

Minimizing Signal Loss and Optimizing Pharmacovigilance in VAERS

Jessica Rose¹

Abstract

Background: The Vaccine Adverse Event Reporting System (VAERS) serves as a critical repository for adverse event (AE) data spanning over three decades, yet its passive reporting nature, outdated infrastructure, data inconsistencies, inability to establish causality, and lack of denominator data for incidence calculations limit its pharmacovigilance utility and necessitate modernization.

Objectives: To evaluate three key limitations of VAERS - data quality and inconsistency, inadequate signal detection with no inherent causality inference, and inability to compute incidence rates—and propose a comprehensive framework for its redesign, incorporating enhanced data cleaning, advanced signal detection, and automated causality assessment.

Methods: An extensive literature review was conducted alongside analysis of the VAERS database, focusing on 2021 data. Data cleaning addressed vaccine lot inconsistencies, free-text extraction, and date standardization. MedDRA Preferred Terms (PTs) were condensed to mitigate coding dispersion and recover obscured signals. Proportional Reporting Ratio (PRR) analyses detected safety signals, followed by causality assessments using the Bradford-Hill Criteria (BHC) to generate a Composite Causality Score. Myocarditis following COVID-19 vaccination (particularly Moderna) was used as a

test case for signal detection and automation. A cutting-edge AI tool was evaluated for automated PRR calculation and BHC-based causality ranking. Temporal signals in early-life vaccination (via CAGE_MO variable) were also examined.

Results: Literature confirmed VAERS limitations, including underreporting and data infidelity. Post-cleaning of 2021 data, the PRR for myocarditis was 23.7 (COVID-19 vs. influenza vaccines), with BHC assessment confirming a strong causal link to Moderna COVID-19 vaccines, particularly in male children with elevated incidence based on CDC dose data. Coding dispersion obscured additional signals, while condensation revealed them. The AI tool successfully flagged myocarditis with PRR >2 and calculated a Composite Causality Score of 9.76/10. Analysis of early-life temporal variables highlighted potential for improved signal detection in birth-dose reports.

Conclusions: VAERS requires fundamental upgrades to support inherent disproportionality analyses, causality assessments, and incidence calculations. Integrating automated statistical methods, AI-driven tools, transparent data handling, and higher-granularity variables would enhance data quality, usability, and reliability, transforming VAERS into a more effective public health pharmacovigilance resource.

¹ Independent Data Analyst, Haifa, Israel

Keywords: COVID-19, AE, VAERS, PRR, BHC, myocarditis, AI tool

Corresponding author:

Jessica Rose

Email: jessicarose1974@protonmail.com

Background

The VAERS system is meant to serve as an early-warning system to detect safety signals from data not identified during pre-licensure clinical trials. (1,2) It was originally designed to meet post-licensure surveillance obligations under the National Childhood Vaccine Injury Act (NCVIA) of 1986 (3) to meet pharmacovigilance standards, and operates according to standard operating procedures (SOPs, 2021), (3–5) as part of this pharmacovigilance.

Although VAERS is functional, this functionality is far from optimal. VAERS lacks a mandated standard for the electronic submission of individual case safety reports and this inevitably leads to data inconsistencies across reports, MedDRA coding variability and reduced data quality. (8–19) Since VAERS is fundamentally a passive safety signal detection system, its inherent effectiveness is limited in the face of suboptimal data quality. In the absence of a standardized data input method, poor quality data driven by human error - both the level of initial reporting and subsequent data processing - inevitable loss of numerous genuine safety signals through dilution, misclassification or burial in noise is anticipated. (19)

VAERS data is entered manually with no option for pull-down menus to ensure data consistency across reports. This is problematic because detection of disproportional reporting signals - as done in PRR analyses - will be flawed. (20,21) In addition, follow-up causality assessments will become almost moot in the absence of a PRR flagging a potential safety signal. (21,22) Good quality data is required for flagging.

Underreporting is the result of a multitude of factors but primarily arises as a by-product of the passive nature of VAERS that requires voluntary submission of reports: only a proportion of people will ever report an AE to a pharmacovigilance database (or have one reported for them). (24,25) Oftentimes individuals don't suspect an emergent AE as associated with a vaccine that they received - or administered - so they might not report it to a pharmacovigilance database. In fact, many health care practitioners simply do not believe that vaccines can cause AEs or serious AEs (SAEs), and many general practitioners will simply forego the duty to report. Most (~67%) VAERS reports are indeed filed by nurse and general practitioners, however. (25)

VAERS reports are submitted through an online public web portal, processed by contractors, coded with MedDRA terminology, and eventually made

available for download on the VAERS website. (1,26) The current VAERS platform relies heavily on free-text narrative fields and requires approximately 30 minutes of time to complete a single report. (1) Additionally, the system also employs a multi-page online submission form that incorporates session timeouts, requiring users to complete each page within a prescribed time limit to prevent data loss, which often ironically leads to report loss.

