7997	0.7926	0.7948	0.7806	0.773	0.798	0.7628	0.7873	0.7882	0.7764
INFECTION OF GASTROSTOMY	0.6811	0.7556	0.7575	0.6441	0.6808	0.8063	0.7889	0.8254	0.8011	0.8102
MECHANICAL COMPLICATION OF GASTROSTOMY	0.6606	0.6755	0.7106	0.7241	0.7396	0.7957	0.7141	0.7875	0.7147	0.6704
OTHER GASTROSTOMY COMPLICATIONS	0.6657	0.5503	0.6653	0.689	0.7287	0.6126	0.6848	0.7199	0.669	0.7416
OTHER COLOSTOMY AND ENTEROSTOMY COMPLICATION	0.6945	0.6914	0.6559	0.7233	0.702	0.6843	0.743	0.7437	0.7123	0.7635
INTEST POSTOP NONABSORB	0.7227	0.7595	0.7187	0.7194	0.7301	0.7043	0.6853	0.7936	0.7209	0.7326
DECUBITUS ULCER, LOWER B	0.7847	0.7591	0.7338	0.7621	0.7807	0.7845	0.7624	0.7875	0.7581	0.7556
DECUBITUS ULCER, BUTTOCK	0.7655	0.7391	0.7315	0.7557	0.8107	0.772	0.7585	0.8096	0.7594	0.7193
MALFUNC VASC DEVICE/GRAF	0.6397	0.6565	0.62	0.6446	0.6254	0.6872	0.6575	0.6983	0.6484	0.6379
MECHANICAL COMPLICATION OF PROSTHETIC GRAFT OF OTH	0.5961	0.7628	0.6142	0.7634	0.7307	0.775	0.8134	0.6566	0.7003	0.9066
MALFUNC OTHER DEVICE/GRA	0.7374	0.7292	0.7242	0.7129	0.7063	0.7863	0.6893	0.7266	0.7361	0.7353
INFECT DUE TO CARDIAC IM	0.7649	0.7341	0.6972	0.7429	0.7552	0.817	0.7696	0.7702	0.7567	0.7578
INFECT DUE TO VASCULAR G	0.7686	0.7618	0.7722	0.7731	0.7663	0.7652	0.7585	0.7763	0.7689	0.7513
INFECT DUE TO GU IMPLANT	0.6601	0.7961	0.6189	0.7676	0.7609	0.7281	0.7523	0.645	0.7436	0.711
COMPL FROM OTH VASCULAR	0.6617	0.6708	0.6801	0.6396	0.6847	0.6545	0.6735	0.6334	0.6634	0.6682
COMPLICATIONS OF TRANSPLANTED KIDNEY	0.772	0.7872	0.7419	0.7568	0.7798	0.806	0.7795	0.842	0.7987	0.8764
COMPLICATIONS OF TRANSPLANTED BONE MARROW	0.7941	0.7825	0.786	0.8065	0.8074	0.7926	0.8436	0.8017	0.84	0.8482
IATROGENIC CEREBRO INFAR	0.751	0.7802	0.7647	0.744	0.7543	0.7804	0.7659	0.7623	0.7561	0.767
AMPUTAT COMPLIC NEC	0.5921	0.6081	0.6218	0.6253	0.6446	0.4591	0.6521	0.6823	0.649	0.6662
HEMORR COMPLIC PROCEDURE	0.7966	0.7898	0.7914	0.7876	0.7891	0.79	0.7768	0.7955	0.7986	0.7705

ACCIDENTAL OP LACERATION	0.7261	0.7244	0.694	0.7201	0.7159	0.7218	0.7202	0.724	0.709	0.6988
DISRUPT INTER OPER WOUND	0.6072	0.5992	0.6002	0.5646	0.601	0.6282	0.5949	0.6372	0.6079	0.6189
DISRUPT EXTER OPER WOUNF	0.6044	0.6211	0.6279	0.6327	0.5917	0.6357	0.6201	0.596	0.5718	0.6126
OTHER POSTOPERATIVE INFECTION	0.7269	0.7201	0.7274	0.744	0.7294	0.7223	0.7351	0.7309	0.7313	0.7161
NON-HEALING SURGICAL WOUND	0.6006	0.6209	0.4986	0.6607	0.5598	0.5659	0.6089	0.5296	0.6387	0.6458
COMPLIC MED CARE NEC/NOS	0.7272	0.7144	0.711	0.7206	0.7632	0.7312	0.7192	0.7324	0.7666	0.7704
ABN REACT-PROCEDURE NEC	0.5917	0.5815	0.5941	0.6161	0.6103	0.5773	0.6266	0.5956	0.6161	0.6035
ADV EFF ANTIBIOTICS NEC	0.8161	0.8555	0.8603	0.7932	0.825	0.8273	0.8036	0.8322	0.8412	0.7933
ADV EFF ANTINEOPLASTIC	0.8368	0.8364	0.8179	0.8491	0.8143	0.8056	0.8248	0.8445	0.8425	0.7923
ADV EFF ANTICOAGULANTS	0.8354	0.8014	0.8168	0.8192	0.788	0.8189	0.8202	0.8139	0.8237	0.7941
ATTENTION TO TRACHEOSTOMY	0.689	0.7266	0.684	0.723	0.7351	0.7527	0.7268	0.7249	0.7066	0.7072
ATTENTION TO ILEOSTOMY	0.5943	0.5756	0.5157	0.5153	0.6216	0.5922	0.7151	0.5494	0.553	0.5893
ATTENTION TO COLOSTOMY	0.6416	0.5516	0.707	0.6578	0.5898	0.5748	0.6337	0.6481	0.6583	0.6107

Table 13. Contributors for each hour

Rank	**Hour 1**	**2**	**3**	**6**	**12**	**18**	**24**	**48**	**72**	**96**
1	Category	Category	Category	Category	Category	Category	Category	Category	Category	Category
2	RDW imum (50444.min	RDW imum 44.min)	Creatinine imum (50090.min)	Creatinine imum 90.max)	Creatinine lian 90.median	Creatinine imum (50090.min	Creatinine imum (50090.min)	Creatinine imum (50090.min)	Creatinine imum 90.max)	Creatinine imum 90.max)
3	50090.min	50090.mi		50444.min	50444.max	50444.min	50444.min	50444.min	50444.min	50444.min
4					50333.median	50444.median	50661.max	50333.median	Sulfameth-trimethoprim.POIV.rx	50019.max
5										Age

6										50090.max
7										50653.median

4.1.5 Iteration 5: Validation AUC for 43 different ICD-9-CM codes.

Below were my observations.

Table 14. Validation AUC for the 43 different ICD-9-CM codes.
AUC 0.422-0.869

description	Hour 1	2	3	6	12	18	24	48	72	Hour 96
ACUTE DELIRIUM	0.6902	0.6947	0.6481	0.6805	0.6917	0.6428	0.6423	0.6303	0.6653	0.6564
METABOLIC ENCEPHALOPATHY	0.7281	0.7084	0.7424	0.7006	0.7269	0.7015	0.7858	0.7869	0.7437	0.7647
TOXIC ENCEPHALOPATHY	0.7097	0.7303	0.7207	0.7529	0.7507	0.7699	0.8003	0.7753	0.747	0.768
IATROGENIC PULMONARY EMBOLISM AND INFARCTION	0.7682	0.7663	0.7818	0.7345	0.7725	0.7497	0.7222	0.7291	0.7715	0.7685
IATROGENIC PNEUMOTHORAX	0.8052	0.786	0.7935	0.7812	0.7571	0.7954	0.7992	0.7814	0.7827	0.7855
STOMATITIS	0.6146	0.5382	0.6115	0.6161	0.6652	0.6356	0.422	0.5512	0.6699	0.6012
INFECTION OF GASTROSTOMY	0.8124	0.644	0.7996	0.7756	0.8667	0.7687	0.8276	0.8142	0.8632	0.8458
MECHANICAL COMPLICATION OF GASTROSTOMY	0.6982	0.5988	0.7057	0.7233	0.8209	0.7822	0.7218	0.7024	0.7786	0.6935
OTHER GASTROSTOMY COMPLICATIONS	0.7329	0.591	0.6127	0.7558	0.6988	0.6097	0.6793	0.6355	0.6835	0.6686
PERITON ADHES (POSTOP)	0.6856	0.6573	0.6495	0.6317	0.6516	0.5965	0.6525	0.6475	0.6637	0.6395
OTHER COLOSTOMY AND ENTEROSTOMY COMPLICATION	0.699	0.7527	0.714	0.6818	0.7263	0.7708	0.7849	0.676	0.7521	0.6996
INTEST POSTOP NONABSORB	0.6942	0.7668	0.7387	0.677	0.6837	0.736	0.7006	0.7144	0.6975	0.7133
DERMATITIS DUE	0.7181	0.711	0.7185	0.708	0.6466	0.7102	0.6904	0.7321	0.7127	0.6999

TO DRUGS AND MEDICINES TAKEN INTERN										
DECUBITUS ULCER	0.7735	0.7832	0.7552	0.7681	0.8402	0.7368	0.8049	0.8106	0.7532	0.7508
DECUBITUS ULCER, LOWER B	0.7679	0.7839	0.7754	0.7597	0.7832	0.7837	0.7714	0.772	0.7268	0.775
DECUBITUS ULCER, BUTTOCK	0.7501	0.7516	0.7629	0.7972	0.7685	0.7696	0.7719	0.7172	0.754	0.7224
MALFUNC VASC DEVICE/GRAF	0.6021	0.6218	0.6827	0.6619	0.6752	0.7184	0.7135	0.6685	0.6283	0.6669
MECHANICAL COMPLICATION OF PROSTHETIC GRAFT OF OTH	0.7036	0.8213	0.6284	0.5811	0.4831	0.6174	0.6677	0.6176	0.5925	0.7386
MALFUNC OTHER DEVICE/GRA	0.7267	0.7304	0.7439	0.7574	0.7162	0.7543	0.76	0.7502	0.7227	0.7583
INFECT DUE TO CARDIAC IM	0.7327	0.7758	0.7439	0.7591	0.704	0.8047	0.7357	0.6412	0.7361	0.7438
INFECT DUE TO VASCULAR G	0.7704	0.772	0.7869	0.7663	0.7695	0.7706	0.7842	0.7835	0.7789	0.7842
INFECT DUE TO GU IMPLANT	0.6879	0.7838	0.6831	0.7749	0.6506	0.7453	0.5861	0.8252	0.6467	0.6598
COMPL FROM OTH VASCULAR	0.6592	0.6605	0.675	0.6425	0.6616	0.6594	0.672	0.6572	0.6669	0.6302
COMPLICATIONS OF TRANSPLANTED KIDNEY	0.7435	0.8241	0.7774	0.7316	0.7859	0.7746	0.8296	0.7705	0.8135	0.7854
COMPLICATIONS OF TRANSPLANTED LIVER	0.6962	0.8045	0.7443	0.6832	0.7283	0.8194	0.6728	0.655	0.6704	0.7372
COMPLICATIONS OF TRANSPLANTED BONE MARROW	0.799	0.828	0.8258	0.7648	0.8282	0.8079	0.8071	0.802	0.7693	0.8483
IATROGENIC CEREBRO INFAR	0.7443	0.7722	0.7736	0.7704	0.7546	0.7728	0.7565	0.7367	0.7552	0.7923
AMPUTAT COMPLIC NEC	0.5901	0.708	0.6296	0.6044	0.6526	0.5481	0.7155	0.679	0.6509	0.603
HEMORR COMPLIC PROCEDURE	0.8045	0.8014	0.7967	0.7968	0.7966	0.7863	0.7971	0.7961	0.7807	0.7879
ACCIDENTAL OP LACERATION	0.7022	0.7211	0.7302	0.7029	0.7164	0.7259	0.722	0.707	0.7038	0.7145
DISRUPT INTER OPER WOUND	0.6092	0.6105	0.6445	0.6009	0.6569	0.591	0.6529	0.635	0.6413	0.6291
DISRUPT EXTER OPER WOUNF	0.6355	0.6429	0.6347	0.6053	0.6229	0.6913	0.6279	0.6145	0.6157	0.632
OTHER POSTOPERATIVE INFECTION	0.7395	0.7686	0.7268	0.7435	0.7337	0.7375	0.7288	0.7309	0.7199	0.7093
PERSISTENT POSTOPERATIVE FISTULA NOT ELSEWHERE CLA	0.6215	0.6403	0.6473	0.6578	0.6422	0.6047	0.6163	0.6431	0.5787	0.721
NON-HEALING SURGICAL WOUND	0.5016	0.6769	0.5772	0.6075	0.5674	0.5982	0.6193	0.7046	0.5657	0.5824
COMPLIC MED CARE NEC/NOS	0.747	0.689	0.7196	0.7567	0.752	0.7234	0.7536	0.7029	0.6989	0.6774

ABN REACT-PROCEDURE NEC	0.6048	0.6322	0.6065	0.6207	0.6288	0.6072	0.5954	0.6288	0.6337	0.6382
ADV EFF ANTIBIOTICS NEC	0.8419	0.7708	0.8079	0.7967	0.7663	0.8166	0.8155	0.773	0.8521	0.8022
ADV EFF ANTINEOPLASTIC	0.8693	0.8299	0.8474	0.8393	0.8041	0.838	0.8196	0.8618	0.8283	0.8246
ADV EFF ANTICOAGULANTS	0.8146	0.7952	0.7915	0.8204	0.7951	0.8427	0.8087	0.784	0.8195	0.8114
ATTENTION TO TRACHEOSTOMY	0.725	0.7059	0.7092	0.7028	0.7541	0.6822	0.7131	0.7163	0.6158	0.7259
ATTENTION TO ILEOSTOMY	0.5438	0.6168	0.542	0.5512	0.6729	0.5049	0.5152	0.6007	0.7207	0.4356
ATTENTION TO COLOSTOMY	0.6021	0.5369	0.6876	0.6312	0.6762	0.6626	0.4726	0.6852	0.6738	0.6377

Table 15. Statistical analysis of AUC

description	Mean AUC	Median	Min	Max	Range	Hour of Min	Hour of Max Prediction	Range of Hours
ACUTE DELIRIUM	0.66423	0.66085	0.6303	0.6947	0.0644	48	2	-46
METABOLIC ENCEPHALOPATHY	0.7389	0.73525	0.7006	0.7869	0.0863	6	48	42
TOXIC ENCEPHALOPATHY	0.75248	0.7518	0.7097	0.8003	0.0906	1	24	23
IATROGENIC PULMONARY EMBOLISM AND INFARCTION	0.75643	0.76725	0.7222	0.7818	0.0596	24	3	-21
IATROGENIC PNEUMOTHORAX	0.78672	0.78575	0.7571	0.8052	0.0481	12	1	-11
STOMATITIS	0.59255	0.61305	0.422	0.6699	0.2479	24	72	48
INFECTION OF GASTROSTOMY	0.80178	0.8133	0.644	0.8667	0.2227	2	12	10
MECHANICAL COMPLICATION OF GASTROSTOMY	0.72254	0.71375	0.5988	0.8209	0.2221	2	12	10
OTHER GASTROSTOMY COMPLICATIONS	0.66678	0.67395	0.591	0.7558	0.1648	2	6	4
PERITON ADHES (POSTOP)	0.64754	0.65055	0.5965	0.6856	0.0891	18	1	-17
OTHER COLOSTOMY AND ENTEROSTOMY COMPLICATION	0.72572	0.72015	0.676	0.7849	0.1089	48	24	-24
INTEST POSTOP NONABSORB	0.71222	0.70695	0.677	0.7668	0.0898	6	2	-4

DERMATITIS DUE TO DRUGS AND MEDICINES TAKEN INTERN	0.70475	0.7106	0.6466	0.7321	0.0855	12	48	36
DECUBITUS ULCER	0.77765	0.7708	0.7368	0.8402	0.1034	18	12	-6
DECUBITUS ULCER, LOWER B	0.7699	0.7735	0.7268	0.7839	0.0571	72	2	-70
DECUBITUS ULCER, BUTTOCK	0.75654	0.75845	0.7172	0.7972	0.08	48	6	-42
MALFUNC VASC DEVICE/GRAF	0.66393	0.6677	0.6021	0.7184	0.1163	1	18	17
MECHANICAL COMPLICATION OF PROSTHETIC GRAFT OF OTH	0.64513	0.623	0.4831	0.8213	0.3382	12	2	-10
MALFUNC OTHER DEVICE/GRA	0.74201	0.74705	0.7162	0.76	0.0438	12	24	12
INFECT DUE TO CARDIAC IM	0.7377	0.73995	0.6412	0.8047	0.1635	48	18	-30
INFECT DUE TO VASCULAR G	0.77665	0.77545	0.7663	0.7869	0.0206	6	3	-3
INFECT DUE TO GU IMPLANT	0.70434	0.6855	0.5861	0.8252	0.2391	24	48	24
COMPL FROM OTH VASCULAR	0.65845	0.65995	0.6302	0.675	0.0448	96	3	-93
COMPLICATION S OF TRANSPLANTE D KIDNEY	0.78361	0.7814	0.7316	0.8296	0.098	6	24	18
COMPLICATION S OF TRANSPLANTE D LIVER	0.72113	0.71225	0.655	0.8194	0.1644	48	18	-30
COMPLICATION S OF TRANSPLANTE D BONE MARROW	0.80804	0.8075	0.7648	0.8483	0.0835	6	96	90
IATROGENIC CEREBRO INFAR	0.76286	0.76345	0.7367	0.7923	0.0556	48	96	48
AMPUTAT COMPLIC NEC	0.63812	0.64025	0.5481	0.7155	0.1674	24	24	0
HEMORR COMPLIC PROCEDURE	0.79441	0.79665	0.7807	0.8045	0.0238	72	1	-71
ACCIDENTAL OP LACERATION	0.7146	0.71545	0.7022	0.7302	0.028	1	3	2
DISRUPT INTER OPER WOUND	0.62713	0.63205	0.591	0.6569	0.0659	18	12	-6
DISRUPT EXTER OPER WOUNF	0.63227	0.62995	0.6053	0.6913	0.086	6	18	12
OTHER POSTOPERATIV E INFECTION	0.73385	0.7323	0.7093	0.7686	0.0593	96	2	-94

PERSISTENT POSTOPERATIVE FISTULA NOT ELSEWHERE CLA	0.63729	0.64125	0.5787	0.721	0.1423	72	96	24
NON-HEALING SURGICAL WOUND	0.60008	0.5903	0.5016	0.7046	0.203	1	48	47
COMPLIC MED CARE NEC/NOS	0.72205	0.7215	0.6774	0.7567	0.0793	96	6	-90
ABN REACT-PROCEDURE NEC	0.61963	0.62475	0.5954	0.6382	0.0428	96	96	0
ADV EFF ANTIBIOTICS NEC	0.8043	0.80505	0.7663	0.8521	0.0858	12	72	60
ADV EFF ANTINEOPLASTIC	0.83623	0.83395	0.8041	0.8693	0.0652	12	1	-11
ADV EFF ANTICOAGULANTS	0.80831	0.81005	0.784	0.8427	0.0587	48	18	-30
ATTENTION TO TRACHEOSTOMY	0.70503	0.71115	0.6158	0.7541	0.1383	12	12	0
ATTENTION TO ILEOSTOMY	0.57038	0.5475	0.4356	0.7207	0.2851	72	72	0
ATTENTION TO COLOSTOMY	0.62659	0.65015	0.4726	0.6876	0.215	24	3	-21

Table 16. Top 5 AUCs

Description	ICD 9 code	AUC	Hour
1. ADV EFF ANTINEOPLASTIC	E933.1	0.8693	1
2. INFECTION OF GASTROSTOMY	536.41	0.8667	12
3. ADV EFF ANTIBIOTICS NEC	E930.8	0.8521	72
4. COMPLICATIONS OF TRANSPLANTED BONE MARROW	996.85	0.8483	96
5. ADV EFF ANTICOAGULANTS	E934.2	0.8427	18

Table 17. Lowest AUCs

Description	ICD 9 code	AUC	Hour
1. STOMATITIS	528	0.422	24
2. ATTENTION TO ILEOSTOMY	V55.2	0.4356	72
3. ATTENTION TO COLOSTOMY	V55.3	0.4726	24
4. MECHANICAL COMPLICATION OF	996.52	0.4831	12

PROSTHETIC GRAFT OF OTH			
5. NON-HEALING SURGICAL WOUND	998.83	0.5016	1

Table 18. Contributors for each hour

Hour	Rank 1	Rank 2	Rank 3			
1	Category	50090.max	50444.max			
2	Category	50444.min	50090.max	age	hadm˙id	
3	Category	50090.max	50444.min	50444.max	age	hadm˙id
6	Category	50090.max				
12	Category	50444.min	50090.median			
18	Category	50090.min	50444.median	50444.min		
24	Category	50090.min	50444.min			
48	Category	50090.min				
72	Category	50090.min	mycophenolate mofetil.POIV.rx			
96	Category	50090.min	50061.max	50444.min	mycophenolate mofetil.POIV.rx	

4.1.6 Iteration 6: Predicting for categories with ICD-9-CM codes

Below were my observations.

Figure 15. Example of Validation ROC Curve for Category 3, Hour 3.
AUC 0.8252- 0.9576

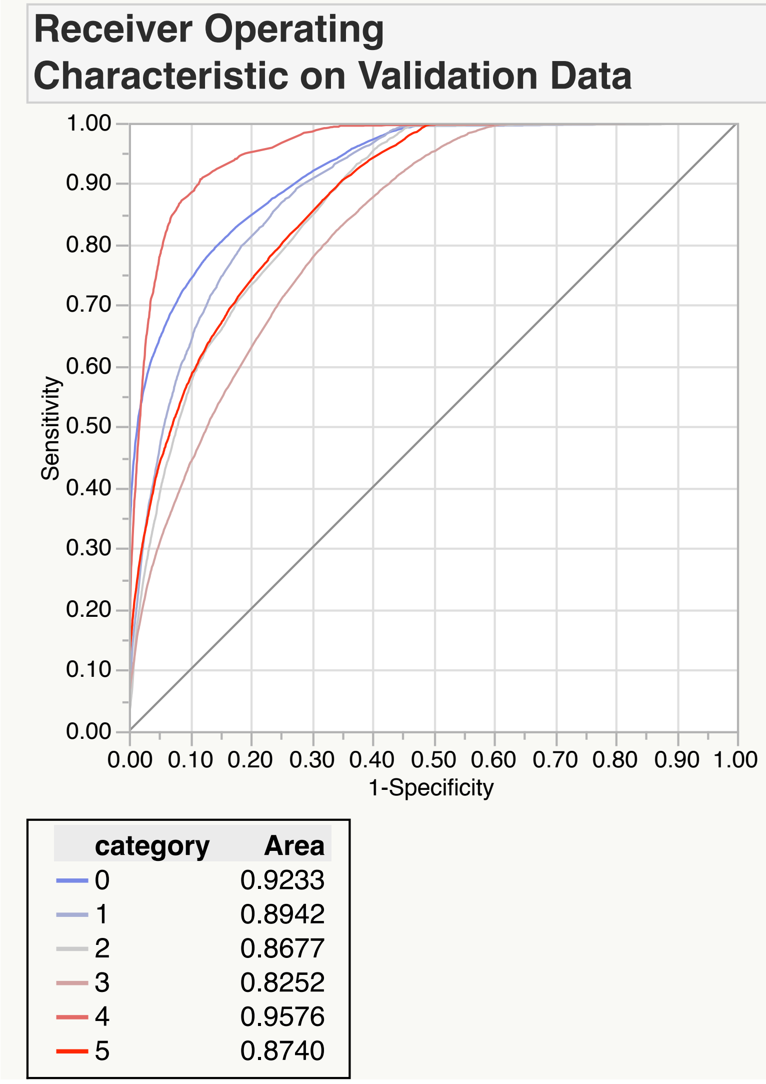


Table 19. Example of Confusion Matrix for Validation Data of Category 3, Hour 3.

Validation	Predicted					
Actual	0	1	2	3	4	5
0	22331	372	370	3917	227	869
1	1238	2729	396	2378	141	584
2	1379	374	2745	3328	109	579
3	2987	1162	1392	13121	282	1481
4	494	156	161	652	1246	279
5	1852	627	667	3815	229	5543

Table 20. Contributors for predicting categories with knowledge of ICD-9-CM codes
For surgical/procedural complications (category 3) hour 3.

Features

1. icd9code
2. age
3. PLT count minimum (50428.min)
4. WBC minimum (50468.min)
5. Glucose minimum (50112.min)
6. PTT minimum (50440.min)
7. RDW minimum (50444.min)
8. Urea N (50177.min)
9. PT minimum (50439.min)
10. Creat minimum (50090.min)
11. 50412.min
12. 50083.min
13. 50383.min
14. 50442.min
15. 50411.min
16. 50149.min
17. 50172.min
18. elix.index
19. 50068.min
20. 50413.min
21. 50386.min
22. 50159.min
23. 50428.median
24. 50399.min
25. ethnicity
26. 50468.median
27. 50383.median
28. 50412.median
29. 50177.median
30. 50444.median
31. 50439.median
32. 50090.median
33. gender
34. 50172.median
35. 50112.median
36. 50383.max
37. 50444.max
38. 50149.median
39. 50440.median

Features (continued)

43. 50149.max
44. 50083.median
45. 50428.max
46. 50442.max
47. 50159.max
48. 50159.median
49. 50068.median
50. 50413.median
51. 50468.max
52. 50386.median
53. 50083.max
54. 50440.max
55. 50386.max
56. 50411.max
57. 50399.max
58. heparin.IV.rx
59. 50112.max
60. 50439.max
61. acetaminophen.POPR.rx
62. line flush.IV.rx
63. nitroprusside.IV.rx
64. atorvastatin.PO.rx
65. pantoprazole.POIV.rx
66. furosemide.POIV.rx
67. metronidazole.POIV.rx
68. 50068.max
69. 50413.max
70. morphine.MULTI.rx
71. 50090.max
72. eptifibatide.IV.rx
73. bisacodyl.POPR.rx
74. nitroglycerine.IV.rx
75. 50177.max
76. tincture of opium.IH.rx
77. RT/LEFT HEART CARD CATH.cpt
78. propofol.IV.rx
79. lorazepam.MULTI.rx
80. lisinopril.PO.rx
81. metoprolol.POIV.rx

40. 50442.median
41. 50411.median
42. 50399.median
82. aspirin.POPR.rx
83. integrelin.IV.rx

4.1.7 Iteration 7: Feature selection/principal components analysis for predicting categories without ICD-9-codes and without identifiers.

