Early trends of adverse event reports associated with COVID-19 vaccines: association rule mining with VAERS data

Thanh M. Nguyen, M.S.[1]; Srijan Sengupta, Ph.D.[2]; Long La, Pharm.D[1]; Allan Fong, M.S.[1]

[1] Medstar Health Research Institute, Washington DC
[2] North Carolina State University, Raleigh, North Carolina

Introduction

- **Vaccine hesitancy** could affect the success of the ongoing COVID-19 vaccination campaign.
- Factors contributing to vaccine hesitancy, including:
  - Vaccine's coincidental temporal relationships to adverse health outcomes
  - Unfamiliarity with vaccine-preventable diseases.

- To address these issues, our paper aims to answer the following questions:
  - **What are the top adverse symptoms associated with COVID-19 vaccines?**
  - **What are the differences in adverse symptoms between Moderna and Pfizer vaccines?**

- Data description: **VAERS** is a national passive surveillance system co-managed by the CDC and the FDA to detect new, unusual, or rare vaccine adverse events. Data cleaning process was summarized in Figure 1.

Methods

- **Association rule mining** (ARM) is a data mining technique to discover relationships (associations, rules) among items in a data set.
- Association rules (AR) are statements A → B, i.e., the antecedent A leads to the subsequent B.
  - For example: the rule chills → headache means that if patients experience chills, they are likely to experience headache.
- We use Apriori algorithm, which computes the strengths of AR via metrics like support, confidence, and lift, as follow:

1. $Support\ (A \rightarrow B) =$
$\frac{number\ of\ patients\ experience\ both\ symptom\ A\ \&\ symptom\ B}{total\ number\ of\ patients}$, i.e.
support calculates the joint probability that symptom A and symptom B appear together in a report

2. Confidence $(A \rightarrow B) =$
$\frac{number\ of\ patients\ experience\ both\ symptom\ A\ \&\ symptom\ B}{total\ number\ of\ patients\ having\ symptom\ A}$,

i.e. confidence measures the conditional probability of symptom B given symptom A

3. Lift $(A \rightarrow B) = \frac{confidence(A \rightarrow B)}{P(B)}$, i.e. lift measures how useful this rule is – how much does A increase the chance of B, while controlling for the chance of B. We only reports rules with lift greater than 50 and support greater than 0.0045.
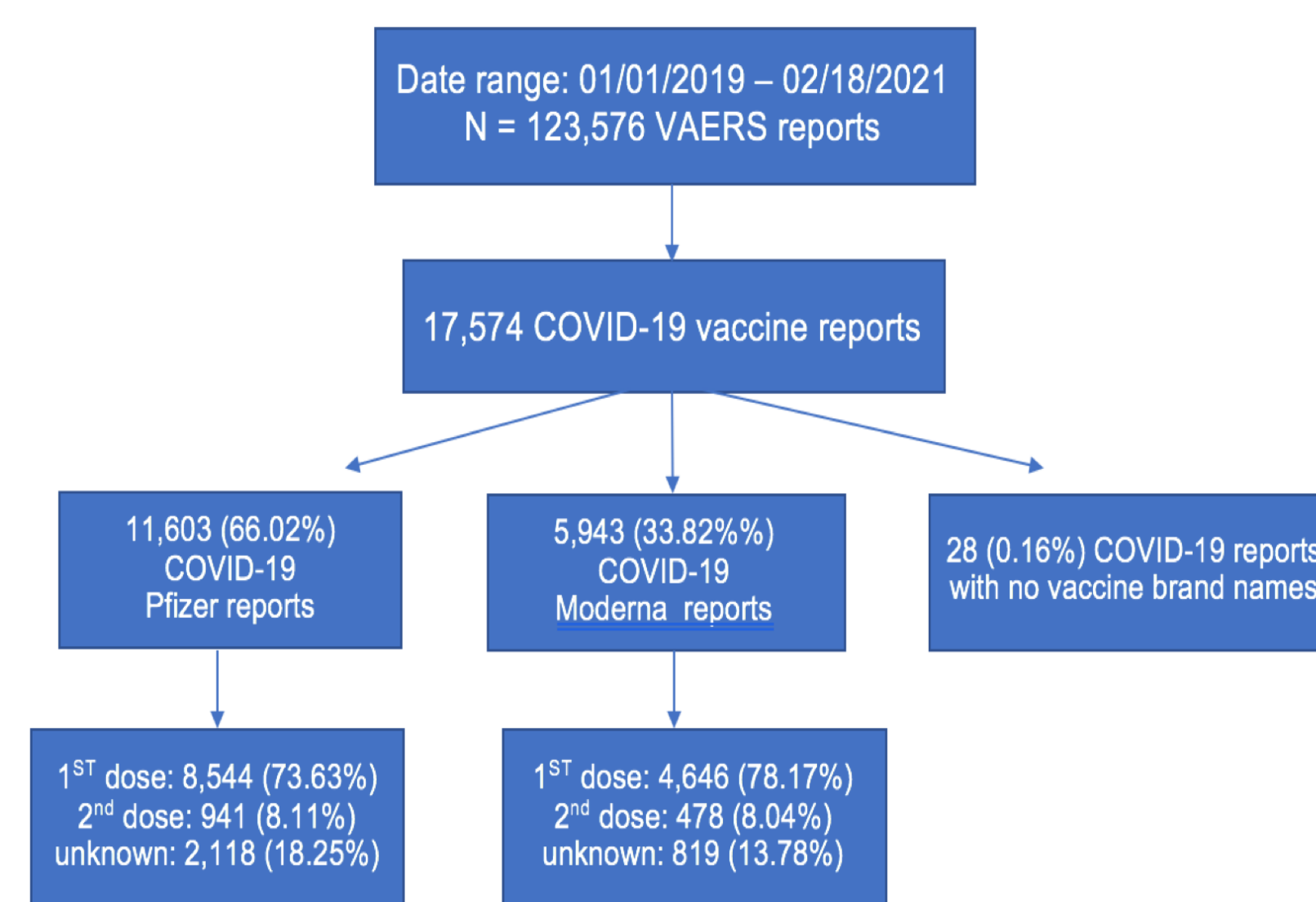
Figure

