CASE STUDY

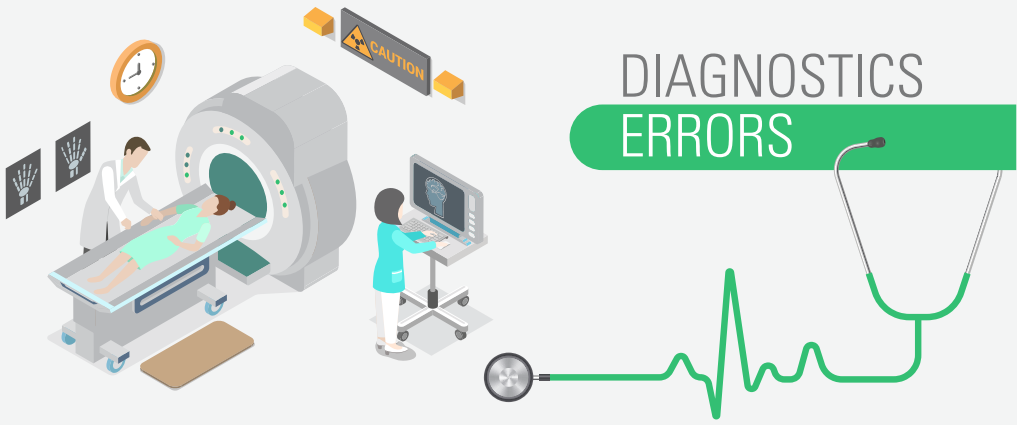
PREDICTIVE ANALYTICS FOR A GLOBAL HEALTHCARE COMPANY TO IMPROVE PATIENT CARE

ENABLING ACCURATE AND EARLY DIAGNOSIS AND TREATMENT OF PATIENTS, WITH REDUCED DANGEROUS MISDIAGNOSIS, USING MACHINE LEARNING (ML) MODELLING

Each year, approximately **12 million** Americans experience some form of diagnostics mistakes while seeking outpatient services

INTRODUCTION

Misdiagnosis of a disease and its conditions can have life-threatening consequences. Each year, approximately 12 million[1] Americans experience some form of diagnostics mistakes while seeking outpatient services.



Some of the other studies also found eye-opening results

A study by the Institute of Medicine (IOM) found that diagnostic errors and other inefficiencies cost the US economy a staggering **$750** billion each year.[2]

As per a Johns Hopkins study in 2018, more than **250,000** Americans die every year due to medical errors/misdiagnoses. Other reports claim the numbers to be as high as **440,000**[3]

Another Johns Hopkins study found that in the US out of **6,000** cases, one out of every **71** cancer cases were misdiagnosed, and up to one out of five cancers were misclassified, resulting in delayed and inadequate treatment.

This illustration shows how Axtria engaged with a global healthcare organization in identifying misdiagnosed disease patients using ML models and helped them identify patients who can be put on drug treatment much earlier..

