Applying Advanced Analytics to Improve Tagrisso Adherence Among Members with Non-Small Cell Lung Cancer

Humana-Mays 2023 Case Competition

Humana

TEXAS A&M UNIVERSITY
Mays Business School

Table of Contents

1. Introduction

1.1 Background

Tagrisso is a drug used to treat Epidermal Growth Factor Receptor-mutated (EGFRm) Non-Small Cell Lung Cancer (NSCLC). Lung cancer is the leading cause of death among men and women with cancer, accounting for 20% of all deaths from cancer. Of patients with lung cancer, 80-85% suffer from NSCLC, while the remaining suffer from small cell lung cancer. Among patients with NSCLC, approximately 10-15% of patients in the US have EGFRm NSCLC, disproportionately affecting women, people of East Asian descent, and the senior population[1]. This yields a large market for Tagrisso, which saw worldwide sales of $5.44 billion in 2022[2]. The cost of a six-month treatment for Tagrisso varies based on a number of factors, but can cost insurers as much as $9,000 per month for each member[3].

The introduction of Tagrisso has drastically improved survival rates among patients with EGFRm NSCLC. The disease is often fatal when treated with chemotherapy, with a four-year survival rate among stage II-IIIA patients of just 29%. However, when Tagrisso is used as indicated, 70% of patients will be alive and disease-free after four years. For this reason, it is vital that patients remain on Tagrisso in spite of relatively minor side effects. Additionally, the drug is extremely effective at preventing disease recurrence, with 3-year recurrence of just 2%, compared to 13% for chemotherapy[1].

Tagrisso is known to have a wide array of side effects, including nausea, fatigue, constipation, and diarrhea, which can cause patients to become non-adherent to their prescribed treatment regimen[4]. However, these side effects are far outweighed by the benefits that Tagrisso provides both in treating EGFRm NSCLC and in preventing recurrence. Encouraging adherence in spite of the side effects has been a challenge that physicians and insurers have faced since the introduction of Tagrisso.

1.2 The Humana-Mays Healthcare Analytics Case Competition

1.2.1 The Business Issue

Lung cancer rates are highest among people 65 and older, with incidence rates ranging from 198 per 100,000 people for 65-year olds to 334 per 100,000 people among those aged 80[5,6]. With over 5 million Medicare Advantage members[7] (18% market share), we estimate that Humana has approximately 13,300 members diagnosed with lung cancer annually, with about 1,100 of those diagnosed with EGFRm NSCLC[1]. At a cost of $9,000 per member per month, an average 6-month treatment of Tagrisso for these members totals nearly $60 million in drug costs alone. This represents a tremendous financial burden, but one that is only exacerbated by members becoming non-adherent. When members become non-adherent, they increase the likelihood of disease progression and recurrence, which first and foremost impacts their health, but also increases long-term costs to both Humana and the member. Supporting members in adherence to their Tagrisso regimen will both improve health outcomes and reduce long-term costs of care associated with disease progression and recurrence.

The development and implementation of a predictive model to predict the likelihood of an adverse drug event (ADE) and subsequent Tagrisso non-adherence represents a strategic investment in Humana's mission to provide high-quality, patient-centric care. As the domestic health system continues to make progress towards the “Triple Aim” of improved health outcomes, reduced spending, and improved member experience, the use of data analytics to tailor strategies will position Humana at the forefront of Value-Based Care. With a commitment to data-driven decision-making, Humana can allocate resources more efficiently, presenting an opportunity for cost savings but also making it possible to respond to members’ needs proactively.

The use of a predictive model goes beyond mere data analytics. It is meant to inform initiatives to proactively address the leading causes of Tagrisso non-adherence and to enable rapid response to detected events that could be a precursor to non-adherence. Ultimately, the model is key to supporting Humana’s progress towards the Triple Aim.

Improved Health Outcomes
Tagrisso is a critical medication for non-small cell lung cancer, and its efficacy is closely tied to treatment adherence. There are a number of side effects that are known to be associated with Tagrisso. After experiencing these side effects, members may decide to stop taking Tagrisso. However, the severity of these side effects does not compare to the harm caused by not taking Tagrisso. Predicting non-adherence allows healthcare teams to optimize patient care plans and be responsive before members become non-adherent. By being proactive in education and outreach, Humana can help members remain adherent and ultimately have better overall health outcomes.

Reduced Spending
Non-adherence with Tagrisso treatment often results in avoidable healthcare costs, including emergency room visits and hospitalizations. By tailoring interventions to the specific needs of the member, the predictive model can lead to substantial short and long-term cost savings for Humana in the form of reduced medical claims. Additionally, by targeting interventions to members most likely to become non-adherent, the model will ensure an efficient allocation of resources by directing them towards members identified to be most in need of proactive support.

Improved Member Experience
By predicting which members are at risk of becoming non-adherent, healthcare providers can proactively engage with these individuals. Timely interventions can be tailored to meet specific member needs, creating an atmosphere of support. This could enhance the overall member experience and significantly improve member satisfaction with Humana's programs and services.

1.2.2 Key Performance Indicators

Two primary measures were used to evaluate the effectiveness of the model:

1. **Receiver Operating Characteristic (ROC) Curve** - We evaluated the performance of the model using the Area Under the Curve (AUC). A higher AUC indicated a better model performance. An example of an ROC Curve is shown in **Figure 1**.
2. **Confusion Matrix** - For any given threshold, the confusion matrix was used to visualize the sensitivity, specificity, and accuracy of the model to evaluate its ability to detect true positive and true negative flags. An example is shown in **Figure 2**.
3. **Disparity Score** - For each sensitive feature (race and sex), the disparity score was calculated. An average disparity score greater than 0.9 indicated an unbiased and fair model.

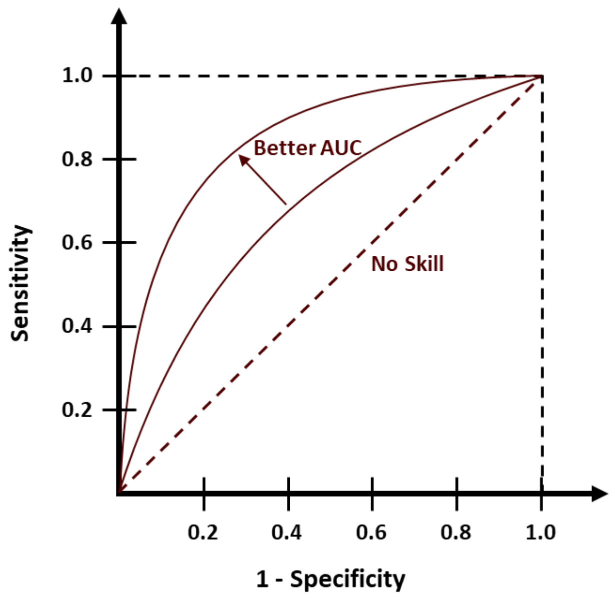


Figure 1. Example ROC Curve

		Predicted Tagrisso ADE & Nonadherence: Low (0)	Predicted Tagrisso ADE & Nonadherence: High (1)	
Actual Tagrisso ADE & Nonadherence	Low (0)	True Negative	False Positive	Total Actual Low
	High (1)	False Negative	True Positive	Total Actual High
		Total Predicted Low	Total Predicted High	Total Observations

Figure 2. Example of a confusion matrix

We evaluated our recommendations through the lens of the Triple Aim:

1. **Improved Health Outcomes** - Recommendations should improve health outcomes for Humana's EGFRm NSCLC member population by reducing cancer progression and improving overall survival.
2. **Reduced Spending** - Recommendations should lead to long-term cost savings in the management of EGFRm NSCLC cases, potentially by reducing hospitalizations, emergency room visits, and medication-related expenses.
3. **Improved Member Satisfaction** - Recommendations should enhance member satisfaction with Humana's programs, contributing to a positive experience and improved engagement.

2. Data Overview

2.1 Model Objective

We created a model with the objective to proactively identify Humana members taking Tagrisso who are at risk of experiencing adverse drug events and subsequently becoming non-adherent to their treatment. While the model was created retrospectively using demographic and claims-based data from Humana members taking Tagrisso in the past five years, we placed emphasis on using features that could in the future be used to identify members prospectively and in real-time who are most likely to be at risk of having an ADE and becoming non-adherent.

Thus, an ideal model for this application serves two purposes: to identify factors that can be overcome with prospective outreach and to identify events that commonly serve as precursors to non-adherence. By emphasizing these two key purposes, our predictive model will serve as a valuable tool for not only predicting non-adherence, but also informing proactive initiatives to effectively increase adherence rates among Humana members.

1. Prospective Outreach to At-Risk Populations
The first key application of our predictive model is to determine what specific characteristics make a Humana member more likely to become non-adherent to Tagrisso treatment. By analyzing demographics, previous healthcare utilization, and prior conditions, we can identify individuals at higher risk of discontinuing their treatment. This prospective approach allows for targeted outreach and support to populations most likely to become non-adherent.

2. Identifying Precursors to Non-Adherence
The second key application of our predictive model is to identify common events that are likely to precede non-adherence with Tagrisso treatment. By extending the model to analyze claims data in near-real-time, our model could detect early warning signs. These might include documentation of concerning ICD-10 diagnosis codes or ADEs, emergency room (ER) utilization, or changes in prescription fill patterns. As soon as such events are identified, existing case management teams can initiate outreach to address concerns and provide education to improve Tagrisso adherence.