Below were my observations.

Table 21. Example of Principal Components Analysis with Bayesian Logistic Regression

Scheme:weka.classifiers.bayes.BayesianLogisticRegression -D -Tl 5.0E-4 -S 0.5 -H 1 -V 0.27 -R R:0.01-316,3.16 -P 1 -F 2 -seed 1 -I 100 –N

Class	True Pos	False Pos	Precision	F-measure	ROC Area
Y	0.496	0.341	0.593	0.54	0.578
N	0.659	0.504	0.567	0.61	0.578
Weighted Average	0.578	0.422	0.58	0.575	0.578

Table 22. Confusion Matrix for Bayesian Logistic

Actual	Predicted Y	Predicted N
Y	2184	2215
N	1499	2900

Table 23. Example of Principal Components Analysis with Naïve Bayes

Class	True Pos	False Pos	Precision	F-measure	ROC Area
Y	0.139	0.101	0.578	0.224	0.557
N	0.899	0.861	0.511	0.651	0.557
Weighted Average	0.519	0.481	0.544	0.437	0.557

Table 24. Confusion Matrix for Naïve Bayes

Actual	Predicted Y	Predicted N
Y	610	3789
N	445	3954

Table 25. Example of Principal Components Analysis with Logistic Regression

Class	True Pos	False Pos	Precision	F-measure	ROC Area
Y	0.639	0.398	0.616	0.628	0.656
N	0.602	0.361	0.625	0.613	0.656
Weighted Average	0.621	0.379	0.621	0.62	0.656

Table 26. Confusion Matrix for Logistic Regression

Actual	Predicted Y	Predicted N
Y	2840	1602
N	1769	2673

Table 27. Contributors

1. RDW Median (50444.median)
2. RBC Minimum (50442.min)
3. RBC Maximum (50442.max)

4.1.8 Iteration 8: Dual-layer prediction

Table 28. Top three features with the highest mean rank over the time intervals for each category. Features were only included if they existed in the data for at least seven of the ten time intervals. Only the first statistic for each lab value is included.

Attribute Name	Hour 1	2	3	6	12	18	24	48	72	96
Infectious complications (Category 1)										
Red Blood Cell Distribution Width (RDW) Median Lab Value	14	7	17	2	3	1	3	3	10	38
Parenteral Infusion Procedure		3	6	38	11	10	25	6	5	9
Vancomycin Medication	37	12	1	3	1	2	1	1	2	108
Hemostatic or hemolytic complications (Category 2)										
Serum Transfusion Procedure	16	3	20	5	27	17	16	14	11	4
Partial Thromboplastin Time (PTT) Maximum Lab Value	49	14	41	2	5	22	10	2	8	7
Hematocrit (HCT) Minimum Lab Value	35	55	34	12	10	4	7	1	4	3
Surgical/procedural complications (Category 3)										
Neostigmine Medication			42	2	1	1	1	1	4	6
Glycopyrrolate Medication			38	1	2	2	2	4	7	9
Ceftriaxone Medication				31	10	10	3	18	2	3
Medical complications (Category 4)										
Ethnicity	4	1	1	4	5	2	1	14	14	104
Red Blood Cell Distribution Width (RDW) Maximum Lab Value	3	4	3	1	3	1	2	1	9	125
Atropine Sulfate Medication	18	15	54	20	51	21	29	7	3	13
Other complications (Category 5)										
Chlorhexidine Medication				1	2	2	1	1	1	1
Skin Suture Procedure	10	7	1	3	15	60	50			
Infusion of Vasopressor Procedure		5	14	33	1	26	16	7	76	10

Table 29. Top two features with the largest increase over the time intervals as measured by the linear regression slope. Features were only included if they existed in the data for at least seven of the 10 time intervals and their max ranking was within the top 10 at some interval. Only the top statistic for each lab value is included.

Attribute Name	Hour 1	2	3	6	12	18	24	48	72	96	Slope
Infectious complications (Category 1)											
Single Internal Mammary-Coronary Artery Bypass Procedure		43	44	72	10	13	10	11	6	3	-0.457
Aortocoronary Bypass Procedure		44	45	73	13	17	9	19	12	11	-0.379
Hemostatic or hemolytic complications (Category 2)											
Corpuscular Hemoglobin Concentration Maximum Lab Value	87	165	135	25	49	131	109	5	18	18	-1.130
Platelet Transfusion Procedure		1	19	150	35	9	9	17	10	6	-0.457
Surgical/procedural complications (Category 3)											
Neosynephrine Medication	99	100	121	25	12	16	16	8	13	8	-0.820
Nitroglycerin Medication	89	145	99	12	5	8	5	9	11	16	-0.754
Medical complications (Category 4)											
Nitroprusside Medication	90	143	122	14	9	8	13	23	12	5	-0.868
IV Line Flush	97	129	128	43	55	33	32	71	5	29	-0.809
Other complications (Category 5)											
Lansoprazole Medication				90	97	11	95	3	20	3	-0.945
Insertion of Intercostal Catheter Procedure	42	61	158	21	9	5	21	5	34	2	-0.584

Table 30. Top two features with the largest decrease over the time intervals as measured by the linear regression slope. Features were only included if they existed in the data for at least seven of the 10 time intervals and their max ranking was within the top 10 at some interval. Only the top statistic for each lab value is included.

Attribute Name	Hour 1	2	3	6	12	18	24	48	72	96	Slope
Infectious complications (Category 1)											
Red Blood Cell Count (RBC) Maximum Lab Value	39	14	12	9	20	24	62	168	148	189	2.010
Hematocrit (HCT) Maximum Lab Value	29	11	7	14	26	21	45	139	147	170	1.862
Hemostatic or hemolytic complications (Category 2)											
Basophils Maximum Lab Value		4	5	34	206	227	176	256	286	370	3.390
Hemoglobin (HGB) Maximum Lab Value	39	53	10	67	58	118	133	150	180	254	2.152
Surgical/procedural complications (Category 3)											
Basophils Median Lab Value		7	3	43	196	243	231	309	366	364	3.731
Monocytes Minimum Lab Value		97	1	114	210	212	245	302	288	309	2.482
Medical complications (Category 4)											
Hemoglobin (HGB) Maximum Lab Value	9	18	22	16	8	35	21	54	149	245	2.235
Insertion of Endotracheal Tube Procedure	115	32	4	151	252	212	78	178	274	295	2.187
Other complications (Category 5)											
Red Blood Cell Distribution Width (RDW) Maximum Lab Value	6	1	5	5	11	21	5	115	230	239	2.815
Hemoglobin (HGB) Maximum Lab Value	22	13	9	13	38	24	105	140	246	249	2.811

These were the charts I put together based on the above information.

Figure 16-20. Top two features with the largest decrease and top two features with the largest increase over the time intervals as measured by the linear regression slope for each category. Features were only included if they existed in the data for at least seven of the 10 time intervals and their max ranking was within the top 10 at some interval. Only the top statistic for each lab value is included.

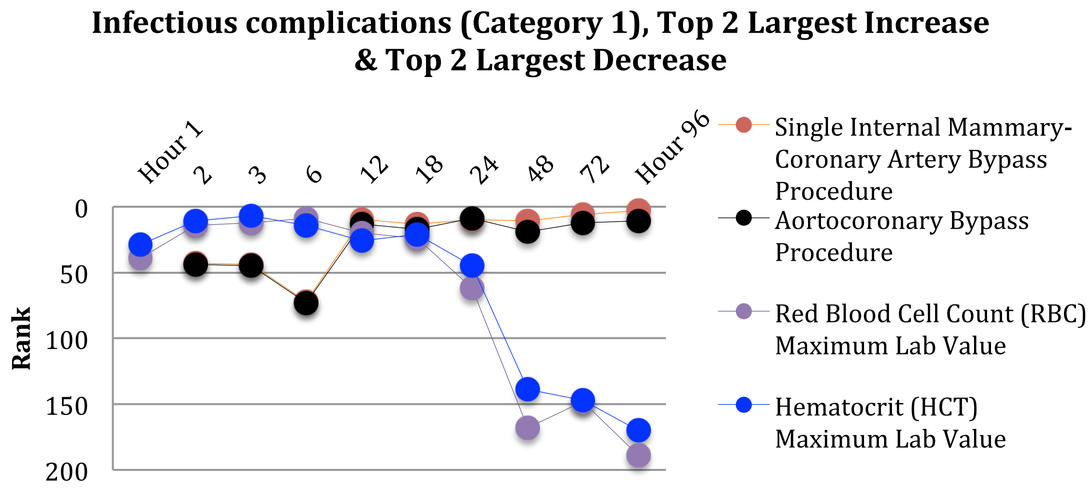


Figure 17

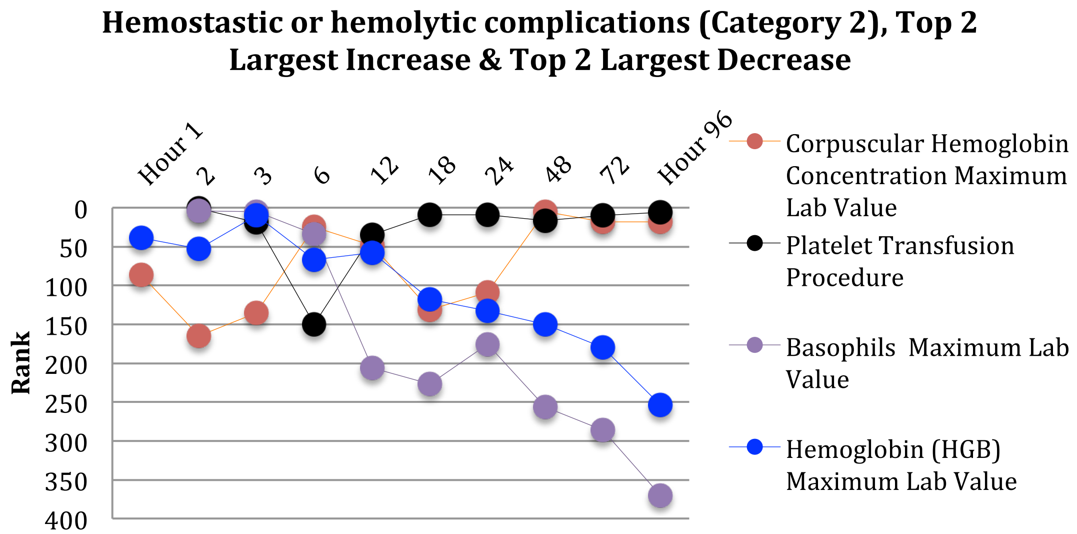


Figure 18

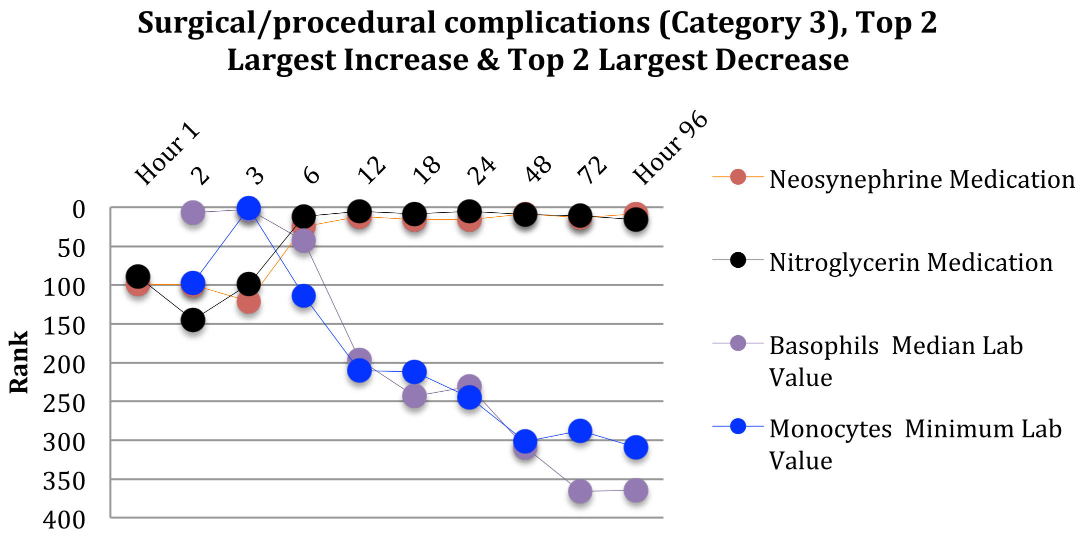


Figure 19

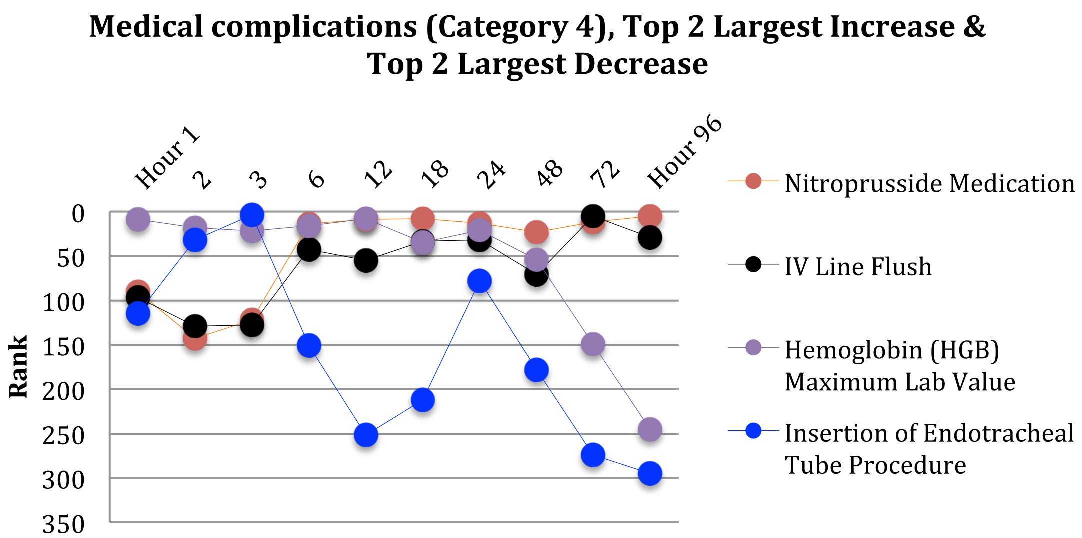


Figure 20

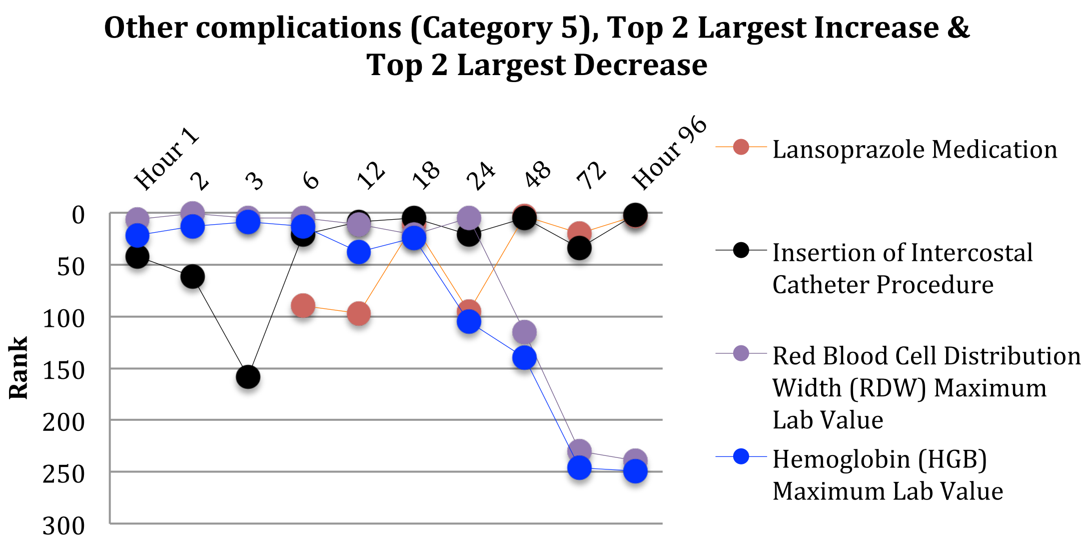
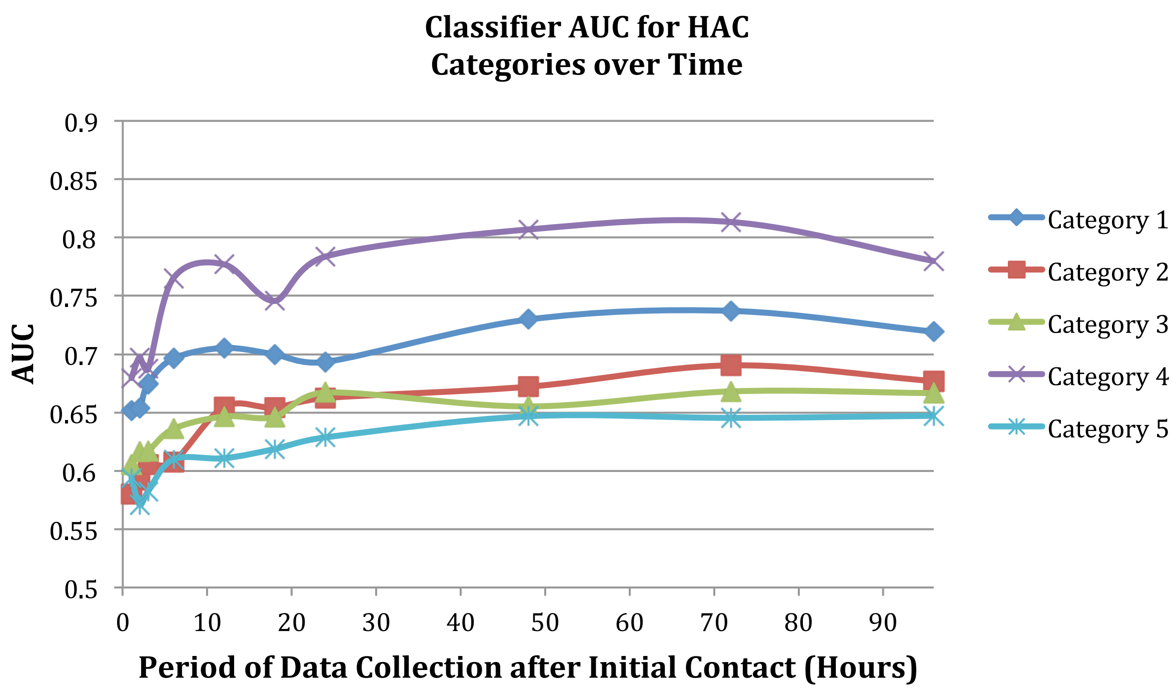


Table 31 reports the sensitivity, specificity, and AUC of our developed two level classifiers for predicting HACs, and Figure 2 visualizes the change in AUC for the classifiers over the increasing time intervals for each category. Full classifier information can be found in Appendix C.

Table 31. Reported sensitivity (True Positive Rate – TPR), specificity (True Negative Rate – TNR), and Area under the Receiver Operating Characteristic Curve (AUC) for the two level classifiers. TPR/TNR/AUC.

	Category 1	Category 2	Category 3	Category 4	Category 5
Hour 1	.602/.625/.652	.568/.560/.580	.571/.582/.605	.643/.604/.680	.608/.511/.594
Hour 2	.615/.616/.654	.551/.598/.592	.561/.601/.616	.685/.637/.697	.615/.487/.571
Hour 3	.635/.621/.675	.545/.620/.606	.561/.600/.617	.663/.607/.688	.599/.513/.583
Hour 6	.649/.631/.696	.570/.573/.608	.593/.615/.636	.705/.688/.765	.600/.561/.610
Hour 12	.664/.634/.705	.619/.611/.655	.611/.600/.647	.729/.688/.777	.571/.578/.611
Hour 18	.663/.628/.700	.611/.614/.654	.609/.612/.646	.672/.678/.746	.579/.580/.619
Hour 24	.655/.630/.693	.601/.631/.662	.621/.618/.667	.731/.703/.784	.588/.591/.629
Hour 48	.672/.657/.730	.605/.649/.672	.599/.614/.655	.724/.707/.807	.612/.608/.647
Hour 72	.698/.673/.737	.634/.661/.690	.612/.630/.668	.747/.738/.813	.610/.597/.645
Hour 96	.653/.655/.719	.612/.655/.677	.621/.621/.667	.706/.715/.780	.609/.609/.647

Figure 21 AUC of HAC predicting classifiers for each category over the data collection periods



4.1.9 Iteration 9: Dual-layer prediction with expanded complication ICD-9-codes (200+)

See Table in Appendix for the entire expanded complication codes (200+).

Table 32. Reported sensitivity (True Positive Rate – TPR), specificity (True Negative Rate – TNR), and Area under the Receiver Operating Characteristic Curve (AUC) for the two level classifiers. TPR/TNR/AUC.

TPR/TNR /AUC	Category 1	Category 2	Category 3	Category 4	Category 5
Hour 1	.570/.649/.645	.585/.520/.577	.515/.609/.584	.633/.610/.662	.621/.492/.580
Hour 2	.561/.646/.642	.550/.636/.613	.535/.628/.600	.643/.639/.708	.545/.594/.593
Hour 3	.554/.669/.654	.610/.538/.592	.577/.590/.603	.723/.636/.752	.580/.547/.592
Hour 6	.652/.618/.681	.575/.618/.638	.532/.643/.622	.723/.731/.786	.560/.558/.587
Hour 12	.633/.651/.695	.582/.619/.629	.551/.657/.634	.667/.713/.751	.617/.531/.610
Hour 18	.640/.630/.677	.601/.638/.670	.547/.652/.631	.698/.762/.789	.543/.616/.608
Hour 24	.662/.602/.685	.620/.661/.684	.538/.667/.644	.669/.760/.771	.475/.708/.629
Hour 48	.653/.632/.693	.664/.645/.700	.564/.662/.649	.703/.766/.813	.536/.644/.630
Hour 72	.603/.686/.691	.659/.668/.719	.536/.685/.661	.751/.820/.845	.520/.690/.629
Hour 96	.652/.666/.704	.660/.671/.721	.602/.627/.646	.749/.645/.767	.526/.732/.660

5 Analysis

5.1 Feature Rankings

5.1.1 Iteration 1

To assess the predictive power of individual features, we computed the information gain ratio of each attribute [18]. The information gain ratio measures the level of dependence between two attributes. Attributes with high information gain ratios with respect to the classification attribute (presence of HAC) are likely to be good predictors of HAC in patients for that dataset. The attributes were ranked by their information gain ratios to identify the relative predictive power of each attribute over each time interval. Since all of our features are either discretized down into four bins or only take on a small number of nominal values, we avoid the major drawbacks of using information gain namely that of overfitting on attributes with large numbers of discrete values [19].

5.1.2 Iteration 2: Prediction for definites vs. maybes

These were my observations.

One of the consistently highly ranked features found in this study is red cell distribution width (RDW) which agrees with previous literature that has found RDW to be a good independent predictor of morbidity [16], mortality [17], and hospital readmission [18]. Disorders, such as anemia, can affect cell size. Another highly ranked feature was urea nitrogen and previous literature has found that it is predictive of long-term mortality in critically ill patients [19].

Vancomycin was in the top 10 list of median predictors and also listed as a predictor with the biggest increase. Previous literature has shown that vancomycin, an antibiotic agent, is often used for surgical-site infections, and that organisms have started to develop resistance against these agents [69]. My hypothesis is that vancomycin has the biggest increase in rank because as a complication worsens, more antibiotic agents are used, and vancomycin, one of the stronger antibiotics becomes crucial over time. Previous literature mentions that vancomycin is often used in more severe cases including hospital acquired Clostridium difficile and for treatment-resistant organisms [70-74].

Integrelin was listed as a predictor with the biggest squared change and also the biggest decrease. Although I could not find previous literature that links integrelin to hospital acquired infections, previous literature has shown that it is used in the hospital with surgical operations [75].

In addition, while lab features tend to dominate the model during early intervals and procedural/medication features for later intervals, there is still a good amount of features that are highly ranked across all time intervals, and the distribution of these features seems to be roughly equal between labs, procedures, and medications values. This indicates that while labs may be more useful for predicting HACs immediately after admission and procedural/medication information may be more useful for predicting HACs several days after admission, all three of these feature types are needed to have accurate prediction of HACs.

5.1.3 Iteration 3: Prediction for categories

These were my observations.

For surgical/procedural complications (category 3), white blood cells (WBC) and subtypes lymphocytes and neutrophils were consistently highly ranked features. Previous literature has shown that WBC are significant features for predicting surgical/procedural complications and mortality [47-48].

5.1.4 Iteration 4: Predicting for 37 out of 43 ICD-9-CM codes with knowledge of categories

These were my observations.

In this initial iteration, I included categories as a feature to see how high of an AUC we could hit. Secondly, I only included a subset of the ICD-9-CM codes.

Again, RDW was a highly ranked feature, which is consistent with what we've seen in iteration 2. Creatinine was also a highly ranked feature, and previous literatures have shown that creatinine is a significant feature for predicting complications [49-53].

Mechanical complication of prosthetic graft of other tissues had the highest AUC at hour 96, but I could not find previous literature linking creatinine or RDW to that complication specifically.

5.1.5 Iteration 5: Predicting for ICD-9 codes with knowledge of categories and hadm˙id

These were my observations.

This time, I wanted to test the full set of ICD-9-codes. Similar to last time, RDW and creatinine were the top two predictors.

