Reactogenicity Prediction after Administration of COVID 19 Vaccines - mRNA Based

Dr. Ankita Gupta (Assistant Professor)
Maharaja Agrasen Institute of Technology, GGSIPU, India

Himanshu Singh Rajawat
Maharaja Agrasen Institute of Technology, GGSIPU, India

Hardik Bachhan
Maharaja Agrasen Institute of Technology, GGSIPU, India

Rajat Goel
Maharaja Agrasen Institute of Technology, GGSIPU, India

Abstract

In December 2020, 2 mRNA-based COVID-19 vaccines (Pfizer-BioNTech and Moderna) were granted Emergency Use Authorization by the US Food and Drug Administration as 2-dose series and recommended for use by the Advisory Committee on Immunization Practices. In late February 2021, the US Food and Drug Administration granted Emergency Use Authorization for a third COVID-19 vaccine, a single-dose adenovirus vector-based vaccine from Janssen (Johnson & Johnson).

In clinical trials of the mRNA-based 2-dose vaccines, participants reported local and systemic reactions (reactogenicity). Frequently reported reactions included injection site pain, fatigue, and headache; greater reactogenicity was reported following the second dose. Continued monitoring of reactogenicity of COVID-19 vaccines outside of clinical trial settings may provide additional information for health care practitioners and the public about transient local and systemic reactions following COVID-19 vaccination.

1. Introduction

In December 2020, the US Food and Drug Administration (FDA) issued the Emergency Use Authorization (EUA) for two mRNA-based COVID-19 vaccines (BNT162b2 Pfizer-BioNTech and the mRNA-1273 vaccine from Moderna) as 2-dose series. Following the implementation of vaccination, local and systemic adverse reactions after receipt of the vaccines began to be reported. As of the Year 2022 till April, reports of 150,395 (0.07%) adverse events after receipt of vaccine had been submitted to the Vaccine Adverse Event Reporting System (VAERS). Although rare, some uncommon allergic reactions can develop and lead to death or disability. For example, from December 14 to 23, 2021, 1,893,360 people in the US received their first dose of the vaccine, and 21 of them reported suffering from anaphylaxis. Continued monitoring and assessing adverse events of these vaccines outside of trial settings could improve our understanding of the safety issues and contribute to the decision-making in terms of the implementation and administration of vaccination. It is also crucial for optimal outcomes of patients to identify patients at risk of severe adverse events in a safe medical environment.

Our goal is to make use of the VAERS data to:

- Predict the onset time of adverse events and recognize the key predictors to inform the medication preparation after vaccination and identify the high-risk population.

We use the baseline characteristics of patients, e.g. sex, age, medication history, etc, as inputs, and used different algorithms, e.8. linear regression, Lasso, Ridge, random forest, etc. to predict the onset time.

2. Literature Survey

There are many works regarding the method for predicting vaccine outcomes and vaccine-associated adverse effects. For example, Gonzalez-Dia et al. have provided a general procedure for predicting vaccine-induced immunity and reactogenicity using machine learning methods and described four basic steps including data processing, feature selection, choosing an algorithm, and testing.

Ahamad et al. have conducted the identification and classification of post-vaccination reactogenicity of COVID-19

vaccination, using the same data source. They used decision trees and random forests. support vector machines and gradient boosting machines as classifiers to find the significant features leading to the hospitalization and death of patients. However, they pre-processed their data to solve the sparse symptom feature problem, by which only 86 most frequently appeared symptoms were selected and combined. This method cannot perform well when encountered with high-dimensional sparse features.

PubMed was searched for articles published up to Dec 29, 2021, using the terms ("BNT162b2" OR "mRNA-1273" OR "mRNA COVID-19 vaccine") AND ("reactogenicity" OR "side-effects" OR "adverse effects" OR "health impact"), not restricted by language or type of publication. Among 429 results, few publications described health impacts following vaccination by BNT162b2 (Pfizer-BioNTech) or mRNA-1273 (Moderna). Available literature included reports of manufacturer-sponsored phase 1–3 clinical trials, observational and cross-sectional studies among specific groups (eg, transplant recipients or employees of a specific health-care system), and reviews of society recommendations that discussed reactogenicity and adverse events following mRNA vaccination.