2.2 Analytical Tools Used

We primarily used Excel, Tableau, Python, and JMP for our analysis and recommendations. Our exploratory analysis of the data was facilitated by Excel and Tableau, which enabled us to gain valuable insights into the structure and characteristics of both the training and holdout data. We used Python as our primary engine for data preprocessing and model fitting. For the data preprocessing step, Python was used to clean the data and derive aggregated variables from the claims-level data. For the model fitting step, Python was used to create and tune the hyperparameters of different models. We used JMP to perform member segmentation, identifying distinct member groups within the data which allowed us to create more targeted recommendations.

2.3 The Dataset

Before creating the model, we sought to gain an understanding of the data with which we had to work. The data was partitioned into a training dataset and a holdout dataset. It is important to note that the training dataset and holdout dataset varied in the average number of medical claims per member (30.1 per member in the training set versus 20.4 in the holdout set for the six months following the beginning of treatment) and the average number of prescription claims (17.0 versus 7.9) per member. We had to consider this data mismatch when designing the model.

- Training Dataset: Demographic and target variable information for 1,232 members, including 100,159 medical claim lines and 32,133 pharmacy claim lines.
- Holdout Dataset: Demographic information for 420 members, including 23,232 medical claim lines and 6,669 pharmacy claim lines.

The data was extremely imbalanced, with only 117 members (9.5%) in the training dataset having an ADE and subsequent non-adherence. The dataset consisted of three separate files, the demographic data, medical claims data, and pharmacy claims data.

- Demographic data - member-level data on the therapy start and end date, race, age, sex, disabled flag, low income flag, and target flag.
- Medical claims data - claim-level data including information on the primary and secondary ICD-10 diagnosis codes, date of service, place of service, utilization category, and flags for diagnoses related to common Tagrisso side effects.
- Pharmacy claims data - claim-level data including information on the national drug code, service date, number of days supply dispensed, prescription cost, drug group and class, drug strength, and flags for whether the drug was a generic, maintenance, mail order drug, or drug relating to treating known side effects of Tagrisso.

2.4 Exploratory Data Analysis

We performed exploratory data analysis on MS Excel. For categorical variables including sex, race, disabled flag, and low-income flag, we analyzed the category as a percentage of total population and the percent flagged within that category. **Figure 3** shows the distribution of total population in maroon and that of the flagged population in gray. In the sex category, 66.2% of members were female, but they accounted for 72.6% of the flagged cases. Analysis across races revealed that although white members account for 56.7% the population, they make up 63.2% of flagged cases. On the other hand, the asian population comprised 12.3% total cases but made up only 7.7% of flagged cases.

In the disability indicator category, we noted that 13.7% of the members in the dataset were classified as disabled and 15.4% of the category was flagged. Finally, we analyzed members with low-income subsidies and found that they made up 36.2% of the total population and 36.8% of flagged cases.

We studied age-wise distribution of the members and not surprisingly, older members make up both the greatest proportion of members and flags. The highest proportion of flagged cases came from 70-79 year old members, with the group making up 43.9% of total cases and 49.1% of cases flagged.

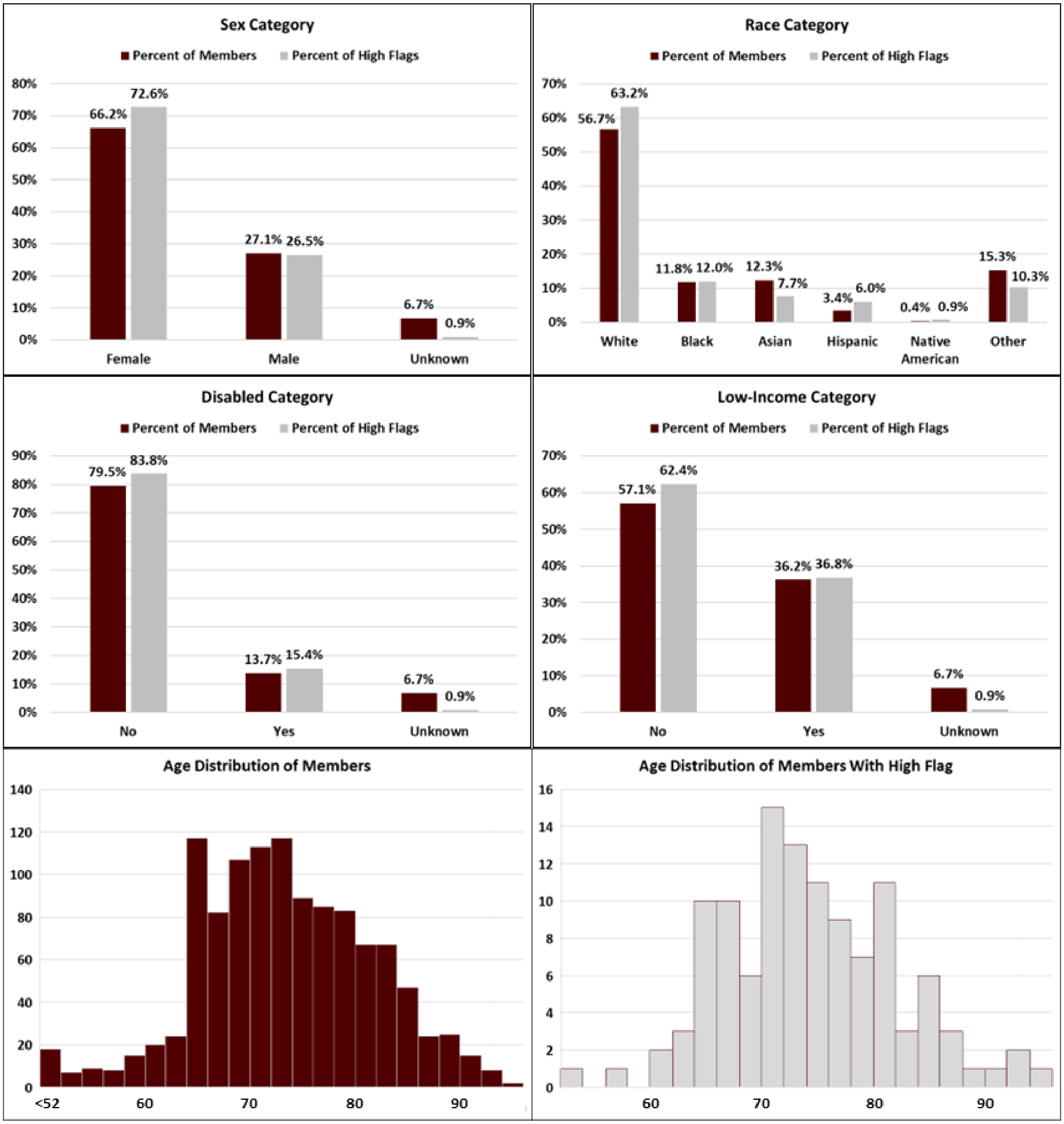


Figure 3. Distribution of total members and flagged members as per various categorical variables and age buckets

3. MODELING

3.1 Data Preprocessing

3.1.1 Handling Features with Null Values

While the medical and prescription claims did not include null values in included features, certain features in the demographic data had a significant number of null values, though never more than 7% of values in the feature column. These features include race, age, sex, disability flag, and low income flag. In categorical variables with an unknown category, null values were assigned to the unknown group (race). In categorical variables without an unknown category they were assigned using mode imputation (sex, disability flag, and low income flag). Continuous variables (age) were imputed using median imputation.

3.1.2 Handling Categorical Features and One Hot Encoding

In the dataset, demographic data was provided as categorical variables. We transformed these variables using one hot encoding, whereby they were converted into *n-1* binary variables, where *n* is the number of categories. **Figure 4** shows a graphic representation of the process.

race_cd
2
1
0
5
2

Transformation to binary variables

race_cd_0	race_cd_2	race_cd_3	race_cd_4	race_cd_5
0	1	0	0	0
0	0	0	0	0
1	0	0	0	0
0	0	0	0	1
0	1	0	0	0

Figure 4. Graphical representation of one hot encoding

3.1.3 Derived Variables

Because the dataset consisted heavily of claims-level data, we derived aggregated variables at the member-level. For each individual, medical and pharmacy claims were filtered to only include claims within one year of the start of therapy. While therapy end date was included in the training set, it was not used because it was not included in the holdout set and would not enable real-time identification of members likely to have an ADE and discontinue treatment. From the one year of claims, the following data were calculated and appended to the test data set at the member-level.

In total, 76 features were used for the initial creation of the model. 12 of these features were from one hot encoding of demographic information, 44 were derived from prescription claims data, and 20 were derived from medical claims data.

- Medical claim-derived variables -
  - Number of medical claims by place of service
  - Number of documented ADEs and ICD-10 diagnosis codes for known Tagrisso side effects
  - Flag documenting whether member was documented as having disease in various disease diagnosis categories
- Pharmacy claim-derived variables
  - Tagrisso Fill Variables (days filled of Tagrisso, number of fills)
  - Number of fills by drug group
  - Number of fills for maintenance and non-maintenance drugs
  - Number of fills for generic and branded drugs
  - Number of fills fulfilled through mail order and non-mail order

3.1.4 Improving Generalizability of the Model

For any model to be most effective, it must be generalizable to unseen data. As noted in Section 2.3, the data is severely imbalanced and there is a mismatch between the training and holdout data, which required us to perform various processing steps to improve the generalizability of the model.

To address the imbalanced data, where only 9.5% of members were flagged for the target variable, we implemented synthetic minority oversampling technique (SMOTE), whereby synthetic samples were created to balance the training dataset with high and low target flags. These synthetic samples were created to match the distribution of the data points that had high flags. In addition to implementing SMOTE, we also kept the imbalanced data in mind in the selection and implementation of our model. We prioritized ensemble models that generally perform well with imbalanced data and utilized cross-validation and regularization techniques to prevent overfitting.