The AUC varies depending on the ICD9 code we are predicting (AUC range 0.422 to 0.869)

The following stood out to me:

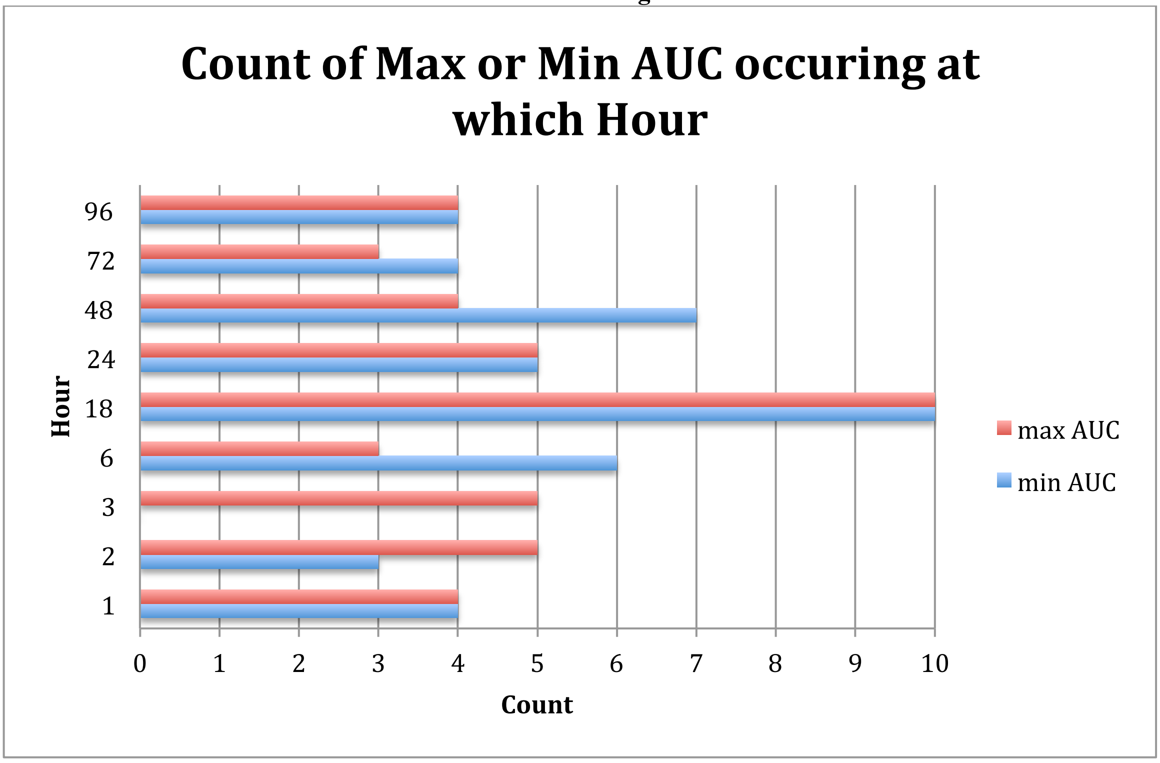
- the max AUC 0.869 for ICD9code E933.1 Adv Eff Antineoplastic occurred at hour 1.
- the min AUC 0.422 for ICD9code 528 Stomatitis occurred at hour 24.

- hour 18 contains the highest count of maximum AUC and minimum AUC
- hour 3 had five maximum AUC counts with 0 minimum AUC counts.

I could not find previous literature linking RDW or creatinine as predictive features for antineoplastic and immunosuppressive drugs causing adverse effects in therapeutic use or stomatitis.

The fact that hour 18 contains the highest count of maximum AUC and minimum AUC, I believe that this data supports the notion that the first 24 hours of a critical illness admission are crucial and that excessive handoffs during this period are likely to be problematic [54].

Figure 32. Count of Maximum or Minimum AUCs occurring at which Hour



It was also interesting to me that even with the categories and hadm˙id as handicaps, the AUC did not surpass that of the previous iteration. Puzzled by this, I chatted with my colleague Peijin, and he pointed out that because each patient could have multiple HACs, I was thereby splitting the AUC among multiple complications and implicitly placing a ceiling on the AUCs. Therefore, I was restricting the maximum AUC that could be achieved. Iteration 8 below was our ultimate and ideal iteration for the use of ICD-9-CM codes.

5.1.6 Iteration 6: Predicting for Categories with knowledge of ICD-9-CM codes

These were my observations.

Out of my own curiosity, I wanted to see how well we could predict categories with the knowledge of ICD-9-CM codes. Given that we bucketed the ICD-9-CM codes to match each of the categories, I was expecting high AUC.

To my surprise, the AUC was sometimes lower than the AUC achieved by our previous iterations.

As expected, the top feature was icd9code followed closely by age. Having seen that, I hypothesize that for future iterations, we should include patient symptoms at admission as a feature. Previous literatures have shown that age is a predictor of mortality [56-60].

I could not find previous literature linking platelet count (ranked number 3) to procedural/surgical complications (Category 3), but did find a paper linking platelet interaction to complications [61].

Glucose was ranked number 5, and previous literature have shown that increased mean blood glucose values were associated with risk of infection for patients undergoing coronary artery bypass graft (CABG) surgery and also considered independent predictors of morbidity and mortality in critically ill patients [62-64].

Partial thromboplastin time (PTT) was ranked number 6 and prothrombin time (PT) was ranked number 9 and have been cited by previous literature that its prolonged presence should raise suspicion for possible complications [65-67].

5.1.7 Iteration 7: Principal Components Analysis (PCA)

These were my observations.

I wanted to play around with feature selection and see if it would help improve our AUC. It only seemed to help our AUC by a little.

However, through PCA, we reduced the number of variables from approximately 300 to 45. The top contributors were RDW median, which we've seen already, and red blood cell (RBC) minimum and RBC maximum [68].

5.1.8 Iteration 8: Dual-layer prediction

These were my observations.

If we look at the top two largest increase and top two largest decrease features for the various categories (Figures 2-6), a common trend is that the initial incline and decline are observed at the same hour. For example, in category 1 the features that measure single internal mammary-coronary artery bypass procedure and aortocoronary bypass procedure has their initial inclines at hour 12. And the initial decline for features that measure red blood cell count (RBC) maximum lab value and hematocrit (HCT) maximum lab value also occurs at hour 12. Looking across all categories, this suggests that hours 6-18 are crucial for observing initial incline and decline of features to predict which category of HAC will occur [20-22, 28-33]. Another interesting commonality is that category 2 (Figure 3), category 4 (Figure 5), and category 5 (Figure 6) share the feature that measures hemoglobin (HGB) maximum lab value as one of the top 2 largest decreasing features. This suggests that hemoglobin (HGB) maximum lab value could be a feature to help predict HAC within the first 18 hours [23-27].

If we look within each of the categories, there are commonalities between the top two largest increase features and top two largest decrease features. For example, category 1 (Figure 2) top two increase features are single internal mammary-coronary artery bypass procedure and aortocoronary bypass overlay one another very closely. The same occurs for red blood cell count (RBC) maximum lab value and hematocrit (HCT) maximum lab value. This suggests that bypass-related procedures [20-22] and red blood cell related lab values are predictors for category 1, which consists of postoperative infections. Similarly, in category 3 (Figure 4) neosynephrine medication and nitroglycerin medication overlay one another very closely. The same occurs for the features that measure basophils median lab value and monocytes minimum lab value. This suggests that vasodilator and vasopressor medications and white blood cell lab values could be features to help predict category 3, which consists of gastronomic complications and operation surgical wounds [34-40]. In category 5, we also observed that the feature that measures chlorhexidine medication is the top median rank feature, the top mean rank feature, and also the number one ranked feature at hour 96. This suggests that chlorhexidine may be useful in predicting category 5 HACs [41-45].

5.1.9 Iteration 9: Dual-layer with expanded complication codes

In this iteration, the feature rankings remained the same, but we saw an improvement in the AUC to 0.8454 occurring at 48 hours.

5.2 Classifiers

The generated classifiers could act as a clinical decision support tool to assist doctors in the early detection of HACs in clinical settings, although any such system would have to be prospectively evaluated. The individual identified features could serve as important markers for clinicians to keep an eye out for as hospitalization progresses, in order to catch early signs of HACs. It is clear, based on the fact that all attributes undergo dynamic rank change as a function of time, that the process of hospitalization for critically ill patients is highly dynamic and likely progresses through distinct epochs.

5.3 HAC Cost Implications

Based on our research and previous literature on medical claims, I estimated that at least $10 billion can be saved per year (Table 33) [77-81]. Depending on which literature I read, there was a wide range of costs (Table Appendix).

Table 33. HAC cost breakdown examined in previous literature.

Category of HACs	% of injuries in our study	Total excessive cost per HAC	Total excessive cost of HAC per year based on our study
1- Infectious Complications	18%	$80,724	$30.5 billion
2- Bleeding or clotting Complications	23%	$10,225	$4.9 billion
3- Surgical Complications	55%	$16,851	$19 billion
4- Medical Complications	6%	$5,674	$710 million
5- Other complications	30%	$16,416	$10 billion
Total unique complications	33%	$13,000	$10 billion

6 Conclusion

Through this project, we were able to identify important features such as red blood cell distribution width (RDW) that could help inform healthcare professionals on decisions that would help reduce HACs from occurring in the hospital setting.

I believe that there is still room for improvement in the AUCs that we have achieved, and that we could increase the AUC by using new extracted data.

7 Future Work

These are my personal opinions:
We should extract new data for our next set of analyses because it feels like our current data limits our AUC.

- Modify the data extractor script to include symptoms at admission since the iteration I performed with ICD-9-CM codes as predictive features had AUC > 0.90.
- Modify the data extractor script to account for other labs that have been shown as strong predictors from previous literature. For example, glycaemic liability index (GLI).

These are the methods we have discussed as a group:

- Data extraction in this study used several summative statistics (min/max/median) for continuous features and only recorded the most recent measurement for nominal features. Such measures of the different lab, procedural, and medication features do not encode information about the history of these values over time for each patient. For example, different patients having measurements of (1, 2, 3), (3, 2, 1), and (1, 1, 3, 3) over time for a given lab test would end up with identical lab features from our extraction pipeline. Future work could go into developing methods for extracting such temporal information into features for use in the downstream models, or adopting methods such as those described by Lasko et al. [55].

- Applying a transformation such as restricted cubic spline, but also maintains continuous nature of features. The benefit here is that transformations can eliminate outliers and is insensitive to Gaussian distribution assumption.
- According to Ploeg et al. [76] modern modeling techniques such as SVM, NN and RF need far more events per variable to achieve a stable AUC-value than classical modeling techniques such as LR and CART. Their estimation is that approximately 200 events are needed per variable so given our approximately 350 variables, we should have about 75,000 patients. One suggestion was to combine our CPT codes and use a higher level-mapping table.
- Change the outcome from nominal Case = Yes or No or multinomial Case = 1 to 5 to a continuous outcome variable of the probability of developing complication (ranging from 0 to 1). The benefit of this is that we could use logistic regression or another model that doesn't suffer from discretization bias.

8 Acknowledgements

I would like to thank Prof. Alterovitz for all his support and guidance throughout this process. Thanks to Prof. Alterovitz's patience and counsel, I have had the opportuntiy to learn magnitudes of information in this area. I have been wanting to work on a health informatics project for a few years now so I'm excited that Prof. Alterovitz gave me the opportunity to do so and help me fulfill my dream. It has also been a pleasure working with Peijin Zhang and Prof. Warner, and they have taught me a lot throughout this experience as well!

Futhermore, I would like to thank my family and friends for their unwavering encouragement.

My brother and best friend are both doctors so I hope this work will help them and their patients in the foreseeable future. I strongly believe that machine learning is a powerful tool, and to have it be used to reduce patients from being in the hospital longer than needed makes me happy. It's wonderful to see all the other healthcare research projects out there, and I am looking forward to my own contribution in the future.

9 References

1 Fuller, R. L., McCullough, E. C., Bao, M. Z. & Averill, R. F. Estimating the costs of potentially preventable hospital acquired complications. *Health care financing review* 30, 17-32 (2009).

2 National Health Expenditures 2012 Highlights. (Centers for Medicare & Medicaid Services, 2012).

3 Klevens, R. M. *et al.* Estimating health care-associated infections and deaths in U.S. hospitals, 2002. *Public health reports* 122, 160-166 (2007).

4 Houle, D., Govindaraju, D. R. & Omholt, S. Phenomics: the next challenge. *Nature reviews. Genetics* 11, 855-866, doi:10.1038/nrg2897 (2010).

5 Kohane, I. S. Using electronic health records to drive discovery in disease genomics. *Nature reviews. Genetics* 12, 417-428, doi:10.1038/nrg2999 (2011).

6 Murphy, S. *et al.* Instrumenting the health care enterprise for discovery research in the genomic era. *Genome research* 19, 1675-1681, doi:10.1101/gr.094615.109 (2009).

7 Warner, J. L. & Alterovitz, G. Phenome based analysis as a means for discovering context dependent clinical reference ranges. *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium* 2012, 1441-1449 (2012).

8 Saeed, M. *et al.* Multiparameter Intelligent Monitoring in Intensive Care II: a public-access intensive care unit database. *Critical care medicine* 39, 952-960, doi:10.1097/CCM.0b013e31820a92c6 (2011).

9 Escobar, G. *et al.* Early Detection of Impending Physiologic Deterioration Among Patients Who Are Not in Intensive Care: Development of Predictive Models Using Data From an Automated Electronic Medical Record. *Journal of Hospital Medicine.* 7, 388-395, 10.1002/jhm.1929 (2012).

10 Tabak YP, Sun X, Nunez CM, *et al.* Using electronic health record data to develop inpatient mortality predictive model: Acute Laboratory Risk of Mortality Score (ALaRMS), *J Am Med Inform Assoc*; 21, 455–463, doi:10.1136/amiajnl-2013- 001790 (2014).

11 W. Dai, *et al.*, Prediction of hospitalization due to heart diseases by supervised learning methods, *Int. J. Med. Inform.*, doi:10.1016/j.ijmedinf.2014.10.002 (2014).

12 Tran, T. *et al.,* Risk stratification using data from electronic medical records better predicts suicide risks than clinician assessments, *BMC Psychiatry* 2014, 14:76, http://www.biomedcentral.com/1471-244X/14/76 (2014).

13 Roubinian, N. *et al.*, Predicting red blood cell transfusion in hospitalized patients: role of hemoglobin level, comorbidities, and illness severity, *BMC Health Services Research* 2014, 14:213 http://www.biomedcentral.com/1472-6963/14/213 .

14 Rana, S. *et al.*, Predicting unplanned readmission after myocardial infarction

from routinely collected administrative hospital data, *Australian Health Review*, 38, 377–382, doi: 10.1071/AH14059 (2014).

15 Cooper, G. F. & Herskovits, E. A Bayesian method for the induction of probabilistic networks from data. *Machine learning* 9, 309-347, doi:10.1007/bf00994110 (1992).

16 Friedman, N., Geiger, D. & Goldszmidt, M. Bayesian Network Classifiers. *Machine learning* 29, 131-163, doi:10.1023/a:1007465528199 (1997).

17 Dupuy, A. & Simon, R. M. Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. *J Natl Cancer Inst* 99, 147-157, doi:10.1093/jnci/djk018 (2007).

18 KENT, J. T. Information gain and a general measure of correlation. *Biometrika* 70, 163-173, doi:10.1093/biomet/70.1.163 (1983).

19. Beier, K., et. al., *Elevation of blood urea nitrogen is predictive of long-term mortality in critically ill patients independent of "normal" creatinine.* Crit Care Med, 20133. 39(2):p.305-13.

20. Raja, S., and M. Rochon, and J. Jarman, *Brompton Harefield Infection Score (BHIS): Development and validation of a stratification tool for predicting risk of surgical site infection after coronary artery bypass grafting*. International Journal of Surgery, 2015. 16(A): p.69-73.

21. Yumung, G., et al., *Deep sternal wound infection after coronary artery bypass surgery: management and risk factor analysis for mortality.* Heart Surg Forum, 2014. 17(4):E212-6.

22. Robich, MP., et al., *Prolonged Effect of Postoperative Infectious Complications on Survival After Cardiac Surgery.* Ann Thorac Surg, 2015.

23. Mostafa, EF., and AN. Mohammad, *Incidence and predictors of rebleeding after band ligation of oesophageal varices.* Arab J Gastroenterol, 2014. 15(3-4): p.135-41.

24. Ramo, BA. et al., *Surgical site infections are posterior spinal fusion for neuromuscular scoliosis: a thirty-year experience at a single institution.* J Bone Joint Surg Am, 2014. 96(24): p.2038-48.

25. Nie, K., et al., *Risk factors of intra-abdominal bacterial infection after liver transplantation in patients with hepatocellular carcinoma.* Chin J Cancer Res, 2014. 26(3): p.309-14.

26. Humphers, J., et al., *The Impact of Glycosylated Hemoglobin and Diabetes Mellitus on Postoperative Wound Healing Complications and Infection Following Foot and Ankle Surgery.* J Am Podiatr Med Assoc, 2014.

27. Chien, CY., et al., *Blood stream infection in patients undergoing systematic off-pump coronary artery bypass: incidence, risk factors, outcome, and associated pathogens.* Surg Infect (Larchmt), 2014. 15(5): p.613-8.

28. Hart, A., et al., *Blood transfusion in primary total hip and knee arthroplasty. Incidence, risk factors, and thirty-day complication rates.* J Bone Joint Surg Am, 2014. 96(23): p.1945-41.

29. Ozkardesler, S., et al., *Effects of blood products on nosocomial infections in liver transplant recipients.* Exp Clin Transplant, 2013. 11(6): p.530-6.

30. Kunac, A., et al., *Bacteremia and ventilator-associated pneumonia: a marker for contemporaneous extra-pulmonic infection.* Surg Infect (Larchmt), 2014. 15(2): p.77-83.
31. Lin, JH., et al., *Prognostic factors and complication rates for double-filtration plasmapheresis in patients with Guillain-Barré syndrome.* Transfus Apher Sci, 2015. 52(1):78-83.
32. Chahal, GS., et al., *A comparison of complications requiring return to theatre in hip and knee arthroplasty patients taking enoxaparin versus rivaroxaban for thromboprophylaxis.* Ortop Traumatol Rehabil, 2013. 15(2): p.125-9.
33. Omoke, NI., and CG. Nwigwe, *An analysis of risk factors associated with traumatic extremity amputation stump wound infection in a Nigerian setting.* Int Orthop, 2012. 36(11): p.2327-32.
34. Jog, M., et al., *Risk of contamination of lidocaine hydrochloride and phenylephrine hydrochloride topical solution: in vivo and in vitro analyses.* J Laryngol Otol, 2013. 127(8): p.799-801.
35. Mariappan, R., et al., *Circulatory collapse after topical application of vancomycin powder during spine surgery.* J Neurosurg Spine, 2013. 19(3): p381-3.
36. Hamaguchi, E., and H. Kawano, *A case of severe hypotension with catecholamine-resistant septic shock in the perioperative period.* Masui, 2012. 6(4): p. 400-3.
37. Blériot, C., et al., *Liver-resident macrophage necroptosis orchestrates type 1 microbicidal inflammation and type-2-mediated tissue repair during bacterial infection.* Immunity, 2015. 42(1): p.145-158.
38. Agarwal, R., et al., *Infection with human rhinovirus 16 promotes enhanced IgE responsiveness in basophils of atopic asthmatics.* Clin Exp Allergy, 2014. 44(1): p.1266-73
39. Ramadan, A., et al., *Activation of basophils by the double-stranded RNA poly(A:U) exacerbates allergic inflammation.* Allergy, 2013. 68(6): p.732-8.
40. Dapunt, U., et al., *The macrophage inflammatory proteins MIP1alpha (CCL3) and MIP2alpha (CXCL2) in implant-associated osteomyelitis: linking inflammation to bone degradation.* Mediators Inflamm, 2014.
41. Tsai, HC., et al., *Central venous catheter-associated bloodstream infections in pediatric hematology-oncology patients and effectiveness of antimicrobial lock therapy.* J Microbiol Immunol Infect, 2014. S1684-1182(14): p.00164-9.
42. Chien, CY., et al., *Care bundle to prevent methicillin-resistant Staphylococcus auereus sternal wound infection after off-pump coronary artery bypass.* Am J Infect Control, 2014. 42(5): p.562-4.
43. Falk-Brynhildsen, K., et al., *Bacterial recolonization of the skin and wound contamination during cardiac surgery: a randomized controlled trial of the use of plastic adhesive drape compared with bare skin.* J Hosp Infect, 2014. 84(2): p. 151-8.
44. Carroll K., et al., *Risk factors for superficial wound complications in hip and knee arthroplasty.* Clin Microbiol Infect, 2014. 20(2): p.130-5.
45. Thompson, P., and S. Houston. *Decreasing methicillin-resistant Staphylococcus aureus surgical site infections with chlorhexidine and mupirocin.* Am J Infect Control, 2013. 41(7): p.629-33.

46. White, A. and W. Liu, *Technical Note: Bias in Information-Based Measures in Decision Tree Induction.* Machine Learning, 1994. 15(3): p. 321-329.
47. Frazee, B., et al., *Community-Acquired Necrotizing Soft Tissue Infections. A Review of 122 Cases Presenting to a Single Emergency Department over 12 Years.* The Journal of Emergency Medicine, 2008. 34(2): p. 139-146.
48. Tran, HS., et al., *Predictors of operative outcome in patients with human immunodeficiency virus infection and acquired immunodeficiency syndrome.* Am J Surg, 2000. 180(3): p.228-33.
49. Lange, N., et al., *Pre-hospital vitamin D concentration, mortality, and bloodstream infection in a hospitalized patient population.* Am J Med, 2013. 126(7): p.640.
50. Gupta, T., et al., *Factors predicting temozolomide induced clinically significant acute hematologic toxicity in patients with high-grade gliomas: a clinical audit.* Clin Neurol Neurosurg, 2013. 115(9(: p.1814-9.
51. Tröger U., et al., *Decreased meorpenem levels in Itensive Care Unit patients with augmented renal clearance: benefit of therapeutic drug monitoring.* Int J Antimicrob Agents, 2012. 40(4): p.370-2.
52. Cillóniz, C., et al., *Pulmonary complications of pneumococcal community-acquired pneumonia: incidence, predictors and outcomes.* Clin Microbiol Infect, 2012. 18(11): p. 1134-42.
53. Srisawat, N., et al., *Plasma neutrophil gelatinase-associated lipocalin predicts recovery from acute kidney injury following community-acquired pneumonia.* Kidney Int, 2011. 80(5): p.545-52.
54. Collins, SA., et al., *In search of common ground in handoff documentation in an Intensive Care Unit.* J Biomed Inform, 2012. 45(2): p.307-15.
55. Lasko, TA., et al., *Computational phenotype discovery using unsupervised feature learning over noisy, sparse, and irregular clinical data.* PLoS One, 2013. 8(6):e66341.
56. Carton, JA., et al., *Infectious endocarditis of the native valve: its epidemiological profile and an analysis of its mortality between the years 1984 and 1993.* Med Clin (Barc), 1995. 104(13): p.493-9.
57. Friedant, AJ., et al., *A simple prediction score for developing a hospital-acquired infection after acute ischemic stroke.* J Stroke Cerebrovasc Dis, 2015. 24(3): p.680-6.
58. Nakaya, A., et al., *Does the hematopoietic cell transplantation specific comorbidity index (HCT-CI) predict transplantation outcomes? A prospective multicenter validation study of the Kanto Study Group for Cell Therapy.* Biol Blood Marrow Transplant, 2014. 20(10): p.1553-9.
59. Claridge, JA., et al., *Bacterial species-specific hospital mortality rate for intra-abdominal infections.* Surg Infect (Larchmt), 2014. 15(3): p.194-9.
60. Yang, BK., et al., *The simple predictors of pseudomembranous colitis in patients with hospital-acquired diarrhea: a prospective observational study.* Gut Liver, 2014. 8(1): p. 41-8.
61. Hincker, A., *Rotational thromboelastometry predicts thromboembolic complications after major non-cardiac surgery.* Crit Care, 2014, 18(5): p.549.