In this analysis of VAERS data from the first 6 months of COVID-19 vaccination rollout in the USA, when over 298 million doses of mRNA vaccines were administered, we found that reactogenicity was similar to what was reported from clinical trials and from early post-authorization monitoring. In VAERS, local injection-site and systemic reactions were commonly reported. participants more frequently reported transient reactions following mRNA-1273 than following BNT162b2, and more frequently following dose two of either vaccine compared with after dose one. Female participants and individuals younger than 65 years reported adverse events and reactions more frequently than male participants and those aged 65 years and older, respectively. Reporting rates for death were higher in older age groups, as expected on the basis of general age-specific mortality in the general adult population.

3. Methodology

A. Dataset

Data were described by the incidence (rate) and the distribution of the adverse events following receipt of COVID-19 vaccinations in the US. The number of adverse events was calculated by adding the people vaccinated on the same day in VAERS. Then this number was divided by the total vaccination on that day to get the rate data. Missing values of vaccination dates in VAERS were imputed by the value of the next record. Bar plot and line charts were used to describe the rate and distribution of the events.

B. Time of Onset Prediction

To predict the time of the onset of adverse events and identify the key predictors, we considered the first event record for each person. The interval (in days) was calculated as the event onset date minus the vaccination date (continuous). Predictors included all 27 (7+17) baseline variables. Data were split into training sets and test sets (8:2). The model was trained by a series of algorithms on the training set (80% sample) as below and Prediction performance was evaluated by mean square error (MSE) on the test set (20% sample).

- **Ordinary least square:** Variable importance was assessed by the sign and magnitude of coefficients.

- **Regularized regression (lasso and ridge):** The optimal regularization parameter is chosen by 10-fold cross-validation. Variable importance was assessed by the same logic above.
- **Random Forest Algorithm:** The number of trees was set to be 500 and the number of predictors sampled for splitting at each node was set to be 8 (p/3). Variable importance was assessed by the mean decrease inaccuracy.
- **Neural network:** For simplification, 2 hidden layers with 2 and 1 neuron in each of the layers and the Sigmoid activation function were used. Variable importance was assessed by the weights of the first layer.

4. Result

A. Data Interpretation

From Dec 20, 2021, to April 22, 2022, a total of 118,746 persons reported adverse events in the VAERS system producing 150,395 adverse events. Of all events, 71605 (47.6%) were reported on the first day of vaccination. The maximum duration between the event and vaccination date was 50 days.

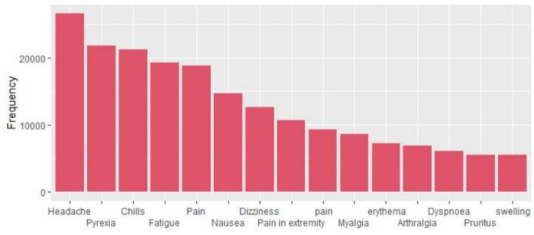

Fig 1: Frequency of Top 15 Symptoms

Figure 14 shows the trend of adverse events over time by the vaccine manufacturer. Two peaks of adverse events showed up in late January and late April. There is no detectable difference between Modena and Pfizer vaccination. Figure 16 shows the rate of adverse events over time. The highest rate occurred last December and early January.

