FDA Drug-Drug Interactions

Authors: Bianca DeHaan, Matias Cinera, Shawn Diaz

***Abstract*—This project aims to create an application that can be utilized to explore drug-drug interactions for certain drugs of interest and adverse events.**

I. Introduction

The FDA's Adverse Event Reporting System (FAERS) database contains adverse events reports, as well as other medical information that is useful for surveilling various drugs and biologic products after they've been approved. The data is submitted both from professional backgrounds (physicians, pharmacists, nurses) and consumers (patients, family members, lawyers), and can be used to update labeling, improve product safety, restrict use, communicate with the public, or even remove a product from the market.

The aim of this paper is to produce a program that can utilize the FAERS database to extract drug adverse events for a specific drug of interest (DOI). From the reports in FAERS, the Apriori algorithm is used for understanding the relationship between our drug of interest and other drugs, in the context of causing a specific drug adverse event. The DDI is calculated by comparing the lift values of the drug alone and the drug with other compounds, which provides meaningful insight on whether the combination of drugs has a higher likelihood of causing the adverse event.

This project aims to provide all of this information to be accessible within a website, so that users can enter their drug of interest and adverse event and find out the corresponding DDI potential of that drug and the adverse event, the number of cases with that adverse event, and the DDI index of the drug combined with other drugs, in context of the adverse event.

II. Background

A. Association Rules

An association rule is essentially an implication of the form $X \rightarrow Y$, where X is the antecedent, and Y is the consequent. These two components, X and Y, are disjoint itemsets, meaning they share no common elements. In a medical scenario, such a rule might look like this: X represents a specific drug or a combination of drugs, and Y denotes an adverse event linked to them. This format helps to clearly establish the relationship between a particular medication or combination of medications (X) and the resulting health outcome (Y).

To effectively gauge the relevance and strength of association rules, especially in the context of drug adverse events reported in the FEARS database, three key metrics are used: support, confidence, and lift. Support represents the prevalence of a particular drug or drugs and adverse event combination, given the reported cases in the FAERS database. In other words, it quantifies how frequently the drug-event pair (X and Y) occur in the data. Confidence assesses the reliability of the implication. It does this by measuring the likelihood of observing event Y given the presence of a drug or drugs X. A higher confidence suggests a stronger correlation between the drugs and the adverse event. Lastly, lift provides insight into the strength of the association between the drug and the event, compared to their independent occurrences. A lift value greater than 1 implies that the drugs in question and the adverse event are more likely to occur together than by chance, indicating a potential causal relationship.

Together, these metrics offer a comprehensive view of the associations found within the data, helping with the process of identifying significant drug-event correlations, which might warrant further investigation in the future.

III. Calculations

For our project, there are multiple parts calculated in order to paint the whole picture of a certain drug-drug interaction.

A. Reporting Odds Ratio

To understand the risk of the adverse event with the drug of interest, versus the risk of the adverse event with other drugs, we calculate the Reporting Odds Ratio (ROR) using the FAERS data. When this value is greater than 1, it is considered a signal that the odds of the adverse event occurring with the drug of interest is that much more likely, than the odds of reports among other drugs.

To calculate the Reporting Odds Ratio, we need four key data points:

A. Number of cases with the adverse event taking the drug of interest
B. Number of cases with other events taking the drug of interest
C. Number of cases with the adverse event not taking the drug of interest
D. Number of cases with other events not taking the drug of interest

These values are directly queried from the FAERS database, and the ROR is calculated as the ratio:

Reporting Odds Ratio = (C / A) / (D / B)

B. DDI Index

Calculating the drug-drug interaction (DDI) Index is a relationship between the lift values of two association rules. For example, let's say we are interested in drug-drug interactions with Drug A, and the adverse event Acute Kidney Injury. The

DDI Index would be defined as the lift value of the association rule {Drug_A = 1, Drug_B = 1} -> {Acute Kidney Injury = 1} in comparison to the lift value of the association rule {Drug_A = 1} -> {Acute Kidney Injury = 1}. In this comparison, higher DDI index values indicate a higher likelihood of potential drug-drug interactions, in the context of the adverse event.

In order to get these association rules, the Apriori association rule mining algorithm was used. This algorithm works by identifying the frequent individual items in the database, and extending them to larger item sets as long as those item sets have a large enough support, defined by the minimum support parameter. After Apriori calculates these association rules, with their antecedents, consequents, support values, confidence values, and lift values, we take those rules that are relevant in calculating the DDI index value.

:

IV. Website

The website was built using Python and Flask. At the home page, the user is greeted with three prompts, namely for the drug of interest, the adverse event, and the minimum support for the Apriori algorithm. With this information, the website queries the data needed from FAERS. From that information, it is pruned to extract what is needed for the Apriori algorithm. This mostly means just the drug name, and if the specific adverse event was recorded in that report, while all other data is not needed anymore. After cleaning this data, more pre-processing happens by using an Encoder, which encodes the data in a way that Apriori can easily read and use. Finally, the Apriori algorithm is run on the cleaned data, and the results are searched for those relevant in calculating the drug-drug interaction indexes.