VAERS staff are tasked with the difficult process of processing submitted free-text form raw VAERS data during the vetting process to minimize fraudulent and duplicate entries. (4,5) Following vetting, valid reports are assigned a permanent VAERS ID and uploaded to the front-end VAERS database. Additional information including hospital records and autopsy reports are intended to be requested by trained staff when appropriate as outlined in the SOP. (4) Data loss can often be significant during this imperfect process and can contribute to inconsistent data entry, duplicate records, fragmented longitudinal follow-up, and unexplained deletions or modifications of submitted reports. Date format inconsistencies often arising from mixed regional conventions (e.g., MM/DD/YYYY versus DD/MM/YYYY or YYYY-MM-DD), ambiguous abbreviations (e.g., “09/04/2025” as April 9 or Sept 4), or improper storage of dates as character or string types rather than native date/datetime data types, can lead to incorrect sorting, faulty date arithmetic, parsing failures, and critical errors in time-series analysis or reporting, and this problem is systemic in VAERS.

Questions pertaining to the fidelity of data entry and transfer are exceedingly pertinent when it comes to high granularity variable data such as for the calculated age (in months) variable in VAERS (CAGE_MO). Three variables are derived (calculated) in VAERS data collection (NUMDAYS, CAGE_YR, CAGE_MO), according to the VAERS Data Use Guide. In reference to CAGE_MO: “The sum of the two variables CAGE_YR and CAGE_MO provide the calculated age of a person. For example, if CAGE_YR = 1 and CAGE_MO = 0.5, then the age of the individual is 1.5 years, or 1 year 6 months.” (2) However, it is unclear if the CAGE_MO variable is meant to be interpreted as an integer or a floating-point value. The former would result in the truncation of all fractional values (0.1–0.9) to 0 and in the landscape of infant AE-reporting for infants less than 1 month old resulting in data and temporal signal loss by erroneously assigning infants aged 1–27 days to “0 months”. This would

obscure granular (weekly or even daily) critical early-life temporal signals.

As part of the VAERS SOP, the CDC and the FDA are meant to perform routine VAERS surveillance to identify potential emergent safety signals. (4) Although VAERS AE reports often capture temporal associations, they cannot confirm causal links making it challenging to differentiate true vaccine reactions from unrelated health issues. This is why follow-up investigation is necessary when a causal effect is suspected. Disproportionality metrics such as the PRR are meant to be used to assess associations from safety signals detected in VAERS, while documented application of established causality frameworks such as the WHO-UMC methodology should be used to strengthen follow-up evaluation of causal relationships. (27) The PRR is the key metric for identifying safety signals in VAERS. Bayesian criteria to assess associations and detect potential safety signals could also be implemented as part ensuring optimizing the pharmacovigilance quality of VAERS.

The VAERS system was indeed used in the removal of the Rotavirus vaccine from the market in 1999 following the detection of a disproportional intussusception signal in children. (28) If data quality is poor or inconsistent, then disproportional reporting signals run the risk of being lost, making causal relationships between emergent safety signals and vaccines difficult to discover. The intussusception example is the exception; not the rule.

When a calculated PRR exceeds 2, a causality assessment using the BHC should be mandated to rigorously evaluate potential causal links, ensuring a robust and systematic approach to vaccine safety monitoring.

Incidence rates cannot be calculated using VAERS data because it is designed to be limited to signal detection, rather than quantitative risk assessment. Comprehensive data on total vaccinations administered during vaccination programs, for example, is currently not linked to VAERS. Thus, the absence of reliable denominator data (total doses administered by age, state, and product) prevents accurate incidence rate calculations.

It is currently a requirement to download the VAERS data from the website as three separate comma-separated files for data merging to properly analyze individual VAERS_IDs across multiple variables. (1) Only basic skills (with R, for example) are required to do this, but they are required.

AI tools are inevitably the way of the future with

regard to improving data quality and analysis. (29–32) They can be used ensure VAERS data quality by exploiting their ability to recognize patterns in big data, and to automate requested tasks. Since VAERS relies on a non-standardized data collection method through web-based forms (or manual entry), it is time-consuming, prone to human error, and difficult to scale. Real-time AI tools could parse free-text narratives upon submission, extract missing structured data, assign accurate MedDRA (and optionally ICD-10) codes, detect duplicates, and link follow-up reports to a single permanent individual VAERS_ID identifier. In addition, AI tools could be used to autodetect potential AE-vaccine associations using statistical methods for subsequent assessment of causal relationships from flagged high-probability associations, thus resolving many data inconsistency issues and ensuring that VAERS is optimally and efficiently being used as a pharmacovigilance system, as previously asserted.

It is important to note that it is often the case that several vaccines are given simultaneously resulting in data convolution thus preventing proper attribution of an AE to a specific vaccine. AI tools could also be used to deconvolute this data - a problem that has proven to be extremely difficult to address to date.

Methods

A comprehensive literature review was done primarily using Pubmed (National Institutes of Health (NIH) National Library of Medicine) to ensure relevant publications were peer-reviewed. A technical analysis of the VAERS platform was completed to assess the efficiency of its current architecture to include data collection processes, data accuracy/dirtiness, and its user interface.