The prioritization of ensemble models and use of cross-validation and regularization techniques were also key to preventing problems caused by data mismatch between the training and holdout datasets. By combining the predictions of multiple models, ensemble models reduce overfitting to a particular dataset and generalize better on unseen data. Likewise, cross-validation and regularization reduce overfitting by evaluating performance across multiple subsets of data and penalizing large feature weights to prevent overly complex models, respectively. We also implemented standardization scaling of continuous variables to transform the feature to have a mean of 0 and a standard deviation of 1. This can improve the model's ability to understand the relative importance of each variable and improve the performance of regularization.

3.2 Model Selection

We initially created five different types of models - Logistic Regression, K-Nearest Neighbors (KNN), XGBoost, and LightGBM. For each model, data was partitioned into training and test sets with a 70:30 split. The model was chosen based on the AUC of each model. The AUC for the Receiver Operating Characteristic (ROC) Curve of each model, along with the advantages and disadvantages of each method is shown in **Table 1**.

Each model was created using the same input data, but with non-important variables removed so the final set of features varied by model. After tuning each model, the maximum AUC of each model was compared. As shown in **Table 1**, the best model was the LightGBM model, with XGBoost second and KNN third. The LightGBM model also had a faster training speed than XGBoost or LightGBM, allowing quicker tuning of the hyperparameters. From the AUC values, we decided to proceed with the LightGBM model for the final model.

	Logistic	KNN	XGBoost	LightGBM
Test Set AUC	0.622	0.802	0.817	0.865
Advantages[8,9]	Easily Interpretable, fast, low computational costs	Better AUC, Increased Tuning capabilities, can handle complex relationships	High AUC, Fast, Considerable Tuning capabilities, resistant to overfitting	Highest AUC, Fastest, Considerable Tuning capabilities
Disadvantages[8,9]	Low AUC relative to the others, limited in capturing non-linear relationships	Takes an extremely long amount of time, very sensitive to parameters	More difficult to interpret, requires tuning of hyperparameters	More difficult to interpret, requires tuning of hyperparameters

Table 1. AUC achieved using each model and the advantages and disadvantages of each model.

3.3 Model Tuning of LightGBM Model

We optimized the hyperparameters using a grid search across a wide range of parameters, using AUC as the evaluation metric. After zeroing in on the optimal parameters, we performed a more granular grid search across a narrower parameter range.

For each grid search iteration, the model was created in two stages. An initial model was fitted to all 76 features of the training set. Of these 76 features, the model identified and removed features that were not important predictors of a member having an ADE and becoming non-adherent to Tagrisso. The remaining features were used to fit a second model and evaluate it against a test set via five-fold cross validation. **Table 2** shows the tuned hyperparameters that resulted from the grid search. Further evaluation of the final model is shown in the next section.

Model	Tuned Hyperparameters	Average AUC Score on Test Set
Final LightGBM	Task = train boosting = 'gbdt' boosting_type = ‘rf’ objective = 'Binary' num_leaves = 12 learning_rate = 0.05 metric = 'auc' verbose = -1 max_depth = 3 max_bin = 15 feature_fraction=0.5 lambda_l1 = 10 lambda_l2 = 85 num_boost_round=1000 min_child_weight = 0.8 bagging_freq = 20	0.865

Table 2. Tuned hyperparameters and AUC scores for the final model.

3.4 Final Model

The tuned LightGBM model was evaluated against the test set. A ROC curve was created by varying the probability threshold at which predictions are set to 0 or 1 and graphing the resulting sensitivity and 1 - specificity. As seen in the ROC curve in **Figure 5**, the AUC of the model was 0.8650, indicating strong predictive power of the model to predict individuals most likely to have an ADE and become non-adherent to their Tagrisso regimen.

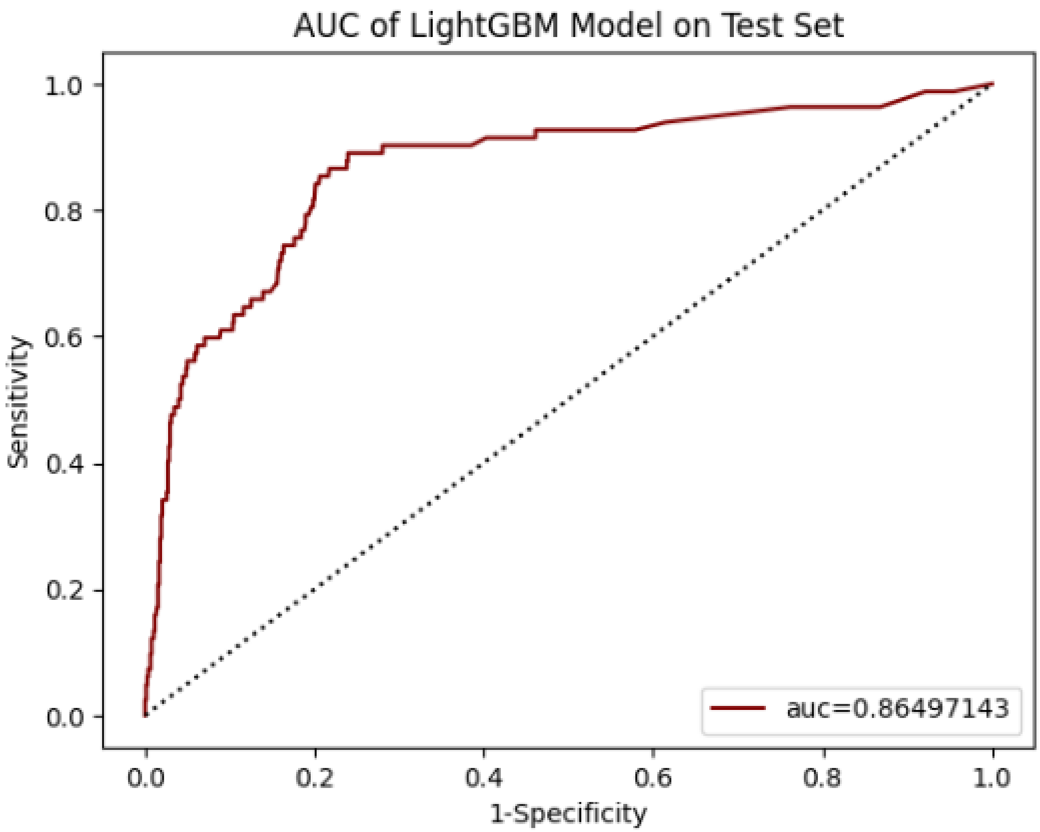


Figure 5. Receiver operating characteristic (ROC) curve of the model predictions of the test set.

The probability density function (PDF) of the model's prediction is shown in **Figure 6**, calculated using kernel density estimation. The blue area shows the probability density of predictions for members who experienced an ADE and subsequently discontinued Tagrisso treatment, while the orange area shows the prediction density for members who did not experience an ADE and discontinue treatment. The PDF indicates strong predictive power if the peak of the blue area is to the right and the peak of the orange area is to the left of the graph. Despite having some overlap between the areas, the density of the blue area is highest at higher probabilities, while the density of the orange area is highest at lower probabilities, indicating that the model is effective. The model predicted an average probability of 0.56 for those that did have an adverse event and discontinued treatment and an average probability of 0.17 for those that did not, further supporting the effectiveness of the model.

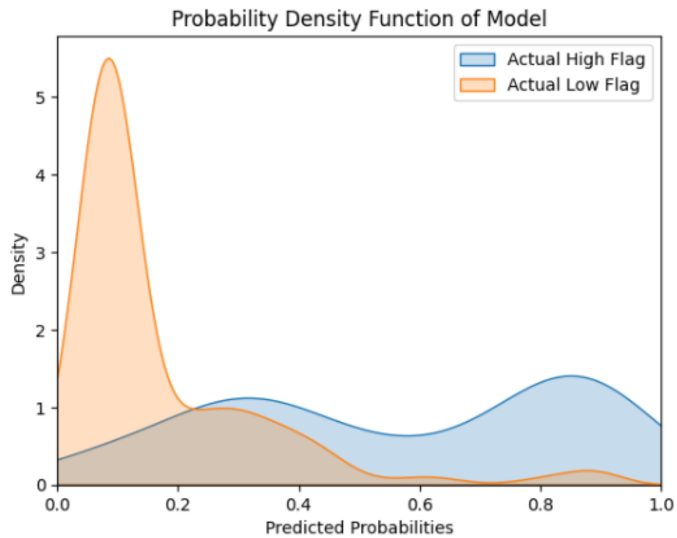


Figure 6. Probability density function of model predictions of the test set.

The probability density function was used to visualize how certain probability thresholds would be expected to impact the sensitivity and specificity of the model. While the exact threshold of choice depends on the cost of missing a true positive and the cost of including a false positive, it is evident that the optimal threshold likely lies between 0.2 and 0.3, based on the left peak of the blue area. The confusion matrix in **Figure 7** shows the confusion matrix of the model with a probability threshold of 0.25. At this threshold, the model has a sensitivity of 86.6%, a specificity of 77.1%, and an overall accuracy of 78.0%.