AXTRIA
INGENIOUS INSIGHTS

BUSINESS SCENARIO

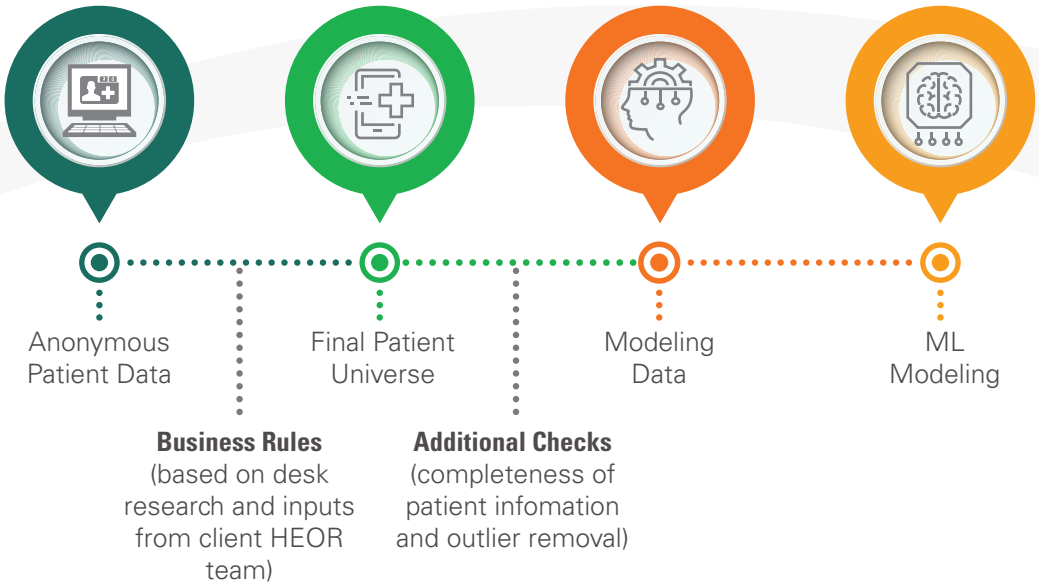
- The drug is indicated for three Disease States – A, B, and D
- Patients may be initially misdiagnosed with Disease State C before being identified to have Disease A or Disease B
- Early identification of misdiagnosed Disease A and Disease B patients is critical for getting patients placed on the right treatment on time and drive growth for the brand

OBJECTIVE

- Profile Disease State C patients misdiagnosed as Disease A or Disease B and compare with correctly diagnosed Disease A or Disease B patients
- Evaluate currently active population who have a Disease C diagnosis, and determine patients likely to be diagnosed with Disease A or Disease B

AXTRIA'S APPROACH

Axtria adopted a four-step approach for this engagement:



- Robust patient universe identified for each of the relevant disease states.
  a. The final patient universe was a combination of misdiagnosed patients of Disease A and Disease B, along with the patients identified as actual Disease C population
- Modeling Data was a patient-level aggregation with ~400 variables/predictors. Y-variable was categorical (1 for misdiagnosed patients)
- Random Forest, XGBoost and Neural network were the three techniques used to determine the patient-level probability of misdiagnosis
- Machine Learning Model intended to predict critical patients' characteristics that help identify misdiagnosed patients, and identify the possibly misdiagnosed patients among the current pool of Disease C patients

Multiple supervised machine learning techniques were used to build a prediction/ classification model that can accurately separate **misdiagnosed Disease A/Disease B patients (positive examples)** from **True Disease C patients (negative examples)**.

Model Selection: Selection of ML algorithms

- A tradeoff between accuracy and interpretability
- Improving accuracy generally means building a more complex model

Model Validation Techniques

- Use the 'Test Data' from Historical population to validate the model (20% of the total Modeling Data)

AND/OR

- Use k – Fold Cross-Validation
  - » Validation is done on different folds/samples of the ''Modeling Data."
  - » The final metrics calculated are averaged over the various folds/samples

ML model identified features that are most predictive of misdiagnosis. Some patient profiles to highlight classifications made by the model:

Sample Patient Profiles

Age	Gender	Comorbidities	Total HCP Visits	Distinguishing Px/Rx	Misdiagnosing Rate of HCP	Prediction
73	Female	Disease D Comorbidity 1	>100 visits (~50 visits post Disease C)	Rx – Drug X	7.7%	Misdiagnosed Disease A ✓
37	Female	Comorbidity 4 Comorbidity 6	>60 visits (~20 visits post Disease C)	Rx – Drug Z	50%	Misdiagnosed Disease A ✓
54	Female	Comorbidity 1	>30 visits (~10 visits post Disease C)	Rx – Drug X	1.6%	Misdiagnosed Disease A ✗
67	Female	Comorbidity 8 Comorbidity 9	>50 visits (~40 visits post Disease C)	Rx – Generic Drug	18.2%	Misdiagnosed Disease B ✓
65	Male	-	~20 visits (8 visits post Disease C)	-	4.4%	True Disease C ✗

Key Insights	• Female patients and patients >60 years of age have a high rate of misdiagnosis • Patients getting Disease C diagnosed by HCPs with a high misdiagnosing rate are more liable to misdiagnosis • High #HCP visits indicate continued incorrect treatment and is a distinguishing characteristic of misdiagnosed patients

