IDENTIFYING KEY PREDICTORS THAT AFFECT THE LENGTH-OF-STAY IN THE EMERGENCY DEPARTMENT OF A LOCAL HOSPITAL WITH EXPLORATORY ANALYTICS

Lee Jinq-Yi
Marcus Tan Jun Liang
Muhammad Faris Bin Abdul Malik
Supervisor: Dr. Kam Tin Seong
Singapore Management University

Abstract – Managing wait times at Emergency Department (ED) is generally challenging because the ED deals with patients without appointment and with a large variety of illnesses and large variance in the time required diagnosing and treating them. In the ED, patients are being classified into four acuity categories, namely P1, P2, P3 and P4, in the decreasing order of their severity. The main bulk of the patients in the ED is made up of P3 patients (less severe patients) and as a result, some patients may encounter a long length-of-stay (LoS) that exceeds the KPI of the hospital. Using the Computerized Physician Order Entry System as a source of data, this study will explore various methods where different investigative tests and treatment ordered for a specific patient can be analyzed to determine the impact on patient's LoS.

Index Terms – Emergency Department, Data analytics, Length-of-Stay

---

1. INTRODUCTION

Singapore's healthcare system is widely regarded as one of the best in the world but despite its high standards, a common gripe of the general public is that of the long waiting times at the Emergency Departments of local hospitals.

While the patients who require immediate medical attention are given priority and are seen almost immediately, the patients who do not require immediate medical attention and are deemed as non-urgent are the most likely to face the brunt of these long waits. This group of non-emergency patients make up more than half of the patients at Emergency Departments and are classified as Priority 3 (P3) patients. In hospitals today, these patients have reported waiting times of up to seven hours, and it is easy to see why their healthcare experience is marred by the long waits cited.

The main motivation of this project is to better understand the factors and interactions that affect the length of stay (LoS) in the emergency department among P3 patients, in a bid to deconstruct the complex problem at hand. This is carried out in conjunction with Professor Tan Kar Way's ongoing research as she is currently working on the optimization of queuing times in the emergency department. It is hoped that a better understanding of these factors would then contribute to the development of a more accurate model for better modelling in her research.

1.1 Understanding the Process

To understand the process in the Emergency Department (ED), the Figure 1 below provides a graphical representation of the process flow in the department that contributes to the patient's Length of Stay- which is defined as the total time spent in the ED, from the time of registration to the time of disposition (which is the point at which the patient leaves the ED to either be discharged or to be warded).

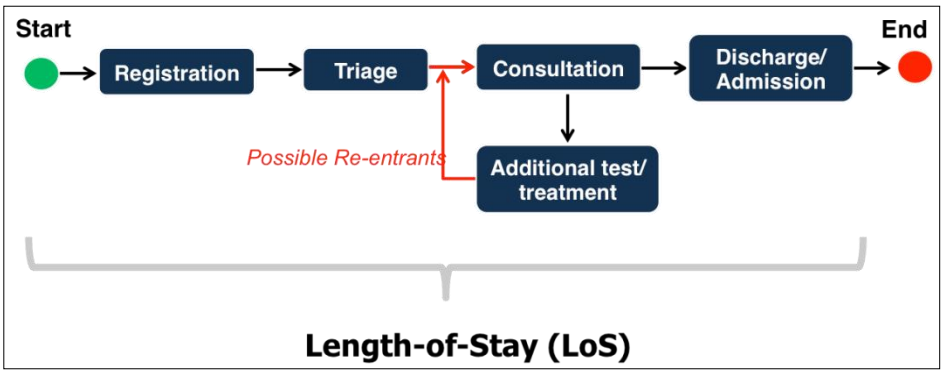


Figure 1: Overview of the Emergency Department Process

1.1.1 Registration

Registration is first carried out when the patient enters the ED. Patients details are recorded into the hospital EMERGE system.

1.1.2 Triage

After registration, the patient then proceeds to the triage where a nurse performs basic diagnostics to determine the severity level of the patient's illness and the urgency of medical attention required. This is where patients are categorized into the various categories – P1, P2, P3 or P4.

1.1.3 Consultation + Additional Tests

Patients then see the doctor for consultation. If required, patients will then go for further tests or treatment. After these tests, patients are then required to revisit the doctor for a re-evaluation with their test results. All tests and treatments that are being ordered are logged into the Computerized Physician Order Entry (CPOE) system.

1.1.4 Discharge/Admission

When the doctor is satisfied with the necessary tests done, the patient would then either be discharged or warded, marking the patient's end of stay in the Emergency Department.

1.1.5 Additional Comments

The main variation in the process comes from step 3 as some patients will have no need to re-enter the system if no further tests are ordered by the doctor, while others may have to re-enter the system multiple times due to various tests ordered one after another in the case that test results are inconclusive.

To address this issue, patients are classified by the number of re-entries into the system, where a patient with zero re-entries would have undergone no further tests or treatment, while another patient that had to undergo two rounds of further tests before being discharged would be classified as a patient with two re-entries.

2. OBJECTIVES

2.1 Business objective

To better understand factors that may affect the length of stay of patients in the Emergency Department. This will then improve the accuracy and introduce new insights for Prof Tan's dynamic queuing modelling

2.2 Technical Objective

To use data to get a better understanding of a complex problem through data overview in Exploratory Data Analysis (EDA), as we explore the various factors such as the number of re-entries, the type of different investigative tests & treatment, and the results of the tests. We will then analyze these various factors and its effect on LoS. After which, we will use statistical methods to explore and understand these interactions better.

3. DATA

3.1 Singapore General Hospital

For this project, we received data from Singapore General Hospital (SGH) with regard to patients in the A&E department. This data spanned over a three month period (January 2013 to March 2013), we were given two files, the "Emerge Dataset" and the "CPOE Dataset".

3.1.1 Emerge Dataset

This first dataset includes the time and date stamps of each unique visit within the time period, and the diagnosis of each patient.

This dataset also included the time at which the patient visited each station within the A&E, but prior research has found these timings to be unreliable. After our own checks, we concurred with these findings, and thus did not include the intermediate splits in our analysis.

3.1.2 CPOE Dataset

The second data set that we have is the Computerized Patient Order Entry file, which includes a list of the tests ordered for each patient during his stay at the Emergency Department, and the results of the tests carried out. This also included the normal range of results expected for each test taken by the patient.

3.1.3 Combined Dataset

From the two datasets received from SGH, we have put together a combined dataset that incorporates relevant details for each patient, for easy access. To do so, we matched the visit ID of each patient from both the CPOE and Emerge datasets.

Furthermore, using the data received, we have also derived new variables during our research, and this list has been attached as Appendix A.

3.2 Scope of the Project

The scope of our project was defined by the dataset that we have received, coupled with various areas of focus that were defined by our sponsor.

For our project, it only covers the three-month timespan, due to the data that was provided to us. Furthermore, we only looked at P3 patients- who are termed as the non-emergency patients- as they are the group with the most variable LoS, given that they do not require urgent medical attention.

4. METHODOLOGY

4.1 Data Cleaning

Having received the data, we used the scope provided by our sponsor to streamline our data set and to focus our research. To do so, we carried out an exclusion analysis which can be seen in Appendix B. This approach explains how and why we excluded various datapoints at each step of the way.

This was a crucial part to start off, as inaccurate data could potentially be a huge hindrance to the accuracy of our results and the understanding of our data.

4.2 Data Preparation

After the streamlining of the datapoints that we wanted to focus on, we then exported the files into two different programs that we used for the analysis to be carried out. These two programs are JMP Pro and Tableau,

4.2.1 JMP Pro

JMP Pro was used to visualize the data and to create quick charts and analysis based on our data. This was the tool that we mainly used in our analysis, and most of our charts have come from the use of JMP.

For us to make use of it, all data had to be formatted and sorted in order, as the program tended to automatically sort data based on alphabetical order, so thus if we wanted to order data based alternative ordering methods, such as chronologically, the variables then had to be derived and subsequently renamed.

4.2.2 Tableau

Tableau was used to create the heatmaps for our data. For us to utilize this, we had to ensure that all data was normalized, and that column headers were formatted for the appropriate visualizations.

4.3 Charts Used

To understand our data, we used several visualization forms, and they are explained below.

4.3.1 Bar Charts

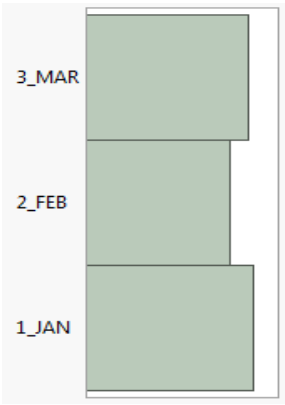


Figure 2: Example of a Bar Chart

To get a clearer idea of the distribution of our data, we frequently used bar charts for a quick overview of data distribution and frequency counts of the various categories. Figure 2 shows an example of a bar chart that we used.