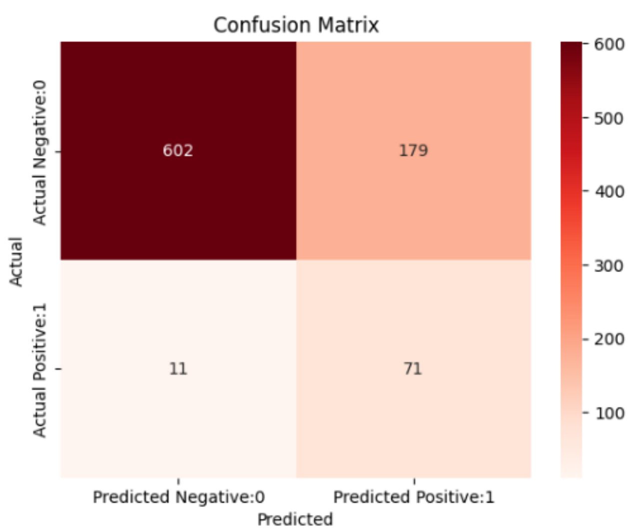


Figure 7. Confusion matrix for predictions of the test set with threshold of 0.25.

We also evaluated the model for fairness using the disparity score, whereby the AUC of the model is evaluated for each separate group of sensitive variables (race and sex). The disparity score of each minority group was calculated as the AUC divided by the AUC of the reference group, defined as white males. The overall disparity score was calculated as the average of the individual disparity scores. The model was considered acceptable if the overall disparity score was greater than 0.9. Our model had a disparity score of 0.934, meaning it was fair to each group and was not significantly biased towards the reference group.

3.5 Feature Importance

Of the 35 features used in the model, the ten most important features identified by the model are shown in **Figure 8**. The features fell broadly into five categories, shown in **Figure 9**. The most important feature was the number of days supply of Tagrisso that a member filled at the pharmacy. A higher number of days supply filled indicated a greater likelihood of adherence to their prescriptions. Six of the other ten most important features had to do with medical utilization and documented diagnosis codes, which were derived from medical claims information. Two of the ten most important variables were derived from pharmacy claims information, including the proportion of prescriptions the member filled via mail order. Demographic information accounted for 15% of total importance in the model, but only a race code of black was in the top ten most important features. The feature importance was used to inform the segmentation and recommendations to address the issue of members experiencing an ADE and subsequently becoming non-adherent to Tagrisso.

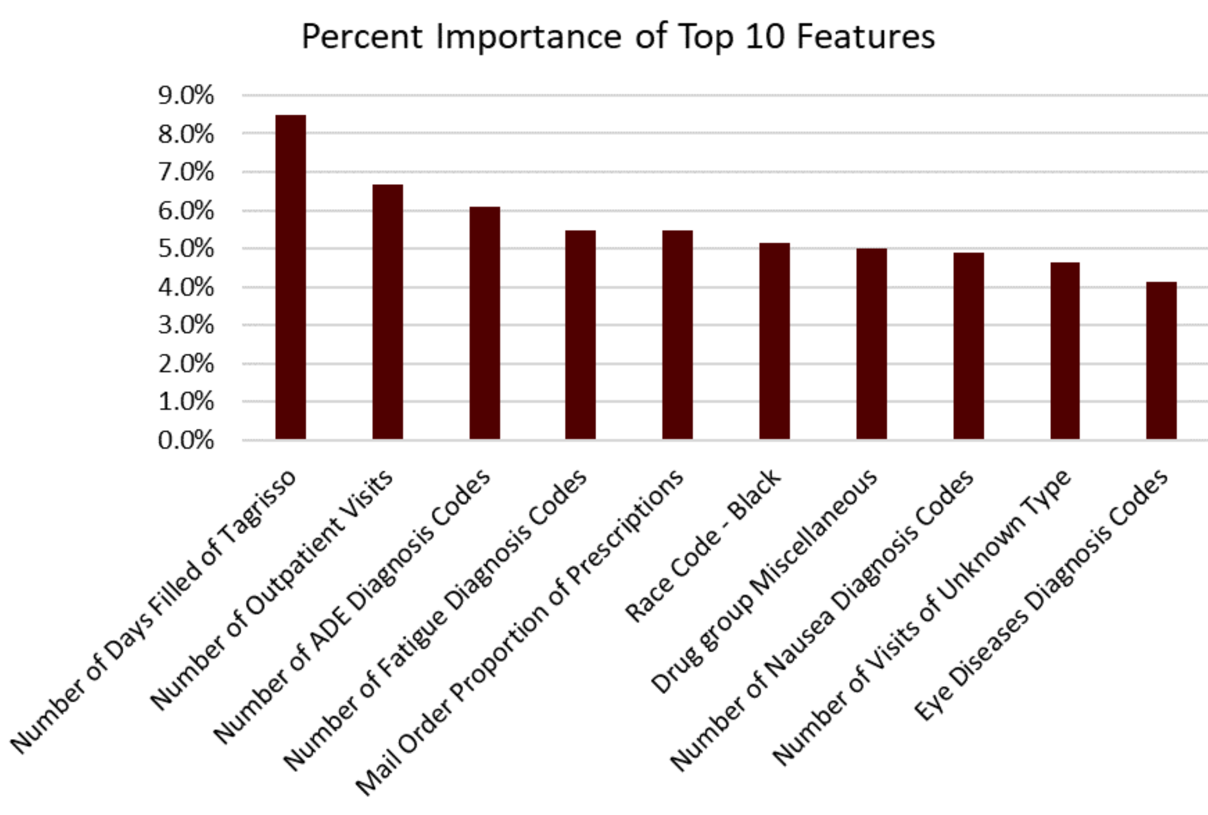


Figure 8. Percent importance of the top 10 features in the model.

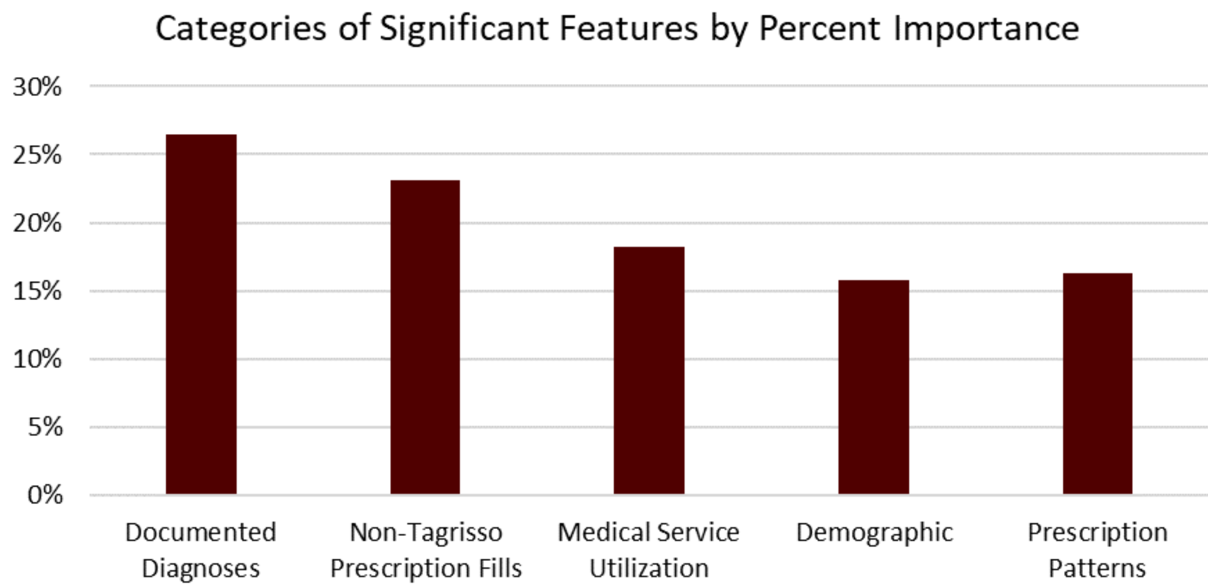


Figure 9. Percent importance of significant features in the model, categorized by variable type.

3.6 Limitations of the Model

The current model predicts the likelihood of a member experiencing an ADE and becoming non-adherent, but does not predict *when* the member will become non-adherent. Additional information on when an individual becomes non-adherent would help to inform the best timeline for action, as well as gain a greater understanding of which events specifically serve as precursors to non-adherence.

The current model also does not use CPT service codes. The data in the medical claims file includes multiple lines for many individual claims which indicates that many claims had multiple service codes. If the service codes had been included in the data, we could have gained insight into what service codes indicate a greater likelihood of the member having an ADE and becoming non-adherent. We suspect that services of greater intensity could indicate a higher degree of severity which could indicate greater risk of the member for the target variable.

4. Population Segmentation

4.1 Approach

Leveraging the features importances identified by the model, we performed exploratory analyses on the data sets to identify distinct population segments at increased risk of having an ADE and becoming non-adherent to Tagrisso treatment. Targeted initiatives can be designed to address the underlying cause of these at-risk populations.

We identified three key population segments who disproportionately experienced ADEs and discontinued Tagrisso therapy. The three segments were 1) members with a low number of Tagrisso days filled, 2) members who filled prescriptions at in-store pharmacies, and 3) members who were documented as having specific disease codes and medical visits. Each segment was defined using primary variables and as necessary secondary and tertiary variables. **Table 3** shows the segments and the variables used to create them. The following sections detail the rationale for choosing the segments, the relative risk associated with each segment, and recommendations for addressing the risk of each segment.

Segment	Representative Variable
Members With a Low Number of Tagrisso Days Filled	Tagrisso Days Filled < 75
Members Who Filled Prescriptions In-Store	Prescriptions Filled Via Mail-Order = 0
Members With Specific Documented Disease Codes and Medical Visits	Number of Medical Claims >= 21 (Primary) Number of ER Visits >= 1 (Secondary) Number of ADE Diagnoses >= 1 Number of Nausea Diagnoses >= 1 Number of Fatigue Diagnoses >= 1

Table 3. The three segments and their defining variables.

