Final Project - VAERS User Guide

Mehreen Mirza   Meaghan Brennan   Noa Gittelman-Egan   Yaretsy Castro

Last updated on 2023-12-05

Introduction

The Vaccine Adverse Event Reporting System (VAERS) is a dataset accessible via the U.S. Department of Health and Human Services HHS.gov website. According to the website, VAERS is co-sponsored by the Center of Disease Control and Prevention (CDC), and the Food and Drug Administration (FDA). The website offers three datasets contained via CSV file, all of which can be downloaded into a single Zip file; VAERS Data, VAERS Symptoms, and VAERS Vaccine. For this user guide, we relied upon the 2020 dataset. Datasets are accessible yearly from 1990-2023, with the last update being December 1, 2023. The webpage includes a VAERS Data Use Guide which claims to contain "important information about VAERS from the FDA, a brief description of VAERS, cautions on interpreting VAERS data, definitions of terms, description of files, and list of commonly used abbreviations".

The VAERS homepage tells the user what to do if they are experiencing a reaction following a vaccination and includes quick links to report an adverse event, search the VAERS databases, review resources, along some learning tools including video demonstrations of how to navigate the VAERS website and database. For a patient to report an adverse event, they must navigate to this page and either fill out the online report or they may report using a writable PDF form. Following their self-reported condition, their data is added to the dataset which is updated publicly every year.

This database relies upon the responsible self-reporting of patients or their medical advisors as the reports are not necessarily confirmed cases of an adverse vaccine event. The purpose of this dataset is to provide the U.S. Department of Health and Human Services with an overview of potential adverse reaction reports. The data collection process thus is largely up to the public and medical professionals, while the data analysis process relies on government-appointed epidemiologists, medical researchers, and public health specialists.

Background

The VAERS Dataset was created and is monitored by the CDC and FDA to track reports of mass adverse health effects possibly linked to vaccines. The dataset consists of individual patient's self-reported symptoms that aggregate to the total population of the data. The VAERS describes the population of the data as "all persons reporting adverse event(s) after vaccination" ("Vaccine Adverse Events Reporting System (VAERS) Help" (n.d.)), and the current dataset is composed of reports received from June 1990 to the present, with reports updating weekly through the preceding Friday. This dataset replaced the previous VAERS-1 Form active from July 1990 to June 2017, and had been updated monthly before December 2020 when it transitioned to weekly.

The context of this dataset is vital in understanding its purpose, as it does not consist of verified adverse reactions to vaccines but rather self-reported alerts of symptoms following vaccinations. The dataset intends to provide the CDC data to identify and track potential shared events of adverse health reactions to vaccines, and therefore multiple considerations must be applied when consuming this data. These considerations include the possible complications of unverified self-reporting, the ambiguity of what counts as an "adverse effect", and overall human error and false reporting. Further elements of complications include the

politicization and controversy of vaccines in American society, as well as moments of mass vaccination rates, including COVID-19 vaccines, resulting in increased data points and threats of coincidental inferences.

The dataset includes both categorical and numeric variables. Categorical variables include Symptoms, Allergies, and Recovered. For example, the Recovered variable included categorical response options divided as N (no), Y (yes), U (unknown), and M (missing) and represent that the patient has not recovered, has recovered, recovery status is unknown, and that recovery status is missing, respectively. Numeric variables include Days (indicating the number of days of prolonged symptoms), and Age (indicating the age of the patient).

How to Use/Structure:

The functionality of the dataset is two-sided; the self-reporting interface used for data collection and the dataset accessibility via the VAERS website. To self-report an adverse vaccine requires basic manual data-entry completed by the patient or medical provider via the VAERS website. There are two ways to submit an online report: report an online report or report using a writable PDF form.

Option 1 consists of an online questionnaire that must be completed in one setting and cannot be saved or returned to. Inactivity for 20 minutes will result in the erasure of any entered information and the user will be warned after 15 minutes.

Option 2 consists of a downloadable writable PDF form to a computer. This option is preferable for users who prefer to report offline and who cannot complete the questionnaire all at once. Once complete, the user must return to the VAERS webpage to upload the completed PDF.