4.3.2 Histograms

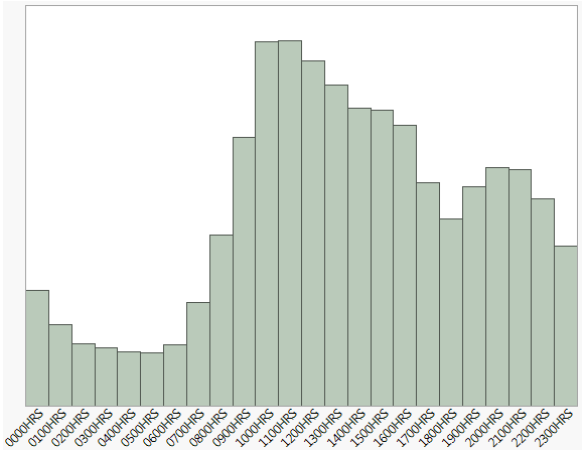


Figure 3: Example of a Histogram

The next tool that we used to visualize our data was the histogram, which was especially useful when the categories were next to each other and this was usually for continuous variables such as time. However, in our project, we found continuous variables to be few and far between, thus although the histogram is a good tool for data visualization, we did not make use of it as much as we did for the bar charts.

4.3.3 Dot Plots

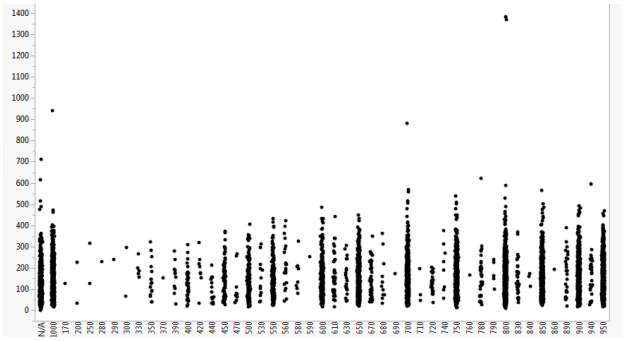
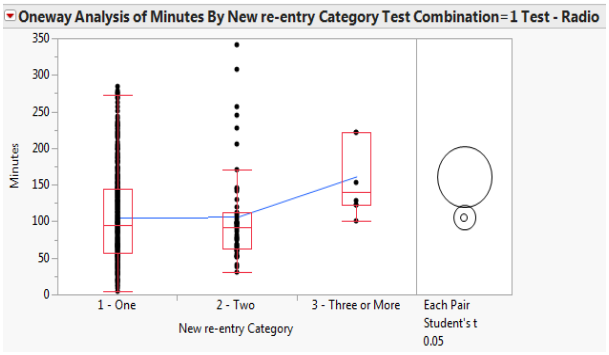


Figure 4: Example of a Dot Plot

We also used dot plots, especially when we plotted a continuous variable like LoS against categorical variables like various pass percentages of the various tests, which we will elaborate on later. This was useful as it allowed us to see the general trend of the datapoints.

However, one disadvantage that we noticed was that the dots tended to group together, which could result in misleading interpretations- if many dots appeared on the same point, it would seem as though there was only one data point at that specific point, when there could have been many more that were unseen. To improve on this, when we included many data points in our analysis, we used other techniques to understand the data trends and distributions.

4.3.4 ANOVA & Tukey Kramer Method



Connecting Letters Report

Level		Mean
3 - Three or More	A	161.50000
1 - One	A	116.35798
2 - Two	A	105.92157

Levels not connected by same letter are significantly different.

Ordered Differences Report

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
3 - Three or More	2 - Two	55.57843	37.12090	-17.3320	128.4889	0.1349
3 - Three or More	1 - One	45.14202	34.78014	-23.1709	113.4549	0.1948
1 - One	2 - Two	10.43641	14.33112	-17.7118	38.5847	0.4668

Figure 5: Example of ANOVA and Tukey Kramer

In cases where we assumed normality for the data, the analysis of variances (ANOVA) test was used in instances where more than two categorical variances were compared against each other, in a bid to determine if they are statistically different enough to be obtained from different populations. This would then elucidate if the difference in the categorical variable resulted in statistically different groups.

The next step to this was to carry out the Tukey Kramer Method Pairs, also known as the Tukey HSD. This tests if the means between the different pairs are significantly different from one another, and if so, exactly which pairs differ from one another.

4.3.5 Kruskal Wallis Test & Steel-Dwass Method

Nonparametric Comparisons For All Pairs Using Steel-Dwass Method

q*	Alpha
2.34370	0.05

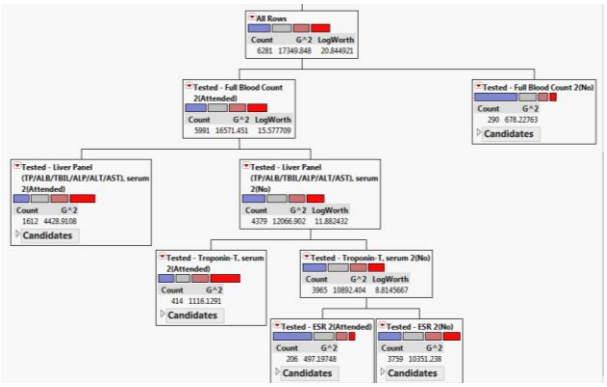
Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
3 - Three or More	1 - One	129.675	53.73922	2.41304	0.0418*	58.0000	2.0000	104.0000
3 - Three or More	2 - Two	19.161	6.53036	2.93407	0.0094*	59.0000	16.0000	123.0000
2 - Two	1 - One	-13.913	23.96549	-0.58053	0.8305	-4.0000	-24.0000	14.0000

Figure 6: Example of Kruskal Wallis & Steel-Dwass

In other cases where normality could not be assumed, non-parametric tests were utilized. In these cases, the tests seen in the previous section would then not be as appropriate, and thus, an equivalent test for non-parametric samples was used. This included the Wilcoxon/Kruscal-Wallis test.

The Wilcoxon/Kruscal-Wallis test is a method to compare several independent random samples and is a non-parametric alternative to the one way ANOVA. After researching on the nonparametric equivalent of Tukey HSD, we examined the difference between these pairs of results using Steel-Dwass Method.

4.3.6 Decision Tree Model



To better understand and to visualize the most important factors that affect LoS, we also used decision tree models for a quick and easy way to identify the most significant factors. This was based on LoS being split into 4 different quartiles.

Essentially, what a decision tree model does it that it branches out further from the most significant effect, and continues to split at the next most significant effect. Ideally, there should be a sizable difference in the length of the different coloured bars in the various branches, to show that the factors do interact significantly with the response variable.

4.3.7 Heat Maps

Re-Entry	Test Combination: 1 Test - Lab	2 Tests - Lab, Non-Parenteral	2 Tests - Lab, Radio	3 Tests - All Three
0 - No Re-Entry	127.7	93.0	144.8	106.3
1 - One Re-Entry	142.9	159.2	146.8	170.1
2 - Two Re-Entries	155.9	161.6	172.9	191.9
3 - Three Re-Entries or More	98.0	156.5	146.0	223.2

Figure 7: Example of a Heat Map

To visualize data, we also used heat mapping to enable us to pinpoint the average of the continuous variable at a glance for each category. From a heat map, the most significant and least significant values are denoted by especially dark or light coloured boxes, and this allows us to identify them at a glance.

Furthermore, when this is used over a spectrum of variables, this allows us to see trends on a broader scale, and is particularly useful in time series analysis.

5. FINDINGS AND ANALYSIS

5.1 Re-entry vs LoS by Test Combination

We conducted statistical tests and methods to find out if there is a significant difference between the number of re-entry and the patient's LoS for each of the **7 different tests combinations**. As mentioned in the chosen solution, 3 types of analysis are being conducted: comparing of means alone, parametric and non-parametric tests. Results shown from the statistical tests and methods are seen to be non-conclusive. This is due to many time variations that are not captured by our data such as the waiting time while waiting for triage, consultation and investigative tests/treatments. Data also does not include the time a patient spent consulting the doctor. These lack of data proved to be crucial in determining the relationship between the number of re-entry and the patient's LoS and has caused the above findings to be inconclusive. Detailed results on our findings of the 19212 patients are shown below.

5.1.1 Combination 1: 1 Test – Non-Parenteral

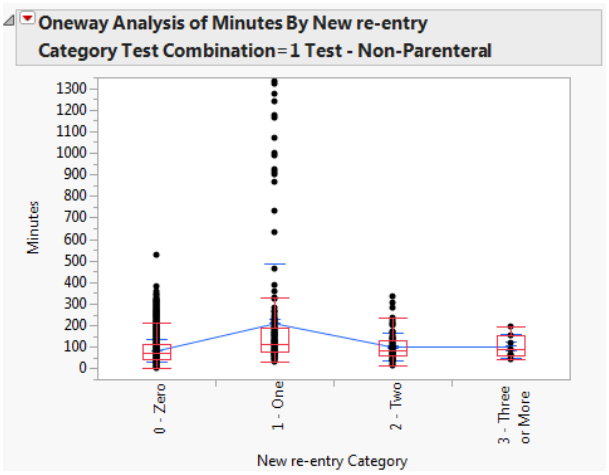


Figure 5.1.1: Boxplot of combination 1

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
0 - Zero	6550	83.827	54.127	0.669	82.52	85.14
1 - One	166	210.584	277.299	21.523	168.09	253.08
2 - Two	93	100.978	62.058	6.435	88.20	113.76
3 - Three or More	7	104.714	55.749	21.071	53.16	156.27

Figure 5.1.2: Means & Standard Dev. of Combination 1

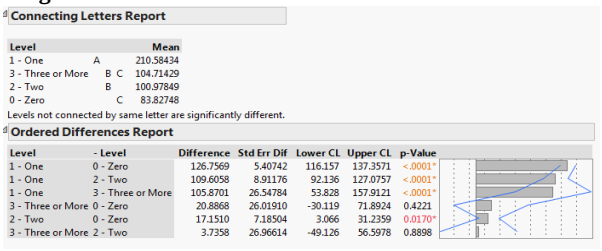


Figure 5.1.3: Tukey-Kramer HSD of Combination 1

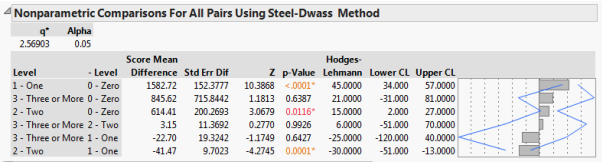


Figure 5.1.4: Steel Dwass Analysis of Combination 1

Out of the 19212 P3 Patients for the month of January to March, 35.5% or 6816 patients take only non-parenteral medication. This is the largest proportion out of the 7 different combinations of tests. Taking only its means into account, it shows that 1 re-entry patients has a very high LoS mean as compared to the rest. To verify this abnormal results, we proceed on statistical tests and methods and results are mentioned below.