4.2 Segment 1: Members With a Low Number of Tagrisso Days Filled

The number of days supply that a member filled of Tagrisso holds substantial weight in our model due to its direct impact on adherence - a member cannot take their medicine if they do not have it. While the length of prescribed therapy varies by member, on average a lower number of supply filled of Tagrisso was a strong predictor of whether a member had an ADE and subsequently discontinued treatment.

Per a study on neoadjuvant therapy for resectable non-small cell lung cancer using Osimertinib, the median dosing time for the participants in the study was 75 days[10]. We used this as our threshold for segmentation. As **Figure 10** illustrates, 55% of members were included in this segment. 15.5% of members in this segment had an ADE and became non-adherent to their medication, compared to only 2.2% who were not included in the segment. Our findings confirmed our hypothesis that a lower number of Tagrisso supply filled was associated with a greater risk of having an ADE and becoming non-adherent.

We concluded that using this feature could provide an opportunity to address one major factor predicting adherence and allows us to tailor and target interventions to members that can ultimately result in better member outcomes and lower costs. Fortunately, this group is easily identifiable both retrospectively and in real time, creating a prime opportunity to craft targeted interventions.

As a countermeasure to member drop-outs as part of this segment, we recommend that an automated personalized outreach program be instituted by Humana any time a member fails to refill their prescription on time. A more detailed proposal for this program is included in Section 5.1.

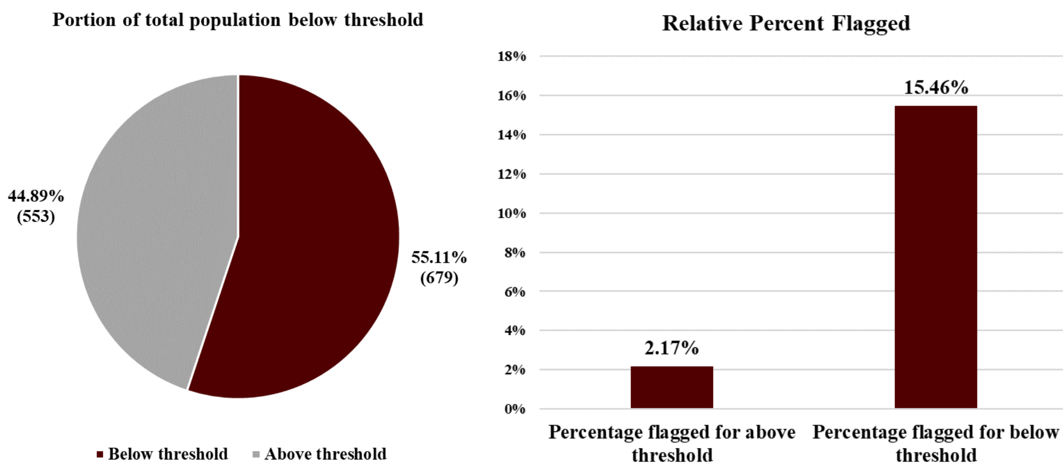


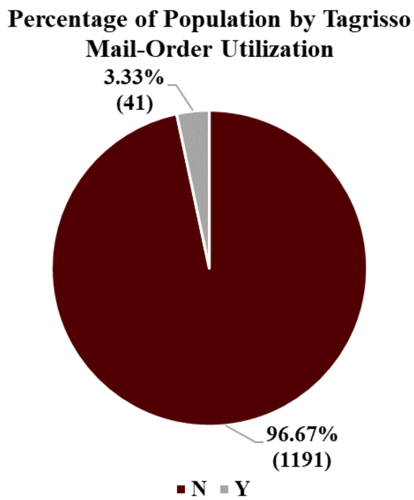
Figure 10. Proportion of members flagged for Tagrisso prescription administered below threshold and above threshold (75 days)

4.3 Segment 2: Members Who Filled Prescriptions In-Store

Another feature identified by the model as being an important predictor of whether a member will remain adherent to their Tagrisso therapy regimen is whether the member utilized a mail-order pharmacy to fill their Tagrisso prescriptions. Multiple studies have found that the use of mail-order pharmacies increases adherence of members for a number of different conditions[11,12]. As such, we hypothesized that users of mail-order pharmacies would have lower rates of the target variable.

This hypothesis was proven true through our analysis. As shown in **Figure 11**, while only a small percentage of members used a mail-order pharmacy to fill their prescription, this group did not have any members who had an ADE and became non-adherent, despite having 14 documented occurrences of ADEs. Though the sample size of members who used a mail-order pharmacy is small, we believe the extensive literature supporting the effect of their use on medication adherence rates supports our finding.

As a result of this finding, we recommend that Humana use financial incentives such as co-pay discounts or gift cards to encourage the adoption of mail-order pharmacies among members taking Tagrisso. This can be initiated before the therapy even begins to serve as a proactive deterrent to drop-outs. As an added benefit, members will be able to track the delivery of their prescriptions and will have increased convenience in obtaining Tagrisso. Removing the inertia posed by the need to pick up prescriptions from an in-store pharmacy could help to increase medication adherence rates among Humana members.



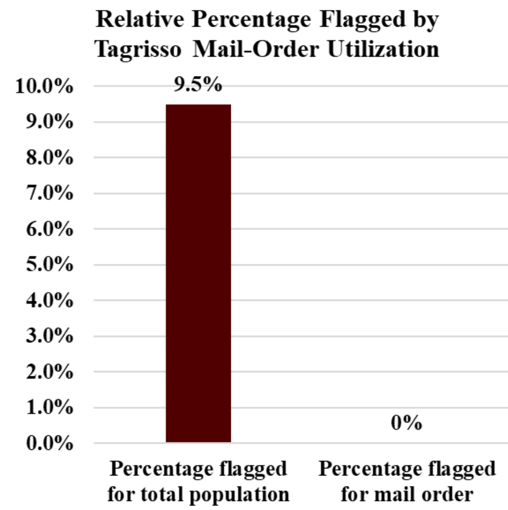


Figure 11. Proportion of members flagged for receiving Tagrisso prescription via mail-orders

4.4 Segment 3: Members With Specific Documented Disease Codes and Medical Visits

Medical care utilization and disease codes documented on medical claims serve as a very significant predictor of the target variable. Tagrisso is known to have side effects such as nausea, fatigue, and constipation. We hypothesized that an increased number of medical visits, ER visits, and documentation of known Tagrisso side effects would increase the likelihood of a member having an ADE and becoming non-adherent to their medication regimen.

To assess this hypothesis, we separated the population into two groups for each hypothesized predictor. The results are shown in **Figure 12**. The primary variable used in this segment was members with a high number of doctor's visits. The average number of visits in the six months after beginning treatment was 21, which was used as the threshold to divide members. 33.3% of members had a number of visits above this threshold, of which 24.4% had an ADE and became non-adherent compared to only 2.1% of members below the threshold. We also observed the effect of ER utilization: of the 8.0% of members who visited an ER, 33.7% were flagged, compared to only 7.4% of non-users of the ER.

Members who had a claim that documented nausea or fatigue were also disproportionately likely to be flagged. Only 4% of members had an ICD code indicating nausea, but of those members, 52% had an ADE and became non-adherent, compared to 7.6% for the remaining population. Likewise, only 9.4% of members had an ICD code indicating fatigue, but of those members, 44.8% were flagged for the target variable compared to 5.8% for the remaining population.

We recommend that an automated system be instituted by Humana to monitor the medical claims feed in real-time to detect for Tagrisso side effects and unusual medical service usage and automatically trigger member referrals to case management. As these are all significant predictors of a member becoming non-adherent, it is key that responses be taken immediately to support these members in their treatment. Immediate outreach to initiate discussion about a member's concerns and educate them about the vitality of the treatment is imperative.