Prior to opening the questionnaire, the webpage outlines what information will be required:

- Patient information (age, date of birth, sex)
- Vaccine information (brand name, dosage)
- Date, time, and location administered
- Date and time when adverse event(s) started
- Symptoms and outcome of the adverse event(s)
- Medical tests and laboratory results
- Physician's contact information

Following the submission of a VAERS report, the anonymity of the user will remain by labeling the entry with a VAERS identification number. VAERS will not conduct any follow-up review unless the report is classified as "serious" by the VAERS program, in which case further obtainment of medical records may be pursued in compliance with medical privacy standards.

Privacy Protection

The VAERS dataset is managed by the CDC and the FDA, both of which are agencies of the U.S. Governments that work extensively with medical records and sensitive information. Both public health authorities strictly act under the Health Insurance Portability and Accountability Act (HIPAA) and the Standards for Privacy of Individually Identifiable Health Information (Privacy Rule). These institutions conduct surveillance for possible adverse events associated with vaccines and thus require the consumption of medical records to carry out this initiative for public health.

As technology develops and more medical information is absorbed via online submission forms, concerns about medical privacy are addressed via a Secure Sockets Layer (SSL) for data transmission, an encryption software used for obtaining confidential user information.

Anonymity is ensured by providing a VAERS identification code in place of a patient's name.

Stakeholders

The stakeholders involved in the VAERS dataset include but are not limited to, government organizations and institutions, vaccine developers and distributors, insurance companies, and medical research institutions. Phases of production for this dataset include data infrastructure creation, data collection, cleaning, aggregation, analysis, dissemination, and consumption. The dataset itself is maintained by the Center for Disease Control and Prevention but is widely consumable and accessible to the public as a governmental database. Thus, this dataset is utilized by multiple actors for research purposes including the National Library of Medicine, and has been relied upon in informing the United States government on the status of public health and possible health crises.

The Center for Disease Control and Prevention (CDC) is one of the main managers of VAERS, and their mission is to, as their title suggests, "develop and apply disease prevention and control" as well as promote health education for people in the United States, so their work with this system is intended to add to their research and knowledge of health/health conditions in the United States. The organization has made a statement promising that the benefits to society will always be placed over benefits to the institution itself. As a federal agency, their work is representative not only of themselves but of the country, so it is overseen by the government and constantly being vetted. If their data were to be mishandled by them, the American population's health at large would be endangered, therefore constant scrutiny is important and in place. As an agency with responsibility for disease control, the CDC has an interest in this data because health conditions are always a concern for them and they want to know how possible solutions are working, as well as the general health of as many people as they can. CDC WONDER is a system put in place by the CDC to make their data very transparent and accessible for both health professionals and the general public to have access to, including VAERS.

The Food and Drug Administration (FDA) is an agency in the Department of Health and Human Services. It consists of the Office of the Commissioner and four directorates. The FDA has stated 3 responsibilities: to protect public health by ensuring the safety, efficacy, and security of human and veterinary drugs, biological products, and medical devices; and by ensuring the safety of our nation's food supply, cosmetics, and products that emit radiation; regulate the manufacturing, marketing, and distribution of tobacco products; and advancing the public health by helping to speed innovations that make medical products more effective, safer, and more affordable. In connection to the VAERS dataset, they along with the CDC co-sponsor the dataset. There has recently been an instance where differing points have been made about the dataset. Surgeon General Joseph Ladapo, who runs the Florida Department of Health has stated that there has been an increase of reports filed that show that there may be something wrong with the COVID-19 shots. The FDA and the CDC released a response to rebut his claim and even claimed that his statement was incorrect. The FDA reported that the adverse events should be "compared to background rates in the population."

The National Library of Medicine is the academic journal for the National Center of Biotechnology Information that aims to "advance science and health by providing access to biomedical and genomic information". This database provides researchers and medical professionals a platform to submit, download, learn, develop, analyze, and explore medical research. As a partner of the CDC and an official website of the United States government, this institution's work is government-vetted and has a public obligation for transparency that holds great weight for legal repercussions in the event of irresponsible data management. The NLM is partnered with over 8,100 members of the National Network of Libraries of Medicine as well as being the funding partner for 16 U.S. universities. The platform is responsible for providing free online access to biomedical research literature, as well as conducting research of its own on "biomedical communications systems, methods, technologies, and networks and information dissemination and utilization among health professionals, patients, and the public". The VAERS dataset relates to the NLM as an uploaded dataset that is downloadable and viewable via the NLM website and is featured in over 18,000 pieces of published medical literature. The NLM appears as a third party to the VAERS dataset, as it serves as a public platform for the collection and dissemination of medical data materials, however as a similar governmental database to the VAERS there is an element of relation between the two.