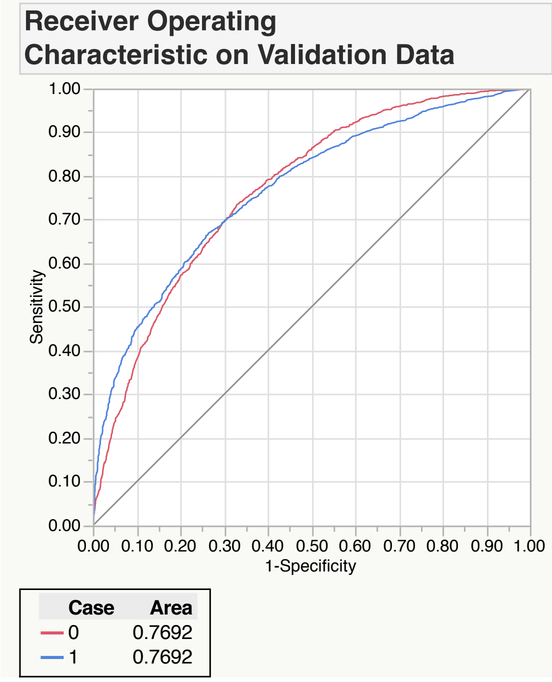
62. Malsa, M, et al., *HbA1c and diabetes predict perioperative hyperglycemia and glycemic variability in on-pump coronary artery bypass graft patients.* J Cardiothorac Vasc Anesth, 2011. 25(5): p.799-803.
63. Castellanos, MR., et al., *Fasting hyperglycemia upon hospital admission is associated with higher pneumonia complication rates among the elderly.* Int Arch Med, 2010. 3: p. 16.
64. Donati, A., et al., *Glycaemic variability, infections and mortality in a medical-surgical intensive care unit.* Crit Care Resusc, 2014. 16(1): p.13-23.
65. Yankol, Y., et al., *Acquired factor XI deficiency: a rare complication after liver transplantation.* Transplant Proc, 2015. 47(1): p.179-81.
66. Siekańska-Cholewa, A., et al., *Acquired factor V inhibitor in a woman following aortic aneurysm surgery.* Blood Coagul Fibrinolysis, 2014. 25(5): p.515-7.
67. Hertfelder, HJ., et al., *Perioperative monitoring of primary and secondary hemostasis in coronary artery bypass grafting.* Sermin Thromb Hemost, 2005. 31(4) p. 426-40.
68. Marks, DJ., et al., *Thoracic empyema: a 12-year study from a UK tertiary cardiothoracic referral centre.* PLoS One, 2012. 7(1): e30074.
69. Lenz, AM., et al., *Resistance profiles in surgical-site infection.* Future Microbiol, 2008. 3(4): p.453-62.
70. Halsey, J., *Current and future treatment modalities for Clostridium difficile-associated disease.* Am J Health Syst Pharm, 2008. 65(8) : p. 705-15.
71. Lundeen, SJ., et al., *Clostridium difficile enteritis: an early postoperative complication in inflammatory bowel disease patients after colectomy.* J Gastrointest Surg, 2007. 11(2): p.138-42.
72. Gallo, J., et al., *In vitro testing of gentamicin-vancomycin loaded bone cement to prevent prosthetic joint infection.* Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub, 2005. 149(1): p.153-8.
73. Bligny, D., et al., *Clostridium difficile infection in a Department of Internal Medicine. A consecutive series of 45 patients.* Ann Med Interne (Paris), 2002. 153(5): p. 291-9.
74. Ludlam, HA., et al., *The prevention of infection with Staphylococcus aureus in continuous ambulatory peritoneal dialysis.* J Hosp Infect, 1989. 14(4): p.293-301.
75. Heath, M., et al., *Using Zero-Balance Ultrafiltration with Dialysate as a Replacement Solution for Toxin and Eptifibatide Removal on a Dialysis-Dependent Patient During Cardiopulmonary Bypass.* J Cardiothorac Vasc Anesth, 2014. S1053-0770(14): p. 00607-7.
76. Ploeg, T., et al., *Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints.* BMC Medical Research Methodology, 2014. **14**(137): p. 1-13.
77. Zhan, C. and M. Miller, *Excessive Length of Stay, Charges and Mortality Attributable to Medical Injuries During Hospitalization.* JAMA, 2003. **290**(14): p.1868-1874.
78. Shreve, C., et al., *Economic Measurement of Medical Errors.* http://www.soa. org/files/pdf/research-econ-measurement.pdf, 2010.

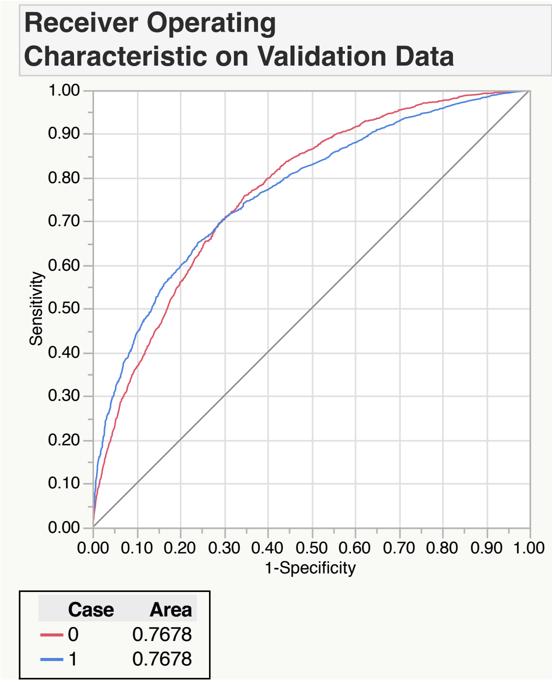
79. Centers for Disease Control and Prevention. *National Hospital Ambulatory Medical Care Survey: 2011 Emergency Department Summary Tables.* http://www.cdc.gov/nchs/data/ahcd/nhamcs˙emergency/2011˙ed˙web˙tables.pdf, 2011.
80. The Agency for Healthcare Research and Quality. *Interim Update on 2013 Annual Hospital-Acquired Condition Rate and Estimates of Cost Savings and Deaths Averted From 2010 to 2013.* http://www.ahrq.gov/professionals/quality-patient-safety/pfp/interimhacrate2013.pdf, 2013.
81. Magill SS, Edwards JR, Bamberg W, et al. *Multistate Point-Prevalence Survey of Health Care–Associated Infections.* *N Engl J Med*, 2014; 370: p.1198-208.

10 Supplementary

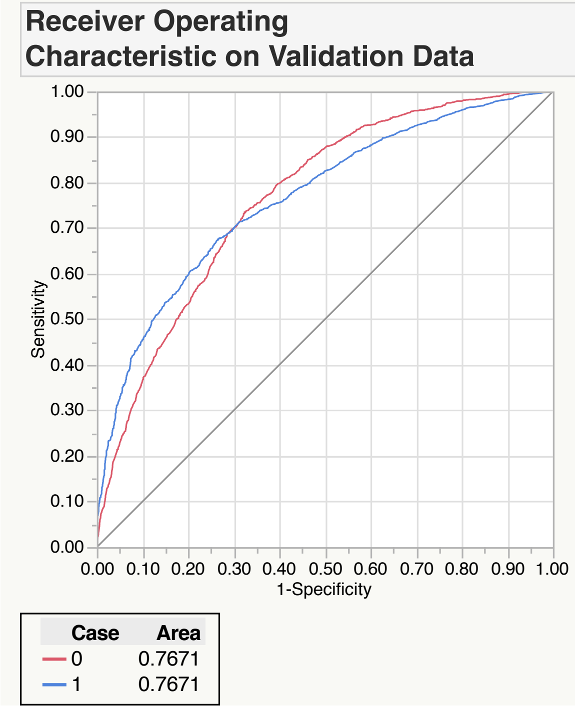
10.1.1 Iteration 2: Maybes + Definites

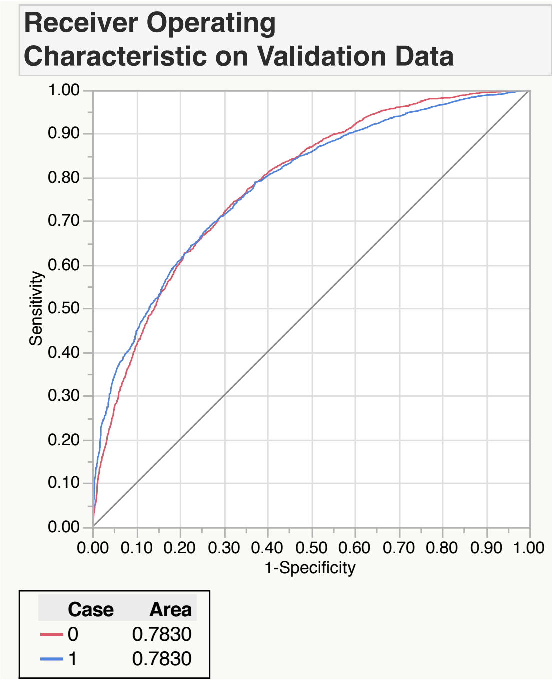
Hour 1- Maybes+Definites **Hour 2-Maybes+Definites**



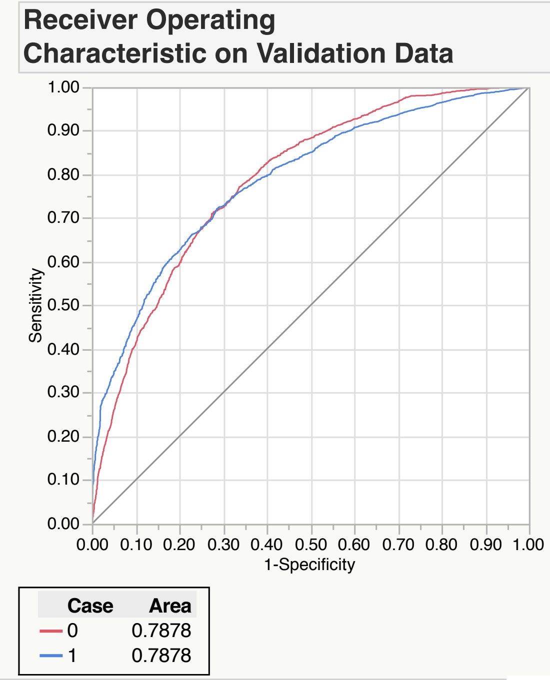


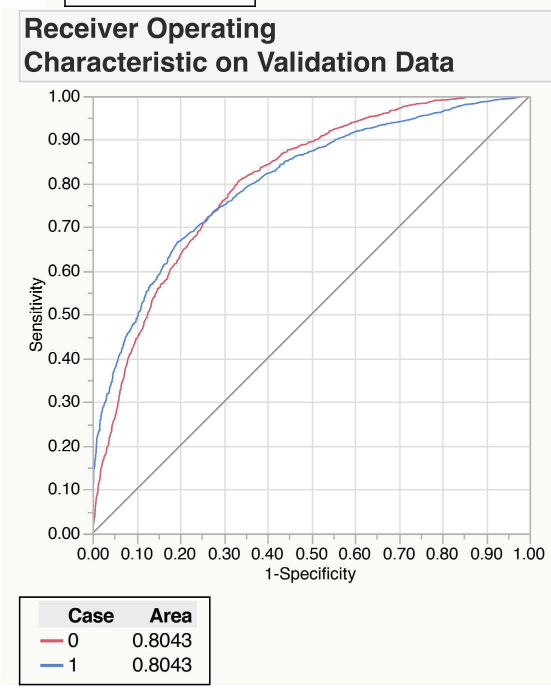
Hour 3 – Maybes + Definites **Hour 6 – Maybes + Definites**



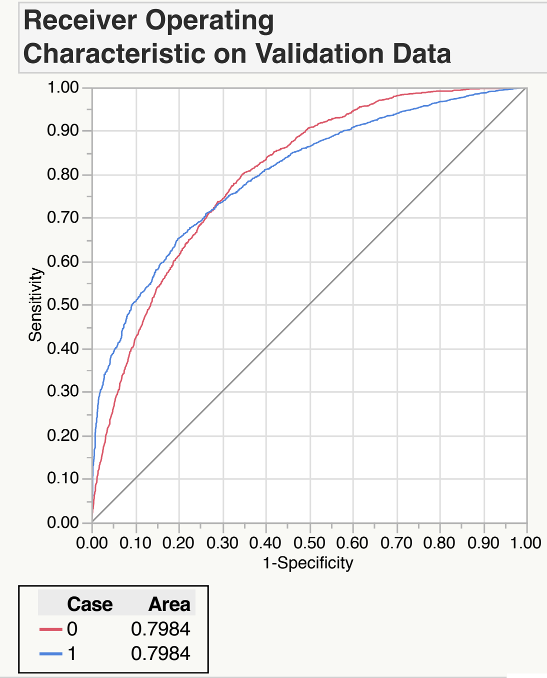


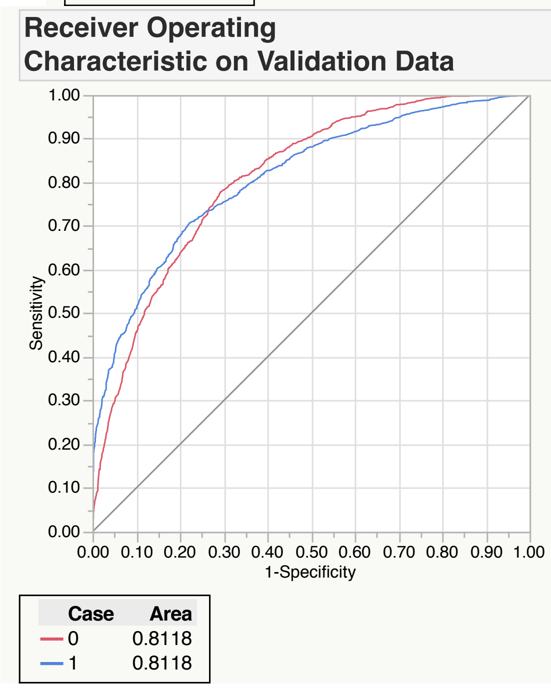
Hour 12 – Maybes + Definites **Hour 18- Maybes + Definites**



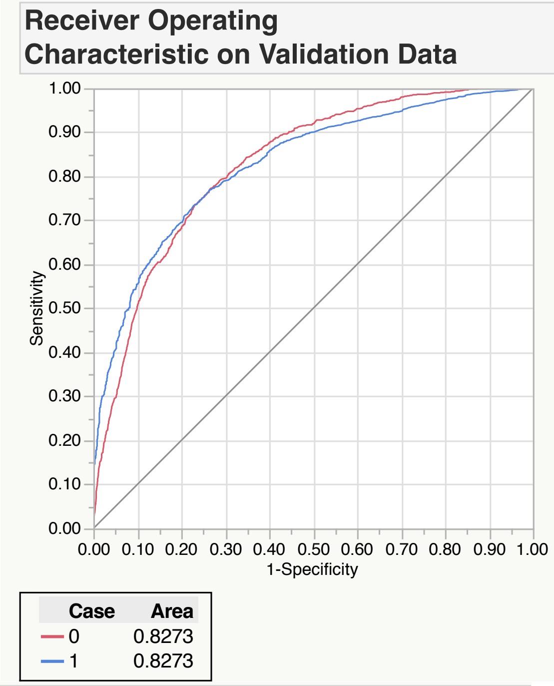


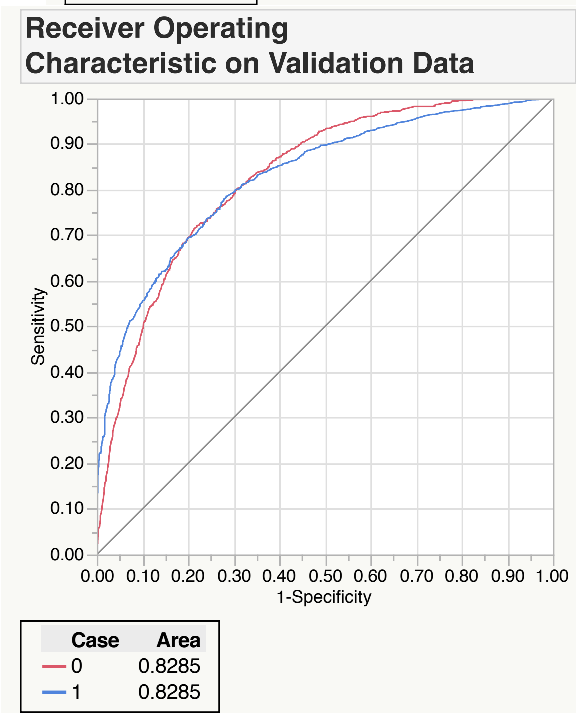
Hour 24 –Maybes + Definites **Hour 48 – Maybes + Definites**



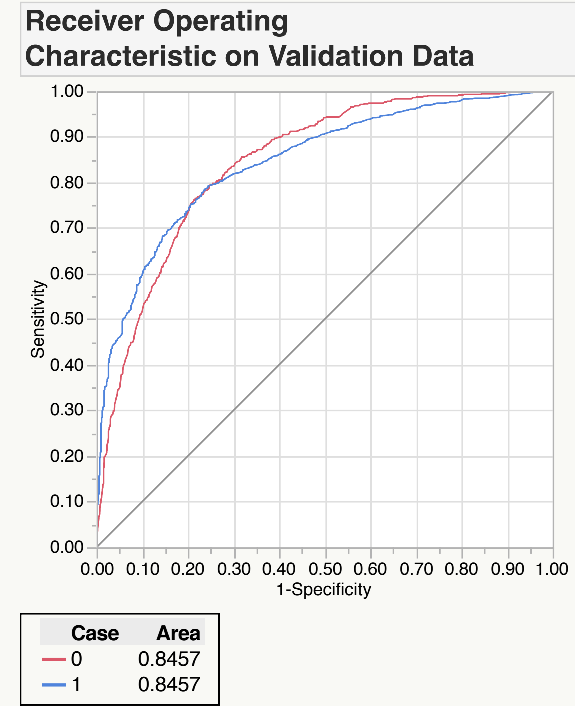


Hour 72 – Maybes + Definites **Hour 96 – Maybes + Definites**

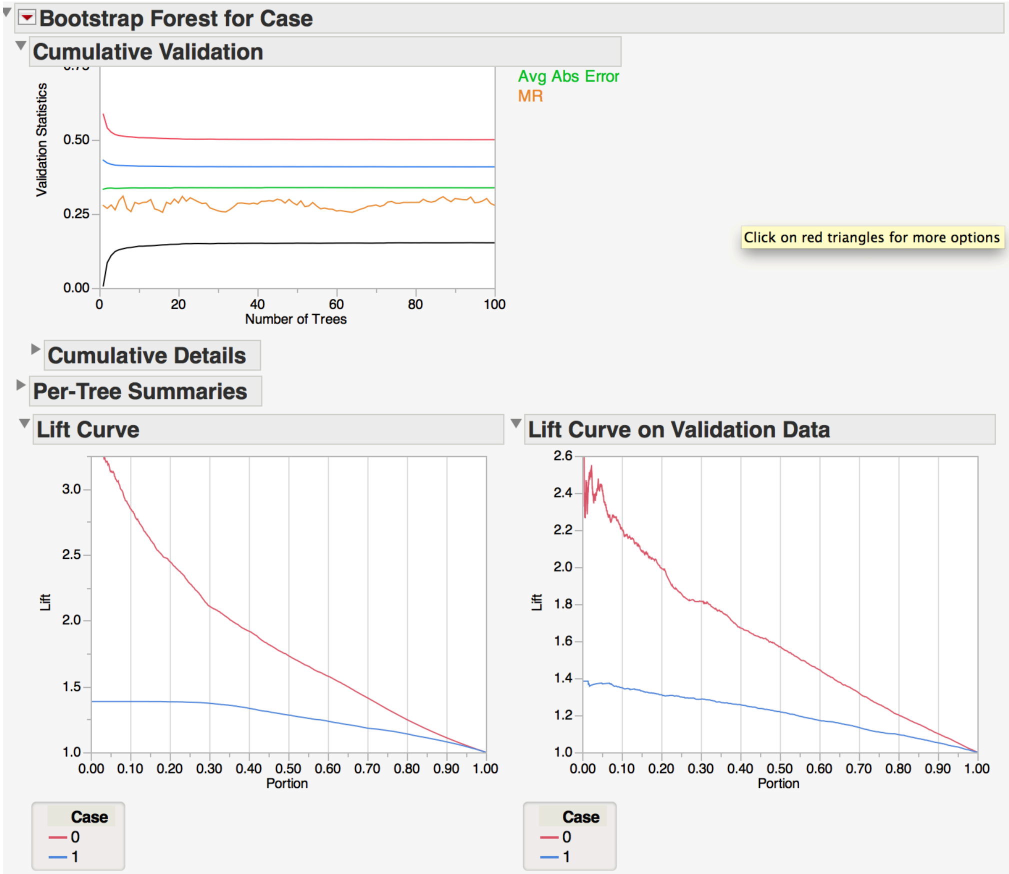




Hour 120 Maybes + Definites

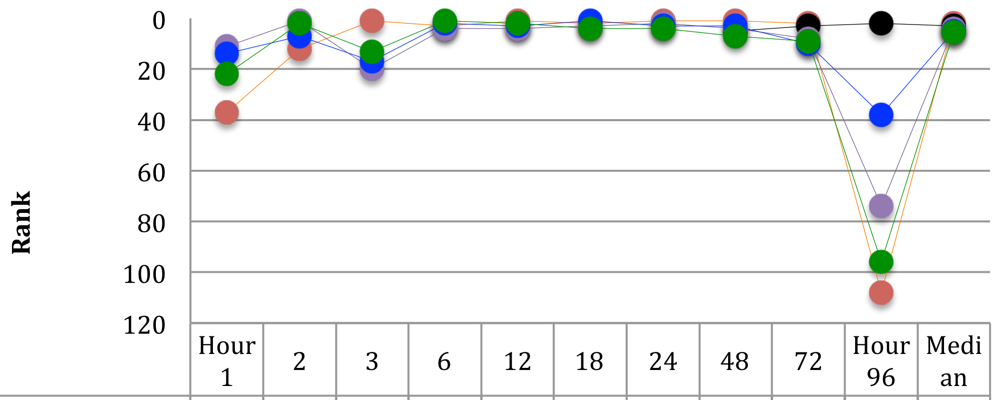


10.1.2 Iteration 3: Categories



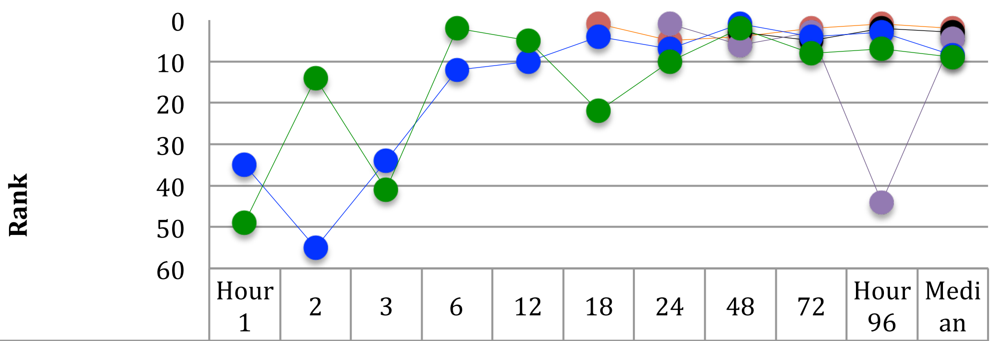
10.1.3 Iteration 8: Dual-layer prediction

Category 1, Top 5 Median Rank Attributes



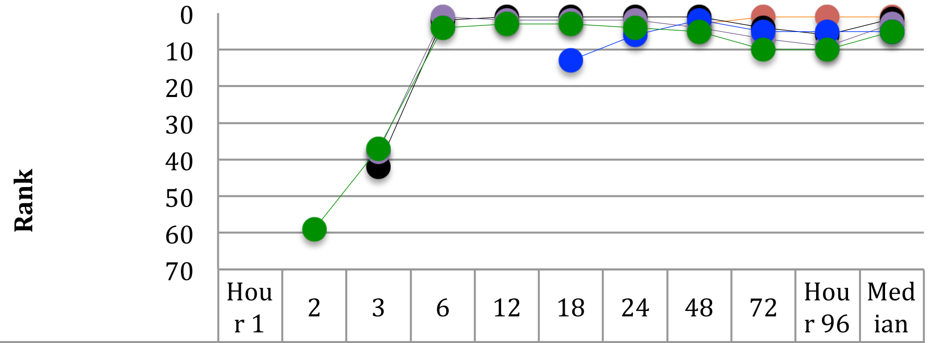
	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
vancomycin.IV.rx	37	12	1	3	1	2	1	1	2	108	2
INCISION W/REMOVAL FOREI.cpt								5	3	2	3
50444.min	11	1	20	4	4	3	2	4	8	74	4
50444.median	14	7	17	2	3	1	3	3	10	38	5
50444.max	22	2	13	1	2	4	4	7	9	96	5.5

Category 2, Top 5 Median Rank Attributes



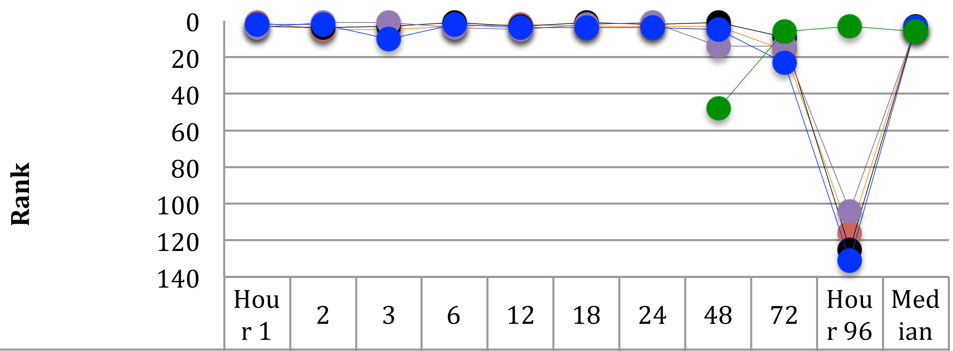
	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
REOPEN THORACOTOMY SITE.cpt						1	5	4	2	1	2
protamine sulfate.IV.rx								3	5	2	3
50650.nn							1	6	3	44	4.5
50383.min	35	55	34	12	10	4	7	1	4	3	8.5
50440.max	49	14	41	2	5	22	10	2	8	7	9

Category 3, Top 5 Median Rank Attributes



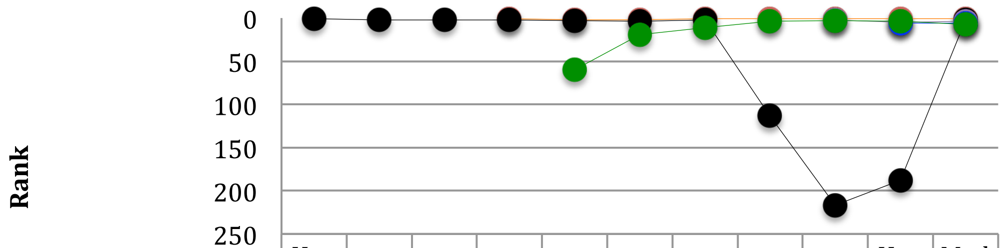
	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
mycophenolate mofetil.POIV.rx								3	1	1	1
neostigmine.IV.rx			42	2	1	1	1	1	4	6	1.5
glycopyrrolate.IV.rx			38	1	2	2	2	4	7	9	3
epinephrine.IV.rx						13	6	2	5	5	5
metoclopramide.POIV.rx		59	37	4	3	3	4	5	10	10	5