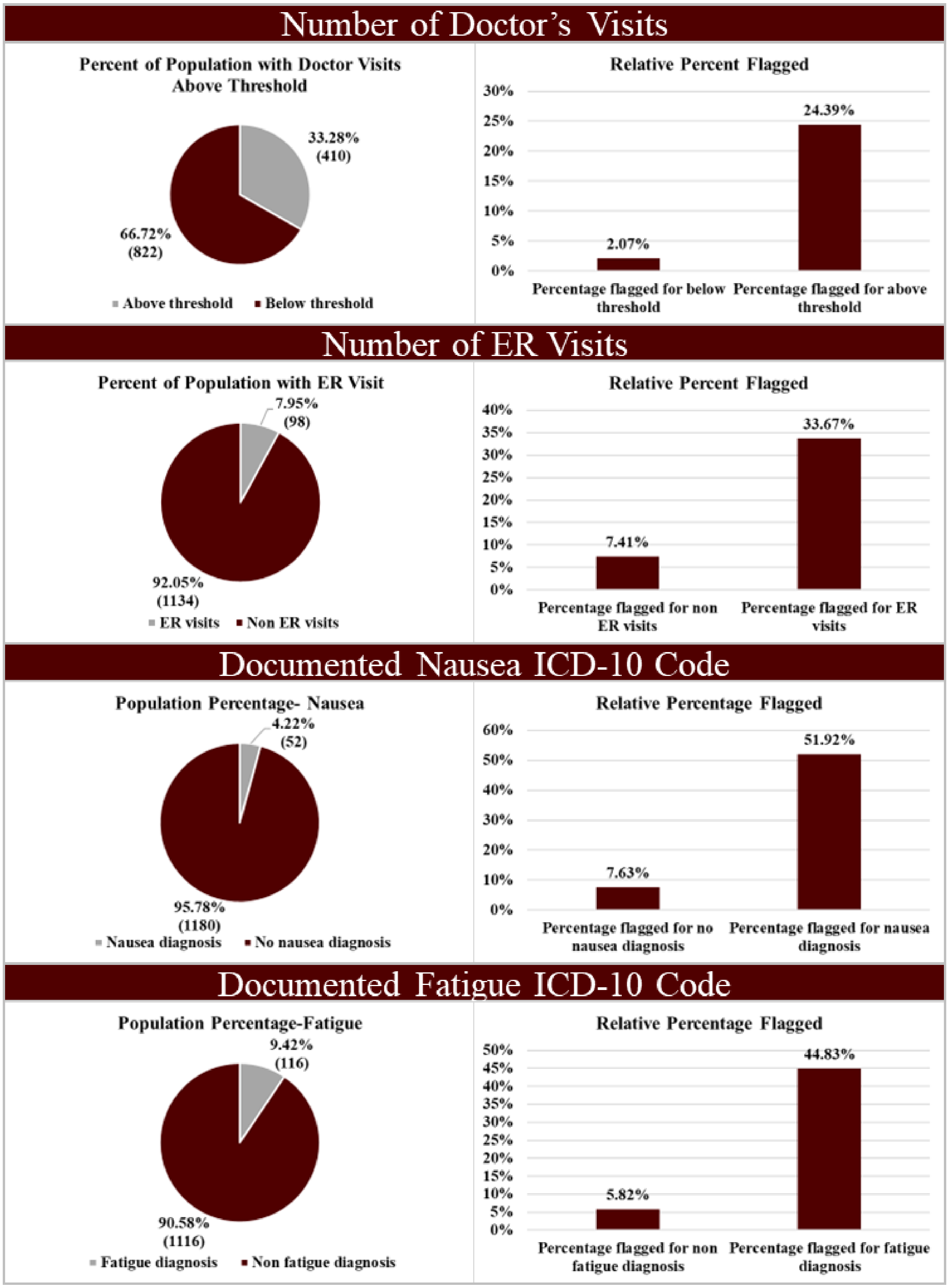


Figure 12. Proportion of members flagged for various precursors of non-adherence and effect on target variable.

5. Recommendations - Analysis And Actionable Insights

In this section, we present a series of recommendations, founded in data-driven insights derived from our model. Building on the population segmentation outlined in Section 4, our approach centers on addressing critical factors that influence member adherence to Tagrisso therapy after having an ADE, ultimately aiming to reduce premature discontinuations. We introduced three key population segments: Segment 1: those with a low number of Tagrisso days filled, Segment 2: individuals who filled prescriptions in-store, and Segment 3: members with specific documented disease codes and medical visits. We have created three recommendations to address the underlying causes of non-adherence among the identified segments. We have separated our recommendations into those that involve prospective outreach to at-risk populations (recommendation 2) and those that involve identifying and responding to precursors of non-adherence (recommendations 1 and 3).

For the first segment, we recommend the implementation of an automated personalized outreach system to proactively remind members of upcoming refills and to track and respond to members who do not fill their prescriptions as expected. This involves proactive outreach to members and allows for a real-time response to potential adherence challenges. For the second segment, we propose a promotional campaign and incentive program to encourage the use of mail order prescriptions, which offer improved convenience and accessibility while ensuring timely delivery to members. For the third segment, we suggest utilizing the medical and pharmacy claims feeds to monitor for events in real-time that indicate a member is at risk of becoming non-adherent, which will automatically trigger referrals to case management to provide tailored education and support to reduce the risk of discontinuation of therapy.

Table 4 includes high-level details on each recommendation and the target segment. In total, we expect to see annual cost savings of nearly $3.2 million and a total net decrease in costs of over $2.5 million. Detailed explanations of each recommendation is included in the following sections.

Recommendation	Targeted Segment	Underlying Cause of Non-Adherence	How Does Recommendation Address Underlying Cause?
Implement Automated Personalized Outreach	Members With a Low Number of Tagrisso Days Filled	Inertia / Lack of Support	Hyperpersonalized reminders, education, and support will increase the reduce the inertia causing delays in refills and provide support for members experiencing ADE's.
Incentivize Utilization of Mail Order Pharmacy	Members Who Filled Prescriptions In-Store	Inertia / Barriers to Travel	Mail-order pharmacies are known to increase medication adherence by reducing the inertia and challenges posed by having to travel to in-store pharmacies.
Implement Automated Case Management Referrals	Members With Specific Documented Disease Codes and Medical Visits	Adverse drug events / side effects	Members are most likely to discontinue treatment when they have increased medical utilization and certain diagnoses. Rapid case management response will support members who suffer adverse events in remaining adherent.

Table 4. The three recommendations, their target segment, and how the recommendation will improve Tagrisso adherence.

5.1 Proposed Solutions

5.1.1 Recommendation 1: Implement Automated Personalized Outreach

As noted previously, the first identified segment is made up of individuals who filled a low supply of Tagrisso prescriptions during their therapy period. In order to address the challenge of member non-adherence due to irregular consumption of Tagrisso, we propose the implementation of an Automated Personalized Outreach (APO) program. This system will leverage advanced analytics and real-time data monitoring to proactively identify instances where members may be at risk of non-adherence.

It is likely that Humana has existing automated outreach systems in place to promote medication adherence in their members. These systems are most commonly used for medications related to quality metrics that influence Star Ratings, such as diabetes, hypertension, and cholesterol medications. Humana could leverage this existing infrastructure to expand these systems to perform more targeted outreach to members using Tagrisso who are at risk of non-adherence.

Our proposal for the APO program will operate by integrating with Humana's existing member data management systems, claims feeds, and electronic health records. It will continuously monitor the prescription claims feed to determine the number of days supply of Tagrisso that a member has filled and determine if they are expected to soon run out of medication. Five days prior to the expected end of Tagrisso supply, the system will send a personalized reminder to the member to refill their prescription. On the day that the supply is expected to be exhausted, the system will send another prompt. If the prescription has not been filled three days after the expected end date of supply, the member will be referred to case management, who can follow up with the member to determine the reason the prescription has not been filled and provide support to the member to help them remain adherent..

There has been a trend in healthcare towards hyperpersonalization, in which even automated outreach is tailored to the preferences and needs of the individual. The automated outreach can be sent in the form of an email, text message, or phone call, depending on the member's stated preferences or their expected preferences given their demographic information. In addition to serving as a reminder of upcoming prescription fills, these alerts will be strategically designed to engage and educate members about the importance of adherence to their therapy. They will include tailored messages that address potential concerns, provide guidance on managing side effects, and emphasize the life-saving benefits of consistent medication intake.

By deploying this Automated Personalized Outreach system, we believe Humana can significantly improve member adherence rates, ultimately leading to more successful therapies and improved member outcomes. In addition, the hyperpersonalized approach of the messages will help the member feel supported by Humana, improving member engagement and satisfaction.

Cost Analysis

As noted previously, Humana likely already has infrastructure in place to send automated outreach to members to encourage medication adherence for quality measures relating to Star Ratings. Leveraging this existing infrastructure may reduce the initial cost of implementing the system. This reduces, but does not eliminate, the setup costs of the program, which we estimate as requiring three months of work from two employees.

Other major cost drivers of implementing automated personalized outreach are the cost to send each message, which we estimate at $0.50 per message, and the cost of case management for members that do not refill on schedule. We estimate that case management costs $146 per interaction[14] and each non-adherent member will require two interactions.

As a result of the system, we expect to see a 40% reduction in the number of members who have an ADE and become non-adherent, based on literature showing similar initiatives resulted in a 42 to 72% reduction in medication non-adherence rates[15,16]. The reduced costs are based on assumptions that non-adherent members have increased medical costs due to an increased number of office and ER visits, costing an average of $200 and $1917 per visit, respectively[17]. Additionally, non-adherent members increase their risk of recurrence from 2% to 13%[1], which leads to higher long-term costs due to having to undergo another round of treatment.

In total, we expect an overall cost reduction of $1.29 million and an overall return on this initiative of $1.11 million given the assumptions and calculations stated in **Figure 13**.

Assumptions		
Total Number of Annual EGFRm NSCLC Members	1100	
Number of Members Reached by Initiative Annually	605	55% of members
Number of Affected Members who Become Non-Adherent	93.8	15.5% of affected members
Reduction in Non-Adherent Members due to Program	40%	
Cost Per Message	$0.50	
Messages Sent Per Non-Adherent Member	10	
Case Management Referrals Per Non-Adherent Member	2	
Cost of Case Management Referral	$146	
Increased Future Doctor's Visits Per Non-Adherent Member	10	
Average Cost of Doctor's Visit	$200	
Increased Costs of Doctor's Visit Per Non-Adherent Member	$2,000	
Increased ER Utilization Per Non-Adherent Member	5	
Average Cost of ER Visit	$1,917	
Increased Costs of ER Utilization Per Non-Adherent Member	$9,585	
Increased Recurrence Rate for Non-Adherent Members	11%	
Cost of Recurrence	$ 208,000	$108,000 for drug cost and $100,000 for hospitalization
Increased Cost Due To Recurrence Per Non-Adherent Member	$ 22,880	

Initial Cost of Program		
Set-Up Cost of Program	$ 100,000	Cost of 2 FTE's for 3 months (Reduced by Leveraging Existing Systems)

Annual Cost of Program	
Cost of Messages	$3,025
Cost of Case Management	$176,660
Total Annual Cost of Program	**$179,685**

Program Outcomes - Reduction in Medical Spending	
Reduced Doctor's Visits	$75,020
Reduced ER Utilization	$359,533
Reduced Recurrence	$858,229
Total Reduction in Medical Spending	**$1,292,782**

Overall Annual Value of Program	**$1,113,097**

Figure 13. Analysis and projection of annual value of Automated Personalized Outreach program.

Anticipated Challenges

While the implementation of an Automated Personalized Outreach system holds immense promise for improving member adherence, we have identified potential challenges, along with proposals to address these challenges.