Data Collection

The VAERS data collection standards are that of expecting honesty from those who send them reports, as reports are sent in and included in the VAERS dataset simply based on the description and submission of the report. The collectors encourage reading their checklist beforehand and expect the information provided to them to be accurate and honest, as well as for people to not submit multiple reports. Additionally, the data collectors abide by government regulation standards due to being a government agency and abide by the standard of making sure that this data is available to the public after it is collected. The tools they use are reporting forms and their coding database, the latter of which is used to put all of the reports into a large, organized database that they can easily read and share with others.

The reporting forms consist of an online reporting form and a writable PDF form, both of which are worked on entirely by the person reporting symptoms and then submitted to the VAERS data collectors. The coding skills of the collectors are also a tool, as it will allow them to input the data and responses to the form into their database for further use. The VAERS reporting database is a public health tracking system that entirely depends on the public's participation to self-report adverse vaccine reactions. Because of this, many data collection rituals that in-person data collection would require are not necessary, as the data collection process is entirely virtual and anonymous for the patient. The data is collected on the VAERS website via an online reporting form or via a writable PDF form also located on the website. Data is collected continuously and the form is accessible 24/7. While the unit of observation of the dataset are patients, reports are submitted by healthcare professionals, caregivers, parents, or the patient themselves. The decision to self-report an adverse vaccine effect would require the patient's sense of an adverse physical effect and does not require medical confirmation of an active adverse vaccine effect. Therefore this dataset is used as an early warning system to the CDC and FDA that could offer evidence of an actual adverse vaccine threat.

Commensurability

There are multiple concerns about the commensurability of the dataset. Changes to the data environment that might make the data collection process incommensurable with previous or future collection processes include the status of public health, the public's perception, and trust for vaccines, as well as domestic changes in vaccine creations, distributions, and efficiencies among medical institutions and the pharmaceutical industry. In addition, conversions from physical paper documentation to the virtual database could make the data incommensurable with data that was initially inputted virtually.

A specific point of focus is the conversion process from physical paper documentation to the virtual database, which involves two distinct reporting methods: an online form and a PDF-writable form. To ensure the accurate inclusion of both formats in the dataset, VAERS employs data collectors responsible for inputting data and responses from the forms. It's noteworthy that the data might exhibit variations if the reporting methods entail different sets of questions. Nevertheless, with data collectors constantly reviewing both forms, discrepancies are minimized.

Another concern involves the submission of multiple reports for the same event. VAERS strongly advises against submitting multiple forms, emphasizing the importance of trust in vaccines and organizations for reporting adverse events. The challenge arises if there is a lack of public trust, leading to multiple reports of the same event, potentially from both a patient and a medical organization. There is also a possibility that the patient does not file a report due to medical misdiagnosis. It has been clarified that a report only necessitates the patient's perception of an adverse physical effect and does not mandate medical confirmation.

Important Social Groups

There are three important social groups to consider with the development of the dataset: data proficiency experts, vaccine safety experts, and ambassadors of VAERS-related companies.

Data Proficiency Experts

To elaborate, behind the VAERS dataset and all datasets are individuals who are trained in the technical data and quantitative skills that allow for all functions of the dataset to be accessible and successful. Roles such as Data Analysts, Medical Coders, Medical Analysts, Health Informatics, and Data Strategists provide vital skill sets and knowledge that run behind the scenes of the VAERS dataset. Skills necessary in these roles are typically data science, computer science, and statistics/biostatistics. On the surface level, the creation of the database requires both a public-facing and a private-facing interface. These technical roles of coder, data scientist, and medical professional work in tandem to inform how the database should be created, what it should look like, the information it requests, and how that data is inevitably consumed and calculated. This social group is forced to consider the balance between a public-facing user-friendly interface while maintaining its medical and sophisticated tone. As data scientists and biostatisticians, the intent in analyzing this data is to inform the medical community in the event of a shared adverse vaccine side effect. The VAERS dataset aggregates these reports, but it requires the combined opinion of a medical professional and data analyst to solidify both the reliability of a claim and the frequency at which the claim is supported.