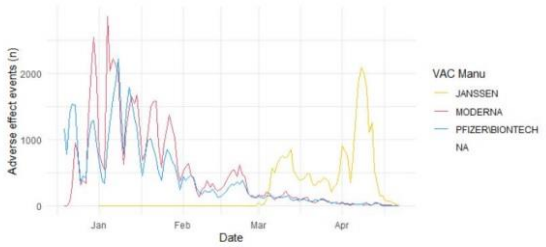

Fig 2: Trends of Adverse Events

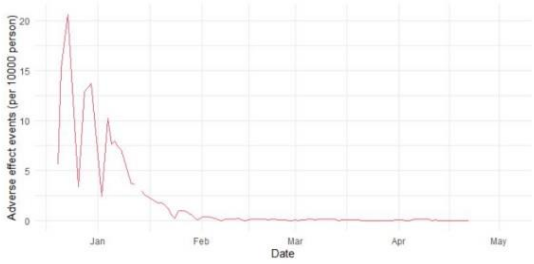

Fig3: Rate of Adverse Events

A. Onset Time Prediction

Data Below shows the performance of different algorithms in predicting the time of the event occurrence. The linear regression (with or without penalization) produced an error of around 5 days and implied the best predictors for a shorter duration of event onset were: anxiety, depression, allergic history, cancer, and diabetes. By contrast, the predictors for a longer duration of event onset were: female sex, thyroid disorder, another medication usage, kidney disease, and anemia. By applying the random forest and Neural network, the error was reduced to less than 3 days. The random forest further showed that age and other medications were of great importance in prediction and the neural network also implied the importance of dementia and kidney disease in prediction. Interestingly, there is evidence indicating that the Moderna was predictive of shorter onset

time while Pfizer was more predictive of longer onset time. This distinction may be due to either the vaccine mechanism or the fact that the waiting time between the first and the second dose of Pfizer is shorter than that of Moderna, allowing for more time for Moderna to develop adverse events.

Table 1: Predicting Event Time by different algorithms

Methods	Training MSE	Test MSE
OLS Method	5.188198	5.298035
Lasso Reg.	5.189760	5.297949
Ridge Reg.	5.215689	5.297750
Random Forest	2.768631	2.944696
Neural Network	2.512959	2.557440

Table 2: Best Predictors for Shorter and Longer duration based on Training and Test MSE

Methods	Best Predictors for Shorter Duration	Best Predictors for Longer Duration
OLS Method	Prevalent CVD Female Sex Arthritis Hypertension Allergic History	Disability Hyperlipidemia Moderna Manufacturer Obesity Depression
Lasso Regression	Prevalent CVD Female Sex Arthritis Asthma	Disability Hyperlipidemia Moderna Obesity
Ridge Regression	Prevalent CVD Pfizer manufacturer Arthritis Asthama Allergic History	Disability Hyperlipidemia Moderna Manufacturer Obesity Depression
Random Forest	Age Allergic History Other Medication use Disability Asthma	Age Allergic History Other Medication use Disability Asthma
Neural Network	Disability Dementia Kidney Disease Hyperlipdidemia Allergic History	Disability Dementia Kidney Disease Hyperlipdidemia Allergic History

5. Conclusion

In this study, the neural network outperformed other algorithms in predicting the event onset time and this may be due to the hidden layer exploiting the interactions between predictors and improving the prediction performance. In Future Work, we will keep exploring machine learning methods that can handle high dimensional sparse features effectively, such as but not limited to Sparse PCA, Naive Bayes Algorithms, and Dense Neural Networks.

7. References

[1] Chapin-Bardales J, Gee J, Myers T. Reactogenicity Following Receipt of mRNA-Based COVID-19 Vaccines. *JAMA*. 2021;325(21):2201–2202. doi:10.1001/jama.2021.5374

[2] Chapin-Bardales J, Myers T, Gee J, Shay DK, Marquez P, Baggs J, Zhang B, Licata C, Shimabukuro TT. Reactogenicity within 2 weeks after mRNA COVID-19 vaccines: Findings from the CDC v-safe surveillance system. Vaccine. 2021 Nov 26;39(48):7066-7073. doi: 10.1016/j.vaccine.2021.10.019. Epub 2021 Oct 16. PMID: 34763946; PMCID: PMC8519847.

[3] Predicting the adverse events following receipt of mRNA-based COVID-19 vaccines: Xiaojuan liu and Yirong Yang

[4] Pilar Vizcarra, Johannes Haemmerle, Hector Velasco, Tamara Velasco, Marina Fernández-Escribano, Alejandro Vallejo, José L. Casado, BNT162b2 mRNA COVID-19 vaccine Reactogenicity: The key role of immunity.

[5] Vaccine side-effects and SARS-CoV-2 infection after vaccination in users of the COVID Symptom Study app in the UK: a prospective observational study: Cristina Menni, Kerstin Klaser, Anna May.