Based on Figure 6.1.3, Tukey HSD tests shows that there are significant differences between **one** vs **zero**, **two** and **three or more** re-entry. There are also significant difference for **two** vs **zero** re-entry at 95% CI level.

Steel-Dwass method however shows that there is only significant difference between **zero** vs **one** and **two** re-entry and **one** vs **two** re-entry.

5.1.2 Combination 2: 1 Test – Radiology

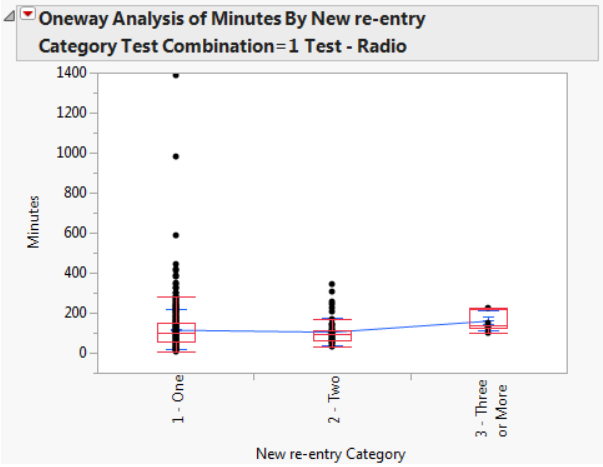


Figure 5.2.1: Boxplot of Combination 2

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
1 - One	514	116.358	100.546	4.435	107.65	125.07
2 - Two	51	105.922	67.315	9.426	86.99	124.85
3 - Three or More	8	161.500	51.766	18.302	118.22	204.78

Figure 5.2.2: Means & Standard Dev. of Combination 2

Connecting Letters Report

Level		Mean
3 - Three or More	A	161.50000
1 - One	A	116.35798
2 - Two	A	105.92157

Levels not connected by same letter are significantly different.

Ordered Differences Report

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
3 - Three or More	2 - Two	55.57843	37.12090	-17.3320	128.4889	0.1349
3 - Three or More	1 - One	45.14202	34.78014	-23.1709	113.4549	0.1948
1 - One	2 - Two	10.43641	14.33112	-17.7118	38.5847	0.4668

Figure 5.2.3: Tukey-Kramer HSD of Combination2

Nonparametric Comparisons For All Pairs Using Steel-Dwass Method

q*	Alpha
2.34370	0.05

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
3 - Three or More	1 - One	129.675	53.73922	2.41304	0.0418*	58.0000	2.0000	104.0000
3 - Three or More	2 - Two	19.161	6.53036	2.93407	0.0094*	59.0000	16.0000	123.0000
2 - Two	1 - One	-13.913	23.96549	-0.58053	0.8305	-4.0000	-24.0000	14.0000

Figure 5.2.4: Steel Dwass Analysis of Combination 2

About 3% of the 19212 P3 patients undergo only radiology tests. Based on the means, it shows that a one re-entry patient has a higher LoS than two and three re-entry.

Based on Figure 6.2.3, Tukey HSD tests shows that there are no significant differences across all pairs of re-entry.

However, Steel-Dwass method however shows that there are significant differences between **three or more vs one** and **two** reentry.

5.1.3 Combination 3: 1 Test – Laboratory

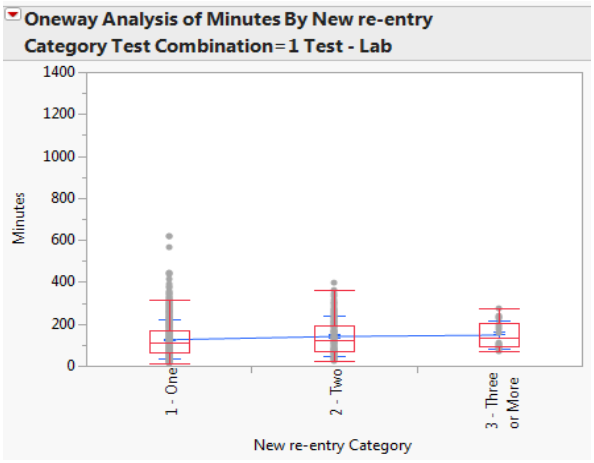


Figure 5.3.1: Boxplot of Combination 3

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
1 - One	771	127.703	92.7296	3.340	121.15	134.26
2 - Two	148	142.939	94.8268	7.795	127.54	158.34
3 - Three or More	18	149.500	65.8521	15.521	116.75	182.25

Figure 5.3.2: Means & Standard Dev. of Combination 3

Connecting Letters Report

Level		Mean
3 - Three or More	A	149.50000
2 - Two	A	142.93919
1 - One	A	127.70298

Levels not connected by same letter are significantly different.

Ordered Differences Report

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
3 - Three or More	1 - One	21.79702	22.09008	-21.5549	65.14895	0.3240
2 - Two	1 - One	15.23621	8.31423	-1.0805	31.55294	0.0672
3 - Three or More	2 - Two	6.56081	23.12646	-38.8250	51.94665	0.7767

Figure 5.3.3: Tukey-Kramer HSD of Combination 3

Nonparametric Comparisons For All Pairs Using Steel-Dwass Method

q*	Alpha
2.34370	0.05

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
3 - Three or More	1 - One	97.75800	54.34088	1.798977	0.1699	28.00000	-10.0000	65.00000
2 - Two	1 - One	45.89040	23.82059	1.926501	0.1312	12.00000	-3.0000	28.00000
3 - Three or More	2 - Two	10.68656	11.99758	0.890727	0.6462	16.00000	-31.0000	57.00000

Figure 5.3.4: Steel Dwass Analysis of Combination 3

About 5% of the 19212 patients undergo only laboratory tests with its means showing that the higher the number of re-entry, the higher the LoS will be.

Based on Figure 5.3.3 and 5.3.4, both post hoc analysis of Tukey HSD and Steel-Dwass all pairs shows no significant differences between the pairs.

5.1.4 Combination 4: 2 Tests – Non-Parenteral & Laboratory

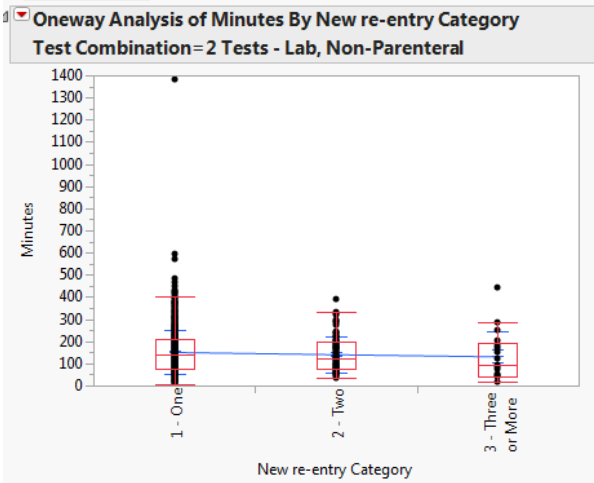


Figure 5.4.1: Boxplot of Combination 4

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
1 - One	1266	152.472	97.245	2.733	147.11	157.83
2 - Two	109	142.422	81.255	7.783	127.00	157.85
3 - Three or More	17	132.882	114.642	27.805	73.94	191.83

Figure 5.4.2: Means & Standard Dev. of Combination 4

Connecting Letters Report

Level		Mean
1 - One	A	152.47235
2 - Two	A	142.42202
3 - Three or More	A	132.88235

Levels not connected by same letter are significantly different.

Ordered Differences Report

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
1 - One	3 - Three or More	19.59000	23.51684	-26.5424	65.72235	0.4050
1 - One	2 - Two	10.05034	9.61453	-8.8102	28.91091	0.2961
2 - Two	3 - Three or More	9.53967	25.11623	-39.7302	58.80950	0.7041

Figure 5.4.3: Tukey-Kramer HSD of Combination 4

Nonparametric Comparisons For All Pairs Using Steel-Dwass Method

q*	Alpha
2.34370	0.05

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
3 - Three or More	2 - Two	-9.690	9.52162	-1.01765	0.5657	-23.0000	-70.0000	37.00000
2 - Two	1 - One	-32.010	39.63572	-0.80760	0.6983	-6.0000	-26.0000	12.00000
3 - Three or More	1 - One	-116.514	90.46338	-1.28797	0.4019	-27.0000	-80.0000	27.00000

Figure 5.4.4: Steel Dwass Analysis of Combination 4

About 22% of the 19212 patients under go through both lab and non-parenteral tests. The results of the mean shows an abnormal decrease in mean as the number of re-entry increase.