For the technical analysis, VAERS domestic data was downloaded as three separate comma separated values (csv) files representing i) general data for each report; ii) the reported AEs or “symptoms”; and iii) vaccine data and three datasets were created: 1. 2021 data only (including all vaccines), 2. 2020-September 2023 data (only COVID-19 vaccine data) for subsequent analyses and 3. all years data for Hepatitis B vaccines.

In each case, using the Language and Environment for Statistical Computing, R, the three csv files were merged using the VAERS ID as a linking variable, and variables of interest selected such as VAERS_ID, VAX_MANU and ONSET_DATE, to

name a few. In addition to merging the three separated data files, an additional variable called “SYMPTOM” was created for each file to concatenate the “SYMPTOM_” data variables (there are up to 20 in some cases) and duplicate VAERS_IDS were removed.

For data cleaning - in lieu of having an Extensible Markup Language (XML) standard as is the case in the FDA Adverse Event Reporting System (FAERS) (33) - date data for the variables “ONSET_DATE”, “DATEDIED”, “VAX_DATE” and “RECVDATE” were converted from character to date format with the month, day, year (“%m/%d/%Y”) structure for consistency. Vaccine lot data were cleaned with respect to consistent vaccine lot nomenclature. For example, Pfizer’s COVID-19 vaccine lots maintain a six-character format: two letters followed directly by four numbers. All letters were converted to uppercase and spaces removed. An algorithm was developed to extract missing age data from the free text variable SYMPTOM_TEXT to feed missing fields in the AGE_YRS variable.

PRR analyses were done by comparing COVID and Influenza vaccine-associated AEs where **A** = Number of reports for the specific drug-event combination (COVID vaccine + specific AE), **B** = Number of reports for the specific drug without the event (COVID vaccine + other AEs), **C** = Number of reports for Influenza vaccine with the event (INFLUENZA vaccine + specific AE) and **D** = Number of reports for Influenza vaccine without the event (INFLUENZA vaccine + other AEs) where the **PRR** = $[A / (A+B)] / [C / (C+D)]$. The Influenza vaccine was chosen as the comparator since it is highly administered and also associated with many AE reports in VAERS. It should be noted that the PRR calculation has limitations, as it does not change when the specific vaccine-related AE event counts are very large or very small. (9) Nonetheless, it is the standard for primary detection of safety signals used by CDC and FDA employees (18) and is used often to assess safety signals from data in peer-reviewed publications. (20,21,34–36)

As proof-of-concept with regard to PRR and causality assessment, the cleaned VAERS 2021 file was fed to the Manus 1.5 Lite AI-tool. (37) The AI tool was prompted to complete a PRR analysis and BHC assessment and to output a value ranging from 1–10 of the likelihood of a causal relationship between Myocarditis and the Moderna vaccine. This BHC assessment was done by generating a Composite Causality Score (CCS) based on a quantitative score

derived from the PRR and chi-squared values, and a qualitative score (QLS) derived from the sum of weighted scores for each of the ten BHC to produce a 0-10 scale causality score. First, the Proportional Reporting Ratio (PRR) and chi-squared statistic are calculated from a 2×2 contingency table of AE reports, then a Quantitative Score (QS) from 0–4 is assigned based on signal strength ($PRR \geq 2$, $\chi^2 \geq 4$, cases ≥ 3). The BHC (Strength, Consistency, Specificity, Temporality, Biological Gradient, Plausibility, Coherence, Experiment, Analogy) are evaluated on a 0-3 scale and multiplied by their assigned weights (Strength, Gradient, Experiment: 1.5; Consistency, Plausibility, Coherence, Reversibility: 1.0; Specificity, Analogy: 0.5) and summed to obtain the QLS (max 28.5), with Temporality serving as a mandatory gatekeeper (score=0 invalidates the assessment). Finally, the QLS is normalized to a 0-6 scale by multiplying $(QLS/28.5) \times 6$, then added to the QS to obtain the final CCS (0–10), which maps to WHO-UMC causality categories: 0=Unlikely, 0.1–3.0=Possible, 3.1–7.0=Probable, 7.1–10.0=Certain/Definite. A full methodology is attached as a supplemental document (**Supplementary Materials**).

Myocarditis was selected as an AE of interest since the association between myocarditis and Moderna is well-established as it appears in the product insert with a warning: “The observed risk has been highest in males 12 years through 24 years of age.” (38) A full report, including a contingency table, PRR calculation, detailed BHC assessment for each criterion was provided with the final likelihood score was produced.

Denominator data stratified by age according to CDC guidelines were collected from their website in order to determine Myocarditis incidence rates. (39) The denominator was the estimated number of people who received at least one dose of a COVID-19 vaccine by the end of September 2023. This particular analysis required that an additional VAERS domestic dataset was created to include only COVID-19 vaccines (filter(grepl(“COVID”, VAX_TYPE))) from 2020-September 2023. Myocarditis as an AE was filtered out (filter(grepl(“Myocarditis”, SYMPTOM))) to yield the final data set used to calculate incidence rates. The data was stratified using CDC age grouping (0–4, 5–11, 12–17, 18–24, 25–39, 40–49, 50–64, 65–74, 75 plus) and normalized to CDC dose data. Manus was independently prompted to collect CDC denominator data in order to confirm observed results and to confirm that age is deterministic for Myocarditis incidence rates.