Category 4, Top 5 Median Rank Attributes



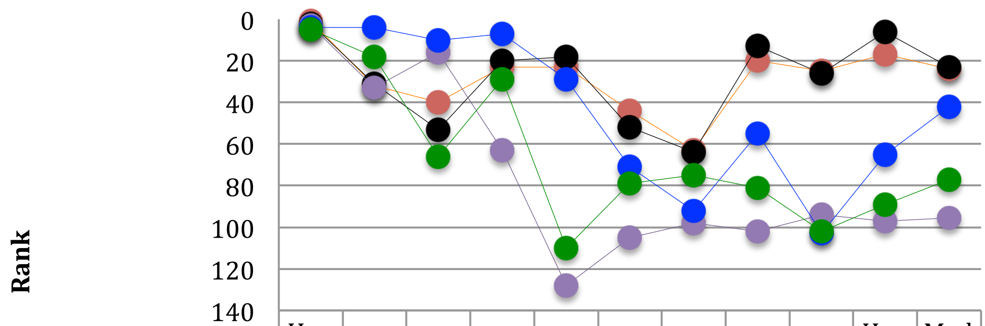
	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
50444.median	1	5	5	3	2	3	3	3	16	116	3
50444.max	3	4	3	1	3	1	2	1	9	125	3
ethnicity	4	1	1	4	5	2	1	14	14	104	4
50444.min	2	2	10	2	4	4	4	5	23	131	4
protamine sulfate.IV.rx								48	6	3	6

Category 5, Top 5 Median Rank Attributes



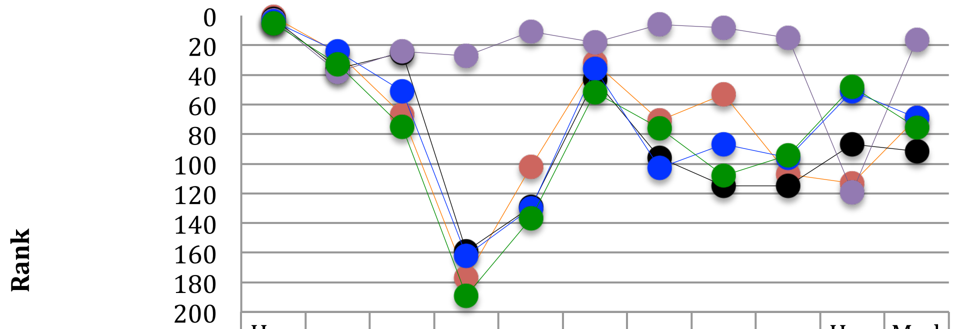
	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
chlorhexidine.PO.rx				1	2	2	1	1	1	1	1
ethnicity	1	2	2	2	3	4	2	113	217	188	2.5
miconazole powder 2%.TP.rx									2	5	3.5
ascorbate.POIV.rx										6	6
hydrocortisone.POIV.rx					60	19	11	4	3	4	7.5

Category 1, Top 5 Hour 1 Attributes



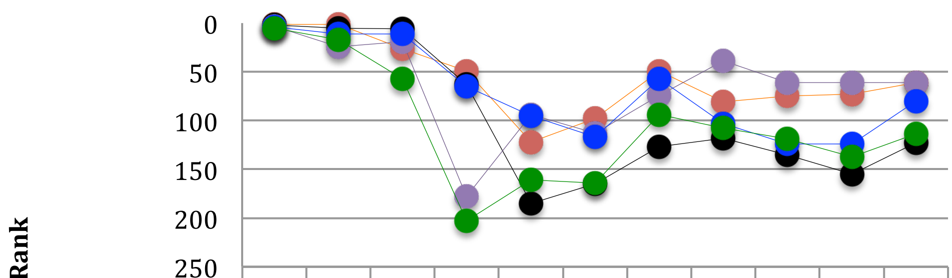
	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
CORONAR ARTERIOGR-2 CATH.cpt	1	32	40	23	23	44	63	20	25	17	24
LT HEART ANGIOCARDIOGRAM.cpt	2	31	53	20	18	52	64	13	26	6	23
ARTERIAL CATHETERIZATION.cpt	3	33	16	63	128	105	98	102	94	97	95.5
RT/LEFT HEART CARD CATH.cpt	4	4	10	7	29	71	92	55	103	65	42
PTCA W/O THROMBOLYTIC AG.cpt	5	18	66	29	110	79	75	81	102	89	77

Category 2, Top 5 Hour 1 Attributes



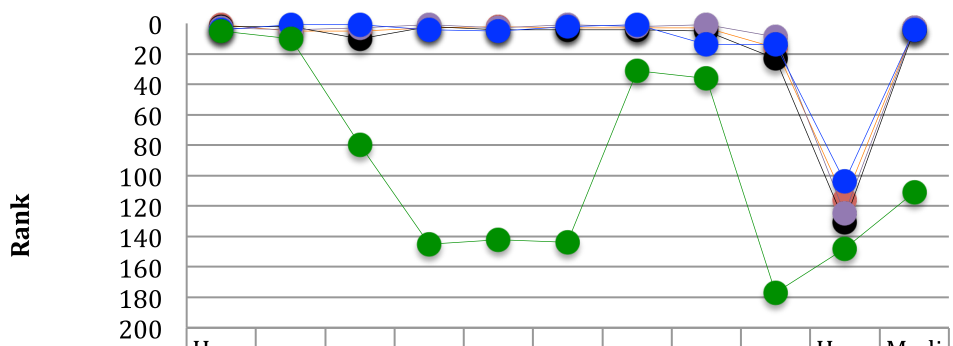
	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
SM BOWEL ENDOSCOPY NEC.cpt	1	25	67	177	102	32	71	53	107	113	69
INJ/INF PLATELET INHIBIT.cpt	2	36	25	159	129	43	96	115	115	87	91.5
EXT INFUS CONC NUTRITION.cpt	3	38	24	27	11	18	6	8	15	119	16.5
PULSATION BALLOON IMPLAN.cpt	4	24	51	162	130	36	103	87	96	51	69
INSERT DRUG-ELUT CORONA.cpt	5	33	75	189	137	52	76	108	94	48	75.5

Category 3, Top 5 Hour 1 Attributes



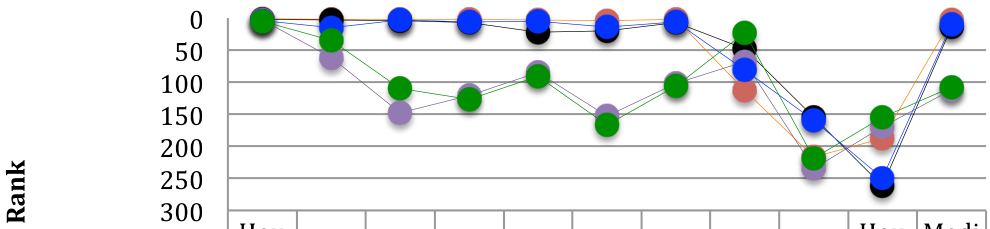
	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
ARTERIAL CATHETERIZATION.cpt	1	1	26	49	122	98	49	81	75	73	61
CONTINUOUS INVASIVE MECH.cpt	2	5	6	63	185	165	127	118	135	155	122.5
CONTR CEREBR ARTERIOGRAM.cpt	3	24	19	178	94	113	74	39	61	61	61
RT/LEFT HEART CARD CATH.cpt	4	11	11	65	95	117	57	103	124	124	80
LT HEART ANGIOCARDIOGRAM.cpt	5	17	57	203	161	164	94	108	119	137	113.5

Category 4, Top 5 Hour 1 Attributes



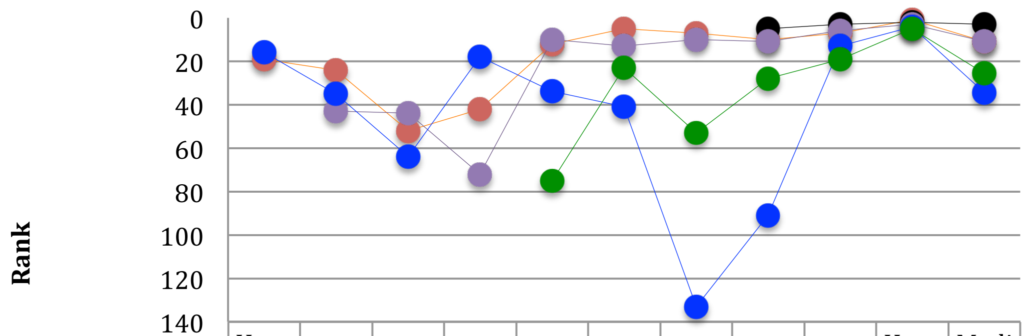
	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
50444.median	1	5	5	3	2	3	3	3	16	116	3
50444.min	2	2	10	2	4	4	4	5	23	131	4
50444.max	3	4	3	1	3	1	2	1	9	125	3
ethnicity	4	1	1	4	5	2	1	14	14	104	4
EXTRACORPOREAL CIRCULAT.cpt	5	10	80	145	142	144	31	36	177	148	111

Category 5, Top 5 Hour 1 Attributes



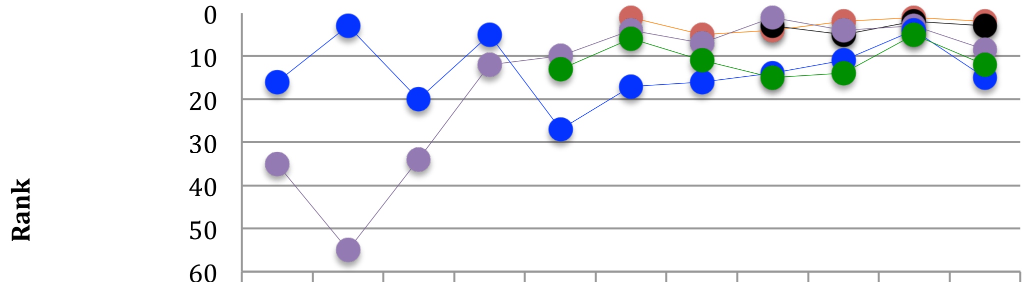
	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
ethnicity	1	2	2	2	3	4	2	113	217	188	2.5
50444.median	2	3	4	7	22	20	7	48	155	262	13.5
DX ULTRASOUND-HEART.cpt	3	62	148	121	85	153	102	68	235	170	111.5
50444.min	4	15	3	6	5	14	6	81	160	250	10
CONTINUOUS INVASIVE MECH.cpt	5	35	110	127	91	167	106	23	219	155	108

Category 1, Top 5 Hour 96 Attributes



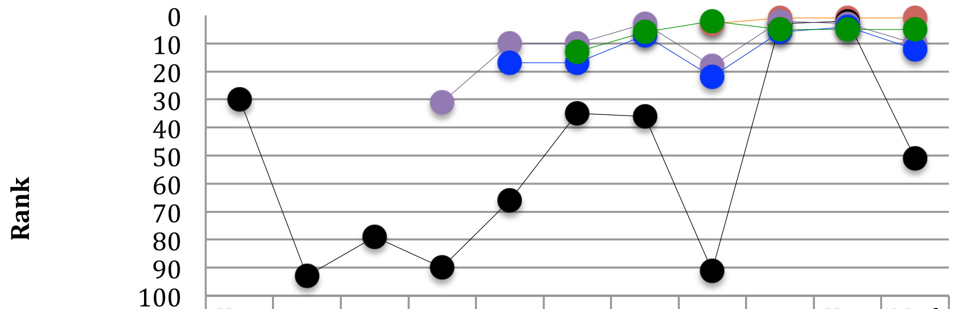
	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
EXTRACORPOREAL CIRCULAT.cpt	19	24	52	42	12	5	7	10	7	1	11
INCISION W/REMOVAL FOREI.cpt								5	3	2	3
1 INT MAM-COR ART BYPASS.cpt		43	44	72	10	13	10	11	6	3	11
HEMODIALYSIS.cpt	16	35	64	18	34	41	133	91	13	4	34.5
CONT INTRA-ARTERIAL BLD.cpt					75	23	53	28	19	5	25.5

Category 2, Top 5 Hour 96 Attributes



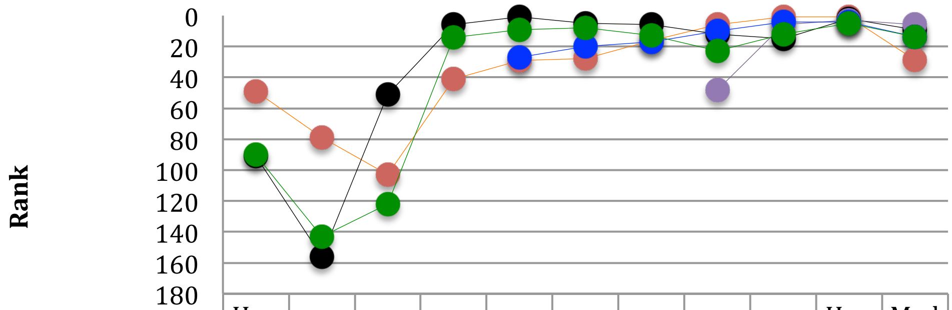
	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
REOPEN THORACOTOMY SITE.cpt						1	5	4	2	1	2
protamine sulfate.IV.rx								3	5	2	3
50383.min	35	55	34	12	10	4	7	1	4	3	8.5
SERUM TRANSFUSION NEC.cpt	16	3	20	5	27	17	16	14	11	4	15
SUTURE OF ARTERY.cpt					13	6	11	15	14	5	12

Category 3, Top 5 Hour 96 Attributes



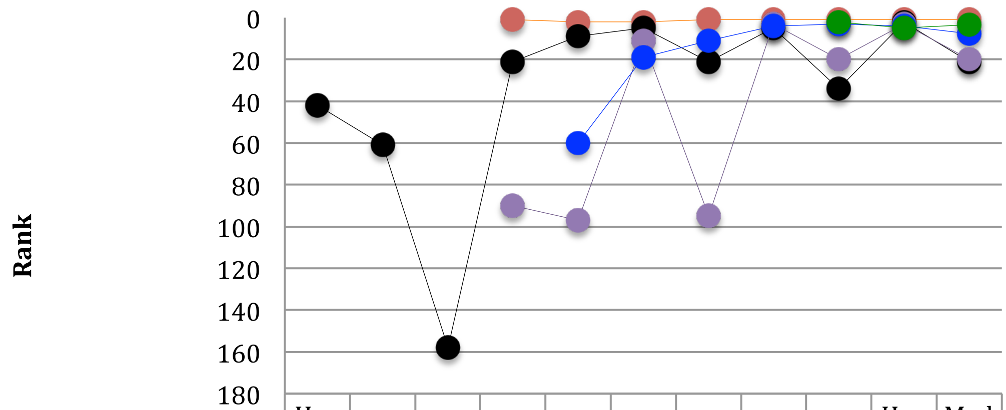
	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
mycophenolate mofetil.POIV.rx								3	1	1	1
amiodarone.POIV.rx	30	93	79	90	66	35	36	91	3	2	51
ceftriaxone.IVIM.rx				31	10	10	3	18	2	3	10
azithromycin.POIV.rx					17	17	7	22	6	4	12
epinephrine.IV.rx						13	6	2	5	5	5

Category 4, Top 5 Hour 96 Attributes



	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
nitroglycerine.IV.rx	49	79	103	41	29	28	16	6	1	1	28.5
cefazolin.IV.rx	91	156	51	6	1	5	6	12	15	2	9
protamine sulfate.IV.rx								48	6	3	6
ketorolac.IVIM.rx					27	20	17	10	4	4	13.5
nitroprusside.IV.rx	90	143	122	14	9	8	13	23	12	5	13.5

Category 5, Top 5 Hour 96 Attributes



	Hour 1	2	3	6	12	18	24	48	72	Hour 96	Median
chlorhexidine.PO.rx				1	2	2	1	1	1	1	1
INSERT INTERCOSTAL CATH.cpt	42	61	158	21	9	5	21	5	34	2	21
lansoprazole.PO.rx				90	97	11	95	3	20	3	20
hydrocortisone.POIV.rx					60	19	11	4	3	4	7.5
miconazole powder 2%.TP.rx									2	5	3.5

Table. List of ICD-9-CM codes used to define definite HACs and associated categories. Not expanded.

ICD-9-CM Code	Description	Category	ICD-9-CM Code	Description	Category
293	ACUTE DELIRIUM	5	996.74	COMPL FROM OTH VASCULAR	3
348.31	METABOLIC ENCEPHALOPATHY	5	996.81	COMPLICATIONS OF TRANSPLANTED KIDNEY	3
349.82	TOXIC ENCEPHALOPATHY	5	996.82	COMPLICATIONS OF TRANSPLANTED LIVER	3
415.11	IATROGENIC PULMONARY EMBOLISM AND INFARCTION	2	996.85	COMPLICATIONS OF TRANSPLANTED BONE MARROW	5
512.1	IATROGENIC PNEUMOTHORAX	5	997.02	IATROGENIC CEREBRO INFAR	2
528	STOMATITIS	5	997.69	AMPUTAT COMPLIC NEC	3
536.41	INFECTION OF GASTROSTOMY	1	998.11	HEMORR COMPLIC PROCEDURE	2
536.42	MECHANICAL COMPLICATION OF GASTROSTOMY	3	998.2	ACCIDENTAL OP LACERATION	3
536.49	OTHER GASTROSTOMY COMPLICATIONS	3	998.31	DISRUPT INTER OPER WOUND	3
568	PERITON ADHES (POSTOP)	3	998.32	DISRUPT EXTER OPER WOUNF	3
569.69	OTHER COLOSTOMY AND ENTEROSTOMY COMPLICATION	3	998.59	OTHER POSTOPERATIVE INFECTION	1
579.3	INTEST POSTOP NONABSORB	3	998.6	PERSISTENT POSTOPERATIVE FISTULA	3
693	DERMATITIS DUE TO DRUGS AND MEDICINES	4	998.83	NON-HEALING SURGICAL WOUND	3
707	DECUBITUS ULCER	5	999.9	COMPLIC MED CARE NEC/NOS	5
707.03	DECUBITUS ULCER, LOWER B	5	E879.8	ABN REACT-PROCEDURE NEC	3
707.05	DECUBITUS ULCER, BUTTOCK	5	E930.8	ADV EFF ANTIBIOTICS NEC	4

996.1	MALFUNC VASC DEVICE/GRAF	3	E933.1	ADV EFF ANTINEOPLASTIC	4
996.52	MECHANICAL COMPLICATION OF PROSTHETIC GRAFT	3	E934.2	ADV EFF ANTICOAGULANTS	4
996.59	MALFUNC OTHER DEVICE/GRA	3	V55.0	ATTENTION TO TRACHEOSTOMY	3
996.61	INFECT DUE TO CARDIAC IM	1	V55.2	ATTENTION TO ILEOSTOMY	3
996.62	INFECT DUE TO VASCULAR G	1	V55.3	ATTENTION TO COLOSTOMY	3
996.65	INFECT DUE TO GU IMPLANT	1			

Table. List of Expanded complications.

Category 1 – Infectious Complications	
ICD-9-CM Code	*Description*
519.01	TRACHEOST INFECTN (Begin 1998)
536.41	INFECTION OF GASTROSTOMY
996.61	INFECT DUE TO CARDIAC IM
996.62	INFECT DUE TO VASCULAR G
996.65	INFECT DUE TO GU IMPLANT
997.62	INFECTION AMPUTAT STUMP
998.51	INFECTED POSTOP SEROMA (Begin 1996)
998.59	OTHER POSTOPERATIVE INFECTION
998.59	OTH POSTOP INFECTION (Begin 1996)
999.3	INFEC COMPL MED CARE NEC (end 2007)
999.39	INFECT FOL INFUS/INJ/VAC (Begin 2007)
Category 2 – Bleeding and Clotting Complications	
ICD-9-CM Code	*Description*
415.11	IATROGENIC PULM EMBOL INFARCT (Begin 1995)
415.11	IATROGENIC PULMONARY EMBOLISM AND INFARCTION
518.7	TRANSFSN REL AC LUNG INJ (Begin 2006)
997.02	IATROGENIC CEREBRO INFAR
998.11	HEMORR COMPLIC PROCEDURE
998.11	HEMORR AS PROC CX (Begin 1996)
998.12	HEMATOMA PROC CX (Begin 1996)

999.5	SERUM REACTION NEC (end 2011)
999.6	ABO INCOMPAT REAC NOS (Begin 2010)
999.6	ABO INCOMPATIBILITY REAC (end 2010)
999.8	TRANSFUSION REACTION NEC (end 2008)
999.8	Transfusion reaction NOS (Begin 2010)
Category 3 – Surgical Complications	
ICD-9-CM Code	*Description*
349.1	COMPLICATION CNS DEVICE
429.4	Functional disturbances following cardiac surgery
519.02	TRACHEOST MECH COMPL (Begin 1998)
519.09	TRACHEOST COMPL OTH (Begin 1998)
530.87	MECH COMP ESOPHAGOSTOMY (Begin 2004)
536.42	MECHANICAL COMPLICATION OF GASTROSTOMY
536.49	OTHER GASTROSTOMY COMPLICATIONS
568	PERITON ADHES (POSTOP)
569.62	MECH COMPL COLOST ENTEROS (Begin 1998)
569.69	OTHER COLOSTOMY AND ENTEROSTOMY COMPLICATION
579.3	INTEST POSTOP NONABSORB
909.3	LATE EFF SURG/MED COMPL
996.1	MALFUNC VASC DEVICE/GRAF
996.52	MECHANICAL COMPLICATION OF PROSTHETIC GRAFT OF OTH
996.59	MALFUNC OTHER DEVICE/GRA
996.74	COMPL FROM OTH VASCULAR
996.81	COMPLICATIONS OF TRANSPLANTED KIDNEY
996.82	COMPLICATIONS OF TRANSPLANTED LIVER
997.02	POSTOP STROKE (Begin 1995)
997.1	SURG COMPL-HEART
997.2	SURG COMP-PERI VASC SYST
997.3	SURG COMPLIC-RESPIR SYST (end 2008)
997.4	SURG COMPLIC-GI TRACT (end 2011)
997.5	SURG COMPL-URINARY TRACT
997.69	AMPUTAT COMPLIC NEC
997.71	VASCULAR CX MESENTERIC ARTERY (Begin 2001)

997.72	VASCULAR CX RENAL ARTERY (Begin 2001)
997.79	VASCULAR CX OTHER VESSELS (Begin 2001)
997.91	HTN AS COMPLIC (Begin 1995)
998.2	ACCIDENTAL OP LACERATION
998.2	ACCIDENTAL OP LACERATION
998.3	POSTOP WOUND DISRUPTION (End 2002)
998.3	WOUND DISRUPTION NOS (Begin 2008)
998.31	DISRUPT INTER OPER WOUND
998.31	DISRUPTION OF INTERNAL OPER WOUND (Begin 2002)
998.32	DISRUPT EXTER OPER WOUNF
998.32	DISRUPTION OF EXTERNAL OPER WOUND (Begin 2002)
998.4	FB LEFT DURING PROCEDURE
998.6	PERSISTENT POSTOPERATIVE FISTULA NOT ELSEWHERE CLA
998.6	PERSIST POSTOP FISTULA
998.81	SUBQ EMPHYSEMA FROM PROC (Begin 1994)
998.83	NON-HEALING SURGICAL WOUND
998.83	NONHEALING SURG WND (Begin 1996)
998.9	SURGICAL COMPLICAT NOS
999.2	VASC COMP MED CARE NEC
E879.8	ABN REACT-PROCEDURE NEC
V55.0	ATTENTION TO TRACHEOSTOMY
V55.2	ATTENTION TO ILEOSTOMY
V55.3	ATTENTION TO COLOSTOMY
Category 4 – Medical Complications	
ICD-9-CM Code	*Description*
693	DERMATITIS DUE TO DRUGS AND MEDICINES TAKEN INTERN
E930.8	ADV EFF ANTIBIOTICS NEC
E933.1	ADV EFF ANTINEOPLASTIC
E934.2	ADV EFF ANTICOAGULANTS
Category 5 – Other Complications	
ICD-9-CM Code	*Description*
245.4	IATROGENIC THYROIDITIS
253.7	IATROGENIC PITUITARY DIS
293	ACUTE DELIRIUM
348.31	METABOLIC ENCEPHALOPATHY

349.82	TOXIC ENCEPHALOPATHY
458.2	IATROGENIC HYPOTENSION (Begin 1995 End 2003)
458.29	OTHER IATROGENIC HYPOTENSION (Begin 2003)
512.1	IATROGENIC PNEUMOTHORAX
528	STOMATITIS
707	DECUBITUS ULCER
707.03	DECUBITUS ULCER LOWER B
707.05	DECUBITUS ULCER BUTTOCK
996.85	COMPLICATIONS OF TRANSPLANTED BONE MARROW
997.01	CNS CX (Begin 1995)
997.09	OTH NERV SYS CX (Begin 1995)
997.99	OTH COMPLIC NOS (Begin 1995)
998.13	SEROMA PROC CX (Begin 1996)
998.89	OTHER COMPLIC NEC (Begin 1994)
999.1	AIR EMBOL COMP MED CARE
999.9	COMPLIC MED CARE NEC/NOS
999.9	COMPLIC MED CARE NEC/NOS

Table. List of Classifiers to Determine our Top Predictor

1. TAN Bayes
2. K2 Bayes P1
3. K2 Bayes P2
4. K2 Bayes P4
5. SVM K0 C0.1
6. SVM K0 C1.0
7. SVM K1 C0.1
8. SVM K1 C1.0
9. SVM K1 C10.0
10. SVM K2 C0.1
11. SVM K2 C1.0
12. SVM K2 C10.0
13. Ridge Logistic Regression
14. Boosted Linear Logistic Regression
15. K Star
16. KNN K1
17. KNN K2
18. KNN K4
19. KNN K8
20. KNN K16
21. KNN K32

22. KNN K64
23. KNN K128
24. KNN K256
25. J48
26. Decision Stump
27. REP Decision Tree
28. Hoeffding Tree
29. Random Tree
30. Random Forest T10
31. Random Forest T20
32. Random Forest T40
33. Random Forest T80
34. Random Forest T160
35. Boosted Random Tree
36. Boosted Decision Stump
37. Boosted REP Tree

10.1.4 Economic Costs

Table 34. Range of data examined in previous literature.

