Using Updated Health Care Data to Build Predictive Model on Hospital Readmission Rate

PSTAT Department Presenter: Chloe(Yuxin) Wang Professor: Ian Duncan
Mentor: Doris Padilla, Ming Yi

What is Hospital Readmission?

Hospital readmission occurs when a patient is discharged from a hospital, and admitted again within a time interval. According to the Hospital Readmissions Reduction Program (HRRP) established by the Department of Health and Human Services in 2012, this specific time interval is 30 days. Hospital readmission has always been a serious problem as it is a waste of money and time for the patients, the hospitals, and also the insurance companies. As a result, reducing hospital readmission rate has always been a hot topic in the health care field.

Introduction

Hospital readmission rate being so high has always been a serious problem. It causes a waste of time and money for the patients, the hospitals, and also, the insurance companies who's paying for most of the costs. If we can build a predictive model of hospital readmission rates for different types of patients, hospitals can use our model to check out the patients with higher probability of getting back to the hospital, and provide extra treatment and care ahead of time. In that case, the hospital readmission rate would hopefully be reduced in general.

Similar work has been done but on an out of date dataset. Indicators been analyzed was also only limited to the patients' demographic characteristics and the duration of admission in the hospital. This time, we are using a set of more recent data with the information from 110,000 patients to build the model again.

Methodology

While demographic characteristics as indicators has been done before, this time, we want to focus on the type of diseases that the patients had using the HCC risk scores. Four big disease groups were categorized by different HCC risk scores: cancer group, diabetes group, cardiac group, and congestive heart failure. Technically, congestive heart failure should count as a cardiac disease as well, but it is a very server problem that we want to look at it separately from others.

HCC or RXC No.	Factor	Platinum	Gold	Silver	Bronze	Catastrophic
	Age 60-64, Female	0.616	0.505	0.379	0.240	0.229
	Diagnosis Factors					
HCC001	HIV/AIDS	0.626	0.529	0.434	0.359	0.352
HCC002	Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock	8.000	7.812	7.688	7.731	7.737
HCC003	Central Nervous System Infections, Except Viral Meningitis	5.750	5.666	5.604	5.625	5.626

Segment of 2019 HCC Risk Score Table

We have 3 sets of data of 110,000 patients' admission information in the years from 2014 to 2016. Each patient has multiple claims, so the overall number is over 3 million claims. For the purpose of working more comfortable with this huge number of data, we built a Markov Chain transitional matrix in R to look at the transitional relationship for the patients' cost levels from one year to another. And also we compared the difference of transitions between the total population (which is the average) and all the other disease groups.

Discussion

From the transitional matrix we made, we have some interesting observations regarding the cost level transitions for different disease groups. The percentage of the transition levels for cancer group and diabetes group are really close to one another, but only a little bit higher than all patients instead of a lot more which is what I expected. The reason for that could be complicated, one of my hypotheses is that cancer and diabetes are both very long term chronic disease, and most likely incurable. A lot of patients decide not to pay a sky-high fee every day in hospital, instead, they could just go home and enjoy their life. Taking medicine, and getting physical examination regularly will be where most of their spending go.

For cardiac group and congestive heart failure. The probability of transitions into a high cost are much higher than the other two groups and also the all patient group. Especially for congestive heart failure, the high-to-high transition is over 6 times greater than the all patients group. I believe one reason is patients getting these cardiac diseases likely need more expensive medical apparatus and equipment to assist in treatments. And as time goes by, while a lot of cancer patients could decide not to get treatments anymore, cardiac group patients are still on the ground because they can still live longer with expensive treatments and medicines.

Result

Here is the result we got from the Markov Chain transitional matrix for all patients, and the four disease groups we categorized. Analyzes can be done by comparing the percentages in each tables.

	LOW	MODERATE	HIGH
LOW	0.539672727	0.142645455	0.005509091
MODERATE	0.114654545	0.167236364	0.013063636
HIGH	0.002545455	0.009809091	0.004863636

Cost Level Transition for All 110,000 Patients

	LOW	MODERATE	HIGH
LOW	0.361487692	0.13993935	0.006421691
MODERATE	0.148590796	0.28041384	0.026132715
HIGH	0.004816268	0.02042455	0.011773100

Cost Level Transition for Cancer Group (10.2% of Total)

	LOW	MODERATE	HIGH
LOW	0.441024882	0.1417654	0.006398104
MODERATE	0.127310427	0.2297986	0.022186019
HIGH	0.003791469	0.0174763	0.010248815

Cost Level Transition for Diabetes Group (30.6% of Total)

	LOW	MODERATE	HIGH
LOW	0.288328833	0.15121512	0.007886503
MODERATE	0.148257683	0.32257511	0.034460589
HIGH	0.005357679	0.02661695	0.015301530

Cost Level Transition for Cardiac Group (21.2% of Total)

	LOW	MODERATE	HIGH
LOW	0.192228596	0.13085832	0.01006602
MODERATE	0.126745319	0.39473969	0.05833965
HIGH	0.007360104	0.04805715	0.03160515

Cost Level Transition for Congestive Heart Failure(8.4% of Total)

Reference: Ian Duncan, *Healthcare Risk Adjustment and Predictive Modeling*; HCC table picture from cms.com