Root Cause Analysis

Background

There was an experiment between clinicians being presented with reasons for a suggested diagnosis vs. not having a reason presented. These reasons were generated from the suggestion models, but curated by people with clinical knowledge to ensure that the reasons were clinically relevant and not just an artifact of the model.

With some clinicians being presented with a reason and others not, we seek to explore if there is any evidence to suggest that there is a difference in confirmation rate between those presented with a reason and this without.

Executive Summary

We find that there is sufficient evidence to suggest that having reasons for our ML generated suggestions in the telehealth application can provide a relative increase in the rates of diagnosis for Disease Groups (CDG) by approximately 4%. Among those CDGs, there was a relative increase in the confirmation rate ranging from 2% for Chronic Kidney Disease up to 14% for Diabetes.

We estimate the increase in revenue these correct confirmations provided to be approximately $270,000 annually. If we were to extend the reasons to all CDGs, we estimate the total increase would be $690,000 annually. These are based on 2021 CA membership of 35,000.

We utilize the chi-square test to evaluate the statistical significance of those suggestions with reasons vs. without reasons.

Goals of Analysis

Estimate the impact of ML-specific reasons in CA.

- Estimate how much the ML-specific reasons improved confirmation (if any) using statistical tests.
- Estimate if impact was found on all or certain disease groups.
- Translate the improved confirmation rate into the expected $ impact on the 2021 MA population (assuming we keep only the current whitelist of reasons vs. expanding the reasons to all diagnosis suggestions).

Numbers at a glance

The control group did not have the ML-specific reasons in the task, and the treatment group did.

Control False	Control True	Treatment False	Treatment True	Control	Treatment
27630	7025	27083	7233	34655	34316

This gives a True rate of 20.27% for the control group and True rate of 21.08% for the treatment group, an increase of 3.98%

By CDG we have:

CDG	treatment	control	percent improvement
Chronic Kidney	34.80%	34.02%	2.28%
Congestive Heart Failure	15.27%	14.08%	8.46%
Diabetes	13.78%	12.07%	14.11%

Heart Disease	35.21%	33.94%	3.74%
Obesity	16.17%	15.69%	3.04%
Vascular Disease	8.86%	8.36%	5.96%

Methodology

We utilize the chi-squared test to assess if there are differences between the control group and treatment group.

Chi Squared Test

We first perform a chi_sq test on the entirety of the data, regardless of CDG. This test yields a p-value of .009, indicating that the control and treatment groups are different enough to reject the null hypothesis that these groups are the same there is some difference between the control and treatment group.

Chi Squared Tests by CDG

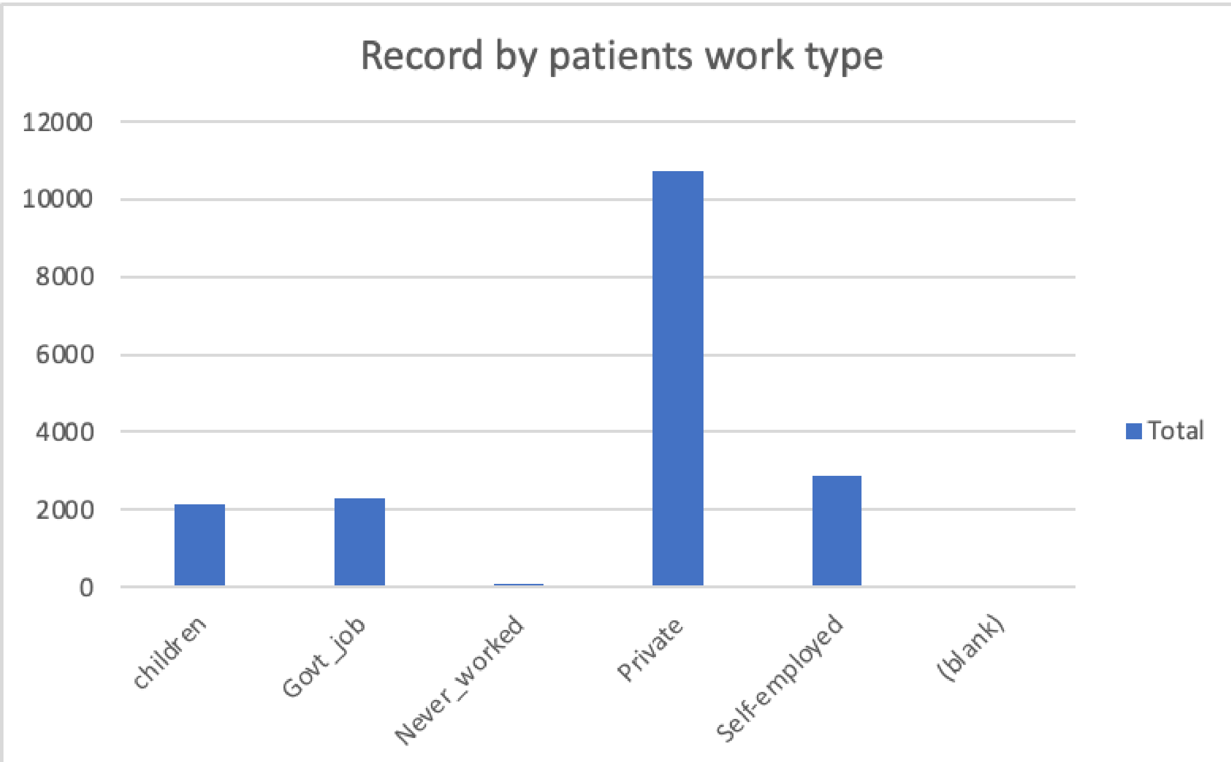
There are 6 CDGs in the relevant data, Chronic Kidney, Congestive Heart Failure, Diabetes, Heart Disease, Obesity, and Vascular Disease.