Introduced biases found at the level of the MedDRA Preferred Term (PT) level of coding were examined by comparing PRRs between condensed and non-condensed AEs and revealed significant increase in PRR values following condensation. Manus performed a comprehensive comparative PRR analysis that demonstrates the impact of MedDRA coding dispersion on pharmacovigilance signal detection in the case of foetal loss and heart attack as a proof-of-concept. The expectation in pharmacovigilance is that condensation should increase the PRR by overcoming the dilution effect of dispersed reporting.

A comparative analysis was also done to demonstrate the difference between the results of a query for infant death reports associated with the Hepatitis B vaccine (`filter(grepl('HepB', VAX_NAME) | grepl('Hep B', VAX_NAME))`) - a birth-dose vaccine - when the floating-point precision of CAGE_MO was ignored and not. If the CAGE_MO field were treated as an integer, all reports with a value greater than 0.0 but less than 1.0 would be incorrectly rounded down to 0.0, effectively losing the temporal information that places them at Day 1, Day 6, Day 9, etc.

As a more concrete example, the decimal value in CAGE_MO (ie: 0.3) represents a fraction of a month when CAGE_YR = 0. Thus, 0.3 months $\approx$ 9 days (0.3×30.44 days), identifying the infant as a 9-day-old newborn. Ignoring fractional-month interpretation would misclassify these events as infants at 0 months completely obscuring safety signals in infants.

Results

Data Cleaning and Validation

The merged and structured file for 2021 VAERS domestic data contained 753,047 unique VAERS_IDs. Vaccine lot data quality was improved substantially simply by converting all letters to uppercase and by removing spaces. As an example, the number of vaccine lot data points for lot EH9899 increased by 7.8%, when just these two simple corrections were applied. This simple vaccine lot cleaning process successfully recovered 8,871 VAERS reports that were previously unusable for a strict, standardized lot number analysis (due to spaces or incorrect casing) yielding a 1.74% increase in the amount of data available for lot-specific signal detection.

Age data were extracted from the SYMPTOM_TEXT field using a custom algorithm, increasing the completeness of the AGE_YRS varia-

ble from 89.2% to 91.5%. A sample of lot numbers from VAERS reports in **Figure 1** demonstrates how lot numbers require standardization prior to analysis, as several versions of the same lot number exist (e.g. letters in upper/lowercase, presence of spaces). **Figure 1** also demonstrates that the dates require conversion from character format (<chr>) to date format (<data>) prior to analysis. This is vital with regard to any analysis pertaining to temporality and causality assessments since if the dates are not in the right format, the differences between ONSET_DATE and VAX_DATE cannot be calculated, for example. Even the NUMDAYS variable calculated by VAERS employees suffer likely from date structure issues.

A lack of standardized reporting may impede the quality of analysis. This inevitably not only results in data point loss but can materially affect analytical results including PRR calculations and correlations. For example, some vaccine lots are more highly associated with certain AEs, like death. (40) If the vaccine lot data is too substandard, that signal might be hidden. When the correlations for chills and the vaccine lot EH9899 pre- and post-vaccine lot data cleaning are calculated this becomes obvious with an increase in Pearson correlation coefficient of +0.00138 with 138 recovered points in vaccine lot data whereby 30 of those recovered reports were associated with the outcome of Chills. Data cleaning not only recovers data but can alter (in this case, strengthen) the statistical signal for specific AEs.

Myocarditis Case Study

Myocarditis is a well-established AE associated with the COVID vaccines, especially in young males. (41–50) The PRR for myocarditis when comparing the case counts in COVID vs. Influenza context is 23.7 (95% CI 1.010–1.011), Fisher's exact $p \leq 0.005$ when comparing COVID-19 and Influenza vaccines for 2021. This significantly high value warrants a causality assessment. It is entirely plausible that a causal link exists since spikeopathy in the myocardium would necessarily lead to myocarditis. (51)

Signal identification and causal assessments can be automated by integrated AI-tools. Myocarditis is an emergent safety signal in young boys in the COVID-19 vaccine context according to VAERS 2021 data, has a PRR of 23.7 thus the requirement for a causality assessment applies (**Figure 2**). The 2021 VAERS data file was uploaded to an AI tool called Manus. Manus was prompted to confirm the

PRR calculation and then to go one step further and perform a causality assessment using the BHC and to return a CCS between 1–10 based on satisfying the 9.76/10 BHC.