Based on Figure 5.4.3 and 5.4.4, both Tukey HSD and Steel-Dwass all pairs methods shows no significant differences between the pairs.

5.1.5 Combination 5: 2 Tests – Non-Parenteral & Radiology

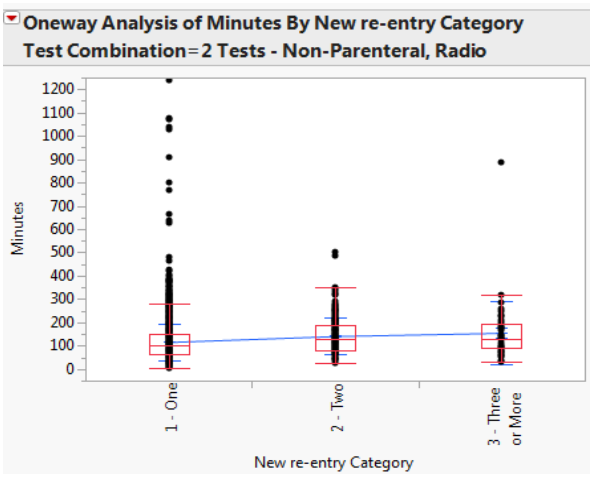


Figure 5.5.1: Boxplot of Combination 5

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
1 - One	3954	116.607	79.385	1.262	114.13	119.08
2 - Two	211	142.261	76.298	5.253	131.91	152.62
3 - Three or More	41	156.049	136.171	21.266	113.07	199.03

Figure 5.5.2: Means & Standard Dev. of Combination 5

Connecting Letters Report

Level		Mean
3 - Three or More	A	156.04878
2 - Two	A	142.26066
1 - One	B	116.60698

Levels not connected by same letter are significantly different.

Ordered Differences Report

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
3 - Three or More	1 - One	39.44180	12.55305	10.0107	68.87285	0.0048*
2 - Two	1 - One	25.65368	5.65001	12.4070	38.90032	<.0001*
3 - Three or More	2 - Two	13.78812	13.64798	-18.2100	45.78626	0.5704

Figure 5.5.3: Tukey-Kramer HSD of Combination 5

Nonparametric Comparisons For All Pairs Using Steel-Dwass Method

q*	Alpha
2.34370	0.05

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
2 - Two	1 - One	481.6368	84.9605	5.668950	<.0001*	25.00000	15.0000	35.00000
3 - Three or More	1 - One	450.9697	181.0594	2.490727	0.0341*	25.00000	2.0000	49.00000
3 - Three or More	2 - Two	1.1215	12.4401	0.090151	0.9955	1.00000	-27.0000	27.00000

Figure 5.5.4: Steel Dwass Analysis of Combination 5

7% of the patients fall into the category of non-parenteral and radio. Means show a normal increase of LoS as the number of re-entry increases.

Based on Figure 5.5.3, Tukey HSD tests shows that there are significant differences between **one** vs **two** and **three or more** re-entry.

Similarly, Steel-Dwass method also shows that there is only significant difference between **one** vs **two** and **one vs three or more** re-entry at 95% CI.

5.1.6 Combination 6: 2 Tests – Radiology & Laboratory

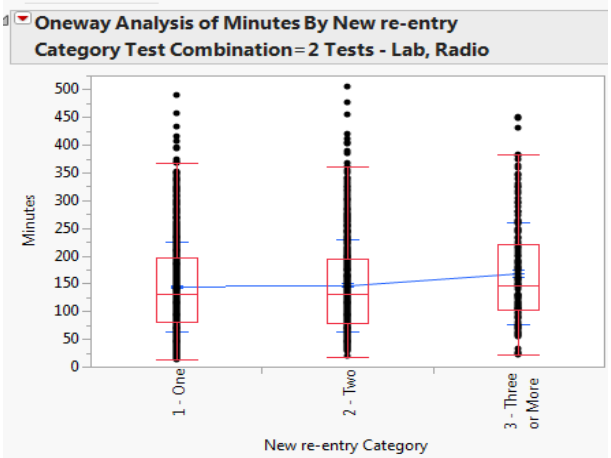


Figure 5.6.1: Boxplot of Combination 6

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
1 - One	1124	144.845	81.3203	2.4256	140.09	149.60
2 - Two	643	146.801	83.1076	3.2774	140.37	153.24
3 - Three or More	184	168.641	91.4541	6.7421	155.34	181.94

Figure 5.6.2: Means & Standard Dev. of Combination 6

Level		Mean
3 - Three or More	A	168.64130
2 - Two	B	146.80093
1 - One	B	144.84520

Levels not connected by same letter are significantly different.

Ordered Differences Report

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
3 - Three or More	1 - One	23.79611	6.593664	10.8647	36.72749	0.0003*
3 - Three or More	2 - Two	21.84037	6.931917	8.2456	35.43513	0.0017*
2 - Two	1 - One	1.95574	4.099630	-6.0844	9.99586	0.6334

Figure 5.6.3: Tukey-Kramer HSD of Combination 6

Nonparametric Comparisons For All Pairs Using Steel-Dwass Method

q*	Alpha
2.34370	0.05

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
3 - Three or More	1 - One	96.92871	30.03928	3.226732	0.0036*	21.00000	6.00000	35.00000
3 - Three or More	2 - Two	57.89811	19.97143	2.899047	0.0105*	20.00000	4.00000	35.00000
2 - Two	1 - One	7.47647	25.22861	0.296349	0.9527	1.00000	-8.00000	10.00000

Figure 5.6.4: Steel Dwass Analysis of Combination 6

1951 patients out of 19212 patients undergo laboratory and radiology tests. Means shows that there is an increase of LoS as the number of re-entry increases.

Based on Figure 5.6.3, Tukey HSD tests shows that there are significant differences between **three or more** re-entry vs **one** and **two re-entry**

Similarly, Steel-Dwass method also shows that there is only significant difference between **three or more** re-entry vs **one** and **three or more** re-entry and **two** re-entry at 95% CI.

5.1.7 Combination 7: All 3 Tests

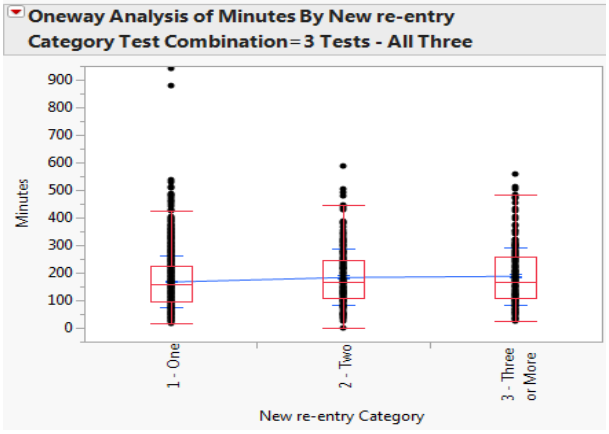


Figure 5.7.1: Boxplot of Combination 7

Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
1 - One	1558	168.684	94.574	2.3960	163.98	173.38
2 - Two	211	184.934	102.949	7.0873	170.96	198.91
3 - Three or More	232	189.552	104.144	6.8374	176.08	203.02

Figure 5.7.2: Means & Standard Dev. of Combination 7

Connecting Letters Report

Level		Mean
3 - Three or More	A	189.55172
2 - Two	A	184.93365
1 - One	B	168.68421

Levels not connected by same letter are significantly different.

Ordered Differences Report

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
3 - Three or More	1 - One	20.86751	6.800264	7.5312	34.20386	0.0022*
2 - Two	1 - One	16.24944	7.088689	2.3474	30.15144	0.0220*
3 - Three or More	2 - Two	4.61807	9.192711	-13.4102	22.64638	0.6155

Figure 5.7.3: Tukey-Kramer HSD of Combination 7

Nonparametric Comparisons For All Pairs Using Steel-Dwass Method

q*	Alpha
2.34370	0.05

Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
3 - Three or More	1 - One	96.99118	36.37298	2.666572	0.0209*	18.00000	2.0000	33.00000
2 - Two	1 - One	74.10430	37.47100	1.977644	0.1177	14.00000	-3.0000	30.00000
3 - Three or More	2 - Two	5.24882	12.17903	0.430972	0.9027	4.00000	-18.0000	26.00000

Figure 5.7.4: Steel Dwass Analysis of Combination 7

10% of patients are being ordered all 3 tests with means showing that there is an increase of LoS as the number of re-entry increases.

Based on Figure 5.7.3, Tukey HSD tests shows that there are significant differences between **one** and **three or more** re-entry and **one** and **two** re-entry at 95% CI.

Steel-Dwass method shows that there is only significant difference only between **three or more** re-entry vs **one** re-entry at 95% CI.

5.2. Results of Test Vs LoS

We also found little to no relationships between the results of tests and the LoS. This was done through the following steps.

5.2.1 Selection of Tests Used

For this portion, we were only able to utilize laboratory tests. This is so as our data was only able to provide the results and acceptable ranges for laboratory tests, and not for radiology or other tests.

