Analyzing Misdiagnosis Compensations

AUTHOR
Vincent Liu

PUBLISHED
April 29, 2025

Introductions

A quick Google search reveals that there is an estimated **7.4 misdiagnosis** each year in the United States. Using the a NPDB data set that gathers all reports on malpractices in the US (including those diagnosis related), I aim to possibly give some insight on this problem through data visualizations and analysis. I will focus on revealing compensation for patients who submit *diagnosis related reports after 2010*.

The data set takes real data collected from The National Practitioner Data Bank (NPDB) who are apart of The U.S. Department of Health & Human Services. The NPDB Public Use Data File includes selected data on malpractice payments and adverse actions (e.g., licensure, clinical privileges, DEA, and Medicare/Medicaid exclusions) involving healthcare practitioners. It is updated quarterly and currently covers reports from September 1, 1990, to December 31, 2024.

Data Cleaning

```
# Loading Data Frame
df <- vroom("NPDB2410.CSV", col_types = cols(BASISCD5 = col_character(),
                                             AACLASS5 = col_character()))

df <- df %>%
  filter(ALGNNATR == "1", # showing only diagnosis related allegations
         ORIGYEAR >= 2010) # showing only reports after 2010

# Cleaning Data Frame
df$PAYMENT <- as.numeric(gsub("[$,]", "", df$PAYMENT))
df$TOTALPMT <- as.numeric(gsub("[$,]", "", df$TOTALPMT))
df <- janitor::clean_names(df) # clean names to lowercase with _ as spaces
```

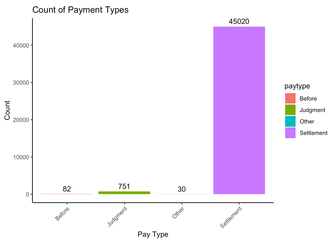
Given the scale of the orignal data set which contained **1,797,419** reports, it was very difficult to run visualizations. Using built-in R commands, we reduced to **144,156** reports when looking at only diagnosis related allegations, and then reduced to **46,304** reports when looking at reports after 2010. This will be the final data set we use to make our analysis.

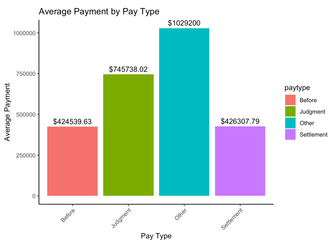
Visualizations

How many patients per payment type and how much?

Many of these reports will result in payment to the patient, lets take a look at how many payments fall under each of the following types. Further more we will look at the average payment this results in.

- 'B' = Before Settlement
- 'J' = Judgment
- 'O' = Other
- 'S' = Settlement
- 'U' = Unknown (Note: We will omit these results as this can either be no payment or a 'B' but are only applicable to certain reports filed electronically in 1995 or later)





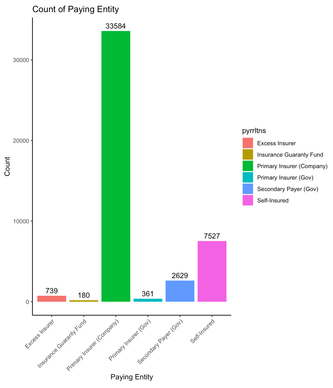
Analysis

Most malpractices case result in settlements which can be the result of multiple factors. These can include but are not limited to trials being costly, emotionally draining, and have unpredictable outcomes. Settlements offer a faster, more certain resolution, especially when liability is clear and expert testimony would otherwise be required.

Taking a look at the first graph we notice that a huge majority of **45020** reports result in a ***settlement*** to the patient; with very little reports in the other categories (even the 'Unknown' category we removed only contained 421 reports). The second graph reveals although considering there are limited amounts of the other payment types, settlements tend to have a much lower average in payments at **$45,020**. This can be a result of the lack of range in the other payment types however, I believe there is another hidden explanation.

Who are paying the patients?

Many paying entities undercut the patient to avoid trials. This can result uncompensated injuries. This data sets also reveals who are the paying entity to patients (eg. insurance). Lets explore who pays for settlements in the next visualization.



Analysis

While insurance companies dominate the payout landscape, self-insured organizations still represent a significant share. As we can see the two main forms of paying entities are ***Insurance Companies*** at **33584** and ***Self-Insured*** at **7527**. Interesting government funded fund appear less frequently. This is due to the majority of hospitals who do are included in reports being taken at the state level, as states have jurisdiction over these cases. On the off chance that the malpractice claim involves the federal government acting through a federally funded clinic or a Veteran's Administration facility, then the action is filed in a federal district court. The majority of doctors have malpractice insurance, so this makes sense that most cases are payed out by insurance companies, those who do not are likely using the government funds.

The large majority of paying entities being private insurance companies and self-insured organizations may contribute to the lower settlement averages seen in the previous visualization. Both these entities have financial incentives to minimize payouts and avoiding litigation, which could lead to the under-compensation of some patients. As such this continues to raise ethical and policy considerations regarding malpractice compensation.

Conclusions

Through this exploration, we find that the majority of diagnosis-related malpractice reports filed after 2010 result in settlements rather than judgments. This aligns with broader trends in malpractice litigation, where the high emotional, financial, and temporal costs of trials encourage both sides to seek faster resolutions. Although settlements offer a quicker pathway to compensation, they also appear to result in lower average payouts for patients when compared to the limited cases that proceed to judgment.

Further, examining the entities responsible for these payments revealed that private insurance companies and self-insured organizations are the predominant payers. This concentration of responsibility among private and internal risk-management systems highlights the financial pressures at play in malpractice resolution. Both types of entities have strong incentives to minimize payout amounts, likely contributing to the overall lower compensation patients receive through settlements.

It is important to note, however, that these findings are based solely on the NPDB public use data set filtered for diagnosis-related allegations after 2010. As such, the trends observed here may not fully capture the broader landscape of all malpractice cases across different specialties, geographic regions, or types of claims. While this data offers valuable insights into compensation patterns, it should be interpreted with the understanding that it reflects only a specific subset of malpractice reports.

Ultimately, while settlements may serve the practical purpose of expediting claims and reducing litigation burdens, they raise ongoing concerns about the adequacy of compensation for injured patients. These findings underscore the need for continued policy discussions around fair malpractice compensation practices, transparency, and the structural forces that influence settlement outcomes.

Data

1. Full Download
2. Format

Analysis Sources

1. https://wilsonlaw.com/blog/do-most-medical-malpractice-cases-settle/
2. https://pmc.ncbi.nlm.nih.gov/articles/PMC2628513/
3. https://www.gallaghermalpractice.com/resources/introduction-to-medical-malpractice/#:~:text=Some%20state%20laws%20require%20physicians,where%20it%20is%20not%20required.