Manus was not only able to confirm the PRR, but output a full report on myocarditis as a safety signal - complete with BHC results, in seconds. The CCS was 9.76/10 - strong evidence for a causal association between COVID-19 vaccines and myocarditis. It is worth emphasizing that, up to this point, the entire analysis - including the design, testing, and execution of all prompts used with the AI tool - was performed single-handedly by the author of this manuscript. The technology now exists to fully automate this process; what remains is to integrate these tools directly into the VAERS platform.

Addressing Signal Loss: The Impact of MedDRA Dispersion

Introduced biases found at the level of the MedDRA Preferred Term (PT) level of coding were examined by comparing PRRs between condensed and non-condensed AEs and revealed a significant increase in PRR values following condensation. Following manual comparative PRR analyses using R, Manus was prompted to perform a comprehensive comparative PRR analysis that confirmed the impact of MedDRA coding dispersion on pharmacovigilance signal detection in the case of foetal loss and heart attack, as a proof-of-concept. The expectation in pharmacovigilance is that condensation should increase the PRR by overcoming the dilution effect of dispersed reporting, and this is exactly what was observed.

When PRRs were calculated using R for “Abortion spontaneous” (pregnancy loss) - the most prevalently used MedDRA code for foetal loss - to include additional seven AEs (“Foetal loss”, “Foetal death”, “Miscarriage”, “Stillbirth”, “Vaginal haemorrhage”, “Abortion”, “Uterine dilation and curettage”), the signal rose by 144% (from PRR = 1.89 to PRR = 4.62). Likewise, for Cardiac arrest (“Heart attack”, “Resuscitation”, “Cardio-respiratory arrest”, “Tachycardia”, “Pulseless electrical activity”, “Sudden cardiac arrest”, “Chest pain” (very typical with myocarditis) and “Coronary infarction”), the PRR increased by 43% (from PRR = 1.8 to PRR = 2.58). In both cases, the PRR changes such that a causality assessment is warranted (PRR > 2) following appropriate coding by removing dispersion of signals.

Condensation of dispersed MedDRA terms can dramatically affect pharmacovigilance signals. When Manus was prompted to complete this analysis, fetal Loss demonstrated the most pronounced effect with a 145.8% increase in PRR when all 8 related terms were combined, while Cardiac Arrest showed a 37.9% increase when the 9 related terms were combined (**Figure 3**).

Manus was prompted to replicate this idea using other AEs known to suffer from MedDRA coding dispersion and concluded that MedDRA coding dispersion has a highly variable, AE-specific impact on pharmacovigilance signals. Interestingly, the same condensation strategy that amplifies the signal for one AE may dilute it for another. It identified the top 5 AEs that suffer from code dispersal as described by the largest PRR changes (**Figure 4**).

Condensation can either recover a dispersed signal or dilute an existing signal if the added terms are reported more frequently in the control group. A result from the AI tool was the discovery of signal amplification leaders (Fetal loss (+146.10%) and Cardiac arrest (+38.16%)) and also previously unidentified signal dilution leaders (Myocarditis/Pericarditis (-72.67%), Thrombosis/Blood Clots (-67.58%), and Fatigue (-39.44%)). What this means is that the non-condensed terms - myocarditis, thrombosis and fatigue are already the primary signal carriers, and additional terms do not affect the signal. This is interesting since these AEs have been flagged as problematic associated AEs in the COVID-19 vaccine context, whereas fetal loss and cardiac arrest reporting is clearly highly dispersed across multiple MedDRA terms in the dataset (strong signal recovery when condensed) and have NOT been flagged as problematic associated AEs in the literature. In fact, claims that COVID-19 injectable products affect fertility have been contested in the literature. (52–55)

Denominator data according to CDC statistics (56) were incorporated and incidence rates calculated. As a case study, myocarditis was selected and rates were calculated according to CDC age grouping. Since CDC data has only been made publicly available until October 2023, a new VAERS domestic data set covering 2021–October 2023 (998.2 MB; 880,960 VAERS_IDs) was made and queried for myocarditis reports (4.8 MB; 1,366 VAERS_IDs). (56) These were then stratified by CDC age group and incidence rates calculated using CDC data per age group (**Figure 5**).

These reporting rates for myocarditis are in alignment with published findings. (41–50)

Finally, a comparative analysis was done for infant death in VAERS in the context of all Hepatitis B vaccines using the CAGE_MO variable to determine the age of death of the infants and toddlers up to 2 years old by week (24.7 MB; 32,504 VAERS_IDs). Many (N = 62 when comparing all infants under the age of 1) death reports in the Hepatitis B vaccine context are lost if the floating-point precision of CAGE_MO is ignored.

The granularity of the data for infants less than 1 month old completely changes the picture of the temporal view of the data when the float data is presented as days (ie: CAGE_MO = 0.3 could indicate 4 months added to CAGE_YRS, or it indicate an infant 9 days old. A month is approximately 30.44 days and thus 0.3 months is approximately 9 days ($0.3 * 30.44$ days)). As shown in **Figure 6**, the picture of child death in the context of the birth dose for Hepatitis B is far more dire when considering the former interpretation of the data. It becomes very clear that when the infants in the 0 months-old bin are properly distributed into a 0–30 days-old bin, that most died on day 1.