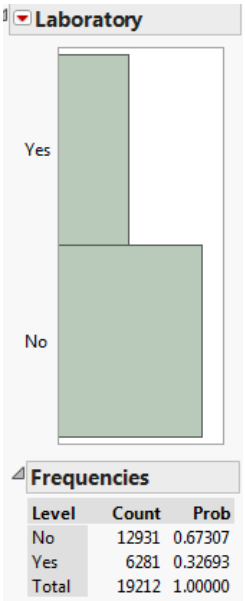


Figure 5.2.1: Distribution of Laboratory Tests

From the bar chart above, it can be seen that there were 6,281 P3 patients that took a laboratory test.

Of these patients, and of the tests taken, not all tests provided a large enough sample size for analysis. Looking at the distribution of patients taking the top six tests below explain why.

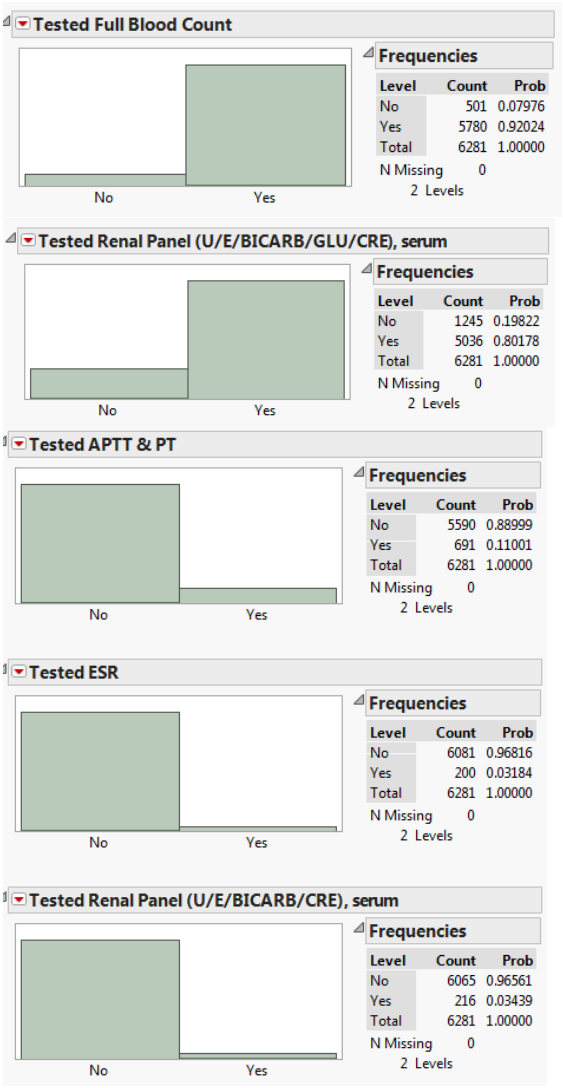


Figure 5.2.2: Top Six most taken Laboratory Tests

From the second most to the third most taken test, there is a drop off of 70% of patients. Thus, we have decided to use the top two tests as a basis to test the hypothesis of test results and LoS.

Furthermore, to ensure internal validity in the sense that all other factors are controlled for, we decided to use patients that are mostly homogenous, apart from their test results. These patients are mostly those that have exhibited one re-entry, only took a single laboratory test, and that laboratory test has to be the test that we are testing the results for.

5.2.3 Calculation of Laboratory Tests Scores

As each laboratory test can give up to 25 individual results, with one sample able to test for multiple items, we have aggregated the results of each patient's test and scored it as Passed or Failed, based on the ideal range given. From this, we calculated the percentage of tests passed or failed within a laboratory test, and gave each patient a score based on this percentage.

5.2.4 Full Blood Count

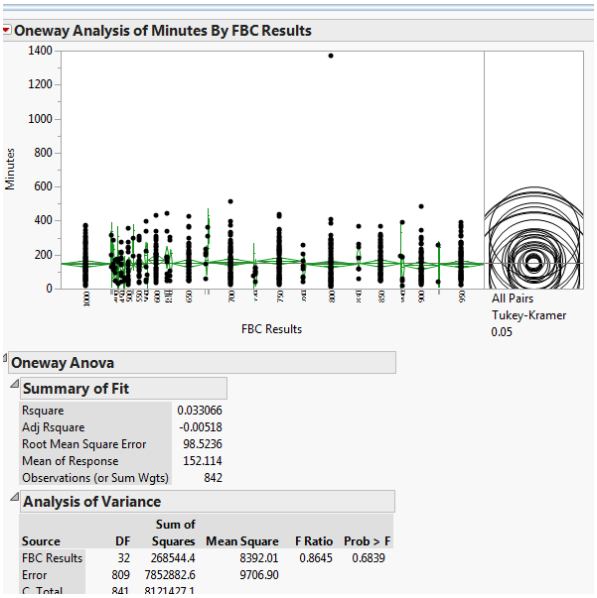


Figure 5.2.4: ANOVA and Tukey-Kramer HSD of Different Test Results (FBC)

To compare against patients with differing test results, we used ANOVA coupled with a Tukey Kramer test.

In this, we observed that across the different levels of test results, there is no trend observed in the dotplot. However, a concern has been raised, with regard to the fact that points in the dot plot may overlap, leading to misleading results as there may be a high concentration of dots in one specific area, showing up as one dot. However, this is disproven by the exceedingly low F statistic.

Furthermore, using the Tukey Kramer on the right as a comparison, the number of overlapping circles coupled with the lack of distinct circles shows that there is indeed no correlation that can be observed between the tbe results of the FBC test and the LoS.

5.2.5 Renal Panel

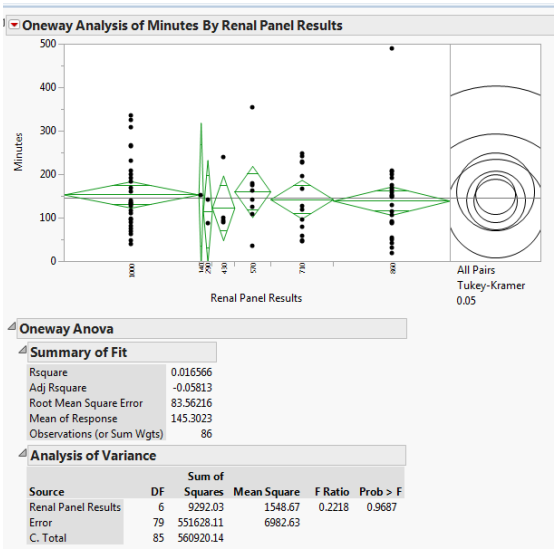


Figure 5.2.5: ANOVA and Tukey-Kramer HSD of Different Test Results (Renal Panel)

The same lack of correlation was again observed for the results to renal panel case, where there was no significant difference between the various proportions of successful tests.

5.3 Types of Laboratory Tests vs LoS

We then explored ways to understand which laboratory tests had a higher impact on LoS for each patient. We used three different methods, namely linear regression, decision tree modelling and heat map charts.

5.3.1 Linear Regression

Taking the tests that had more than 30 patients, we then used linear regression on all the laboratory tests to see if the type of tests taken was able to predict the LoS.

To do so, each test was assigned a binary value, a one or zero whether the test was taken or not, respectively. These were all then put in as predictor variables into the model, with LoS as the response variable.

Fit Group

Response Minutes

Summary of Fit

RSquare	0.057118
RSquare Adj	0.055464
Root Mean Square Error	90.55885
Mean of Response	153.9809
Observations (or Sum Wgts)	6281

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	11	3114431	283130	34.5242
Error	6269	51411475	8201	Prob > F
C. Total	6280	54525906		<.0001*

Lack Of Fit

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	93.087712	11.86246	7.85	<.0001*
Tested - Liver Panel (TP/ALB/TBIL/ALP/ALT/AST/GGT), serum[No]	32.781712	5.054621	6.49	<.0001*
Tested - Swab (aerobic) Culture[No]	25.418785	7.462208	3.41	0.0007*
Tested - Gram Stain[No]	20.98759	4.766316	4.40	<.0001*
Tested - ESR[No]	17.385177	3.110408	5.59	<.0001*
Tested - Blood Culture (aerobic)[No]	9.0244456	1.70403	5.30	<.0001*
Tested - APTT & PT[No]	7.4060199	1.776474	4.17	<.0001*
Tested - Amylase, serum[No]	-9.834368	1.512535	-6.50	<.0001*
Tested - Renal Panel (U/E/BICARB/GLU/CRE), serum[No]	-11.91893	2.217321	-5.38	<.0001*
Tested - Renal Panel (U/E/BICARB/CRE), serum[No]	-13.92075	3.550808	-3.92	<.0001*
Tested - Troponin-T, serum[No]	-14.23679	1.982867	-7.18	<.0001*
Tested - Potassium, serum[No]	-24.20571	3.821835	-6.33	<.0001*

Figure 5.3.1: Linear Regression of Laboratory Tests

As seen from Figure 5.3.1, there are about 11 laboratory tests that show significant impacts on the LoS based on their p-values. However, this model proves to have low predictive value, as seen from the R-square value of 6%. This could have been due to the fact that the use of linear regression with the input solely comprising of categorical variables might not be as appropriate.