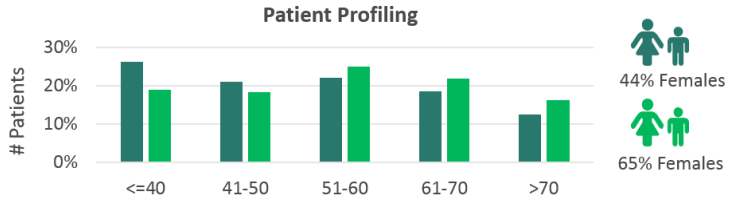
Diagnosed with Disease C Diagnosed with Disease A Diagnosed with Disease B

KEY TAKEAWAYS

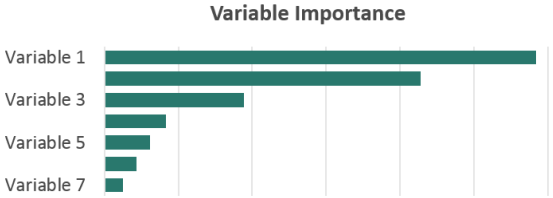
- Axtria recommended machine learning model to predict misdiagnosis among patients with a recall* value of 87% for Disease A patients and 78% for Disease B patients
- Identified nine key features each to identify misdiagnosis of Disease A and Disease B
- Predicted 9.9% of current Disease C patients to be misdiagnosed – 5.4% Disease A and 4.5% Disease B patients
- Recommended a target HCP list by identifying the HCPs who diagnosed the predicted misdiagnosed patients, including a misdiagnosis rate for each HCP

* Recall of a model is the number of correct misdiagnosis predictions Disease B a ratio of total misdiagnosed patients

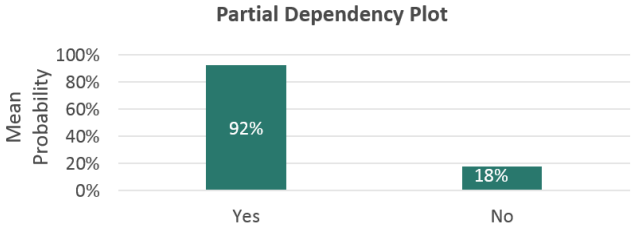
Sample Deliverables



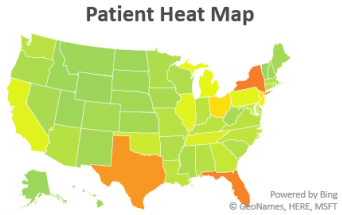
Profiling the selected patient universe for various variables gives an intuitive idea of the variables and their variation for misdiagnosed patients



Variable importance for the model measures the relative effect of that variable on the models



Partial dependency plot shows how the probability changes with changes in each variable



US heat map for predicted misdiagnosed patients. States with high concentration of patients are dark red

CONCLUSION

- Analysis using patient-level data and machine learning modeling technique confirms there is a sizable population of patients who get misdiagnosed as Disease C before getting diagnosed as Disease A/Disease B
- Real Commercial Opportunity lies in the patient segment that is Misdiagnosed Disease A/Disease B but sitting as Disease C for more than 12 months
- Machine Learning technique can predict the likelihood of misdiagnosis with a high degree of accuracy
  - » 87% of misdiagnosed Disease A patients
  - » 78% of misdiagnosed Disease B patients

The early and accurate diagnosis resulted in a potentially significant reduction in patient cost burden and improved the ability of the HCPs to provide better treatment and improve patient prognosis.

Axtria also recommended further analyses:

- Inform messaging and targeting strategy for HCPs treating these patients
- Conduct HEOR study to assess the burden of misdiagnosis – burden on the patient, the cost to the healthcare system, etc.
- Evaluate the feasibility of ongoing refresh and retraining of model Disease B more recent data is available

REFERENCES

1. https://www.cbsnews.com/news/12-million-americans-misdiagnosed-each-year-study-says/
2. https://www.pinnaclecare.com/forms/download/Human-Cost-Financial-Impact-Whitepaper.pdf
3. https://www.cnbc.com/2018/02/22/medical-errors-third-leading-cause-of-death-in-america.html