We subset the data by CDG and run similar chi_sq tests. The p-values for the CDGs are:

CDG	p-value
Obesity	0.5620
CHF	0.0420
Vascular Disease	0.2773
Chronic Kidney Disease	0.3993
Heart Disease	0.1134
Diabetes	0.1173

Key Insights

- The p-value for CHF is under the (admittedly arbitrary) value of .05.
- However, due to multiple hypothesis testing, a reasonable interpretation of the results is that this value being below .05 is potentially due to random chance.
- For example, using Bonferroni correction we would need a p-value under .0083 to consider rejecting the null.



Patients work type	No	Yes
(blank)		
1-20	3444	5
21-40	4535	137
41-60	4889	639
61-80	3254	867
81-100	273	85
Grand Total	16395	1733

# of Employees that left	
Mean	55.125
Standard Error	5.4484782
Median	61
Mode	89
Standard Deviation	26.691983
Sample Variance	712.46196
Kurtosis	-1.4481107
Skewness	-0.0849649
Range	84
Minimum	11
Maximum	95
Sum	1323
Count	24

	forecast	actual	variance	% var	Growth %
jan	2000	1200	-800	-40%	
feb	2000	2000	0	0%	67%
mar	2000	1500	-500	-25%	-25%
apr	2000	1300	-700	-35%	-13%
may	2000	2000	0	0%	54%
jun	2000	800	-1200	-60%	-60%
jul	2000	2200	200	10%	175%
aug	2000	2600	600	30%	18%

sep	2000	2600	600	30%	0%
oct	2000	2600	600	30%	0%
nov	2000	2600	600	30%	0%
dec	2000	2600	600	30%	0%

Recommendation

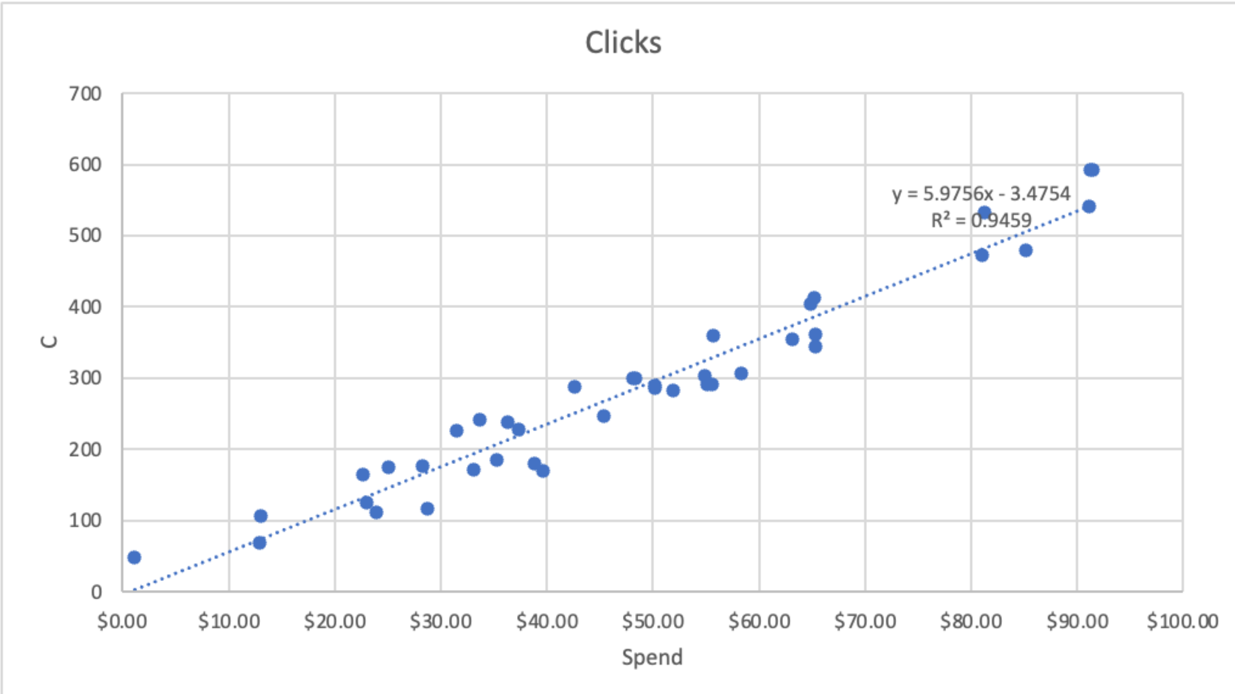
To translate improved confirmation rate into $, we can take the
total value ML suggestions in 2020 in RAF / member
* relative confirmation rate improvement
* fraction of ML diagnosis suggestions that have reasons
* member count with CA visit in 2021 (assume 35k)
* RAF to $ conversion (assume 1 RAF point= $7800)

We find this to be: .0635*.0398*.388*35,000*7,800 ~ $270,000 annually

If we were to assume that all the ML suggestions had reasons, we might expect a revenue impact of .0635*.0398*35,000*7,800 ~ $690,000 annually

Table of relevant data

We present the relevant data that went into the linear regression



Appendix/References

- Tableau Dashboard of exploratory analysis for the data.
- Jira ticket