The main graph in **Figure 6** shows the number of death reports by composite age in months. The distribution confirms that a large number of reports are concentrated in the first month of life (Month 0), which is highlighted in red. The total count for Month 0 is 165 reports. The magnified inset plot provides the fine-grained temporal resolution of the Month 0 bar. It displays the distribution of the 165 reports by Time-to-Event in Days (calculated by converting the fractional CAGE_MO value to days: $CAGE_MO * 30.44$). The inset plot reveals that the reports are not uniformly distributed across the first month but are highly clustered, with a distinct peak of 18 reports occurring at day 1 of life (highlighted in black). All reports with a value greater than 0.0 but less than 1.0 would be incorrectly rounded down to 0.0, effectively losing the temporal information that places them at Day 1, Day 6, Day 9, etc.

When assessing safety signals in VAERS in newborns (less than 1 year old), the correct methodology would involve converting the CAGE_MO variable to days or weeks and this becomes clear when plotting the deaths per week using the AGE_YRS variable (**Figure 7 - left**) and the CAGE_MO variable (**Figure 7 - right**). 50 infant deaths are lost on week 1 alone when using the AGE_YRS variable.

Relying on the AGE_YRS field for fine-grained filtering of infants is a critical error and materially changes the output as shown in **Figure 7**, because of all entries are rounded down to 1 in the

AGE_YRS variable. This is especially relevant in the context of infant deaths reported in the context of temporal associations with vaccine administration. It is also clear from **Figure 7** that both the correct age field (CAGE_MO) and the correct data type (floating-point precision) are necessary to conduct a valid analysis of safety signals in the youngest population.

Discussion

In lieu of VAERS having an eXtensible Markup Language (XML) infrastructure – which refers to the foundational systems and technologies that support the creation, storage, management and exchange of data, a user account set-up should precede user/reporter reporting sessions such as is done in the case of FAERS. (See Safety Reporting Portal (SRP)). VAERS online reporting sessions should be savable, where updates can be made by the user/reporter to existing reports in a timely fashion. Each VAERS report should be assigned a unique VAERS ID and sessions should be saved during/throughout the reporting process to prevent data loss during submission and for subsequent submissions as updates. Currently, multiple VAERS IDs are created for a single individual in the case where a report is submitted in the context of an SAE (like myocarditis) that may be followed by subsequent death. A single VAERS ID should be assigned per person and a system where the user/reporter has access via a user account would ensure this. Timeouts should be either lengthened or removed entirely. Users with an account could log back into a saved report to create updates. In addition, online communication between VAERS employees (vetting team) and the user/reporter would improve if the user had an account available to VAERS employees.

Pull-down menus with pre-loaded data collection fields are essential for ensuring data consistency with regard to facilitating data searches and prevention of typos. This is especially important in the context of drug names, vaccine lot numbers and manufacturer names. These should be entered according to manufacturer's pre-designated lot format. For example, in the case of Pfizer/BioNTech/Comirnaty COVID-19 vaccines, the format for a lot is 6 characters: two alpha-numeric followed by 4 digits. VAERS vaccine lot data is very inconsistently reported effectively ensuring a loss of ability to use this variable for data analytics as demonstrated herein. Dates should also be formatted to allow only a single date structure (DD/MM/YYYY) using a date selector if manual reporting persists.

AI tools offer transformative potential for data refinement within VAERS. As proposed, these tools can extract information from free-text fields (e.g., SYMPTOM_TEXT) to populate missing fields, such as manufacturer data or drug details, generating a standardized spreadsheet with a comprehensive variable list. In addition, challenges in attributing adverse events to specific vaccines (in the case where multiple vaccines are administered at once) or even individual antigens, could potentially be addressed through advanced analytical approaches, including comparative disproportionality analyses across similar vaccines or AI/machine learning methods for improved pattern detection and covariate adjustment in complex reporting data (though no standardized protocol currently exists for reliable deconvolution of component-specific safety signals, and this is not detailed herein).

Precision AI tools could enhance the efficiency of VAERS as a reliable public health tool by parsing free text and moving this data into distinct fields. They could also be used to update the existing VAERS infrastructure by using an XML structure. This is important in terms of schema matching. It is often the case between different data sets that schemas do not match and this could easily be resolved. As previously discussed, VAERS IDs are often reused and/or removed due to duplicate reports and because follow-up reports are hard to link automatically. An AI tool that addressed this issue would resolve this problem. Each record should have a distinct VAERS IDS that autoincrements. If a report is deemed non-releasable, then that VAERS ID would be unpublished with a written note by the ID, as opposed to potentially being removed and/or reused. It is my belief that the full dataset – including demographic not available in the VAERS data available for download - should be made available to the public. This includes fields collected for race, pregnancy and type of reporting party.