5.3.2 Decision Tree Modelling

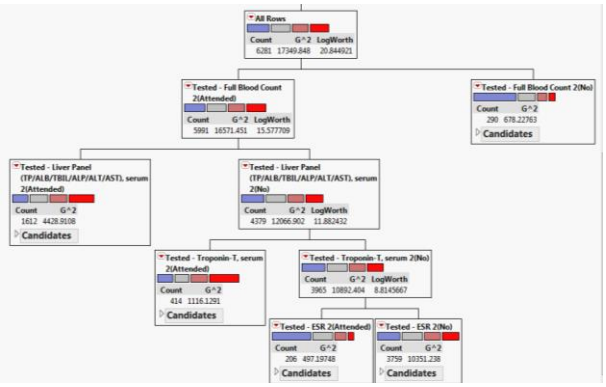


Figure 5.3.2: Decision Tree of Laboratory Tests

To come up with a decision tree model, the first step was to band the LoS based on 4 quartiles of the LoS among patients who took laboratory tests. This then to be banded into four bands, and was used as our response variable in the decision tree.

Again, this showed very little significant results as the difference between the different time bands across levels was not discriminatory enough, and thus we were not able to get clear results from the decision tree model either.

5.3.3 Heat Map

	Minute Band			
Lab Test Taken	0-79min	80-139min	140-199m..	>200min
Tested - (CK, MB,TNT)	26.92%	15.38%	21.15%	36.54%
Tested - Aerobic Culture	26.32%	33.33%	29.82%	10.53%
Tested - Albumin, serum	15.22%	19.57%	30.43%	34.78%
Tested - Amylase, serum	18.55%	21.92%	22.52%	37.01%
Tested - APTT & PT	27.01%	29.55%	21.52%	21.93%
Tested - Blood Culture (aerobic)	22.73%	33.01%	25.00%	19.26%
Tested - Blood Culture (anaerobic)	22.73%	33.01%	25.00%	19.26%
Tested - C-Reactive Protein, serum	31.65%	29.50%	20.43%	18.42%
Tested - Ca/PO4/Mg, serum	36.84%	28.95%	23.68%	10.53%
Tested - Calcium Total, serum	14.06%	18.75%	26.56%	40.63%
Tested - Creatine Kinase-MB (Mass), serum	17.50%	19.44%	22.22%	40.83%
Tested - Creatine Kinase, serum	17.59%	19.69%	21.52%	41.21%
Tested - D-Dimer Quantitation	17.86%	22.32%	25.00%	34.82%
Tested - ESR	42.54%	28.51%	16.23%	12.72%
Tested - Eye (aerobic) Culture	51.25%	36.25%	7.50%	5.00%
Tested - Eye (anaerobic) Culture	55.26%	31.58%	7.89%	5.26%
Tested - Eye (Fungal) Culture	52.63%	31.58%	8.77%	7.02%
Tested - Full Blood Count	22.72%	26.22%	22.80%	28.26%
Tested - Gram Stain	47.17%	35.85%	12.26%	4.72%
Tested - HBA1c, blood	37.84%	35.14%	10.81%	16.22%
Tested - HIV Screen	52.94%	26.47%	8.82%	11.76%
Tested - Liver Function Test	21.88%	27.08%	17.71%	33.33%
Tested - Liver Panel (TP/ALB/TBIL/ALP/ALT/AST), serum	16.99%	23.80%	24.05%	35.15%
Tested - Liver Panel (TP/ALB/TBIL/ALP/ALT/AST/GGT), serum	53.01%	22.89%	10.84%	13.25%
Tested - Magnesium, serum	15.00%	21.67%	30.00%	33.33%
Tested - Malaria Parasite, blood film	15.79%	42.11%	13.16%	28.95%
Tested - NT-proBNP, serum	18.12%	21.01%	26.09%	34.78%
Tested - Phosphate I0rganic, serum	16.33%	18.37%	30.61%	34.69%
Tested - Potassium, serum	9.72%	20.14%	22.22%	47.92%
Tested - Procalcitonin	32.38%	34.29%	20.95%	12.38%
Tested - Procalcitonin, serum	35.14%	37.84%	18.92%	8.11%
Tested - PT & INR	26.06%	21.13%	24.65%	28.17%
Tested - Renal Panel (U/E/BICARB/CRE), serum	27.64%	25.61%	14.63%	32.11%
Tested - Renal Panel (U/E/BICARB/GLU/CRE), serum	21.95%	26.15%	23.35%	28.56%
Tested - Swab (aerobic) Culture	52.63%	34.21%	10.53%	2.63%
Tested - Thyroid Panel (FT4/TSH)	23.75%	27.50%	18.75%	30.00%
Tested - Thyroid Stimulating Hormone, serum	22.64%	5.66%	32.08%	39.62%
Tested - Thyroxine (T4) Free, serum	30.30%	6.06%	30.30%	33.33%
Tested - Troponin-T, serum	15.58%	19.86%	22.60%	41.95%
Tested - Uric Acid, serum	22.53%	22.53%	26.92%	28.02%

Figure 5.3.3: Heat Map of Laboratory Tests

Lastly, using the banded quartiles based on the LoS that we used in our decision tree, we then put it into a heat map, in a bid to be able to see the significant LoS trends at a glance, based on the types of tests carried out. This can be seen in 5.3.3, and yielded slightly clearer results than the previous two methods.

This can be understood in the way that the tests with darker green boxes on the right hand sides contribute more significantly to a longer LoS.

5.4 Time Series

Lastly, we looked at the impact of time and the time at which the patients entered on the LoS. This was done in two ways, a heat map and a decision tree model.

5.4.1 Heat Map

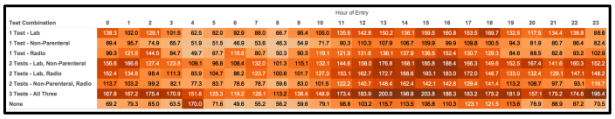


Figure 5.4.1: Heat Map of LoS based on Registration Time of the Day

Looking at the heat map seen above, there can be seen that from 10am to 6pm, there is a significantly longer LoS for patients entering during that period compared to other times of the day. This trend can be seen by the darker patch in the middle compared to other parts of the day.

5.4.2 Decision Tree Model

Based on 5.4.1 and the heat map seen, our next step was to explore if the time series and other factors related to time indeed affected LoS more significantly than those that we had explored before. This included variables like the number of doctors per hour and whether the day was a weekday or weekend.

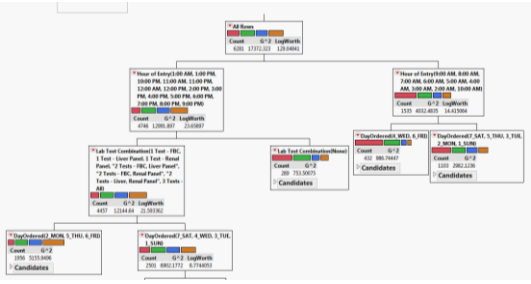


Figure 5.4.2: Decision Tree Model

Putting all the variables into a decision tree model, it can be seen that while the unique tests were all input, they did not appear until the 5th level, which means that the time series may be a more significant factor than that of the type of tests taken, and subsequently the results of these tests. Therefore, this decision tree does show that the hour and day in which a patient enters do indeed make up the most important factors as to determining whether the patient's LoS is long.

6. LIMITATIONS AND ASSUMPTIONS

6.1 Vague Response Variable

During the course of our research, a key limitation that we were faced with was that the response variable that we used was Length of Stay (LoS), which is the aggregate time that spans from the point from which the patient enters the A&E department to the point that the patient leaves the system, either to be discharged or to be admitted into the hospital. This was due to the data that we received from the hospital, as the most accurate timings within the system are the entry and disposition time. The inaccuracies seen in the timings between the various split times were observed not only by our group, but also by previous teams who worked on the same dataset. Due to a non-standardized way in which the timings were taken, they were deemed as inaccurate and thus were largely not taken into account.

However, this posed massive problems in our data analysis as the vagueness of the response variable contributed to our inability to construct a model that was reliable and that high predictive ability. This was partially explained by the example that during field observations, patients with a higher percentage of passed tests cleared consultation measurably quicker than those that failed multiple measures within each laboratory test, however, this failed to show up during our analysis with LoS. This is due to the fact that although the test results may indeed have significantly affected one segment of the LoS process, its effects were reduced by the noise and the other portions of the LoS- which prevented us from truly understanding its effects on the system.

Thus, this was a sizable limitation to our project as unless an effect was large enough to be seen across the whole process, it would not be seen as significant – which was what we saw in most of our analysis.

6.1.1 Learning Points

From this, we learnt that in order for us to get a clearer picture of future systems, we should then be more discerning of the data that we use, as only if we identify the specific area that we want to improve or look at, can we then accurately and effectively identify the levers that most significantly contribute to it.

Also, we have observed that as a data scientist, it is imperative to thoroughly understand the system and the data given to ensure that the data obtained is reliable enough of use for our analysis. This is so as our team did assume that we could still make do with the aggregated response variable as we believed that any change present would still be seen in the overall scale of things. However, from our results and this whole project, we learnt that unless we look at the right areas, the effects of various variables will not necessarily be seen on a larger scale, thus causing results seen and the subsequent predictive model built to be of low predictive value.

Lastly, from this, we also realized the value of understanding and refining the process data collection, where a data scientist's role is not merely confined to the EDA or the analysis of the statistics, but also in giving input on how to improve the data collection process in terms of accuracy and efficacy. Had the data collected been more accurate and more representative of the current situation, this would have benefited our analysis more. However, it is also not enough for us to sit back and lament on the lack of data, but it is also our job to suggest ways on how to improve on this failing within the dataset so that we can help the hospital and subsequent researchers to find the best solution to the current state of events.

