

Improving patient outcomes and controlling costs

- Intro
- Situation
- Background
- Assessment
 - Recommend quality providers to patients
 - identify high risk patients
 - analyze factors that contribute to better health outcomes
- Recommendation

Intro

I have selected the “Improving patient outcomes and controlling costs (https://www.sas.com/en_us/customers/healthpartners.html)” as the topic of my project. This was something that HealthPartners did to reduce yearly costs by \$137 million. I selected this project because of the fact that I work in the healthcare IT industry and have relevant industry knowledge to contribute to additional ideas. The Challenge they posted on their page is: Identify high-quality, cost-efficient providers; pinpoint patients who need case management; and analyze, connect data from medical records and lifestyle characteristics to spot trends and factors that suggest better health outcomes.

Situation

Healthpartners recently started seeing an increase in healthcare cost overall and bad patient reviews. To tackle this problem, they initiated a 3 step program.

- recommend high-quality and cost-efficient providers to patients
- identify high risk patients and assign case managers to them to improve outcomes
- analyze factors and workflows that may contribute to better health outcomes

They are looking to utilize analytics and machine learning integrated into their EHR (Electronic Health Record) to accomplish all three of these goals.

Background

HealthPartners is a nonprofit health care provider and health insurance company located in Bloomington, Minnesota. They consist of 7 hospitals, 55 primary care clinics, 23 urgent care clinics, 27 in-clinic pharmacies, 15 eye care centers, and 22 dental clinics. They have over 1,700 physicians that serves more than 1.2 million patients. They are an enterprise customer of EPIC and has a strong suite of reporting functionalities from the software company; this includes registries, reporting workbench, radar, predictive analytics, and healthy planet. They are an ACO and have started Meaningful Use reporting since the first year it was out. As a transformative organization, they are constantly undergoing improvements to increase the quality of healthcare for their patients.

Assessment

Recommend quality providers to patients

To recommend high quality providers to patients you must first provide a way to assess the quality of a provider. Currently, there's a national push for value over volume so shifting from the previous model of how many patients a provider sees, we will focus our metrics on the quality of care a provider gives. To build this model, we need a large number of quality metrics. Luckily, there's already government initiatives that require a yearly reporting of this: Meaningful Use/MIPS and ACO. Because there's a large list of metrics in there, I won't list it here, but they include:

- % of visits where provider reviewed the patient's allergies
- % of results pushed to patients online chart
- % of medications electronically prescribed
- % of patients 18 years or older who were screened for tobacco use

We can combine these metrics with internally derived metrics in the EHR such as:

- wait time for each patients
- Provider response time for online messaging including RX Refill
- Provider efficiency in the EHR system

Use K-means clustering to identify 5 different clusters and determine whether this relates to values of the following:

- patient satisfaction scores
- % of readmissions
- treatment scores for chronic diseases (derived from outcomes compared to other providers)

This will allow us to assess what level the provider is at concerning their quality of care. Once we have this score, we can start making recommendations to the patients. Patients may be seeing a variety of providers including primary care, specialties, dentists, ophthalmologists, urgent care. We will need to categorize the provider in each of these buckets before making the recommendation so that we do not recommend a dentist when a patient is looking for a primary care physician. Once we've separated this out, we need to determine at what point is the patient selecting a provider. This can happen either while they're scheduling their appointment online, over the phone, or at the front desk. If they are scheduling online, we can add a recommendation after they have selected their time slot and suggest the highest quality provider open during that time slot. If there are providers who are 2 standard deviation quality wise away at nearby time slots from all the providers during that time slot, we can also make the recommendation too.

Identify high risk patients

To identify high risk patients, we will need to implement multiple models for multiple risk types.

First, for the patients that already have a certain disease, we can already place them in the disease registry. For example, if a patient has diabetes, we can flag them in the EHR as a diabetic patient and calculate multiple metrics on them on an ongoing basis so that we know they are a diabetic patient with certain symptoms without needing to requery the database again for this information. Next, we can predict their risk level for various other dangers as a diabetic patient. For example, using the metrics calculated in the registry, we can predict what their risk level is for a hospital admission within the next year. For each disease, the included metrics will be different but for this assignment, I will stick with Diabetes as my example and below are some examples of metrics we can include in a model:

- Patient Age
- BP

- Weight
- HDL/LDL levels
- Genomics data

Once we have these metrics for historical patients, we can pull them into a logistic regression with the dependent variable being whether a patient has been admitted into a hospital within the last year. This can give us a model that will provide a probability that a patient will be at risk of being admitted into the hospital. We do not have to provide a threshold for the logistic regression to actually determine whether they will or not because we want this to be fed to the users as a metric to determine whether they need a care manager or not. Within a few cases, we can see end users of the system actually making that threshold determination themselves for different cases which may not be captured by the EHR.

Second, for the patients who do not yet have a disease, we can build a model to identify whether they are at risk for a disease. Using data from patients who did not have a certain disease but contracted it at a later time, we can use that historical data to build a logistic regression model as the same as above to provide us with a probability of being at risk for diabetes.

Caveat, there is a large problem in these kind of predictive analytics because once you've predicted an outcome and displayed it to an end user, their actions can very well change the actual outcome and this is to be expected because our goal is to decrease that readmission risk or decrease the risk of obtaining that disease. But as we keep moving, they can skew our model. This is where Cubic Spline Regression may be able to give us an edge on the classic regression. But this is still a field being researched.

analyze factors that contribute to better health outcomes

Factors that contribute to better health outcomes can be obtained through understanding correlation analysis, but this requires an art component. Healthcare organizations usually gather various data about their community and also about their patient. A perfect example of this is Dr. Mona Hanna-Attisha from Hurley Medical Center who discovered the increase of blood lead levels in children nearly doubled after the city started using Flint River as its water source. By using their EHR in combination with community data, this correlation was found and corrected to improve patient outcomes. (source (<https://www.healthdatamanagement.com/news/hurley-medicals-ehr-analysis-is-playing-a-key-role-in-flints-water-crisis>))

For our case, it would be helpful to gather some of these external data sources such as:

- zip code
- air pollution levels
- locations of standing water
- topography (marsh, dry areas, etc)
- water sources
- average temperatures

From there, we can generate a correlation matrix where the columns are each of the factors that may affect a health outcome and the rows are the outcome. Each point in the matrix can be the correlation between each factor and the health outcome of a patient. Doing this can give us an idea of what may be strongly correlated and we can dig further into the details of the data to determine causation. With this data, we can then provide recommendations on actionable steps for the organization to suggest to their patients.

Recommendation

To suggest providers to patients, we should compile a list of quality metrics and implement a physician scoring system based on this. Once we have the scoring system, we can suggest providers based on their practice(family med, ortho, dentistry, etc) to the time slot a patient has selected. If we see a provider significantly better during a close time slot, we will offer that provider.

To help identify high risk patients, we should implement logistic regression to each specific disease case to determine the probability that a patient is at risk for that disease or whether that disease will cause a high impact negative outcome. This risk score should be displayed to care providers to help them determine the right course of action, whether that's to assign a care manager or to alter treatment methods

To provide better outcomes, we should conduct a large scale correlation analysis for various community factors on the local population by finding correlations between the community's health data and various local factors that may have changed. From there, we can dig further into highly correlated factors to determine causation and correct for negative effects or promote positive ones.