Disproportionality analysis of VAERS vaccination records using an empirical Bayes estimator

Dániel Pfeifer

Vaccines in the 21st century are one of the most succesful solutions for helping the immune system develop a protection against diseases. However, some still deny their effectiveness. In this study we will be showing what symptoms and illnesses vaccines can cause using the VAERS (Vaccine Adverse Event Reporting System) database, which collected infromation about patients that went back to their respective doctors a couple days or weeks after vaccination, and reported some kind of illness. This data has been preserved from 1990 to 2019, and it is available for free on the VAERS website.

There are multiple challenges with determining the connection between vaccines and symptoms:

- Only the data of those patients is available that actually went back to report an illness. Data for patients that reported no symptoms after vaccination isn't available.

- If there were multiple vaccines given to the same patient, it is possible that the vaccines alone don't cause and symptoms, but their combined effect does.

- It's possible that the vaccine is ineffective and patients come back with the exact same illness they were trying to cure. This doesn't mean that the vaccine caused the illness, it just means it can't cure it.

We will be using two models to determine these connections. The first is a Backtracking Model, connecting the issue back to Association Rules used in customer and product grouping. We will calculate the *support, confidence*, and a self-defined *diversity* value for each (vaccine, symptom) pair to determine ones that are common.

The second model is a more traditional solution, called the Gamma-Poisson Shrinker. The basic idea is that we are trying to estimate a λ parameter, denoting the probability that a vaccine causes a certain symptom in theory.

$$\lambda = \frac{C}{N} = \frac{\text{number of times the vaccine causes the symptom}}{\text{number of cases we observed}},$$

where $C \sim \text{Poi}(\lambda N)$, We will calculate the conjugate prior of λ and determine directly from the database, which vaccines are likely to cause which symptom.

All my calculations and code will be shown and will be reporoducible from the document. All my results will be shown as a Heatmap using Python's Seaborn package.

An example is shown to the right. You can see that the HBPV

vaccine caused Pulmonary fibrosis quite frequently, for example.