6.2 Lack of Other Predictors

During our analysis to understand the predictors to the LoS in an Emergency Department, we also noticed that there were many other areas that we would have loved to look into, but for the lack of data. For some, we were not able to explore its effects on the patients and LoS, but for others we attempted to deduce the value. For instance, in our analysis, we mentioned that the number of re-entries for non-parenteral patients was not clear, and thus a statistical estimate was used to infer the patients whom we suppose did not go in for another consultation, and were thus classed as zero re-entries.

This estimation technique was suggested by the team from SGH, and they noted that it was the best way to deduce if the patient who had undergone non-parenteral treatment had seen the doctor again. This may be the best way at the moment, but it also brings up certain inaccuracies. Again, this leads to the same issue of data clarity and availability that we raised in the last point.

6.1.2 Learning Points

One more thing that we have learnt besides those stated in the previous point is that when using real-world data, inaccuracies are bound to be present, unlike academic data that is given in the pursuit of learning. This also means that not only are the tools to analyze the data important, so are the tools to extrapolate and to accurately deduce certain unknown factors from the data.

7. FUTURE WORK & RECOMMENDATIONS

During the duration of this project, the hospital has also engaged a master's course student to work with them on queueing theory – similar to what we have done. Our findings will then function as the base on which further work can be done, with regards to the general descriptive findings and more specifically, the areas in which further work can yield more positive and significant results.

7.1 Time Series Based Analysis

From our descriptive analysis, it appears that while looking at the situation from a test-based approach may not be as effective, a time series approach could possibly yield better results, seeing that trends can be observed in our descriptive research. From this, a more in-depth analysis can be taken towards building a predictive model based on the times of the days and the days that the patients enter the A&E, among other factors. Other factors that could be considered in further analysis could

also include an analysis on the distribution and frequency of P1 and P2 patients with regard to the time series analysis, as increases in the number of instances seen would inevitability affect the length of stay of P3 patients as the P1 and P2 patients have priority due to the severity of their conditions.

We hope that this can help the hospital in proactively predicting the demand and the strain on medical resources during that period, thus taking pre-emptive measures on both the supply and demand end. This can be done by either ensuring that more doctors are on duty or on call during that period, or making it clear to patients that the period is especially busy- making it known that if the condition that patients are suffering is not life-threatening, going to a general practitioner the next day may benefit both the patient and the A&E department.

7.2 Improved Data Collection Systems

Furthermore, as a major limitation that our team faced was the inability to accurately pinpoint and segregate the different portions of the LoS, we also feel that an increased ability to do so would definitely allow for more potential improvements in the system.

This is so as each factor's effects can then clearly be identified. By looking at each phase of the LoS more specifically, more targeted approaches to understanding and tackling a problem of this scale could then be used to greater effect. For example, in the case of the test results and its results on the patients' LoS, the extent of its effects could possibly be seen more clearly, had the length of the consultation time been isolated. This would then help in the subsequent formulation and more accurate simulation of the resultant queue algorithm, where a shortest-consultation time first method is currently being explored.

To do so, we feel that the use of RFIDs could be explored in the hospital's A&E department, where patients could be tagged and tracked as they enter and leave each station within the A&E department. This is so as the current time-stamps in the data are unreliable, due to human error as previously discussed, and this would then possibly simplify the way that these timestamps are obtained.

In this, our team also acknowledges that operationally, such an implementation would have challenges especially where situations are life-threatening and the medical staff's priorities take precedence over certain processes. Thus, we propose that this could be explored among the P3 patients, as per the scope of our project, as these patients suffer from conditions and ailments that are non-life threatening. Insights from this group will definitely have knock-on benefits for the other patients in the A&E department and we feel that this is a possible avenue for the hospital to look into.

8. CONCLUSIONS

This paper has discussed the analysis into the various predictors that could affect a patients' length of stay in the emergency department of a local hospital. More specifically, we looked at the P3 patients, where their conditions were non-life threating, and were at the risk of long waiting times due to the low priority of treatment.

We have looked at a few key predictors based on the data obtained, and this included areas such as the number of re-entries, the type of tests ordered, the type of laboratory tests ordered, the results of laboratory tests and time series related factors. In doing so, we have used tools such as JMP Pro to carry out statistical analysis such as ANOVA and Kruskal Wallis to compare different groups of patients within the dataset.

From our analysis, we have found that the predictor that shows the most significant trends are those that are related to the time series within the A&E, and this was done through heat-map visualization. The type of laboratory test also showed results based on the heat-map charted out, but using decision tree modelling and linear regression to verify and understand the validity of these results, they do not prove to have a high predictive ability.

Lastly, using an overall heat map, we have noticed that the factors that relate to time are more significant and impactful than that of the types of individual tests. This corroborates with our hypothesis that the LoS has been largely affected by the aggregated values that we have examined, and thus a large part of a patient's waiting time is being decided by the state of the A&E and the queues in front of him, rather than his actual situation.

Throughout this project, our team faced the limitation of the use of an aggregated LoS as a response variable. We feel that this could have affected our findings, and thus a clearer picture of the individual portions within the LoS would definitely aid in a subsequent understanding and analysis of the system, where it is not just determined by the supply factors, as shown in our final decision tree model, but also by the demand factors, which will encompass the needs of the patients. Therefore, our group recommends improved data collection systems such as RFIDs to improve the quality and specificity of data for subsequent projects.

9. ACKNOWLEDGEMENTS

The authors wish to thank Singapore General Hospital, especially to Dr. Lam Shao Wei, Mr. Nguyen and Ngoc Hoang Long Francis for guiding and working closely with us. We also like to thank Dr. Tan Kar Way, our sponsor for giving us this opportunity to work on this project.

10. REFERENCES

Jones P. and Schimanski K. (2010). EMA. *The four hour target to reduce emergency department 'waiting time' : A systematic review of clinical outcome.* Retrieved February 5 2015, from http://www.medschool.lsuhsc.edu/emergency_medicine/docs/4%20hour%20rule.pdf

L. Mayhew & D. Smith (2007). Using queuing theory to analyse the Government's 4-h completion time target in Accident and Emergency departments.

National Audit Office (2013). Department of Health. *Emergency admissions to hospital: managing the demand.* Retrieved February 2015, from http://www.nao.org.uk/wp-content/uploads/2013/10/10288-001-Emergency-admissions.pdf

National Library for Health. (June, 2006). Health Management Specialist Library: Management Briefing. *Achieving and Maintaining the 4 hour Emergency Access Target.* Retrieved February 5 2015, from http://www2.warwick.ac.uk/fac/med/research/hsri/emergencycare/research/edorganisation/briefing_4htarget__20sept06.pdf

Pang M. (April 2, 2013). The Straits Times. *A&E units flooded with non-emergency cases.* Retrieved on 23 February 2015, from http://www.sgh.com.sg/about-us/newsroom/News-Articles-Reports/Pages/ae-units-flooded-emergency-cases.aspx

Tan K. W. and Lau H. C. (2013). *Improving Patient Length-of-Stay in Emergency Department Through Dynamic Queue Management* from http://ink.library.smu.edu.sg/cgi/viewcontent.cgi?article=2934&context=sis_research

APPENDIX A

We were given three months' worth of data by SGH, which included two different files, namely Emerge Case file and Computerized Patients Order Entry (CPOE) file.

Emerge Case File

Metadata Table 1:

Field	Description
Visit ID	A unique number that each patient is given upon entering the A&E per visit
Account Number	A unique number that each patient is given within the system
Registration Date	The date and timestamp recorded when a patient enters the A&E department
Triage Date	The date recorded when a patient entering the Triage in the A&E department
Triage Category	Category assigned to a patient based on severity level
Time of Attendance	Timestamp recorded when a patient enters the A&E department
Chief Complaint	The symptoms of a patient as recorded by doctor
Primary Diagnosis	The ICD9 code assigned to a patient based on their symptoms
Disposition	
Disposition Time	Timestamp of patient's disposition
Diagnosis Recorded Time	Timestamp of patient's diagnosis

We added new columns to find out the LOS. In general, ***LOS = Disposition Time – Registration Date***

Metadata Table 2:

Field	Derived From	Description
Month	Registration Date	The month when a patient enters the A&E
Week No.	Registration Date	The week of the year when a patient enters the A&E
Day of the Week	Registration Date	The day when a patient enters the A&E
Period of the Day	Registration Date	The period of the day when a patient enters the A&E dept. It is classified by the following categories: Early – 0000hrs to 0559hrs Morning – 0600hrs to 1159hrs Afternoon – 1200hrs to 1759hrs Night – 1800hrs to 2359hrs
Hour of Entry	Registration Date	The hour (rounded down to the nearest hour) when a patient enters the A&E
Time Taken	Registration Date, Disposition Date	The total time that a patient spend in the A&E dept.
Minutes	Registration Date, Disposition Date	The total time (in minutes) that a patient spend in the A&E dept

This file was a log of the patients entering the accident and emergency department (A&E) for the months of January to March in 2013. This comprised of 37,255 unique data points over the three month time span, with the various fields as shown in *Metadata Table 1*.

From this data given, we further focused on the various areas that we were specifically looking at.

