2020 MMA Online Datathon

Case Challenge: Medical Insurance Fraud Investigation

Medical provider fraud is one of the biggest problems facing the US public healthcare system. Medicare, one of the largest public insurance programs primarily for people aged 65 or older, accounts for a majority of public healthcare spending and has been the target of many fraud schemes[1]. Extensive records are collected by Medicare on its utilization and costs; however, it is still vulnerable to fraud with fewer than 5% of Medicare claims being audited.

Analysis of Medicare data has shown that many cases of fraud involved physicians and associated providers. They adopt ways in which an ambiguous diagnosis code is used to charge Medicare for unnecessary procedures and drugs. This results in increased costs, raising insurance premiums for all Americans.

You and your team have been brought in to investigate this fraud and provide data-driven recommendations on how to combat Medicare fraud. A dataset is provided that contains a sample of anonymized data for beneficiaries, as well as their inpatient and outpatient insurance claims between Dec 2008 and Dec 2009. A separate table of health care provider flags show those who appear to have engaged in fraud. A data dictionary describing the fields in each table is also provided.

Note that when a provider is flagged as committing potential fraud, all claims involving this provider are audited. This is an extremely costly, time consuming, and potentially adversarial process. Thus, accurate identification of providers who should be targeted with an audit is of utmost importance to Medicare.

Your task is to analyze the data (using the tools of your choice) and make a recommendation on the criteria that should be used to select which medical providers are audited, given the limited audit resources available. The primary consideration should be the total size of potential fraud with additional considerations being the ease of enforcement and timely reimbursement for non-fraudulent claims. Provide your reasoning and any insights into the data that led you to your recommendation. Note that these insights are particularly important: Medicare staff must be able to defend their decision to audit a particular provider should the case end up in courts. The audience for the presentation will be senior directors from the Center for Medicare & Medicaid services who administer the Medicare programs.

Background: Medicare and Insurance Fraud

The United States spends 16.9% of its GDP on healthcare, almost double the amount of the average OECD country[2]. Despite the high spending, it also has some of the worst outcomes

[1] https://www.washingtonpost.com/wp-dyn/content/article/2008/06/12/AR2008061203915.html
[2] https://www.commonwealthfund.org/publications/issue-briefs/2020/jan/us-health-care-global-perspective-2019

with life expectancy of two years less than the OECD country average. The healthcare system consists of a mix of public and private insurance with public spending accounting for roughly half of the total mix. One of the biggest public insurance programs is called Medicare that provides health insurance primarily to Americans who are aged 65 and older. Medicare has already surpassed 15% of federal spending and is projected to increase over the next decade.

Medicare involves a complex ecosystem of actors who participate in this public healthcare system. **Beneficiaries** are the individuals who are enrolled in Medicare. When a beneficiary receives healthcare, their provider typically submits a claim on their behalf to Medicare for the services. The beneficiary often is required to pay a minimum fee called a **deductible** after which Medicare covers the remaining portion. A **physician(s)** is the medical doctor who diagnoses and treats the beneficiary for a particular disease or issue. To receive reimbursement from insurance for their services, physicians (and other medical personnel) record their diagnoses or procedures using billing codes, which are processed and paid out by the insurance companies. A **provider** is the organization who provides the care, which can range from individual health care professionals to entire healthcare facilities.

Medicare is divided into four parts. For the purposes of this case, only Part A and Part B are relevant. Part A covers hospital, skilled nurse and hospice services. These services are generally referred to as **inpatient** services where the patient is admitted into the facility (e.g. admitted as a patient to the hospital). Part B covers **outpatient** services, which typically consists of visits to a physician where the patient is not admitted to the hospital (although the physician visit may take place at a clinic inside a hospital).

Healthcare fraud and abuse can take many forms. Some of the most common types of fraud by providers are:

- Billing for services that were not provided
- Duplicate submission of a claim for the same service
- Misrepresenting the service provided
- Charging for a more complex or expensive service than was actually provided
- Billing for a covered service when the service actually provided was not covered

Recent laws have aimed to control Medicare fraud and abuse by instituting a number of measures such as longer oversight periods, provider screenings, stronger standards for certain providers, the creation of databases to share data between federal and state agencies, and stiffer penalties for violators. With better technological support, experts are optimistic on the ability to simultaneously reduce costs while improving healthcare quality.

Deliverables

- A presentation of up to 20 slides in PPT or PDF format with appendix. While the Appendix is not limited in size, the relevance of every page in the Appendix and how it supports the main body of your presentation should be very clear to the audience (reader). Note that your presentation should be self-contained: the only reason a reader would open the appendix is to get more details on some of the results/visualizations

presented in the main body; the reader should never need to open the appendix to understand the statements made in the main body of the presentation

- You can use analytical tools of your choice; if you would like to include your code, please do so in the Appendix
- Submit presentation via Quercus website