1. **Overview**

LGI-Flag 3.1 has been deployed at Geisinger IT systems. This plan is performed to evaluate the LGI-Flag 3.1 performance on the Geisinger population. This analysis will be run during the silence mode period (where LGI-Flag 3.1 is operated for 6 weeks without interacting with patients).  In the meantime, the existing LGI-Flag 3.0 will continue to work

1. **ANALYSIS OBJECTIVES**

The purpose of the analysis is to:

* Characterize Geisinger population with the additional age group
* Evaluate the input population to verify that anomalous sub-populations aren’t included or something in the integration pipeline caused an issue
* Evaluate the input signals and values to understand population characteristics in comparison to the Geisinger population in the intersecting age groups;
* Verify that the expected Positivity Rate of 3% is met

## **Analysis Plan**

* 1. Data

The data that will be used for the analysis is data that Geisinger will pull as a one-time query that will pull at least 1500 patients aged 45-75, non-compliant to the Geisinger screening plan who received a new CBC result during the silence run. For each patient, all CBC from the last 3 years that exist within the system will also be provided.

The data set will include the entire input data fed into LGI Flag in staging environment with flag results and feature matrix for each patient. The staging environment will be configured to store also the feature matrix in order to perform deeper tests on the data. Data from the deployed AlgoMarker LGI-Flag 3.0 will be also fetched without any change – input date and output the contains the flags for comparison.

The data will be compared to the MHS dataset which was used to train and validate LGI-Flag – a dataset which includes over 1M patients with data between the years 2001-2015. This dataset will be referred to throughout this document as the ‘Comparison Dataset’.

The dataset will be also compared to the existing LGI-Flag 3.0.

* 1. Analysis Method

The Geisinger data will be loaded Geisinger research environment for the data analysis.

* 1. *Geisinger Population Analysis*

1. Compare the most important features distribution between Geisinger LGI-Flag 3.1, , Geisinger LGI-Flag 3.0 and comparison dataset. Those are the features to compare:
   1. Age
   2. Sex
   3. Last Hgb value in the last year
   4. Last MCH value in the last year
   5. Hgb slope in last 1000 days
   6. MCH slope in last 1000 days
   7. #CBC in last 2 years

Calculate mean, standard deviation for each feature. Calculate relative entropy between the distributions. Analyze the main differences and asset the excepted performance for this population using propensity model by matching those features.

1. Compare the score distribution by mean, standard deviation and plot the distribution graph.
   1. Calculate cutoffs for each positive rate between 0.5% till 5% by 0.5% gaps and calculated the expected number of flags in each cutoff.
   2. Compare score distribution in comparison dataset after matching to the score distribution
2. Coverage sanity test for known risk groups:
   1. Define a cohort of anemic patients with decreasing Hgb of at least 1 point in last year that are aged 60-75. Calculate the prevalence of this group and the coverage, RR (Relative risk) of the model in score cutoff of 1,3,5,10% on this group.
3. Apply the model directly on input data from LGI-Flag 3.0 and validate who got flagged.
4. Compare AlgoMarker scores from staging environment LGI-Flag 3.1. Date shifting of +-14 days (about 2 weeks) can affect slightly the prediction on December/January due to age shift.
5. Compare direct usage of the model on input data from LGI-Flag 3.1 and compare to the scores from this AlgoMarker

1. **Overview**

Lung-Flag has been deployed at Geisinger IT systems. This plan is performed to evaluate the Lung-Flag performance on the Geisinger population. This analysis will be run during the silence mode period (where Lung-Flag is operated for 6 weeks without interacting with patients).

1. **ANALYSIS OBJECTIVES**

The purpose of the analysis is to:

* Characterize Geisinger population with the additional age group
* Evaluate the input population to verify that anomalous sub-populations aren’t included or something in the integration pipeline caused an issue
* Evaluate the input signals and values to understand population characteristics in comparison to the Geisinger population in the intersecting age groups;
* Verify that the expected Positivity Rate of 5% is met

## **Analysis Plan**

Data

The data that will be used for the analysis is data that Geisinger will pull as a one-time query that will pull at least 1500 patients aged 50-80, non-compliant to the Geisinger screening plan are eligible by the USPSTF criteria for lung cancer screening. For each patient, all AlgoMarker data from the last 5 years that exist within the system will be provided.

The data set will include the entire input data fed into Lung Flag in staging environment with flag results and feature matrix for each patient. The staging environment will be configured to store also the feature matrix in order to perform deeper tests on the data.

The data will be compared to the Geisinger dataset which was used to validate Lung-Flag in the retrospective analysis. This dataset will be referred to throughout this document as the ‘Comparison Dataset’.