# of HACs (millions)	% Correctly identified HAC	Total correctly identified (millions)	Cost / HAC	Total (millions)	Total ($billions)	Average pivoted on cost / HAC ($billions)
0.7218	0.51	0.37	12,000	4417.416	$4.42	$25.05
0.7218	0.75	0.54	12,000	6496.2	$6.50	
0.7218	0.8	0.58	12,000	6929.28	$6.93	
1.5	0.51	0.77	12,000	9180	$9.18	
1.5	0.75	1.13	12,000	13500	$13.50	
1.5	0.8	1.20	12,000	14400	$14.40	
3.96	0.51	2.02	12,000	24235.2	$24.24	
3.96	0.75	2.97	12,000	35640	$35.64	
3.96	0.8	3.17	12,000	38016	$38.02	
5.98	0.51	3.05	12,000	36597.6	$36.60	
5.98	0.75	4.49	12,000	53820	$53.82	
5.98	0.8	4.78	12,000	57408	$57.41	
0.7218	0.51	0.37	13,000	4785.534	$4.79	$27.14
0.7218	0.75	0.54	13,000	7037.55	$7.04	

0.7218	0.8	0.58	13,000	7506.72	$7.51	
1.5	0.51	0.77	13,000	9945	$9.95	
1.5	0.75	1.13	13,000	14625	$14.63	
1.5	0.8	1.20	13,000	15600	$15.60	
3.96	0.51	2.02	13,000	26254.8	$26.25	
3.96	0.75	2.97	13,000	38610	$38.61	
3.96	0.8	3.17	13,000	41184	$41.18	
5.98	0.51	3.05	13,000	39647.4	$39.65	
5.98	0.75	4.49	13,000	58305	$58.31	
5.98	0.8	4.78	13,000	62192	$62.19	
0.7218	0.51	0.37	17,000	6258.006	$6.26	$35.49
0.7218	0.75	0.54	17,000	9202.95	$9.20	
0.7218	0.8	0.58	17,000	9816.48	$9.82	
1.5	0.51	0.77	17,000	13005	$13.01	
1.5	0.75	1.13	17,000	19125	$19.13	
1.5	0.8	1.20	17,000	20400	$20.40	
3.96	0.51	2.02	17,000	34333.2	$34.33	
3.96	0.75	2.97	17,000	50490	$50.49	
3.96	0.8	3.17	17,000	53856	$53.86	
5.98	0.51	3.05	17,000	51846.6	$51.85	
5.98	0.75	4.49	17,000	76245	$76.25	
5.98	0.8	4.78	17,000	81328	$81.33	
0.7218	0.51	0.37	24,700	9092.5146	$9.09	$51.57
0.7218	0.75	0.54	24,700	13371.345	$13.37	
0.7218	0.8	0.58	24,700	14262.768	$14.26	
1.5	0.51	0.77	24,700	18895.5	$18.90	
1.5	0.75	1.13	24,700	27787.5	$27.79	
1.5	0.8	1.20	24,700	29640	$29.64	
3.96	0.51	2.02	24,700	49884.12	$49.88	
3.96	0.75	2.97	24,700	73359	$73.36	
3.96	0.8	3.17	24,700	78249.6	$78.25	
5.98	0.51	3.05	24,700	75330.06	$75.33	

5.98	0.75	4.49	24,700	110779.5	$110.78	
5.98	0.8	4.78	24,700	118164.8	$118.16	
0.7218	0.51	0.37	38,656	14229.96941	$14.23	$80.71
0.7218	0.75	0.54	38,656	20926.4256	$20.93	
0.7218	0.8	0.58	38,656	22321.52064	$22.32	
1.5	0.51	0.77	38,656	29571.84	$29.57	
1.5	0.75	1.13	38,656	43488	$43.49	
1.5	0.8	1.20	38,656	46387.2	$46.39	
3.96	0.51	2.02	38,656	78069.6576	$78.07	
3.96	0.75	2.97	38,656	114808.32	$114.81	
3.96	0.8	3.17	38,656	122462.208	$122.46	
5.98	0.51	3.05	38,656	117893.0688	$117.89	
5.98	0.75	4.49	38,656	173372.16	$173.37	
5.98	0.8	4.78	38,656	184930.304	$184.93	
0.7218	0.51	0.37	50,000	18405.9	$18.41	$104.39
0.7218	0.75	0.54	50,000	27067.5	$27.07	
0.7218	0.8	0.58	50,000	28872	$28.87	
1.5	0.51	0.77	50,000	38250	$38.25	
1.5	0.75	1.13	50,000	56250	$56.25	
1.5	0.8	1.20	50,000	60000	$60.00	
3.96	0.51	2.02	50,000	100980	$100.98	
3.96	0.75	2.97	50,000	148500	$148.50	
3.96	0.8	3.17	50,000	158400	$158.40	
5.98	0.51	3.05	50,000	152490	$152.49	
5.98	0.75	4.49	50,000	224250	$224.25	
5.98	0.8	4.78	50,000	239200	$239.20	

10.1.5 Iteration 7: Principal Components Analysis

PRINCIPAL COMPONENTS ANALYSIS WITH BAYESIAN LOGISTIC REGRESSION

Scheme:weka.classifiers.bayes.BayesianLogisticRegression -D -Tl 5.0E-4 -S 0.5 -H 1 -V 0.27 -R R:0.01-316,3.16 -P 1 -F 2 -seed 1 -I 100 -N

Relation: Ethnicity_Casenotconverted.temporal.extract.14.2.first.1.hours.category.3.healthy.controls-weka.filters.unsupervised.attribute.Remove-R1-3_principal components-weka.filters.unsupervised.attribute.PrincipalComponents-R0.95-A5-M-1

Instances: 8798

Attributes: 45

0.14 piperacillin.tazobactam.IV.rx+0.14 bisacodyl.POPR.rx+0.14 levothyroxine.POIV.rx+0.14 ondansetron.POIV.rx+0.14 cefazolin.IV.rx...

0.146X50413.max+0.146X50413.median+0.146X50413.min+0.146X50412.max+0.146 X50412.median...

0.194PULSATION.BALLOON.IMPLAN..cpt+0.194INSERT.DRUG.ELUT.CORONA.cpt+0.194ENDOSCOPIC.BRONCHIAL.BX...cpt+0.194DX.ULTRASOUND.HEART.cpt+0.193CONTR.CEREBR.ARTERIOGRAM..cpt...

-0.235X50439.max-0.234X50439.median-0.234X50439.min-0.231X50399.max-0.231X50399.median...

0.241X50399.max+0.241X50399.median+0.241X50399.min+0.235X50439.max+0.235 X50439.median...

0.333X50177.max+0.333X50177.min+0.333X50177.median+0.323X50090.min+0.323X50090.median...

0.41 X50468.min+0.41 X50468.median+0.409X50468.max+0.228X50112.max+0.228X50112.median...

0.669ethnicity=WHITE-0.389ethnicity=UNKNOWN/NOT SPECIFIED-0.353ethnicity=BLACK/AFRICAN AMERICAN-0.171X50112.max-0.17X50112.median...

0.427X50112.max+0.426X50112.median+0.425X50112.min+0.291ethnicity=WHITE-0.236ethnicity=UNKNOWN/NOT SPECIFIED...

-0.407X50428.min-0.407X50428.median-0.407X50428.max+0.265X50468.max+0.264X50468.median...

0.429X50440.max+0.429X50440.median+0.427X50440.min+0.203X50428.max+0.203 X50428.median...

0.578age+0.449ethnicity=UNKNOWN/NOT SPECIFIED-0.314ethnicity=BLACK/AFRICAN AMERICAN-0.221ethnicity=HISPANIC OR LATINO-0.178gender...

-0.478gender+0.368ethnicity=BLACK/AFRICAN AMERICAN-0.367ethnicity=UNKNOWN/NOT SPECIFIED+0.25 elix.index+0.241ethnicity=UNABLE TO OBTAIN...

-0.629elix.index-0.282gender-0.269ethnicity=BLACK/CAPE VERDEAN+0.246ethnicity=HISPANIC OR LATINO+0.227ethnicity=WHITE - EASTERN EUROPEAN...

0.498ethnicity=ASIAN-0.422ethnicity=BLACK/AFRICAN AMERICAN+0.414ethnicity=OTHER+0.347ethnicity=HISPANIC OR LATINO-0.265ethnicity=UNKNOWN/NOT SPECIFIED...

0.692ethnicity=OTHER-0.412ethnicity=HISPANIC OR LATINO-0.194ethnicity=BLACK/CAPE VERDEAN-0.192ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.17 ethnicity=WHITE - EASTERN EUROPEAN...

-0.641ethnicity=ASIAN+0.4 ethnicity=HISPANIC OR LATINO+0.38 ethnicity=OTHER+0.25 ethnicity=WHITE - RUSSIAN-0.147ethnicity=BLACK/AFRICAN AMERICAN...

-0.646ethnicity=PATIENT DECLINED TO ANSWER-0.332ethnicity=WHITE - RUSSIAN-0.257ethnicity=UNABLE TO OBTAIN+0.256ethnicity=MULTI RACE ETHNICITY-0.21ethnicity=WHITE - OTHER EUROPEAN...

-0.451ethnicity=UNABLE TO OBTAIN+0.414ethnicity=MULTI RACE ETHNICITY+0.407ethnicity=PATIENT DECLINED TO ANSWER+0.291ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.253ethnicity=BLACK/HAITIAN...

0.378ethnicity=WHITE - OTHER EUROPEAN+0.353ethnicity=ASIAN+0.33 ethnicity=WHITE - RUSSIAN-0.318ethnicity=MIDDLE EASTERN-0.311ethnicity=PATIENT DECLINED TO ANSWER...

0.42 ethnicity=ASIAN - CHINESE+0.416ethnicity=AMERICAN INDIAN/ALASKA NATIVE+0.391ethnicity=PORTUGUESE+0.308ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.292ethnicity=WHITE - BRAZILIAN...

-0.376ethnicity=UNABLE TO OBTAIN-0.36ethnicity=WHITE - EASTERN EUROPEAN-0.318ethnicity=MULTI RACE ETHNICITY+0.298ethnicity=ASIAN - KOREAN+0.298ethnicity=PATIENT DECLINED TO ANSWER...

-0.421ethnicity=BLACK/HAITIAN+0.365ethnicity=ASIAN - KOREAN-0.362ethnicity=AMERICAN INDIAN/ALASKA NATIVE+0.275ethnicity=WHITE - RUSSIAN+0.263ethnicity=HISPANIC/LATINO - PUERTO RICAN...

-0.494ethnicity=ASIAN - ASIAN INDIAN-0.442ethnicity=SOUTH AMERICAN+0.279ethnicity=ASIAN - KOREAN+0.266ethnicity=BLACK/CAPE VERDEAN-0.254ethnicity=HISPANIC/LATINO - PUERTO RICAN...

0.424ethnicity=ASIAN - CHINESE+0.357ethnicity=SOUTH AMERICAN-0.297ethnicity=BLACK/HAITIAN-0.286ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.274ethnicity=ASIAN - KOREAN...

-0.553ethnicity=BLACK/CAPE VERDEAN+0.361ethnicity=WHITE - BRAZILIAN+0.349ethnicity=WHITE - OTHER EUROPEAN-0.27ethnicity=BLACK/HAITIAN+0.245ethnicity=ASIAN - ASIAN INDIAN...

0.721ethnicity=WHITE - BRAZILIAN-0.332ethnicity=PORTUGUESE+0.328ethnicity=BLACK/HAITIAN-0.224ethnicity=ASIAN - ASIAN INDIAN+0.192ethnicity=HISPANIC/LATINO - DOMINICAN...

-0.412ethnicity=HISPANIC/LATINO - GUATEMALAN-0.329ethnicity=SOUTH AMERICAN-0.316ethnicity=ASIAN - VIETNAMESE+0.314ethnicity=ASIAN - CHINESE-0.303ethnicity=ASIAN - KOREAN...

0.399ethnicity=ASIAN - OTHER-0.38ethnicity=HISPANIC/LATINO - CUBAN+0.335ethnicity=ASIAN - ASIAN INDIAN-0.299ethnicity=ASIAN - THAI-0.265ethnicity=SOUTH AMERICAN...

0.486ethnicity=ASIAN - OTHER+0.478ethnicity=ASIAN - FILIPINO+0.389ethnicity=HISPANIC/LATINO - CUBAN+0.257ethnicity=HISPANIC/LATINO - COLOMBIAN-0.211ethnicity=BLACK/AFRICAN...

0.518ethnicity=HISPANIC/LATINO - GUATEMALAN-0.47ethnicity=HISPANIC/LATINO - DOMINICAN-0.431ethnicity=ASIAN - VIETNAMESE+0.288ethnicity=HISPANIC/LATINO - COLOMBIAN-0.271ethnicity=PORTUGUESE...

-0.724ethnicity=ASIAN - THAI-0.336ethnicity=BLACK/AFRICAN-0.242ethnicity=BLACK/HAITIAN-0.215ethnicity=ASIAN - OTHER+0.205ethnicity=PORTUGUESE...

-0.573ethnicity=HISPANIC/LATINO - COLOMBIAN+0.483ethnicity=WHITE - EASTERN EUROPEAN+0.359ethnicity=BLACK/AFRICAN+0.237ethnicity=ASIAN - FILIPINO-0.206ethnicity=HISPANIC/LATINO - DOMINICAN...

-0.344ethnicity=ASIAN - ASIAN INDIAN+0.314ethnicity=ASIAN - VIETNAMESE+0.295ethnicity=HISPANIC/LATINO - PUERTO RICAN-0.287ethnicity=BLACK/HAITIAN-0.28ethnicity=HISPANIC/LATINO - COLOMBIAN...

-0.425ethnicity=HISPANIC/LATINO - CUBAN+0.404ethnicity=ASIAN - FILIPINO-0.396ethnicity=WHITE - OTHER EUROPEAN+0.282ethnicity=PORTUGUESE-0.277ethnicity=ASIAN - OTHER...

-0.45ethnicity=HISPANIC/LATINO - DOMINICAN+0.316ethnicity=SOUTH AMERICAN+0.29 ethnicity=PORTUGUESE+0.264ethnicity=CARIBBEAN ISLAND+0.263ethnicity=ASIAN - VIETNAMESE...

-0.429ethnicity=HISPANIC/LATINO - COLOMBIAN-0.428ethnicity=HISPANIC/LATINO - CUBAN-0.361ethnicity=BLACK/AFRICAN-0.282ethnicity=ASIAN - VIETNAMESE+0.271ethnicity=ASIAN - FILIPINO...

0.493ethnicity=HISPANIC/LATINO - GUATEMALAN+0.385ethnicity=WHITE - RUSSIAN-0.304ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.303ethnicity=PORTUGUESE-0.273ethnicity=ASIAN - KOREAN...

-0.432ethnicity=ASIAN - THAI+0.402ethnicity=BLACK/AFRICAN-0.401ethnicity=BLACK/CAPE VERDEAN-0.271ethnicity=ASIAN - ASIAN INDIAN+0.221ethnicity=UNABLE TO OBTAIN...

-0.436ethnicity=ASIAN - CHINESE-0.311ethnicity=WHITE - OTHER EUROPEAN-0.309ethnicity=ASIAN - FILIPINO-0.278ethnicity=WHITE - EASTERN EUROPEAN+0.27 ethnicity=CARIBBEAN ISLAND...

-0.617ethnicity=MIDDLE EASTERN-0.425ethnicity=AMERICAN INDIAN/ALASKA NATIVE+0.251ethnicity=HISPANIC/LATINO - GUATEMALAN+0.245ethnicity=CARIBBEAN ISLAND+0.216ethnicity=SOUTH AMERICAN...

-0.68ethnicity=CARIBBEAN ISLAND+0.334ethnicity=UNABLE TO OBTAIN-0.328ethnicity=HISPANIC/LATINO - DOMINICAN+0.187ethnicity=BLACK/CAPE VERDEAN-0.165ethnicity=ASIAN - CHINESE...

-0.416gender-0.374elix.index-0.317ethnicity=WHITE - RUSSIAN+0.228ethnicity=SOUTH AMERICAN+0.227ethnicity=PORTUGUESE...

-0.605elix.index-0.361ethnicity=HISPANIC OR LATINO+0.235ethnicity=WHITE - RUSSIAN+0.22 gender+0.215ethnicity=ASIAN - FILIPINO...

Case

Test mode:10-fold cross-validation

=== Classifier model (full training set) ===

Norm-Based Hyperparameter Selection: 0.2833023219069017

Regression Coefficients

========================

(intercept) : -0.0011095142925906154

0.14 piperacillin.tazobactam.IV.rx+0.14 bisacodyl.POPR.rx+0.14 levothyroxine.POIV.rx+0.14 ondansetron.POIV.rx+0.14 cefazolin.IV.rx... : 0.013933780815802332

0.146X50413.max+0.146X50413.median+0.146X50413.min+0.146X50412.max+0.146X50412.median... : 0.044422268904560885

0.194PULSATION.BALLOON.IMPLAN..cpt+0.194INSERT.DRUG.ELUT.CORONA.cpt+0.194ENDOSCOPIC.BRONCHIAL.BX...cpt+0.194DX.ULTRASOUND.HEART.cpt+0.193CONTR.CEREBR.ARTERIOGRAM..cpt... : 0.010430962594178273

-0.235X50439.max-0.234X50439.median-0.234X50439.min-0.231X50399.max-0.231X50399.median... : -0.04214852967703312

0.241X50399.max+0.241X50399.median+0.241X50399.min+0.235X50439.max+0.235X50439.median... : -0.0809924152154734

0.333X50177.max+0.333X50177.min+0.333X50177.median+0.323X50090.min+0.323X50090.median... : 0.02672733139832329

0.41 X50468.min+0.41 X50468.median+0.409X50468.max+0.228X50112.max+0.228X50112.median... : 0.03580010715494692

0.669ethnicity=WHITE-0.389ethnicity=UNKNOWN/NOT SPECIFIED-0.353ethnicity=BLACK/AFRICAN AMERICAN-0.171X50112.max-0.17X50112.median... : -0.061106883519354636

0.427X50112.max+0.426X50112.median+0.425X50112.min+0.291ethnicity=WHITE-0.236ethnicity=UNKNOWN/NOT SPECIFIED... : 0.05462958592373866

-0.407X50428.min-0.407X50428.median-0.407X50428.max+0.265X50468.max+0.264X50468.median... : 0.01990794002775642

0.429X50440.max+0.429X50440.median+0.427X50440.min+0.203X50428.max+0.203X50428.median... : -0.10758585459944073

0.578age+0.449ethnicity=UNKNOWN/NOT SPECIFIED-0.314ethnicity=BLACK/AFRICAN AMERICAN-0.221ethnicity=HISPANIC OR LATINO-0.178gender... : -0.018620421177229257

-0.478gender+0.368ethnicity=BLACK/AFRICAN AMERICAN-0.367ethnicity=UNKNOWN/NOT SPECIFIED+0.25 elix.index+0.241ethnicity=UNABLE TO OBTAIN... : 0.04740967191327524

-0.629elix.index-0.282gender-0.269ethnicity=BLACK/CAPE VERDEAN+0.246ethnicity=HISPANIC OR LATINO+0.227ethnicity=WHITE - EASTERN EUROPEAN... : 0.03371321311647879

0.498ethnicity=ASIAN-0.422ethnicity=BLACK/AFRICAN AMERICAN+0.414ethnicity=OTHER+0.347ethnicity=HISPANIC OR LATINO-0.265ethnicity=UNKNOWN/NOT SPECIFIED... : -0.023987556228045947

0.692ethnicity=OTHER-0.412ethnicity=HISPANIC OR LATINO-0.194ethnicity=BLACK/CAPE VERDEAN-0.192ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.17 ethnicity=WHITE - EASTERN EUROPEAN... : 0.04163360516115232

-0.641ethnicity=ASIAN+0.4 ethnicity=HISPANIC OR LATINO+0.38 ethnicity=OTHER+0.25 ethnicity=WHITE - RUSSIAN-0.147ethnicity=BLACK/AFRICAN AMERICAN... : -0.007432928634646981

-0.646ethnicity=PATIENT DECLINED TO ANSWER-0.332ethnicity=WHITE - RUSSIAN-0.257ethnicity=UNABLE TO OBTAIN+0.256ethnicity=MULTI RACE ETHNICITY-0.21ethnicity=WHITE - OTHER EUROPEAN... : 0.003718208731237804

-0.451ethnicity=UNABLE TO OBTAIN+0.414ethnicity=MULTI RACE ETHNICITY+0.407ethnicity=PATIENT DECLINED TO ANSWER+0.291ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.253ethnicity=BLACK/HAITIAN... : -0.02835750161057092

0.378ethnicity=WHITE - OTHER EUROPEAN+0.353ethnicity=ASIAN+0.33 ethnicity=WHITE - RUSSIAN-0.318ethnicity=MIDDLE EASTERN-0.311ethnicity=PATIENT DECLINED TO ANSWER... : 0.035380172001200985

0.42 ethnicity=ASIAN - CHINESE+0.416ethnicity=AMERICAN INDIAN/ALASKA NATIVE+0.391ethnicity=PORTUGUESE+0.308ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.292ethnicity=WHITE - BRAZILIAN... : -0.04202107000233811

-0.376ethnicity=UNABLE TO OBTAIN-0.36ethnicity=WHITE - EASTERN EUROPEAN-0.318ethnicity=MULTI RACE ETHNICITY+0.298ethnicity=ASIAN - KOREAN+0.298ethnicity=PATIENT DECLINED TO ANSWER... : 0.017011627463149555

-0.421ethnicity=BLACK/HAITIAN+0.365ethnicity=ASIAN - KOREAN-0.362ethnicity=AMERICAN INDIAN/ALASKA NATIVE+0.275ethnicity=WHITE - RUSSIAN+0.263ethnicity=HISPANIC/LATINO - PUERTO RICAN... : 0.03676055340528281

-0.494ethnicity=ASIAN - ASIAN INDIAN-0.442ethnicity=SOUTH AMERICAN+0.279ethnicity=ASIAN - KOREAN+0.266ethnicity=BLACK/CAPE VERDEAN-0.254ethnicity=HISPANIC/LATINO - PUERTO RICAN... : -4.019823730749742E-4

0.424ethnicity=ASIAN - CHINESE+0.357ethnicity=SOUTH AMERICAN-0.297ethnicity=BLACK/HAITIAN-0.286ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.274ethnicity=ASIAN - KOREAN... : 0.05747196845414819

-0.553ethnicity=BLACK/CAPE VERDEAN+0.361ethnicity=WHITE - BRAZILIAN+0.349ethnicity=WHITE - OTHER EUROPEAN-

0.27ethnicity=BLACK/HAITIAN+0.245ethnicity=ASIAN - ASIAN INDIAN... : -0.04305616209738552

0.721ethnicity=WHITE - BRAZILIAN-0.332ethnicity=PORTUGUESE+0.328ethnicity=BLACK/HAITIAN-0.224ethnicity=ASIAN - ASIAN INDIAN+0.192ethnicity=HISPANIC/LATINO - DOMINICAN... : 0.010385613078950917

-0.412ethnicity=HISPANIC/LATINO - GUATEMALAN-0.329ethnicity=SOUTH AMERICAN-0.316ethnicity=ASIAN - VIETNAMESE+0.314ethnicity=ASIAN - CHINESE-0.303ethnicity=ASIAN - KOREAN... : -0.01688114010410964

0.399ethnicity=ASIAN - OTHER-0.38ethnicity=HISPANIC/LATINO - CUBAN+0.335ethnicity=ASIAN - ASIAN INDIAN-0.299ethnicity=ASIAN - THAI-0.265ethnicity=SOUTH AMERICAN... : -0.04384466581306348

0.486ethnicity=ASIAN - OTHER+0.478ethnicity=ASIAN - FILIPINO+0.389ethnicity=HISPANIC/LATINO - CUBAN+0.257ethnicity=HISPANIC/LATINO - COLOMBIAN-0.211ethnicity=BLACK/AFRICAN... : -0.05725465983699726

0.518ethnicity=HISPANIC/LATINO - GUATEMALAN-0.47ethnicity=HISPANIC/LATINO - DOMINICAN-0.431ethnicity=ASIAN - VIETNAMESE+0.288ethnicity=HISPANIC/LATINO - COLOMBIAN-0.271ethnicity=PORTUGUESE... : -0.008459864215016133

-0.724ethnicity=ASIAN - THAI-0.336ethnicity=BLACK/AFRICAN-0.242ethnicity=BLACK/HAITIAN-0.215ethnicity=ASIAN - OTHER+0.205ethnicity=PORTUGUESE... : -0.03921811985789975

-0.573ethnicity=HISPANIC/LATINO - COLOMBIAN+0.483ethnicity=WHITE - EASTERN EUROPEAN+0.359ethnicity=BLACK/AFRICAN+0.237ethnicity=ASIAN - FILIPINO-0.206ethnicity=HISPANIC/LATINO - DOMINICAN... : -0.012779197725797286

-0.344ethnicity=ASIAN - ASIAN INDIAN+0.314ethnicity=ASIAN - VIETNAMESE+0.295ethnicity=HISPANIC/LATINO - PUERTO RICAN-0.287ethnicity=BLACK/HAITIAN-0.28ethnicity=HISPANIC/LATINO - COLOMBIAN... : 0.021736415165302746

-0.425ethnicity=HISPANIC/LATINO - CUBAN+0.404ethnicity=ASIAN - FILIPINO-0.396ethnicity=WHITE - OTHER EUROPEAN+0.282ethnicity=PORTUGUESE-0.277ethnicity=ASIAN - OTHER... : 0.015675715158805715