V. Conclusions

A. Limitations

One of the major limitations of this problem is computational resources. For example, when querying the FAERS database, only 1000 API results / reports can be returned at a time. Therefore, the tool had to utilize pagination. Even then, there are millions upon millions of reports in FAERS, collected since 1970, which would prove challenging to utilize all the data in the calculation of the drug-drug interaction indexes. Therefore, we only use a recent subset of that data.

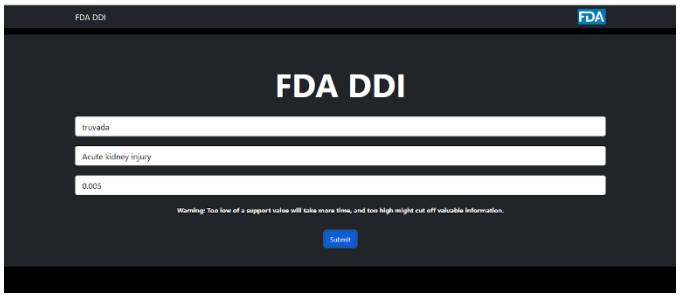
A similar limitation relating to computational resources is in regards to the Apriori algorithm, and specifically, the parameter called minimum_support. The support is how many instances of the data contain all instances of the association rule. Instead of calculating every possible association rule, the Apriori algorithm only calculates those with a minimum support value, set by the user. However, if this minimum support value is too low, too many results are not filtered, leading to the Apriori algorithm taking both too long to calculate the association rules, and taking up more memory than that is available. This would often lead to the program crashing, and being unable to produce any results.

On the other hand, by setting the minimum support too high, useful data is pruned and it's possible not to extract all possibly useful association rules. This is especially relevant for an adverse effect that is uncommon, since the support is prematurely low as that adverse effect only shows up in a minimum amount of records. We made a decision to allow the user to set their own minimum support value, since every drug, event combination will have different ideal values for this parameter. In the future, we could research how to maximize the efficiency of the minimum support value ourselves, to alleviate any further work on the user.

Another limitation has to do with the FAERS data itself. There are multiple warnings, both when interacting with the API, on the FAERS dashboard, and on other pages, that the data in FAERS is not accurate for providing medical diagnosis. This is because there is no vetting process of the data - i.e., anyone can submit data, there can be duplicate data, or even erroneous data. There is no guarantee in a report that the adverse event was caused due to the product. From FAERS FAQ, there doesn't even need to be any proof for a report to be submitted. Therefore, some conclusions that are gleaned from the FAERS data, and more specifically, our tool, should be taken as advisory and with caution. Instead, the tool might be better used to hint at a possible drug-drug interaction, rather than guarantee it.

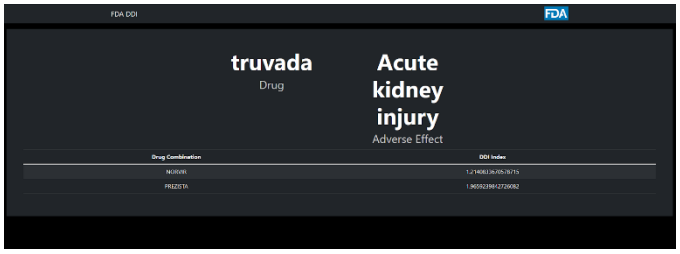
B. Conclusions

After building the website, it was time to set the functionality and information extracted. When you are on the home page, you are met with a page similar to the screenshot shown below. We decided to test the drug 'Truvada' and it's reaction with Acute Kidney Injury.



As a reminder, the support value is very important. If it is too low, the site will take too long and might not return results. Similarly, if it is too high, it prunes useful information and also returns nothing.

After a few moments of processing, the results of this query are shown below. The drugs Norvir and Prezista had a drug-drug interaction with Truvada, and their DDI indexes are displayed to the right.



C. Discussion

In the future, it might be useful to explore what other data we could extract from the FAERS database. One observation made was the use of generic versus brand names in

the data of FAERS, and often adverse event reports use either interchangeably. Since our website queries for one drug name, it could be improved by associating that drug with all other names it can go by - i.e. generic or brand names. Therefore, we could combine all the results of queries from each name the drug of interest goes by, and in turn, have more accurate results. This would require more processing power, because we are combining results from multiple queries, as well as additional logic to remove duplicate entries and ensure that the brand and generic names are indeed referencing the same drug.

REFERENCES

[1] Center for Drug Evaluation and Research, "Questions and answers on FDA's adverse event reporting system (FAERS)," U.S. Food and Drug Administration, https://www.fda.gov/drugs/surveillance/questions-and-answers-fdas-adverse-event-reporting-system-faers (accessed Nov. 24, 2023).