Figure 1: Schematic of data cleaning process

Results

- **The significance of anosmia and ageusia** in the context of COVID-19 vaccines is unclear, due to the presence of 'sars-cov-2 test positive'.
- **COVID-19 test positive** could be a **confounding variable**
- **Anosmia** and **ageusia** are very closely related as **a pair of symptoms** rather than separate symptoms.
- Complete Blood Count (CBC) labs seem to be more lifted within Moderna reports
- Liver testing seems to be more lifted within Pfizer reports. The clinical significance of this difference is unclear at this time

Limitations and next steps

- VAERS is the voluntary and passive surveillance system, thus there could be under-reporting, lack important details (ethnicity), and data quality issues.

- In the future, we could apply some statistical tests to find the significant rules such as Fisher exact test.

Acknowledgements

Conclusion

- **Adverse symptom reports associated with COVID-19 vaccines are broadly in line with expectations.**
- **The results should be interpreted as broad-based statistical patterns rather than rigorous clinical inference.**

antecedents	consequents	support	lift	leverage	confidence
sars-cov-2 test positive, ageusia	anosmia	0.0042	109.70	0.0042	0.93
anosmia, sars-cov-2 test positive	ageusia	0.0042	79.10	0.0042	0.74
anosmia	ageusia	0.0062	78.22	0.01	0.73
metabolic function test normal	full blood count normal	0.0066	69.92	0.0061	0.81
resuscitation	cardiac arrest	0.0035	60.81	0.0065	0.47

Table 1: top rules for COVID-19 data

antecedents	consequents	support	lift	leverage	confidence
haematocrit normal	haemoglobin normal	0.0034	182.92	0.0034	0.89
haematocrit decreased	haemoglobin decreased	0.0034	114.22	0.0034	0.94
metabolic function test normal	full blood count normal	0.0059	81.96	0.0058	0.85
death, cardiac arrest	resuscitation	0.0038	61.91	0.0037	0.56
resuscitation	cardiac arrest	0.0051	53.91	0.0050	0.56
death, resuscitation	cardiac arrest	0.0038	51.14	0.0037	0.53

Table 2: top rules for Moderna vaccine

antecedents	consequents	support	lift	leverage	confidence
aspartate aminotransferase increased	alanine aminotransferase increased	0.0031	262.01	0.0031	0.90
sars-cov-2 test positive, ageusia	anosmia	0.0061	79.49	0.0061	0.93
metabolic function test normal	full blood count normal	0.0066	69.18	0.0065	0.80
anosmia, sars-cov-2 test positive	ageusia	0.0066	63.59	0.0065	0.74
ageusia	anosmia	0.0084	61.39	0.0082	0.72
chest x-ray, full blood count	metabolic function test	0.0033	56.29	0.0032	0.90
full blood count, troponin	metabolic function test	0.0031	55.87	0.0030	0.90

Table 3: top rules for Pfizer vaccine