-0.45ethnicity=HISPANIC/LATINO - DOMINICAN+0.316ethnicity=SOUTH AMERICAN+0.29 ethnicity=PORTUGUESE+0.264ethnicity=CARIBBEAN ISLAND+0.263ethnicity=ASIAN - VIETNAMESE... : 0.040674237178365026

-0.429ethnicity=HISPANIC/LATINO - COLOMBIAN-0.428ethnicity=HISPANIC/LATINO - CUBAN-0.361ethnicity=BLACK/AFRICAN-

0.282ethnicity=ASIAN - VIETNAMESE+0.271ethnicity=ASIAN - FILIPINO... : 0.039462566524274004

0.493ethnicity=HISPANIC/LATINO - GUATEMALAN+0.385ethnicity=WHITE - RUSSIAN-0.304ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.303ethnicity=PORTUGUESE-0.273ethnicity=ASIAN - KOREAN... : 0.024529853298398

-0.432ethnicity=ASIAN - THAI+0.402ethnicity=BLACK/AFRICAN-0.401ethnicity=BLACK/CAPE VERDEAN-0.271ethnicity=ASIAN - ASIAN INDIAN+0.221ethnicity=UNABLE TO OBTAIN... : -0.004281595131810247

-0.436ethnicity=ASIAN - CHINESE-0.311ethnicity=WHITE - OTHER EUROPEAN-0.309ethnicity=ASIAN - FILIPINO-0.278ethnicity=WHITE - EASTERN EUROPEAN+0.27 ethnicity=CARIBBEAN ISLAND... : -0.011011134577230243

-0.617ethnicity=MIDDLE EASTERN-0.425ethnicity=AMERICAN INDIAN/ALASKA NATIVE+0.251ethnicity=HISPANIC/LATINO - GUATEMALAN+0.245ethnicity=CARIBBEAN ISLAND+0.216ethnicity=SOUTH AMERICAN... : 0.07845506996782464

-0.68ethnicity=CARIBBEAN ISLAND+0.334ethnicity=UNABLE TO OBTAIN-0.328ethnicity=HISPANIC/LATINO - DOMINICAN+0.187ethnicity=BLACK/CAPE VERDEAN-0.165ethnicity=ASIAN - CHINESE... : -0.07463394757524655

-0.416gender-0.374elix.index-0.317ethnicity=WHITE - RUSSIAN+0.228ethnicity=SOUTH AMERICAN+0.227ethnicity=PORTUGUESE... : 0.03454661184851193

-0.605elix.index-0.361ethnicity=HISPANIC OR LATINO+0.235ethnicity=WHITE - RUSSIAN+0.22 gender+0.215ethnicity=ASIAN - FILIPINO... : 0.025684788424796795

==========================

Likelihood: -15260.313881194275

Penalty: -13.392610353755153

Regularized Log Posterior: -15273.706491548031

==========================

Time taken to build model: 0.58 seconds

=== Stratified cross-validation ===

```
=== Summary ===

Correctly Classified Instances        5084               57.7859 %
Incorrectly Classified Instances      3714               42.2141 %
Kappa statistic                          0.1557
Mean absolute error                      0.4221
Root mean squared error                  0.6497
Relative absolute error                 84.4283 %
Root relative squared error            129.9448 %
Total Number of Instances             8798

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0.496     0.341      0.593     0.496     0.54       0.578    Y
                 0.659     0.504      0.567     0.659     0.61       0.578    N
Weighted Avg.    0.578     0.422      0.58      0.578     0.575      0.578

=== Confusion Matrix ===

    a    b   <-- classified as
 2184 2215 |    a = Y
 1499 2900 |    b = N
```

NAÏVE BAYES WITH PRINCIPAL COMPONENTS ANALYSIS

Scheme:weka.classifiers.bayes.NaiveBayes

Relation: Ethnicity_Casenotconverted.temporal.extract.14.2.first.1.hours.category.3.healthy.controls-weka.filters.unsupervised.attribute.Remove-R1-3_principal components-weka.filters.unsupervised.attribute.PrincipalComponents-R0.95-A5-M-1

Instances: 8798

Attributes: 45

0.14 piperacillin.tazobactam.IV.rx+0.14 bisacodyl.POPR.rx+0.14 levothyroxine.POIV.rx+0.14 ondansetron.POIV.rx+0.14 cefazolin.IV.rx...

0.146X50413.max+0.146X50413.median+0.146X50413.min+0.146X50412.max+0.146 X50412.median...

0.194PULSATION.BALLOON.IMPLAN..cpt+0.194INSERT.DRUG.ELUT.CORONA.cpt+0.194ENDOSCOPIC.BRONCHIAL.BX...cpt+0.194DX.ULTRASOUND.HEART.cpt+0.193CONTR.CEREBR.ARTERIOGRAM..cpt...

-0.235X50439.max-0.234X50439.median-0.234X50439.min-0.231X50399.max-0.231X50399.median...

0.241X50399.max+0.241X50399.median+0.241X50399.min+0.235X50439.max+0.235 X50439.median...

0.333X50177.max+0.333X50177.min+0.333X50177.median+0.323X50090.min+0.323X50090.median...

0.41 X50468.min+0.41 X50468.median+0.409X50468.max+0.228X50112.max+0.228X50112.median...

0.669ethnicity=WHITE-0.389ethnicity=UNKNOWN/NOT SPECIFIED-0.353ethnicity=BLACK/AFRICAN AMERICAN-0.171X50112.max-0.17X50112.median...

0.427X50112.max+0.426X50112.median+0.425X50112.min+0.291ethnicity=WHITE-0.236ethnicity=UNKNOWN/NOT SPECIFIED...

-0.407X50428.min-0.407X50428.median-0.407X50428.max+0.265X50468.max+0.264X50468.median...

0.429X50440.max+0.429X50440.median+0.427X50440.min+0.203X50428.max+0.203 X50428.median...

0.578age+0.449ethnicity=UNKNOWN/NOT SPECIFIED-0.314ethnicity=BLACK/AFRICAN AMERICAN-0.221ethnicity=HISPANIC OR LATINO-0.178gender...

-0.478gender+0.368ethnicity=BLACK/AFRICAN AMERICAN-0.367ethnicity=UNKNOWN/NOT SPECIFIED+0.25 elix.index+0.241ethnicity=UNABLE TO OBTAIN...

-0.629elix.index-0.282gender-0.269ethnicity=BLACK/CAPE VERDEAN+0.246ethnicity=HISPANIC OR LATINO+0.227ethnicity=WHITE - EASTERN EUROPEAN...

0.498ethnicity=ASIAN-0.422ethnicity=BLACK/AFRICAN AMERICAN+0.414ethnicity=OTHER+0.347ethnicity=HISPANIC OR LATINO-0.265ethnicity=UNKNOWN/NOT SPECIFIED...

0.692ethnicity=OTHER-0.412ethnicity=HISPANIC OR LATINO-0.194ethnicity=BLACK/CAPE VERDEAN-0.192ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.17 ethnicity=WHITE - EASTERN EUROPEAN...

-0.641ethnicity=ASIAN+0.4 ethnicity=HISPANIC OR LATINO+0.38 ethnicity=OTHER+0.25 ethnicity=WHITE - RUSSIAN-0.147ethnicity=BLACK/AFRICAN AMERICAN...

-0.646ethnicity=PATIENT DECLINED TO ANSWER-0.332ethnicity=WHITE - RUSSIAN-0.257ethnicity=UNABLE TO OBTAIN+0.256ethnicity=MULTI RACE ETHNICITY-0.21ethnicity=WHITE - OTHER EUROPEAN...

-0.451ethnicity=UNABLE TO OBTAIN+0.414ethnicity=MULTI RACE ETHNICITY+0.407ethnicity=PATIENT DECLINED TO ANSWER+0.291ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.253ethnicity=BLACK/HAITIAN...

0.378ethnicity=WHITE - OTHER EUROPEAN+0.353ethnicity=ASIAN+0.33 ethnicity=WHITE - RUSSIAN-0.318ethnicity=MIDDLE EASTERN-0.311ethnicity=PATIENT DECLINED TO ANSWER...

0.42 ethnicity=ASIAN - CHINESE+0.416ethnicity=AMERICAN INDIAN/ALASKA NATIVE+0.391ethnicity=PORTUGUESE+0.308ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.292ethnicity=WHITE - BRAZILIAN...

-0.376ethnicity=UNABLE TO OBTAIN-0.36ethnicity=WHITE - EASTERN EUROPEAN-0.318ethnicity=MULTI RACE ETHNICITY+0.298ethnicity=ASIAN - KOREAN+0.298ethnicity=PATIENT DECLINED TO ANSWER...

-0.421ethnicity=BLACK/HAITIAN+0.365ethnicity=ASIAN - KOREAN-0.362ethnicity=AMERICAN INDIAN/ALASKA NATIVE+0.275ethnicity=WHITE - RUSSIAN+0.263ethnicity=HISPANIC/LATINO - PUERTO RICAN...

-0.494ethnicity=ASIAN - ASIAN INDIAN-0.442ethnicity=SOUTH AMERICAN+0.279ethnicity=ASIAN - KOREAN+0.266ethnicity=BLACK/CAPE VERDEAN-0.254ethnicity=HISPANIC/LATINO - PUERTO RICAN...

0.424ethnicity=ASIAN - CHINESE+0.357ethnicity=SOUTH AMERICAN-0.297ethnicity=BLACK/HAITIAN-0.286ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.274ethnicity=ASIAN - KOREAN...

-0.553ethnicity=BLACK/CAPE VERDEAN+0.361ethnicity=WHITE - BRAZILIAN+0.349ethnicity=WHITE - OTHER EUROPEAN-0.27ethnicity=BLACK/HAITIAN+0.245ethnicity=ASIAN - ASIAN INDIAN...

0.721ethnicity=WHITE - BRAZILIAN-0.332ethnicity=PORTUGUESE+0.328ethnicity=BLACK/HAITIAN-0.224ethnicity=ASIAN - ASIAN INDIAN+0.192ethnicity=HISPANIC/LATINO - DOMINICAN...

-0.412ethnicity=HISPANIC/LATINO - GUATEMALAN-0.329ethnicity=SOUTH AMERICAN-0.316ethnicity=ASIAN - VIETNAMESE+0.314ethnicity=ASIAN - CHINESE-0.303ethnicity=ASIAN - KOREAN...

0.399ethnicity=ASIAN - OTHER-0.38ethnicity=HISPANIC/LATINO - CUBAN+0.335ethnicity=ASIAN - ASIAN INDIAN-0.299ethnicity=ASIAN - THAI-0.265ethnicity=SOUTH AMERICAN...

0.486ethnicity=ASIAN - OTHER+0.478ethnicity=ASIAN - FILIPINO+0.389ethnicity=HISPANIC/LATINO - CUBAN+0.257ethnicity=HISPANIC/LATINO - COLOMBIAN-0.211ethnicity=BLACK/AFRICAN...

0.518ethnicity=HISPANIC/LATINO - GUATEMALAN-0.47ethnicity=HISPANIC/LATINO - DOMINICAN-0.431ethnicity=ASIAN - VIETNAMESE+0.288ethnicity=HISPANIC/LATINO - COLOMBIAN-0.271ethnicity=PORTUGUESE...

-0.724ethnicity=ASIAN - THAI-0.336ethnicity=BLACK/AFRICAN-0.242ethnicity=BLACK/HAITIAN-0.215ethnicity=ASIAN - OTHER+0.205ethnicity=PORTUGUESE...

-0.573ethnicity=HISPANIC/LATINO - COLOMBIAN+0.483ethnicity=WHITE - EASTERN EUROPEAN+0.359ethnicity=BLACK/AFRICAN+0.237ethnicity=ASIAN - FILIPINO-0.206ethnicity=HISPANIC/LATINO - DOMINICAN...

-0.344ethnicity=ASIAN - ASIAN INDIAN+0.314ethnicity=ASIAN - VIETNAMESE+0.295ethnicity=HISPANIC/LATINO - PUERTO RICAN-0.287ethnicity=BLACK/HAITIAN-0.28ethnicity=HISPANIC/LATINO - COLOMBIAN...

-0.425ethnicity=HISPANIC/LATINO - CUBAN+0.404ethnicity=ASIAN - FILIPINO-0.396ethnicity=WHITE - OTHER EUROPEAN+0.282ethnicity=PORTUGUESE-0.277ethnicity=ASIAN - OTHER...

-0.45ethnicity=HISPANIC/LATINO - DOMINICAN+0.316ethnicity=SOUTH AMERICAN+0.29 ethnicity=PORTUGUESE+0.264ethnicity=CARIBBEAN ISLAND+0.263ethnicity=ASIAN - VIETNAMESE...

-0.429ethnicity=HISPANIC/LATINO - COLOMBIAN-0.428ethnicity=HISPANIC/LATINO - CUBAN-0.361ethnicity=BLACK/AFRICAN-0.282ethnicity=ASIAN - VIETNAMESE+0.271ethnicity=ASIAN - FILIPINO...

0.493ethnicity=HISPANIC/LATINO - GUATEMALAN+0.385ethnicity=WHITE - RUSSIAN-0.304ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.303ethnicity=PORTUGUESE-0.273ethnicity=ASIAN - KOREAN...

-0.432ethnicity=ASIAN - THAI+0.402ethnicity=BLACK/AFRICAN-0.401ethnicity=BLACK/CAPE VERDEAN-0.271ethnicity=ASIAN - ASIAN INDIAN+0.221ethnicity=UNABLE TO OBTAIN...

-0.436ethnicity=ASIAN - CHINESE-0.311ethnicity=WHITE - OTHER EUROPEAN-0.309ethnicity=ASIAN - FILIPINO-0.278ethnicity=WHITE - EASTERN EUROPEAN+0.27 ethnicity=CARIBBEAN ISLAND...

-0.617ethnicity=MIDDLE EASTERN-0.425ethnicity=AMERICAN INDIAN/ALASKA NATIVE+0.251ethnicity=HISPANIC/LATINO - GUATEMALAN+0.245ethnicity=CARIBBEAN ISLAND+0.216ethnicity=SOUTH AMERICAN...

-0.68ethnicity=CARIBBEAN ISLAND+0.334ethnicity=UNABLE TO OBTAIN-0.328ethnicity=HISPANIC/LATINO - DOMINICAN+0.187ethnicity=BLACK/CAPE VERDEAN-0.165ethnicity=ASIAN - CHINESE...

-0.416gender-0.374elix.index-0.317ethnicity=WHITE - RUSSIAN+0.228ethnicity=SOUTH AMERICAN+0.227ethnicity=PORTUGUESE...

-0.605elix.index-0.361ethnicity=HISPANIC OR LATINO+0.235ethnicity=WHITE - RUSSIAN+0.22 gender+0.215ethnicity=ASIAN - FILIPINO...

Case

Test mode:10-fold cross-validation

=== Classifier model (full training set) ===

Naive Bayes Classifier

Class

Attribute

Y N

(0.5) (0.5)

===

0.14 piperacillin.tazobactam.IV.rx+0.14 bisacodyl.POPR.rx+0.14 levothyroxine.POIV.rx+0.14 ondansetron.POIV.rx+0.14 cefazolin.IV.rx...

mean
-0.2969 0.2969
std. dev.
6.8424 6.5321
weight sum
4399 4399
precision
0.0033 0.0033

0.146X50413.max+0.146X50413.median+0.146X50413.min+0.146X50412.max+0.146X50412.median...
mean
-0.7553 0.7553
std. dev.
6.2311 5.5237
weight sum
4399 4399
precision
0.0034 0.0034

0.194PULSATION.BALLOON.IMPLAN..cpt+0.194INSERT.DRUG.ELUT.CORONA.cpt+0.194ENDOSCOPIC.BRONCHIAL.BX...cpt+0.194DX.ULTRASOUND.HEART.cpt+0.193CONTR.CEREBR.ARTERIOGRAM..cpt...
mean
-0.1217 0.1216
std. dev.
4.2995 5.6941
weight sum
4399 4399
precision
0.0055 0.0055

-0.235X50439.max-0.234X50439.median-0.234X50439.min-0.231X50399.max-0.231X50399.median...
mean
0.1143 -0.1143
std. dev.
2.21 2.5533

weight sum
4399 4399
precision
0.0033 0.0033

0.241X50399.max+0.241X50399.median+0.241X50399.min+0.235X50439.max+0.235X50439.median...
mean
0.1927 -0.1927
std. dev.
2.157 2.317
weight sum
4399 4399
precision
0.0037 0.0037

0.333X50177.max+0.333X50177.min+0.333X50177.median+0.323X50090.min+0.323X50090.median...
mean
-0.0376 0.0377
std. dev.
1.7126 1.843
weight sum
4399 4399
precision
0.0028 0.0028

0.41 X50468.min+0.41 X50468.median+0.409X50468.max+0.228X50112.max+0.228X50112.median...
mean
-0.0311 0.0311
std. dev.
1.4577 1.5465
weight sum
4399 4399
precision
0.0046 0.0046

0.669ethnicity=WHITE-0.389ethnicity=UNKNOWN/NOT SPECIFIED-0.353ethnicity=BLACK/AFRICAN AMERICAN-0.171X50112.max-0.17X50112.median...

mean
0.0544 -0.0544

std. dev.
1.2965 1.4038

weight sum
4399 4399

precision
0.0018 0.0018

0.427X50112.max+0.426X50112.median+0.425X50112.min+0.291ethnicity=WHITE-0.236ethnicity=UNKNOWN/NOT SPECIFIED...

mean
-0.0433 0.0433

std. dev.
1.223 1.3401

weight sum
4399 4399

precision
0.0036 0.0036

-0.407X50428.min-0.407X50428.median-0.407X50428.max+0.265X50468.max+0.264X50468.median...

mean
-0.0086 0.0086

std. dev.
1.1988 1.2292

weight sum
4399 4399

precision
0.0043 0.0043

0.429X50440.max+0.429X50440.median+0.427X50440.min+0.203X50428.max+0.203X50428.median...

mean
0.0665 -0.0666

std. dev.
1.2044 1.0857
weight sum
4399 4399
precision
0.0026 0.0026

0.578age+0.449ethnicity=UNKNOWN/NOT SPECIFIED-0.314ethnicity=BLACK/AFRICAN AMERICAN-0.221ethnicity=HISPANIC OR LATINO-0.178gender...
mean
0.0105 -0.0105
std. dev.
1.074 1.1595
weight sum
4399 4399
precision
0.0019 0.0019

-0.478gender+0.368ethnicity=BLACK/AFRICAN AMERICAN-0.367ethnicity=UNKNOWN/NOT SPECIFIED+0.25 elix.index+0.241ethnicity=UNABLE TO OBTAIN...
mean
-0.0244 0.0244
std. dev.
1.009 1.079
weight sum
4399 4399
precision
0.002 0.002

-0.629elix.index-0.282gender-0.269ethnicity=BLACK/CAPE VERDEAN+0.246ethnicity=HISPANIC OR LATINO+0.227ethnicity=WHITE - EASTERN EUROPEAN...
mean
-0.0196 0.0196
std. dev.
1.0562 1.0027

weight sum
4399 4399
precision
0.0043 0.0043

0.498ethnicity=ASIAN-0.422ethnicity=BLACK/AFRICAN AMERICAN+0.414ethnicity=OTHER+0.347ethnicity=HISPANIC OR LATINO-0.265ethnicity=UNKNOWN/NOT SPECIFIED...
mean
0.0103 -0.0103
std. dev.
0.9887 1.0484
weight sum
4399 4399
precision
0.0016 0.0016

0.692ethnicity=OTHER-0.412ethnicity=HISPANIC OR LATINO-0.194ethnicity=BLACK/CAPE VERDEAN-0.192ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.17 ethnicity=WHITE - EASTERN EUROPEAN...
mean
-0.021 0.021
std. dev.
0.9891 1.028
weight sum
4399 4399
precision
0.003 0.003

-0.641ethnicity=ASIAN+0.4 ethnicity=HISPANIC OR LATINO+0.38 ethnicity=OTHER+0.25 ethnicity=WHITE - RUSSIAN-0.147ethnicity=BLACK/AFRICAN AMERICAN...
mean
0.0017 -0.0017
std. dev.
0.9412 1.0694
weight sum
4399 4399

precision
0.0025 0.0025

-0.646ethnicity=PATIENT DECLINED TO ANSWER-0.332ethnicity=WHITE - RUSSIAN-0.257ethnicity=UNABLE TO OBTAIN+0.256ethnicity=MULTI RACE ETHNICITY-0.21ethnicity=WHITE - OTHER EUROPEAN...

mean
-0.0053 0.0052
std. dev.
1.0484 0.959
weight sum
4399 4399
precision
0.0025 0.0025

-0.451ethnicity=UNABLE TO OBTAIN+0.414ethnicity=MULTI RACE ETHNICITY+0.407ethnicity=PATIENT DECLINED TO ANSWER+0.291ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.253ethnicity=BLACK/HAITIAN...

mean
0.0152 -0.0152
std. dev.
1.0429 0.9623
weight sum
4399 4399
precision
0.0036 0.0036

0.378ethnicity=WHITE - OTHER EUROPEAN+0.353ethnicity=ASIAN+0.33 ethnicity=WHITE - RUSSIAN-0.318ethnicity=MIDDLE EASTERN-0.311ethnicity=PATIENT DECLINED TO ANSWER...

mean
-0.0156 0.0157
std. dev.
1.0534 0.9501
weight sum
4399 4399

precision
0.0051 0.0051

0.42 ethnicity=ASIAN - CHINESE+0.416ethnicity=AMERICAN INDIAN/ALASKA NATIVE+0.391ethnicity=PORTUGUESE+0.308ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.292ethnicity=WHITE - BRAZILIAN...
mean
0.0147 -0.0147
std. dev.
1.0853 0.9107
weight sum
4399 4399
precision
0.0043 0.0043

-0.376ethnicity=UNABLE TO OBTAIN-0.36ethnicity=WHITE - EASTERN EUROPEAN-0.318ethnicity=MULTI RACE ETHNICITY+0.298ethnicity=ASIAN - KOREAN+0.298ethnicity=PATIENT DECLINED TO ANSWER...
mean
-0.0087 0.0087
std. dev.
1.0554 0.9447
weight sum
4399 4399
precision
0.0059 0.0059

-0.421ethnicity=BLACK/HAITIAN+0.365ethnicity=ASIAN - KOREAN-0.362ethnicity=AMERICAN INDIAN/ALASKA NATIVE+0.275ethnicity=WHITE - RUSSIAN+0.263ethnicity=HISPANIC/LATINO - PUERTO RICAN...
mean
-0.0161 0.0161
std. dev.
1.0812 0.914
weight sum
4399 4399
precision
0.0066 0.0066

-0.494ethnicity=ASIAN - ASIAN INDIAN-0.442ethnicity=SOUTH AMERICAN+0.279ethnicity=ASIAN - KOREAN+0.266ethnicity=BLACK/CAPE VERDEAN-0.254ethnicity=HISPANIC/LATINO - PUERTO RICAN...

mean
-0.0004 0.0004

std. dev.
0.9048 1.0887

weight sum
4399 4399

precision
0.0063 0.0063

0.424ethnicity=ASIAN - CHINESE+0.357ethnicity=SOUTH AMERICAN-0.297ethnicity=BLACK/HAITIAN-0.286ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.274ethnicity=ASIAN - KOREAN...

mean
-0.0198 0.0196

std. dev.
1.0141 0.9869

weight sum
4399 4399

precision
0.0049 0.0049

-0.553ethnicity=BLACK/CAPE VERDEAN+0.361ethnicity=WHITE - BRAZILIAN+0.349ethnicity=WHITE - OTHER EUROPEAN-0.27ethnicity=BLACK/HAITIAN+0.245ethnicity=ASIAN - ASIAN INDIAN...

mean
0.0216 -0.0216

std. dev.
1.0923 0.8993

weight sum
4399 4399

precision
0.0044 0.0044

0.721ethnicity=WHITE - BRAZILIAN-0.332ethnicity=PORTUGUESE+0.328ethnicity=BLACK/HAITIAN-0.224ethnicity=ASIAN - ASIAN INDIAN+0.192ethnicity=HISPANIC/LATINO - DOMINICAN...

mean
-0.0022 0.0022

std. dev.
1.1193 0.8656

weight sum
4399 4399

precision
0.0044 0.0044

-0.412ethnicity=HISPANIC/LATINO - GUATEMALAN-0.329ethnicity=SOUTH AMERICAN-0.316ethnicity=ASIAN - VIETNAMESE+0.314ethnicity=ASIAN - CHINESE-0.303ethnicity=ASIAN - KOREAN...

mean
0.0023 -0.0023

std. dev.
0.8685 1.1167

weight sum
4399 4399

precision
0.0051 0.0051

0.399ethnicity=ASIAN - OTHER-0.38ethnicity=HISPANIC/LATINO - CUBAN+0.335ethnicity=ASIAN - ASIAN INDIAN-0.299ethnicity=ASIAN - THAI-0.265ethnicity=SOUTH AMERICAN...

mean
0.0137 -0.0136

std. dev.
1.0933 0.8977

weight sum
4399 4399

precision
0.0074 0.0074

0.486ethnicity=ASIAN - OTHER+0.478ethnicity=ASIAN - FILIPINO+0.389ethnicity=HISPANIC/LATINO -

CUBAN+0.257ethnicity=HISPANIC/LATINO - COLOMBIAN-0.211ethnicity=BLACK/AFRICAN...

mean

0.0161 -0.0161

std. dev.

1.2818 0.5978

weight sum

4399 4399

precision

0.0074 0.0074

0.518ethnicity=HISPANIC/LATINO - GUATEMALAN-0.47ethnicity=HISPANIC/LATINO - DOMINICAN-0.431ethnicity=ASIAN - VIETNAMESE+0.288ethnicity=HISPANIC/LATINO - COLOMBIAN-0.271ethnicity=PORTUGUESE...

mean

0.005 -0.005

std. dev.

0.8523 1.1288

weight sum

4399 4399

precision

0.0064 0.0064

-0.724ethnicity=ASIAN - THAI-0.336ethnicity=BLACK/AFRICAN-0.242ethnicity=BLACK/HAITIAN-0.215ethnicity=ASIAN - OTHER+0.205ethnicity=PORTUGUESE...

mean

0.0117 -0.0117

std. dev.