If an electronic connection to VAERS was developed to facilitate the reporting from large medical systems so that bulk uploads were possible, this would indeed alleviate the problem of siloed data (centralized data). Centralized data is data that is under the control and maintenance of specific organizations or entities. It can lead to problems such as data non-transparency, user non-accessibility and data mutability issues whereby there is a single controller of the database. If there are no checks and balances with regard to preventing biases in analysis, it would be careless to assume that biases wouldn't occur. It would also solve the float-versus-

integer data issue, and data origin (API or other interface) issues. It is very arguable that the quality of VAERS data and subsequent analyses would increase significantly if it were made more accessible and transparent for the users/reporters and thus, not be centrally-controlled or siloed.

Data siloing is a very important consideration when efficient health data monitoring is the goal. Interoperable connections between both local and state data collection systems and VAERS would ensure more comprehensive and user-friendly pharmacovigilance, while automating multi-directional data exchange in real time, and could also solve the lack of denominator issue for calculating rates.

Denominator data should be included to develop a comprehensive, age-stratified dataset of total vaccinations administered per state. VAERS reports allow users to specify the state where a vaccine was administered, and this information can be seamlessly matched with the state-level vaccination dataset aligned with CDC state-level records. By leveraging decentralized platforms and AI tools, both vaccination and AE data can be made available in real time, enabling dynamic, transparent, and precise monitoring of vaccine safety across populations.

The recommendations herein are indeed conceptual, but they leverage emergent technologies such as AI tools (Manus) to improve data quality and analysis. This was clearly demonstrated in the data-driven validation of this concept demonstrating the dilution effect of dispersed AEs to accurately assess signals and pharmacovigilance itself. These recommendations are in full compliance with HHS guidelines for the use and interpretation of VAERS data, and also incorporate the idea of leveraging VAERS' data collection and processing methodologies.

Conclusion

Revamping VAERS to minimize signal loss and optimize pharmacovigilance will require integration of its existing infrastructure with new AI-driven solutions. Ultimately, a modernized VAERS would ensure robust vaccine safety monitoring that would foster transparent, efficient health data management. VAERS is a powerful tool for early detection of unusual patterns, temporal clusters or disproportionate clusters of reports of specific AEs following administration of a particular vaccine, but it could be improved.

Rapid identification of statistical associations be-

tween specific AEs, vaccine characteristics and individual demographic data could be easily parsed out using AI tools. Automated clustering, disproportionality analyses and stratification could be done followed by formal epidemiological investigation or targeted follow-up. These tools would not replace causality assessments performed by CDC and FDA, but could prompt a more timely and efficient system.

As evidenced throughout this paper, the existing VAERS infrastructure suffers from many limitations: many of which reside at the data entry level,

but most of which can be resolved as outlined herein. VAERS would heal nicely from a face-lift, and the public - and public health itself - would all benefit from a higher standard of VAERS functionality and credibility.

Acknowledgment and funding information

The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Figure 1. Demonstration of VAERS vaccine lot (VAX_LOT/VAX_LOT_UPPER (cleaned)) and date (RECVDATE/TODAYS_DATE) data showing the requirement for data cleaning

AGE_YRS <dbl>	VAX_LOT <chr>	VAX_LOT_UPPER <chr>	RECVDATE <date>	TODAYS_DATE <chr>
65	EW0186	EW0186	2023-01-26	01/26/2023
34	013H22A	013H22A	2023-02-02	02/02/2023
12	GJ6743	GJ6743	2023-02-05	02/05/2023
76	ew0176	EW0176	2023-02-07	02/07/2023
49	EN6199	EN6199	2023-02-07	02/07/2023
49	GL0447	GL0447	2023-02-10	02/10/2023
74	EW0184	EW0184	2023-02-15	02/15/2023
74	EM 9810	EM9810	2023-02-21	02/21/2023
65	027L20A	027L20A	2023-02-23	02/23/2023
74	AS7143C	AS7143C	2023-03-03	03/03/2023

Figure 2. Validation of Myocarditis as a Safety Signal via Proportional Reporting Ratio (PRR) (when comparing COVID-19 and Influenza vaccines) with Final CCS

	Myocarditis (AE)	Non-myocarditis (Other AEs)	Total
COVID-19 Vaccine (Drug)	A = 7,566	B = 703,148	A+B = 710,714
Influenza Vaccine (Other Drugs)	C = 1	D = 2,225	C+D = 2,226
Total	A+C = 7,567	B+D = 705,373	N = 712,940

#	Criterion	Evidence for COVID-19 Vaccine & Myocarditis	Score
1	Strength	PRR = 23.71 (Very Strong). Confirmed by high Relative Risk in cohort studies.	3
2	Consistency	Confirmed across multiple international databases (VAERS, EudraVigilance) and cohort studies in various countries.	3
3	Specificity	Highly specific to mRNA vaccines (Pfizer/Moderna) and specific age/sex groups (young males).	3
4	Temporality	Clear onset within 7 days of vaccination; dechallenge/rechallenge documented in case reports.	3
5	Biological Gradient	Higher incidence after the second dose than the first, and higher incidence in younger age groups.	3
6	Plausibility	Multiple plausible mechanisms identified: molecular mimicry, direct spike protein toxicity, and immune-mediated inflammation.	3
7	Coherence	Coherent with the known function of the vaccine (strong immune stimulation) and the pathology of immune-mediated myocarditis.	3
8	Experiment	Experimental evidence from animal models and <i>in vitro</i> studies showing spike protein can induce inflammation in cardiomyocytes.	2
9	Analogy	Weak analogy to other vaccines, but strong analogy within the mRNA platform (Moderna also shows signal).	2