One of the primary concerns lies in striking the delicate balance between automated outreach and preserving the personal touch of patient care. The system must be designed with empathy and sensitivity, ensuring that automated messages do not come across as impersonal or intrusive. Moreover, there may be technological hurdles in integrating the outreach system with existing member data management platforms. Fortunately, Humana likely already has a system in place to perform automated outreach for other medication adherence initiatives that is designed to find the right balance between automation and personal touch. Even if Humana does not have these systems in place, they could work with smaller companies such as Upfront[13] Health that specialize in creating hyperpersonalized outreach campaigns to increase medication adherence.

Furthermore, Humana's Medicare Advantage population skews heavily towards older ages and includes members from diverse demographics. As such, some members may be resistant to digital communication channels or have specific cultural or linguistic needs. An outreach system must account for alternative modes of contact and adjustments in the method of delivery may be needed to address the specific needs of members to ensure the effectiveness of the strategy for all member populations.

5.1.2 Recommendation 2: Incentivize Utilization of Mail Order Pharmacy

As discussed in Section 4, the second segment is members who do not use mail-order pharmacy to fill their prescriptions. Users of mail-order pharmacy were far less likely to become non-adherent to treatment after having an ADE. This is not surprising, given that for many different medications, users of mail-order pharmacies are known to have higher medication adherence rates than users of traditional in-store pharmacies[11,12]. Studies also show that members who use mail-order pharmacies tend to have lower medical spending costs, primarily due to the increased adherence rates[11].

To bolster therapy adherence, we recommend implementing an incentive program aimed at promoting the utilization of mail order pharmacy services. This initiative will provide a high initial monetary reward to the member for their first use of a mail order pharmacy, as well as smaller subsequent rewards, in the aim that members will discover the convenience and accessibility provided by mail-order pharmacies. Ultimately, this will serve to both improve health outcomes and reduce long-term costs through improved Tagrisso adherence, as well as improve member experience through the provided convenience.

We propose a specific schedule of incentives for members to use mail-order pharmacies. On their first use of mail-order to fill a Tagrisso prescription, the member will receive $100 through either reduced co-pays or a gift card. On the second use of mail-order, they will receive $50. On all subsequent uses of mail-order to fill their Tagrisso prescriptions, they will receive $25.

By providing tangible benefits for choosing this mode of prescription acquisition, members are more likely to adopt and adhere to the mail order process. Additionally, the program will be complemented by a comprehensive communication campaign, educating members about the advantages of mail order services, including the added convenience and time-saving benefits. This communication campaign could also be included along with the first recommendation to specifically target members who are less likely to refill their prescriptions in a timely manner.

In the view of the Triple Aim, this recommendation addresses every goal. By incentivizing members to utilize mail-order pharmacies, Humana can increase Tagrisso adherence rates, which will improve health outcomes for members and reduce long-term medical costs for both Humana and the member. Additionally, by streamlining the process of filling prescriptions, the convenience that mail-order pharmacies provide to the member is expected to increase the member's experience with Humana.

Furthermore, this recommendation is supported by data indicating a correlation between timely access to medications and increased adherence rates. Members who receive their prescriptions through mail order are more likely to consistently adhere to their medication regimen. By offering incentives, we seek to reinforce this positive behavior and create a feedback loop of improved adherence and better health outcomes.

Cost Analysis

The primary cost drivers of instituting a mail-order incentive program lie in the program management costs and direct cost of incentives. We estimate that the uptake rate among members will be 60% in the first incentive month with a 80% renewal rate in the second month and a 85% renewal rate in the third month, for a total of 40% of members switching over to using mail-order pharmacies for their source of Tagrisso. Of this 40%, we estimate that this will lead to a reduction in non-adherence among members who have had an ADE of 50%, given literature showing an average reduction in non-adherence among mail-order users of 50%[11]. This results in a total reduction of members becoming non-adherent after an ADE of 20%. Our assumptions of the costs savings of reducing a non-adherent member are the same as for the costs analysis of our first recommendation. As shown in **Figure 14**, the program is expected to result in a overall cost reduction of nearly $700,000 for a total return to Humana of $425,000.

Assumptions		
Total Number of Annual EGFRm NSCLC Members	1100	
Number of Members Reached by Initiative Annually	1067	97% of members
Number of Affected Members who Become Non-Adherent	101.4	9.5% of affected Members
Reduction in Non-Adherent Members due to Program	20%	
Average Cost Per Affected Member Member	$0.50	
First-Month Incentive	$64,020.00	$100 at 60% uptake rate among eligible members
Second-Month Incentive	$25,608.00	$50 at 80% renewal rate
Subsequent-Month Incentive	$108,834.00	$25 at 85% renewal rate (for 10 months)
Total Stickiness	40%	
Increased Future Doctor's Visits Per Non-Adherent Member	10	
Average Cost of Doctor's Visit	$200	
Increased Costs of Doctor's Visit Per Non-Adherent Member	$2,000	
Increased ER Utilization Per Non-Adherent Member	5	
Average Cost of ER Visit	$1,917	
Increased Costs of ER Utilization Per Non-Adherent Member	$9,585	
Increased Recurrence Rate for Non-Adherent Members	11%	
Cost of Recurrence	$ 208,000	$108,000 for drug cost and $100,000 for hospitalization
Increased Cost Due To Recurrence Per Non-Adherent Member	$ 22,880	

Annual Cost of Program		
Cost of Incentives	$198,462	
Cost of Program Management	$75,000	0.5 FTE for 1 year
Total Annual Cost of Program	**$273,462**	

Program Outcomes - Reduction in Medical Spending	
Reduced Doctor's Visits	$40,546
Reduced ER Utilization	$194,317
Reduced Recurrence	$463,846
Total Reduction in Medical Spending	**$698,709**

Overall Annual Value of Program	**$425,247**

Figure 14. Analysis and projection of annual value of Mail-Order Incentive program.

Anticipated Challenges

While the adoption of mail-order pharmacies among members presents an opportunity for improved outcomes for both the member and Humana, we expect to face two main challenges. First, there may be resistance from members who are accustomed to traditional in-store pharmacy visits. Influencing individuals to transition to a mail order system may require comprehensive education and support to address any concerns or uncertainties they may have. Additionally, some members may value the face-to-face interactions with pharmacists or prefer the immediate access to their medication. The goal of the initiative is that the financial incentive will overcome this inertia to help members discover if the use of mail-order pharmacies is right for them. We have proposed a specific incentive structure that we expect to generate buy-in from members, but research and experimentation may be needed to determine the most effective incentive structure to encourage change. It may be a complex task to strike the right balance between providing incentives and maintaining cost-effectiveness of the program.

Another consideration is that logistical challenges in the implementation of the incentive program may arise. Implementing the program will require considerable effort to create the administrative processes and data management systems to track program success. The successful implementation of this recommendation will require a well-coordinated effort across multiple stakeholders.

5.1.3 Recommendation 3: Implement Automated Case Management Referrals

To address the challenge of member drop-outs attributed to adverse drug events (ADEs) and associated symptoms, we recommend the implementation of an Automated Case Management Referral (ACMR) system. This proactive approach aims to identify members at risk of discontinuation due to ADEs and promptly connect them with specialized case management resources.

The system will integrate with Humana's claims feed to leverage advanced analytics and real-time data monitoring to assess members' medical claims, diagnoses, and visit patterns. When indicators of potential ADE related challenges emerge, the system will automatically trigger referrals to dedicated case managers. These triggers will be based on events that were identified by the model to be significant indicators of a patient suffering and ADE and subsequently becoming non-adherent. The most important of these features were an increased number of doctors office visits and ER visits, as well as diagnosis codes for an ADE, nausea, or fatigue. All of these features are included in claims submitted to Humana by physicians so by integrating with Humana's claims feed, the ACMR system can automatically process the claim information and decide whether a member should be referred to case management.

After referral, the case manager can provide personal support to the member to provide tailored education and guidance to help patients manage the side effects of Tagrisso and remain adherent to their treatment regimen. Additionally, the case management program will incorporate a multidisciplinary approach, involving collaboration with healthcare providers, pharmacists, and case managers. This collaborative effort will ensure a holistic and patient-centric approach to addressing ADEs and other side-effects to improve adherence to therapy.

By implementing an ACMR system, we aim to proactively address the underlying causes of premature therapy discontinuation. This intervention is expected to significantly improve member experiences and outcomes by providing timely, targeted support to individuals facing ADE-related challenges when taking Tagrisso.

Cost Analysis

The main cost drivers of the ACMR program are the set-up cost of the program to integrate into the claims line feed, as well as annual costs of monitoring the feed and case management for patients that are identified to be at risk of having an ADE and discontinuing Tagrisso therapy. We estimate the cost of case management to be $146 per encounter[14].

We expect the program to result in a 30% reduction in the number of members who have an ADE and become non-adherent. Our calculations of the reduction in spending are based on assumptions that non-adherent members have increased medical costs of $200 and $1917 for an increased number of doctors' office and ER visits[17]. Additionally, non-adherent members see an increase in their risk of recurrence from 2% to 13%[1], which increases the long-term costs of the member.

As shown in **Figure 15**, we expect the program to yield an annual overall cost reduction of nearly $1.2 million and an overall return on the program of nearly $1.0 million.