Firstly, as our research is mainly based on patients which are in less life-threatening conditions, as stated in our proposal, our scope will only include patients that are classified within the P3 and P3F triage levels. This first step of filtering resulted in a removal of 17,331 data points, leaving us with 19,924 data points to work with.

Next, we then checked the values of the LOS, which is our main response variable. This is derived from the calculation of the length of time between the time of registration at the A&E and the disposition time, which is the time that the patient leaves the A&E department. Upon checking through the values obtained, 6 data points were observed to have negative values, thus these were also removed from the dataset. This left us with 19,918 data points.

Lastly, to further refine the dataset that we are working with, we also noticed that upon inspection of the values given, there was a group of patients that were deemed to have reneged, which also means to have run away- thus avoiding the full treatment within the A&E department, falling into the category of Left Without Being Seen (LWBS). As these cases are clearly incomplete, the LoS would then not be representative of the patients and thus we have also removed all

instances of this disposition, which amounted to 330 data points- leaving us with a final number of **19,588 data points** over three months.

Computerized Patients Order Entry (CPOE) File

Using Account Number as unique key, we matched it against the Account Number in Emerge file to identify the type of test ordered for each patient and the test results in a bid to merge the two files and to match all attributes of the patient.

Metadata Table 3:

Field	Description
Account Number	A unique number that each patient is given within the system
Test Ordered	The test ordered by doctor
Test Code	Category of test • Blood products • Cardiovascular • Endoscopy Center • Laboratory • Medication (Non-parenteral) • Medication (Parenteral) • Nuclear Medicine • Obstetrics Gynaecology • Operating Theatre • Pharmacy • Radiology • Unknown
Test Requested Date Time	Timestamp recorded when the test is ordered
Value	Test results for each patient

As we filtered the data by test code, a key finding is that although there are many categories of tests, the distribution of patients that have undertaken these tests is very varied.

For the number of patients that underwent tests classified under Nuclear Medicine, Endoscopy, Other Non-Med Supplies, there is only one record for each test type over the three months for patients at the P3 triage level. As a result, we have decided to remove these categories.

Aggregated Data File

After merging the data into one data file, we have also done checks on the re-entrant data, as outlined earlier. Upon consultation with the hospital, we learnt that the hospital classifies tests ordered for the same patient during the same visit, but more than 30 minutes apart as a way to identify if patients have re-entered the system. With this information, we used the time of tests ordered in the CPOE file to derive the number of patients' re-entry into the system, adding this as a variable within our dataset.

After which, we did validation checks based on the Emerge file, to check if the re-entry data is reliable. To do so, we multiplied the re-entry number for each patient by 30 min, and used that as the baseline for the LoS of each patient. In other words, if a patient has two re-entries, his LoS should be longer than 1 hour (2* 30 min) at the very least as the tests ordered are half an hour apart. If the LoS is less than this minimum baseline time, the entry is then deemed as no good.

The table below shows the data validation results.

Month	**Good**	**No Good**	**N/A**	**Total**
January	6354	114	22	6488
Feb	5439	150	15	5602
Mar	6106	115	23	6242

As can be seen, most data corroborates with the basic checks carried out, with the percentage of no good entries ranging from between 2-3% per month. This is a small percentage and thus, we have removed the data is “no good” as this would affect the accuracy of our findings.

APPENDIX B

Exclusion analysis for Re-entry Patients

1. All patients in the three months (Jan, Feb, Mar) in the ER (n = 37255)
    a. P1 Patients (n = 6119)
    b. P2 Patients (n = 11166)
    c. P3 Patients (n = 19924)
    d. P4 Patients (n = 47)

2. P3 Patients (n = 19924)
    a. LWBS (n = 461)
    b. Preliminarily deemed as unreliable data points (n = 251)
        i. Patients that have consultation times less than the intervals between ordered times

3. All P3 Patients (n = 19212)
    a. 1 test lab (n = 937)
    P3 patients who are being ordered only laboratory tests upon visit to A&E
    b. 1 test radio (n = 573)
    P3 patients who are being ordered only radiology tests upon visit to A&E
    c. 1 test non-parenteral (n = 6816)
    P3 patients who are being ordered only non-parenteral medications
    d. 2 tests: lab & non-parenteral (n = 1392)
    P3 patients who are ordered only lab tests and non-parenteral medication
    e. 2 tests: radio & non-parenteral (n = 4206)
    P3 patients who are ordered radiology tests & non-parenteral medication
    f. 2 test: lab & radio (n = 1951)
    P3 patients who are ordered laboratory tests and radiology tests
    g. All 3 tests (n = 2001)
    P3 patients who are ordered all tests
    h. None (n = 1336)
        i. P3 Patients who has no tests ordered and not reflected in the CPOE list (n = 1315)
        ii. P3 patients who are ordered only tests that are not reflected under lab, radio or non-parenteral (n=21)

4. 1 Test Lab (n=937)
    i. 1 re-entry (n=771)
        i. Patients who have no 30 minutes interval between tests ordered
        ii. 97 patients with minutes = 0
        iii. 219 patients with negative minutes value (RequestedDTM > Disposition Time)
        iv. 455 patients with minutes < 30
    j. 2 re-entry (n=148)
        i. Patients who have 1 >30 minutes interval between tests ordered.
    k. 3 or more re-entry (n= 18)
        i. Patients who have more than 1 >30 minutes interval between tests ordered

5. 1 Test radio (n=573)
    l. 1 re-entry (n=514)
        i. Patients who have no 30 minutes interval between tests ordered
        ii. 139 patients with minutes = 0
        iii. 77 patients with negative minutes value (RequestedDTM > Disposition Time)
        iv. 298 patients with minutes < 30 minutes
    m. 2 re-entry (n=51)
        i. Patients who have 1 >30 minutes interval between tests ordered.
    n. 3 or more re-entry (n= 8)

i. Patients who have more than 1 >30minutes interval between tests ordered

6. 1 Test Non-Parenteral (n=6816)
   For all patients who are given non-parenteral tests,
   o. 0 re-entry (n = 6550)
      i. Patients who has a disposition of less than 16 minutes (97.5 percentile of all patients) after being ordered non-parenteral treatment
      ii. 1872 patients with minutes = 0
      iii. 1982 patients with negative minutes value (RequestedDTM > Disposition Time)
      iv. 2696 patients with minutes < 30 minutes
   p. 1 re-entry (n=771)
      i. Patients who has a disposition time of more than 16 minutes after being ordered non-parenteral treatment
   q. 2 re-entry (n=148)
      i. Patients who have 2 >30 minutes interval between tests ordered.
   r. 3 or more re-entry (n= 18)
      i. Patients who have more than 2 >30minutes interval between tests ordered

7. 2 Tests Lab & Non-Parenteral (n=1392)
   s. 1 re-entry (n = 1266)
      i. Patients who have no 30 minutes interval between tests ordered.
      ii. 210 patients with minutes = 0
      iii. 355 patients with negative minutes value (RequestedDTM > Disposition Time)
      iv. 701 patients with minutes < 30 minutes
      v. Check if non-parenteral is ordered after lab test. If yes, time between ordered time and disposition time must be less than 16 minutes. Else, considered 2 re-entry
   t. 2 re-entry (n = 109)
      i. Patients who have at least 1 >30 minutes interval between test ordered.
      ii. Check if non-parenteral is ordered after lab test. If yes, time between ordered time and disposition time must be more than 16 minutes. Else, considered 1 re-entry
   u. 3 more re-entry (n= 17)
      i. Patients who have more than 2 >30minutes interval between tests ordered

8. 2 Tests radio & Non-Parenteral (n=4206)
   v. 1 re-entry (n = 3954)
      i. Patients who have no 30 minutes interval between tests ordered.
      ii. 355 patients with minutes = 0
      iii. 1899 patients with negative minutes value (RequestedDTM > Disposition Time)
      iv. 1700 patients with minutes < 30 minutes
      v. Check if non-parenteral is ordered after radio test. If yes, time between ordered time and disposition time must be less than 16 minutes. Else, considered 2 re-entry
   w. 2 re-entry (n = 211)
      i. Patients who have at least 1 >30 minutes interval between test ordered.
      ii. Check if non-parenteral is ordered after radio test. If yes, time between ordered time and disposition time must be more than 16 minutes. Else, considered 1 re-entry
   x. 3 more re-entry (n= 41)
      i. Patients who have more than 2 >30minutes interval between tests ordered

9. 2 Test lab & radio (n = 1951)
   y. 1 re-entry (n=1124)
      i. Patients who have no 30 minutes interval between tests ordered
      ii. 57 patients with minutes = 0

iii. 291 patients with negative minutes value (RequestedDTM > Disposition Time)
iv. 776 patients with minutes 1 > 30 minutes.

z. 2 re-entry (n=643)
i. Patients who have 1 >30 minutes interval between tests ordered.

aa. 3 or more re-entry (n= 184)
i. Patients who have more than 1 >30minutes interval between tests ordered

10. All three tests (n = 2001)
bb. 1 re-entry (n=1558)
i. Patients who have all 3 tests ordered without any 30 minutes interval between test ordered
ii. 489 patients with minutes = 0
iii. 687 patients with negative minutes value (RequestedDTM > Disposition Time)
iv. 582 patients with minutes < 30 minutes.

cc. 2 re-entry (n = 211)
i. Patients who have 1 >30 min interval between tests ordered

dd. 3 or more re-entry (n = 232)
i. Patients who have more than 1 >30 min interval between tests ordered