0.7091 1.2236

weight sum

4399 4399

precision

0.009 0.009

-0.573ethnicity=HISPANIC/LATINO - COLOMBIAN+0.483ethnicity=WHITE - EASTERN EUROPEAN+0.359ethnicity=BLACK/AFRICAN+0.237ethnicity=ASIAN - FILIPINO-0.206ethnicity=HISPANIC/LATINO - DOMINICAN...

mean
0.0054 -0.0055

std. dev.
1.1999 0.7485

weight sum
4399 4399

precision
0.0099 0.0099

-0.344ethnicity=ASIAN - ASIAN INDIAN+0.314ethnicity=ASIAN - VIETNAMESE+0.295ethnicity=HISPANIC/LATINO - PUERTO RICAN-0.287ethnicity=BLACK/HAITIAN-0.28ethnicity=HISPANIC/LATINO - COLOMBIAN...

mean
-0.0103 0.0104

std. dev.
1.1061 0.8809

weight sum
4399 4399

precision
0.0056 0.0056

-0.425ethnicity=HISPANIC/LATINO - CUBAN+0.404ethnicity=ASIAN - FILIPINO-0.396ethnicity=WHITE - OTHER EUROPEAN+0.282ethnicity=PORTUGUESE-0.277ethnicity=ASIAN - OTHER...

mean
-0.0055 0.0054

std. dev.
1.237 0.685

weight sum
4399 4399

precision
0.0075 0.0075

-0.45ethnicity=HISPANIC/LATINO - DOMINICAN+0.316ethnicity=SOUTH AMERICAN+0.29 ethnicity=PORTUGUESE+0.264ethnicity=CARIBBEAN ISLAND+0.263ethnicity=ASIAN - VIETNAMESE...

mean
-0.0123 0.0123

std. dev.
0.9636 1.0342

weight sum
4399 4399

precision
0.0048 0.0048

-0.429ethnicity=HISPANIC/LATINO - COLOMBIAN-0.428ethnicity=HISPANIC/LATINO - CUBAN-0.361ethnicity=BLACK/AFRICAN-0.282ethnicity=ASIAN - VIETNAMESE+0.271ethnicity=ASIAN - FILIPINO...

mean
-0.0103 0.0103

std. dev.
1.1989 0.7485

weight sum
4399 4399

precision
0.0075 0.0075

0.493ethnicity=HISPANIC/LATINO - GUATEMALAN+0.385ethnicity=WHITE - RUSSIAN-0.304ethnicity=HISPANIC/LATINO - PUERTO RICAN+0.303ethnicity=PORTUGUESE-0.273ethnicity=ASIAN - KOREAN...

mean
-0.012 0.012

std. dev.
0.887 1.0994

weight sum
4399 4399

precision
0.0066 0.0066

-0.432ethnicity=ASIAN - THAI+0.402ethnicity=BLACK/AFRICAN-0.401ethnicity=BLACK/CAPE VERDEAN-0.271ethnicity=ASIAN - ASIAN INDIAN+0.221ethnicity=UNABLE TO OBTAIN...

mean

0.0035 -0.0035

std. dev.

0.8748 1.1089

weight sum

4399 4399

precision

0.0089 0.0089

-0.436ethnicity=ASIAN - CHINESE-0.311ethnicity=WHITE - OTHER EUROPEAN-0.309ethnicity=ASIAN - FILIPINO-0.278ethnicity=WHITE - EASTERN EUROPEAN+0.27 ethnicity=CARIBBEAN ISLAND...

mean

0.0082 -0.0082

std. dev.

1.1108 0.8714

weight sum

4399 4399

precision

0.0055 0.0055

-0.617ethnicity=MIDDLE EASTERN-0.425ethnicity=AMERICAN INDIAN/ALASKA NATIVE+0.251ethnicity=HISPANIC/LATINO - GUATEMALAN+0.245ethnicity=CARIBBEAN ISLAND+0.216ethnicity=SOUTH AMERICAN...

mean

-0.0196 0.0196

std. dev.

1.2165 0.7145

weight sum

4399 4399

precision

0.0065 0.0065

-0.68ethnicity=CARIBBEAN ISLAND+0.334ethnicity=UNABLE TO OBTAIN-0.328ethnicity=HISPANIC/LATINO - DOMINICAN+0.187ethnicity=BLACK/CAPE VERDEAN-0.165ethnicity=ASIAN - CHINESE...

mean

0.0276 -0.0276

std. dev.

0.7497 1.1906

weight sum

4399 4399

precision

0.0066 0.0066

-0.416gender-0.374elix.index-0.317ethnicity=WHITE - RUSSIAN+0.228ethnicity=SOUTH AMERICAN+0.227ethnicity=PORTUGUESE...

mean

-0.012 0.012

std. dev.

0.9262 1.0164

weight sum

4399 4399

precision

0.0029 0.0029

-0.605elix.index-0.361ethnicity=HISPANIC OR LATINO+0.235ethnicity=WHITE - RUSSIAN+0.22 gender+0.215ethnicity=ASIAN - FILIPINO...

mean

-0.0104 0.0104

std. dev.

0.9957 0.9402

weight sum

4399 4399

precision

0.0036 0.0036

Time taken to build model: 0.18 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	4564	51.8754 %
Incorrectly Classified Instances	4234	48.1246 %
Kappa statistic	0.0375	
Mean absolute error	0.4795	
Root mean squared error	0.6023	
Relative absolute error	95.9041 %	
Root relative squared error	120.4626 %	
Total Number of Instances	8798	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.139	0.101	0.578	0.139	0.224	0.557	Y
	0.899	0.861	0.511	0.899	0.651	0.557	N
Weighted Avg.	0.519	0.481	0.544	0.519	0.437	0.557	

=== Confusion Matrix ===

```
  a    b   <-- classified as
 610 3789 |    a = Y
 445 3954 |    b = N
```

PRINCIPAL COMPONENTS ANALYSIS WITH LOGISTIC REGRESSION

=== Run information ===

Scheme:weka.classifiers.functions.Logistic -R 1.0E-8 -M -1

Relation: temporal.extract.14.2.first.3.hours.category.3.healthy.controls-weka.filters.unsupervised.attribute.Remove-R1-2-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.FirstOrder-weka.filters.unsupervised.attribute.AddExpression-Ea1^2-Nexpression

Instances: 8884

Attributes: 186

[list of attributes omitted]

Test mode:10-fold cross-validation

=== Classifier model (full training set) ===

Logistic Regression with ridge parameter of 1.0E-8

Coefficients...

	Class
Variable	Y
gender	0.0028
age	0.0026
ethnicity=WHITE	-0.005
ethnicity=UNKNOWN/NOT SPECIFIED	-0.0425
ethnicity=PATIENT DECLINED TO ANSWER	0.1634
ethnicity=OTHER	-0.128
ethnicity=HISPANIC OR LATINO	-0.0674
ethnicity=BLACK/AFRICAN AMERICAN	-0.3238
ethnicity=ASIAN	-0.2038
ethnicity=MULTI RACE ETHNICITY	0.4627
ethnicity=HISPANIC/LATINO - PUERTO RICAN	0.2853
ethnicity=WHITE - BRAZILIAN	1.2617
ethnicity=BLACK/HAITIAN	0.1049
ethnicity=ASIAN - ASIAN INDIAN	0.3543
ethnicity=BLACK/CAPE VERDEAN	-0.1676
ethnicity=PORTUGUESE	-0.3704
ethnicity=WHITE - RUSSIAN	-0.5783
ethnicity=ASIAN - CHINESE	-0.3284
ethnicity=ASIAN - VIETNAMESE	-0.2869
ethnicity=WHITE - EASTERN EUROPEAN	74.6904
ethnicity=UNABLE TO OBTAIN	0.4007
ethnicity=MIDDLE EASTERN	0.8845
ethnicity=AMERICAN INDIAN/ALASKA NATIVE	48.8833
ethnicity=HISPANIC/LATINO - COLOMBIAN	162.7291
ethnicity=WHITE - OTHER EUROPEAN	2.0461
ethnicity=ASIAN - FILIPINO	39.4316
ethnicity=ASIAN - OTHER	0.5426
ethnicity=HISPANIC/LATINO - CUBAN	44.2399
ethnicity=HISPANIC/LATINO - DOMINICAN	-2.1098

ethnicity=HISPANIC/LATINO - GUATEMALAN	-21.9841
ethnicity=ASIAN - THAI	-52.6914
ethnicity=BLACK/AFRICAN	-23.0124
ethnicity=CARIBBEAN ISLAND	-69.3287
elix.index	0.0183
50112.min	0.3502
50112.median	-0.6962
50112.max	0.3463
50399.min	-0.344
50399.median	1.1306
50399.max	-0.7766
50440.min	0.0168
50440.median	-0.0371
50440.max	0.0227
50439.min	0.2359
50439.median	-0.5516
50439.max	0.3011
50428.min	0.0008
50428.median	-0.0004
50428.max	0.0003
50444.min	-457.0172
50444.median	912.7171
50444.max	-455.647
50412.min	-132.7201
50412.median	265.4178
50412.max	-132.6842
50411.min	-86.6882
50411.median	172.4623
50411.max	-85.8719
50413.min	-81.8759
50413.median	163.8057
50413.max	-81.9074
50383.min	-0.4545
50383.max	-0.3954
50383.median	0.8446
50386.min	-304.2088
50386.median	609.475
50386.max	-305.188

50442.min	1305.2194
50442.median	-2612.8315
50442.max	1307.3494
50468.min	-6.1433
50468.median	12.2851
50468.max	-6.149
50068.min	-65.6424
50068.median	131.317
50068.max	-65.6537
50172.min	-56.0744
50172.median	112.1506
50172.max	-56.0114
50083.min	-46.2092
50083.median	92.4812
50083.max	-46.1975
50149.min	-0.8967
50149.median	1.9282
50149.max	-1.2076
50159.min	76.9439
50159.median	-154.0526
50159.max	77.0248
50090.min	-21.6816
50090.median	42.4865
50090.max	-20.6694
50177.min	-25.8253
50177.median	51.7809
50177.max	-25.9564
50333.min	3.8819
50333.median	-0.0802
50333.max	-4.0338
50373.min	-0.0079
50373.median	0.0127
50373.max	0.0333
50417.min	547.4492
50417.median	-1094.8889
50417.max	547.4495
50408.min	231.3717
50408.median	-463.0485

50408.max	231.6688
50419.min	-74.5436
50419.median	149.1198
50419.max	-74.5758
line flush.IV.rx	0.264
heparin.IV.rx	-0.2826
docusate.PO.rx	0.091
acetaminophen.POPR.rx	0.0426
insulin.SCIV.rx	0.1031
pantoprazole.POIV.rx	0.1757
atorvastatin.PO.rx	0.0139
aspirin.POPR.rx	0.0914
senna.PO.rx	0.2116
bisacodyl.POPR.rx	0.269
clopidogrel.PO.rx	-0.1025
kcl.POIV.rx	0.182
nitroglycerine.IV.rx	-0.3531
nitroprusside.IV.rx	-0.5596
furosemide.POIV.rx	0.2204
oxycodone.PO.rx	0.4424
metoprolol.POIV.rx	0.2455
phenytoin.POIV.rx	-0.3524
cefazolin.IV.rx	0.2797
lisinopril.PO.rx	0.2069
propofol.IV.rx	-0.2435
nitroglycerin sl.PO.rx	0.3219
atropine sulfate.IV.rx	0.3219
zolpidem tartrate.PO.rx	0.0509
morphine.MULTI.rx	-0.2944
eptifibatide.IV.rx	0.1129
levofloxacin.POIV.rx	-0.1437
lansoprazole.PO.rx	-0.3234
calcium gluconate.IV.rx	0.0029
magnesium sulfate.IV.rx	0.3765
levothyroxine.POIV.rx	0.1159
glycopyrrolate.IV.rx	1.6233
neostigmine.IV.rx	-0.9749
metoclopramide.POIV.rx	0.3883

midazolam.IV.rx	-0.3331
ranitidine.PO.rx	0.2425
nitroglycerin.IV.rx	-0.3641
phenylephrine.IV.rx	0.1134
meperidine.IVIM.rx	-0.3501
sucralfate.PO.rx	-0.3751
acetylcysteine.POIV.rx	0.2252
fentanyl.IV.rx	-0.3038
ondansetron.POIV.rx	0.0101
dopamine.IV.rx	-0.1779
vancomycin.IV.rx	0.6978
levophed.IV.rx	0.0185
integrelin.IV.rx	-0.2219
lorazepam.MULTI.rx	-0.2705
methylprednisolone.POIV.rx	0.1563
neosynephrine.IV.rx	0.2209
amiodarone.POIV.rx	0.494
heparin.SC.rx	0.2666
metronidazole.POIV.rx	0.4654
tincture of opium.IH.rx	0.1438
albuterol.IH.rx	0.1704
phytonadione.MULTI.rx	-0.482
folic acid.POIV.rx	-0.3342
multivitamins.POIV.rx	0.7737
famotidine.POIV.rx	-0.6269
piperacillin-tazobactam.IV.rx	0.6883
simvastatin.PO.rx	-0.1707
dextrose 50%.IV.rx	-0.504
influenza virus vaccine.IM.rx	-0.3033
RT/LEFT HEART CARD CATH.cpt	0.6325
CORONAR ARTERIOGR-2 CATH .cpt	-1.4488
LT HEART ANGIOCARDIOGRAM .cpt	-0.0483
CONTINUOUS INVASIVE MECH .cpt	-0.3286
SPINAL TAP.cpt	-0.7461
INSERT ENDOTRACHEAL TUBE .cpt	-0.5354
ARTERIAL CATHETERIZATION .cpt	-0.6968
LEFT HEART CARDIAC CATH.cpt	0.2117
DX ULTRASOUND-HEART.cpt	0.0737

PERCUTANEOUS ABDOM DRAIN .cpt	1.3058
VENOUS CATHETER NEC.cpt	-0.0827
ENDOVASCULAR EMBOL/OCCLU .cpt	3.5989
PULMON ART WEDGE MONITOR .cpt	-0.6086
PERITONEAL ADHESIOLYSIS.cpt	2.3743
1 INT MAM-COR ART BYPASS .cpt	-2.4016
(AORTO)CORONARY BYPASS T .cpt	-0.3879
EXTRACORPOREAL CIRCULAT.cpt	2.4938
PULSATION BALLOON IMPLAN .cpt	0.5237
INJ/INF PLATELET INHIBIT .cpt	0.3478
PARENTERAL INFUS CONC NU .cpt	0.1838
SERUM TRANSFUSION NEC.cpt	-0.1122
EXT INFUS CONC NUTRITION .cpt	0.5276
PTCA W/O THROMBOLYTIC AG .cpt	-3.3556
INSERT CORON ARTER STENT .cpt	2.0654
HEMODIALYSIS.cpt	0.2547
SPINAL CANAL EXPLOR NEC.cpt	0.9757
PACKED CELL TRANSFUSION.cpt	0.3497
ANGIOPLASTY OR ATHERECTO .cpt	1.9605
CONTR CEREBR ARTERIOGRAM .cpt	-0.2949
INSERT INTERCOSTAL CATH.cpt	0.5933
PLATELET TRANSFUSION.cpt	-0.8716
SKIN SUTURE NEC.cpt	-0.4587
INFUSION OF VASOPRESSOR.cpt	0.4166
CONTRAST ARTERIOGRAM-LEG .cpt	0.6817
SM BOWEL ENDOSCOPY NEC.cpt	-0.7847
ENDOSCOPIC BRONCHIAL BX- .cpt	1.1458
RESP TRACT INTUBAT NEC.cpt	-1.4524
NON-INVASIVE MECHANICAL.cpt	-0.1957
CORONAR ARTERIOGR-1 CATH .cpt	-0.3732
THORACENTESIS.cpt	-0.1535
PERCUTANEOUS TRANSLUMINA .cpt	-3.4693
INSERT DRUG-ELUT CORONA.cpt	0.9872
a1^2	0.0028
Intercept	2.1367

Odds Ratios...

Variable	Class Y
===	
gender	1.0028
age	1.0026
ethnicity=WHITE	0.995
ethnicity=UNKNOWN/NOT SPECIFIED	0.9584
ethnicity=PATIENT DECLINED TO ANSWER	1.1775
ethnicity=OTHER	0.8799
ethnicity=HISPANIC OR LATINO	0.9348
ethnicity=BLACK/AFRICAN AMERICAN	0.7234
ethnicity=ASIAN	0.8156
ethnicity=MULTI RACE ETHNICITY	1.5883
ethnicity=HISPANIC/LATINO - PUERTO RICAN	1.3302
ethnicity=WHITE - BRAZILIAN	3.5314
ethnicity=BLACK/HAITIAN	1.1106
ethnicity=ASIAN - ASIAN INDIAN	1.4252
ethnicity=BLACK/CAPE VERDEAN	0.8457
ethnicity=PORTUGUESE	0.6905
ethnicity=WHITE - RUSSIAN	0.5608
ethnicity=ASIAN - CHINESE	0.7201
ethnicity=ASIAN - VIETNAMESE	0.7506
ethnicity=WHITE - EASTERN EUROPEAN	2.7393495333445685E32
ethnicity=UNABLE TO OBTAIN	1.4929
ethnicity=MIDDLE EASTERN	2.4217
ethnicity=AMERICAN INDIAN/ALASKA NATIVE	1.6973104979974168E21
ethnicity=HISPANIC/LATINO - COLOMBIAN	4.70262951723222E70
ethnicity=WHITE - OTHER EUROPEAN	7.738
ethnicity=ASIAN - FILIPINO	1.3332316912849928E17
ethnicity=ASIAN - OTHER	1.7205
ethnicity=HISPANIC/LATINO - CUBAN	1.6335459810172172E19
ethnicity=HISPANIC/LATINO - DOMINICAN	0.1213
ethnicity=HISPANIC/LATINO - GUATEMALAN	0
ethnicity=ASIAN - THAI	0
ethnicity=BLACK/AFRICAN	0
ethnicity=CARIBBEAN ISLAND	0
elix.index	1.0185
50112.min	1.4193

50112.median	0.4985
50112.max	1.4138
50399.min	0.7089
50399.median	3.0975
50399.max	0.46
50440.min	1.0169
50440.median	0.9635
50440.max	1.023
50439.min	1.266
50439.median	0.576
50439.max	1.3514
50428.min	1.0008
50428.median	0.9996
50428.max	1.0003
50444.min	0
50444.median	Infinity
50444.max	0
50412.min	0
50412.median	1.859894774137894E115
50412.max	0
50411.min	0
50411.median	7.932419695010202E74
50411.max	0
50413.min	0
50413.median	1.380039442304165E71
50413.max	0
50383.min	0.6348
50383.max	0.6734
50383.median	2.327
50386.min	0
50386.median	4.916273999493743E264
50386.max	0
50442.min	Infinity
50442.median	0
50442.max	Infinity
50468.min	0.0021
50468.median	216437.4516
50468.max	0.0021

50068.min	0
50068.median	1.0721521592697196E57
50068.max	0
50172.min	0
50172.median	5.085892594792186E48
50172.max	0
50083.min	0
50083.median	1.459024450288768E40
50083.max	0
50149.min	0.4079
50149.median	6.8769
50149.max	0.2989
50159.min	2.608099430024337E33
50159.median	0
50159.max	2.827856712826816E33
50090.min	0
50090.median	2.8290541146466115E18
50090.max	0
50177.min	0
50177.median	3.0771472234637156E22
50177.max	0
50333.min	48.5146
50333.median	0.923
50333.max	0.0177
50373.min	0.9921
50373.median	1.0128
50373.max	1.0339
50417.min	5.6778182782971354E237
50417.median	0
50417.max	5.679363258301366E237
50408.min	3.044156124895856E100
50408.median	0
50408.max	4.097191004235487E100
50419.min	0
50419.median	5.779741295950238E64
50419.max	0
line flush.IV.rx	1.3021
heparin.IV.rx	0.7538

docusate.PO.rx	1.0953
acetaminophen.POPR.rx	1.0436
insulin.SCIV.rx	1.1086
pantoprazole.POIV.rx	1.192
atorvastatin.PO.rx	1.014
aspirin.POPR.rx	1.0957
senna.PO.rx	1.2356
bisacodyl.POPR.rx	1.3087
clopidogrel.PO.rx	0.9026
kcl.POIV.rx	1.1996
nitroglycerine.IV.rx	0.7025
nitroprusside.IV.rx	0.5714
furosemide.POIV.rx	1.2466
oxycodone.PO.rx	1.5565
metoprolol.POIV.rx	1.2783
phenytoin.POIV.rx	0.703
cefazolin.IV.rx	1.3227
lisinopril.PO.rx	1.2298
propofol.IV.rx	0.7839
nitroglycerin sl.PO.rx	1.3797
atropine sulfate.IV.rx	1.3798
zolpidem tartrate.PO.rx	1.0522
morphine.MULTI.rx	0.745
eptifibatide.IV.rx	1.1195
levofloxacin.POIV.rx	0.8661
lansoprazole.PO.rx	0.7237
calcium gluconate.IV.rx	1.0029
magnesium sulfate.IV.rx	1.4572
levothyroxine.POIV.rx	1.1229
glycopyrrolate.IV.rx	5.0697
neostigmine.IV.rx	0.3772
metoclopramide.POIV.rx	1.4744
midazolam.IV.rx	0.7167
ranitidine.PO.rx	1.2744
nitroglycerin.IV.rx	0.6948
phenylephrine.IV.rx	1.1201
meperidine.IVIM.rx	0.7046
sucralfate.PO.rx	0.6872

acetylcysteine.POIV.rx	1.2526
fentanyl.IV.rx	0.738
ondansetron.POIV.rx	1.0102
dopamine.IV.rx	0.8371
vancomycin.IV.rx	2.0093
levophed.IV.rx	1.0187
integrelin.IV.rx	0.801
lorazepam.MULTI.rx	0.763
methylprednisolone.POIV.rx	1.1692
neosynephrine.IV.rx	1.2473
amiodarone.POIV.rx	1.6388
heparin.SC.rx	1.3055
metronidazole.POIV.rx	1.5926
tincture of opium.IH.rx	1.1547
albuterol.IH.rx	1.1858
phytonadione.MULTI.rx	0.6176
folic acid.POIV.rx	0.7159
multivitamins.POIV.rx	2.1677
famotidine.POIV.rx	0.5342
piperacillin-tazobactam.IV.rx	1.9904
simvastatin.PO.rx	0.8431
dextrose 50%.IV.rx	0.6041
influenza virus vaccine.IM.rx	0.7384
RT/LEFT HEART CARD CATH.cpt	1.8822
CORONAR ARTERIOGR-2 CATH .cpt	0.2348
LT HEART ANGIOCARDIOGRAM .cpt	0.9529
CONTINUOUS INVASIVE MECH .cpt	0.72
SPINAL TAP.cpt	0.4742
INSERT ENDOTRACHEAL TUBE .cpt	0.5855
ARTERIAL CATHETERIZATION .cpt	0.4982
LEFT HEART CARDIAC CATH.cpt	1.2358
DX ULTRASOUND-HEART.cpt	1.0765
PERCUTANEOUS ABDOM DRAIN .cpt	3.6905
VENOUS CATHETER NEC.cpt	0.9206
ENDOVASCULAR EMBOL/OCCLU .cpt	36.5562
PULMON ART WEDGE MONITOR .cpt	0.5441
PERITONEAL ADHESIOLYSIS.cpt	10.7438
1 INT MAM-COR ART BYPASS .cpt	0.0906

(AORTO)CORONARY BYPASS T .cpt	0.6785
EXTRACORPOREAL CIRCULAT.cpt	12.1075
PULSATION BALLOON IMPLAN .cpt	1.6883
INJ/INF PLATELET INHIBIT .cpt	1.4159
PARENTERAL INFUS CONC NU .cpt	1.2018
SERUM TRANSFUSION NEC.cpt	0.8938
EXT INFUS CONC NUTRITION .cpt	1.6949
PTCA W/O THROMBOLYTIC AG .cpt	0.0349
INSERT CORON ARTER STENT .cpt	7.8886
HEMODIALYSIS.cpt	1.2901
SPINAL CANAL EXPLOR NEC.cpt	2.653
PACKED CELL TRANSFUSION.cpt	1.4186
ANGIOPLASTY OR ATHERECTO .cpt	7.1031
CONTR CEREBR ARTERIOGRAM .cpt	0.7446
INSERT INTERCOSTAL CATH.cpt	1.8099
PLATELET TRANSFUSION.cpt	0.4183
SKIN SUTURE NEC.cpt	0.6321
INFUSION OF VASOPRESSOR.cpt	1.5168
CONTRAST ARTERIOGRAM-LEG .cpt	1.9773
SM BOWEL ENDOSCOPY NEC.cpt	0.4562
ENDOSCOPIC BRONCHIAL BX- .cpt	3.145
RESP TRACT INTUBAT NEC.cpt	0.234
NON-INVASIVE MECHANICAL.cpt	0.8222
CORONAR ARTERIOGR-1 CATH .cpt	0.6885
THORACENTESIS.cpt	0.8577
PERCUTANEOUS TRANSLUMINA .cpt	0.0311
INSERT DRUG-ELUT CORONA.cpt	2.6837
a1^2	1.0028

Time taken to build model: 2174.73 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	5513	62.0554 %
Incorrectly Classified Instances	3371	37.9446 %
Kappa statistic	0.2411	

Mean absolute error 0.4563

Root mean squared error 0.4843

Relative absolute error 91.2547 %

Root relative squared error 96.8521 %

Total Number of Instances 8884

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.639	0.398	0.616	0.639	0.628	0.656	Y
	0.602	0.361	0.625	0.602	0.613	0.656	N
Weighted Avg.	0.621	0.379	0.621	0.621	0.62	0.656	

=== Confusion Matrix ===

```
   a    b   <-- classified as
 2840 1602 |    a = Y
 1769 2673 |    b = N
```