Assumptions		
Total Number of Annual EGFRm NSCLC Members	1100	
Number of Members Reached by Initiative Annually	385	40% of members
Number of Affected Members who Become Non-Adherent	115.5	30% of affected Members
Reduction in Non-Adherent Members due to Program	30%	
Case Management Referrals Per Non-Adherent Member	2	
Cost of Case Management Referral	$146	
Increased Future Doctor's Visits Per Non-Adherent Member	10	
Average Cost of Doctor's Visit	$200	
Increased Costs of Doctor's Visit Per Non-Adherent Member	$2,000	
Increased ER Utilization Per Non-Adherent Member	5	
Average Cost of ER Visit	$1,917	
Increased Costs of ER Utilization Per Non-Adherent Member	$9,585	
Increased Recurrence Rate for Non-Adherent Members	11%	
Cost of Recurrence	$ 208,000	$108,000 for drug cost and $100,000 for hospitalization
Increased Cost Due To Recurrence Per Non-Adherent Member	$ 22,880	

Initial Cost of Program		
Set-Up Cost of Program	$ 400,000	Cost of 2 FTE's for 6 months (Reduced by Leveraging Existing Systems)

Annual Cost of Program	
Claims Feed Monitoring	$100,000
Cost of Case Management	$112,420
Total Annual Cost of Program	**$212,420**

Program Outcomes - Reduction in Medical Spending	
Reduced Doctor's Visits	$69,300
Reduced ER Utilization	$332,120
Reduced Recurrence	$792,792
Total Reduction in Medical Spending	**$1,194,212**

Overall Annual Value of Program	**$981,792**

Figure 15. Analysis and projection of annual value of Automated Case Management Referral program.

Anticipated Challenges

While the implementation of an ACMR system offers substantial potential for improving member adherence, several challenges must be anticipated. One significant concern is the need for robust data integration and interoperability between various healthcare systems and platforms. Depending on whether Humana's existing infrastructure could be leveraged, it could take a significant investment to create the coordination and communication between electronic health records, medical claims data, and case management resources needed for the program to be a success.

Even if the infrastructure is in place, ensuring timely and accurate identification of members experiencing ADEs will require sophisticated data analytics and algorithms. Our model provides the starting point for this. However, the system will need to transition from identifying at-risk members retrospectively to identifying them in real-time. Striking the right balance between sensitivity and specificity in identifying at-risk individuals is a difficult task as false positives or negatives could impact the efficacy and cost-effectiveness of the referral system.

5.2 Measuring Progress - Evaluating Success Through the Triple Aim Framework

To measure the progress and impact of recommendation will require that key performance indicators be constantly monitored in alignment with the Triple Aim framework. One of the primary metrics will be the Month-over-Month (MoM) and Year-over-Year (YoY) reduction in non-adherence rates among Humana's EGFRm NSCLC member population who have suffered an ADE. This reduction will serve as a direct indicator of the recommendations' effectiveness in improving health outcomes. By proactively identifying and supporting members at risk of non-adherence, we anticipate a steady decline in drop-out percentages over time.

Additionally, Humana should track the average cost per member associated with the management of EGFRm NSCLC cases. This metric will be monitored on a biannual and annual basis to assess the recommendations' impact on reducing long-term healthcare expenses. By targeting interventions towards high-risk individuals, we aim to mitigate avoidable costs related to emergency room visits and hospitalizations. Through this approach, we expect to demonstrate substantial cost savings for Humana, further emphasizing the financial benefits of our predictive model.

To complement these quantitative measures, Humana should also conduct member satisfaction surveys to gauge the overall experience of members on Tagrisso. Member feedback will provide valuable insights into the effectiveness of the initiatives and the perceived level of support received. This qualitative assessment will be integral in understanding the holistic impact of the recommendations' on member experience and engagement.

6. CONCLUSION

The predictive model we created is effective at predicting the likelihood of Humana's EGFRm NSCLC members of experiencing an ADE and becoming non-adherent to their Tagrisso treatment regimen. Through rigorous data exploration and using features identified as most important by the model, we identified three distinct population segments: individuals who have filled a low amount of Tagrisso prescriptions, individuals who use in-store pharmacies to fill their prescriptions, and members experiencing adverse drug events and other side effects. These segments were pivotal in tailoring our recommendations, ensuring a patient-centric approach to care.

Our recommendations are strategically designed to address each segment's unique challenges. The implementation of automated personalized outreach will proactively alert members, healthcare providers, and case managers to potential adherence issues, allowing for timely intervention. Encouraging the utilization of mail order pharmacies through financial incentives aims to provide convenient access to medications, reducing barriers and inertia for members to fill their prescription. The introduction of automated case management referrals will allow for targeted support by case managers to those exhibiting signs of non-adherence.

While each recommendation alone would improve adherence rates among members, in combination they will provide a comprehensive solution to many of the common challenges members face. From improved convenience of receiving prescriptions to experiencing Humana's patient-centered approach through interactions with case managers, our recommendations will improve the member experience while also improving health outcomes and reducing costs.

Throughout this process, we remained dedicated to evaluating our progress against the Triple Aim framework. By tracking MoM and YoY reductions in Tagrisso adherence rates, monitoring average costs per member, and seeking member feedback through surveys, our recommendations will demonstrate tangible improvements in health outcomes, cost savings, and member satisfaction.

7. REFERENCES

1. Astrazeneca. (2022, September 11). *Tagrisso demonstrated 5.5-year median disease-free ... - astrazeneca*. Tagrisso demonstrated 5.5-year median disease-free survival in the adjuvant treatment of patients with EGFR-mutated lung cancer. https://www.astrazeneca.com/media-centre/press-releases/2022/tagrisso-demonstrated-5-year-median-disease-free-survival-in-the-adjuvant-treatment-of-patients-with-egfr-mutated-lung-cancer.html

2. Adams, B. (2023, October 12). *AstraZeneca's Tagrisso to lead niche lung cancer market with $7B sales potential: Report*. AstraZeneca's Tagrisso to lead niche lung cancer market with $7B sales potential: report. https://www.fiercepharma.com/marketing/astrazenecas-tagrisso-lead-niche-lung-cancer-market-7b-sales-potential-report

3. CADTH Reimbursement Recommendation Osimertinib (Tagrisso). (2022). *Canadian Journal of Health Technologies*, *2*(1). https://doi.org/14 October 2023

4. HIGHLIGHTS OF PRESCRIBING INFORMATION. (n.d.). https://den8dhaj6zs0e.cloudfront.net/50fd68b9-106b-4550-b5d0-12b045f8b184/52503580-1192-44f7-a05e-5743159ed19b/52503580-1192-44f7-a05e-5743159ed19b_viewable_rendition__v.pdf

5. Centers for Disease Control and Prevention. (n.d.). *USCS data visualizations - CDC*. Centers for Disease Control and Prevention. https://gis.cdc.gov/Cancer/USCS/#/Demographics/

6. Analysis Details Incidence, Prevalence of NSCLC in the United States. (2022, January 10). AJMC. https://www.ajmc.com/view/analysis-details-incidence-prevalence-of-nsclc-in-the-united-states

7. Humana. (n.d.). *Form 10-K for Humana Inc filed 02/16/2023*. Form 10-K . https://humana.gcs-web.com/static-files/f4043f22-fc0d-46d4-8fe9-98d2a3aa26c6

8. Das, V. K. (n.d.). Logistic Regression vs K-Nearest Neighbours vs Support Vector Machine. https://www.globaltechcouncil.org/machine-learning/logistic-regression-vs-k-nearest-neighbours-vs-support-vector-machine/

9. XGBoost vs LightGBM: How Are They Different. (2021, November 28). Neptune.ai. https://neptune.ai/blog/xgboost-vs-lightgbm

10. Hy, Y., Ren, S., & Yang, L. (2022, April 28). *Osimertinib as Neoadjuvant Therapy for Resectable Non-Small Cell Lung Cancer: A Case Series*. National Library of Medicine. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9096023/

11. Devine, S., Vlahoitis, A., & Sundar, H. (2010, March 29). *A comparison of diabetes medication adherence and healthcare costs in patients using mail order pharmacy and retail pharmacy*. Taylor & Francis online. https://www.tandfonline.com/doi/full/10.3111/13696991003741801

12. Ridgway, J. P., Friedman, E. E., Choe, J., Nguyen, C. T., Schuble, T., & Pettit, N. N. (2020). Impact of mail order pharmacy use and travel time to pharmacy on viral suppression among people living with HIV. *AIDS Care*, *32*(11), 1372–1378. https://doi.org/10.1080/09540121.2020.1757019

13. Patient Engagement Solutions | Upfront Healthcare. (n.d.). Upfronthealthcare.com. Retrieved

October 16, 2023, from https://upfronthealthcare.com/solutions/

14. Texas Health and Human Services. (n.d.). *Children and pregnant women: Case management*. Children and Pregnant Women: Case Management | Provider Finance Department. https://pfd.hhs.texas.gov/acute-care/children-and-pregnant-women-case-management

15. Khonsari, S., Subramanian, P., Chinna, K., Latif, L. A., Ling, L. W., & Gholami, O. (2014). Effect of a reminder system using an automated short message service on medication adherence following acute coronary syndrome. *European Journal of Cardiovascular Nursing*, *14*(2), 170–179. https://doi.org/10.1177/1474515114521910

16. Hwang, D., Chang, J. W., Benjafield, A. V., Crocker, M. E., Kelly, C., Becker, K. A., Kim, J. B., Woodrum, R. R., Liang, J., & Derose, S. F. (2018). Effect of Telemedicine Education and Telemonitoring on Continuous Positive Airway Pressure Adherence. The Tele-OSA Randomized Trial. *American Journal of Respiratory and Critical Care Medicine*, *197*(1), 117–126. https://doi.org/10.1164/rccm.201703-0582oc

17. *Medicare and emergency room visits: Coverage and limits*. (2020, March 16). Www.medicalnewstoday.com. https://www.medicalnewstoday.com/articles/does-medicare-cover-emergency-room-visits#costs-and-considerations