Academics & Researchers at Universities who are considered experts on Vaccine Safety

More specifically, university academics and researchers in vaccine safety, e.g. Dean of Medicine, Director of the Institute for Vaccine Safety at Johns Hopkins University, etc., play a pivotal role in public health. Their daily activities include rigorous research, data analysis, and scholarly publication, focusing on the efficacy and safety of vaccines. These experts are instrumental in shaping public health policies and understanding vaccines, leveraging their extensive training in research methodologies and statistical analysis. While striving for scientific objectivity, they can face potential biases due to funding sources or public health agendas. These influences might subtly sway the direction or interpretation of their research. Moreover, despite their expertise, they sometimes grapple with communicating complex scientific information to the public, which can lead to misunderstandings or mistrust of vaccines. There is also an underrepresentation of diverse perspectives within this group, which can limit the scope and applicability of their research. Ensuring diversity in terms of gender, ethnicity, and socio-economic background is crucial for a comprehensive understanding of vaccine safety across different populations. Still, their role remains critical in the validation and dissemination of vaccine safety data, maintaining high standards of scientific integrity and ethical conduct. Their contributions are central to the development of effective public health strategies and maintaining trust in vaccination programs.

Public facing/public relations individuals for VAERS-related companies

Most noticeably, people with public relations jobs and those who are company faces are very relevant in terms of VAERS data as well as its related companies, if only through the constant controversy surrounding vaccines and general medical practices. Dr. Mandy K. Cohen, as the CDC Director, is an example of a "face," someone who has to answer in the most public-appeasing way possible while still getting across company intentions and answers to the many questions they receive. They are incentivized to speak on VAERS data in ways that will be most beneficial to their company overall and ideally satisfy the majority of their listeners. The way "faces" and those dealing with the public are viewed is generally through the lens of them being "in charge" of their respective companies – and with Cohen, this is true – and as such they have to report on the VAERS data and related data in an honest enough way that does not get them in trouble but also appeals to the general public as much as it possibly can. They are expected to abide by protocols of keeping the images of the companies they represent positive, as well as not break any major regulations that government agencies they represent.

Common Discourses around the Dataset

Given that VAERS is vaccine-related, there are many discourses surrounding it and its level of responsibility in the trustworthiness of vaccines. Many of the first articles to appear when searching for the VAERS dataset are anti-vaccine in nature and also often anti-CDC for their framing of vaccines as exclusively positive. The VAERS dataset's credibility is doubted by those who are anti-vaccine because of its connection to the CDC

and arguments that it doesn't do enough, while those who are pro-vaccine believe that the data it holds is often inaccurate as the reports themselves are not heavily moderated or filtered in case of fakes. The dataset is what one may consider controversial, even for those whom it may seem to benefit. Often, both sides of the "vaccine debate" use it to fulfill their arguments in opposing ways.(Journal n.d.)

The word choice and tone of articles discussing the VAERS dataset imply that VAERS is failing to do what is required – that it's not user-friendly or responsive, and that this is bad. The author(s) do not view VAERS as a good dataset in its collection practices, and later quotes also show that most of the accounts they got from users seem to agree with them. It is seen as confusing and not fully transparent, thus leading to a lot of confusion among the public during the COVID-19 pandemic due to polarization among the public arising from ignorance. The name of the dataset itself is confusing as "adverse" effects are not clearly defined. ("Why the Vaccine Safety Reporting System Should Be Renamed" 2023)

Conclusion

While the accuracy and credibility of the VAERS dataset are questionable and its contents should not be taken at face value on their own, there is a lot of value in the dataset through how it involves reports from real people and there are certainly at least a few honest observations. Additionally, it provides a great resource for seeing discourses around vaccines in the way one can look into how many reports there are at a given time – during a rise in a certain illness, perhaps – and for anyone looking to see some of the public's reactions to vaccines. Navigation of the dataset itself is somewhat confusing due to the number of variables and their shortened titles, as well as the disorganized storage of user comments on their symptoms, but the documentation can still be used and understood adequately for many purposes and provides itself as a passable resource for vaccine discourses on either side of the debate.

References

Journal, British Medical. n.d. "Is the US Reporting System for Vaccine Safety Broken?" Accessed December 4, 2023. https://medicalxpress.com/news/2023-11-vaccine-safety-broken.html.

"Vaccine Adverse Events Reporting System (VAERS) Help." n.d. *Vaccine Adverse Events Reporting System (VAERS) Help.* Centers for Disease Control; Prevention. https://wonder.cdc.gov/wonder/help/vaers.html#Source.

"Why the Vaccine Safety Reporting System Should Be Renamed." 2023. *Penn Today.* https://penntoday.upenn.edu/news/appc-why-vaccine-safety-reporting-system-should-be-renamed.