* 1. Analysis Method

The Geisinger data will be loaded Geisinger research environment for the data analysis.

1. Test how many diagnosis codes ICD9/ICD10 doesn’t exist and improve the dictionaries when needed.
2. Compare the most important features distribution between Geisinger Lung-Flag and comparison dataset. Those are the features to compare:
   1. Age
   2. Sex
   3. Smoking; Smoking\_Years, Pack\_Years, Smoking\_Status, Time\_Since\_Quit
   4. WBC value in last year
   5. Platelets value in last year
   6. BMI value in last 3 years
   7. # of ICD10:786 diagnosis in the last 3 years
   8. Race, Hispanic – Y/N
   9. # smoking registration records in last 2 years

Print table with differences between mean, std for each feature (and KLD between distribution of features). Analysis of population difference by model – what is the excepted performance for this population using propensity model on those most important features.

Calculate mean, standard deviation for each feature. Calculate relative entropy between the distributions. Analyze the main differences and asset the excepted performance for this population using propensity model by matching those features.

1. Compare the score distribution by mean, standard deviation and plot the distribution graph.
   1. Calculate cutoffs for each positive rate between 0.5% till 5% by 0.5% gaps and calculated the expected number of flags in each cutoff.
   2. Compare score distribution in comparison dataset after matching to the score distribution
2. Coverage sanity test for known risk groups:
   1. Define a cohort of COPD patients aged 70-75 with at least 30 pack years and at least 30 smoking years. Define COPD patient as a patient with at least 1 COPD diagnosis in last 3 years. Calculate the prevalence of this group and the coverage, RR (Relative risk) of the model in score cutoff of 1,3,5,10% on this group
3. **Overview**

AAA-Flag has been deployed at Geisinger IT systems. This plan is performed to evaluate the AAA-Flag performance on the Geisinger population. This analysis will be run during the silence mode period (where AAA-Flag is operated for 6 weeks without interacting with patients).

1. **ANALYSIS OBJECTIVES**

The purpose of the analysis is to:

* Characterize Geisinger population with the additional age group
* Evaluate the input population to verify that anomalous sub-populations aren’t included or something in the integration pipeline caused an issue
* Evaluate the input signals and values to understand population characteristics in comparison to the Geisinger population in the intersecting age groups;
* Verify that the expected Positivity Rate of 9% is met

## **Analysis Plan**

Data

The data that will be used for the analysis is data that Geisinger will pull as a one-time query that will pull at least 1500 patients aged 65-75, smokers and non-compliant to the Geisinger screening plan for AAA. For each patient, all AlgoMarker data from the last 5 years that exist within the system will be provided.

The data set will include the entire input data fed into AAA Flag in staging environment with flag results and feature matrix for each patient. The staging environment will be configured to store also the feature matrix in order to perform deeper tests on the data.

The data will be compared to the Geisinger dataset which was used to validate AAA-Flag in the retrospective analysis. This dataset will be referred to throughout this document as the ‘Comparison Dataset’.

* 1. Analysis Method

The Geisinger data will be loaded Geisinger research environment for the data analysis.

1. Test how many diagnosis codes ICD9/ICD10 doesn’t exist and improve the dictionaries when needed.
2. Test how many drug codes ATC/RX\_NORM doesn’t exist and improve the dictionaries when needed.
3. Compare the most important features distribution between Geisinger Lung-Flag and comparison dataset. Those are the features to compare:
   1. Age
   2. Sex
   3. Smoking: Smoking\_Years, Pack\_Years, Smoking\_Status, Time\_Since\_Quit
   4. Last Height
   5. # of ICD10:I25 diangoses in last 3 years
   6. RDW value in last year
   7. HDL value in last year
   8. eGFR\_CKD\_EPI value in last year
   9. Platelets value in last year
   10. # smoking registration records in last 2 years

Print table with differences between mean, std foreach feature (and KLD between distribution of features). Analysis of population difference by model – what is the excepted performance for this population using propensity model on those most important features.

Calculate mean, standard deviation for each feature. Calculate relative entropy between the distributions. Analyze the main differences and asset the excepted performance for this population using propensity model by matching those features.

1. Compare the score distribution by mean, standard deviation and plot the distribution graph.
   1. Calculate cutoffs for each positive rate between 0.5% till 5% by 0.5% gaps and calculated the expected number of flags in each cutoff.
   2. Compare score distribution in comparison dataset after matching to the score distribution
2. Coverage sanity test for known risk groups:
   1. Define a cohort of male patients aged 70-75 with at least 30 pack years and at least 30 smoking years. Calculate the prevalence of this group and the coverage, RR (Relative risk) of the model in score cutoff of 1,3,5,10% on this groups