Criterion	Value	Threshold Met
PRR	23.71	≥ 10
Chi-squared (χ^2)	20.95	≥ 4
Case Count (A)	7,566	≥ 3

Metric	Value
Quantitative Score (QS)	4 (Max)
Qualitative Score (QLS)	24.5 / 25.5
Final Composite Causality Score (CCS)	9.76
WHO-UMC Category	Certain/Definite

Figure 3. Visualizing Signal Loss and Recovery through MedDRA Condensation. MedDRA coding dispersion impact showing very high signal amplification with respect to fetal loss and cardiac arrest

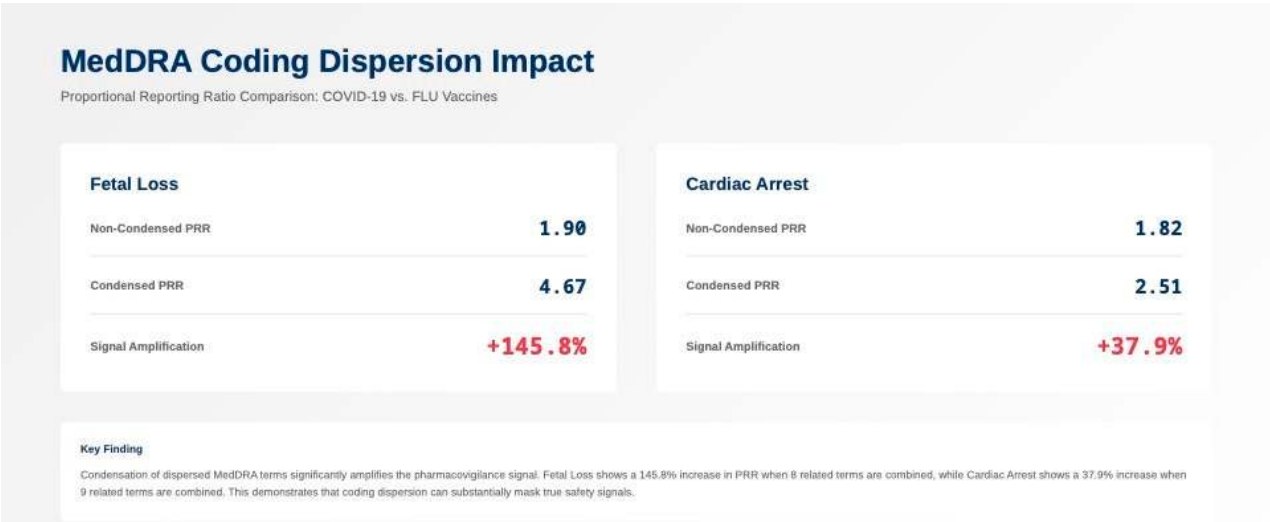
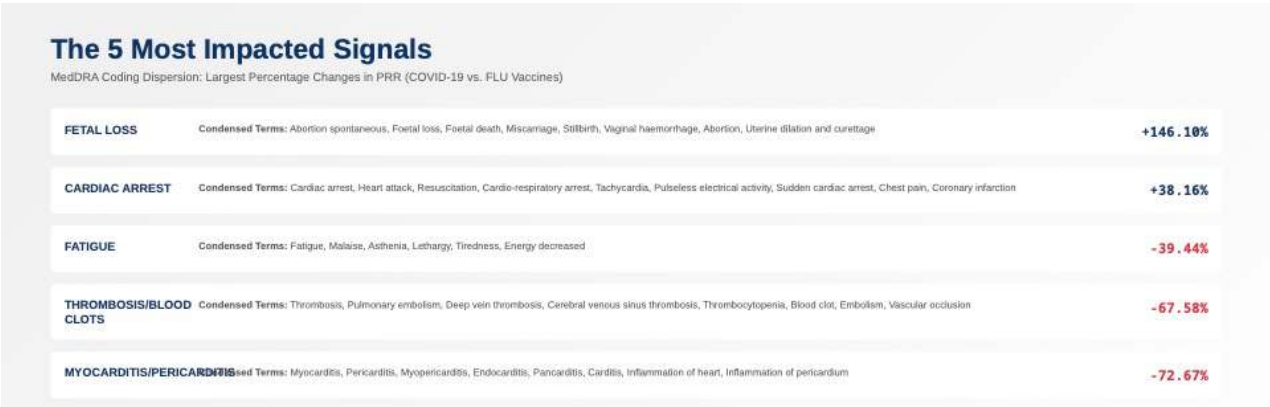


Figure 4. The 5 most impacted signals with regard to MedDRA coding dispersal



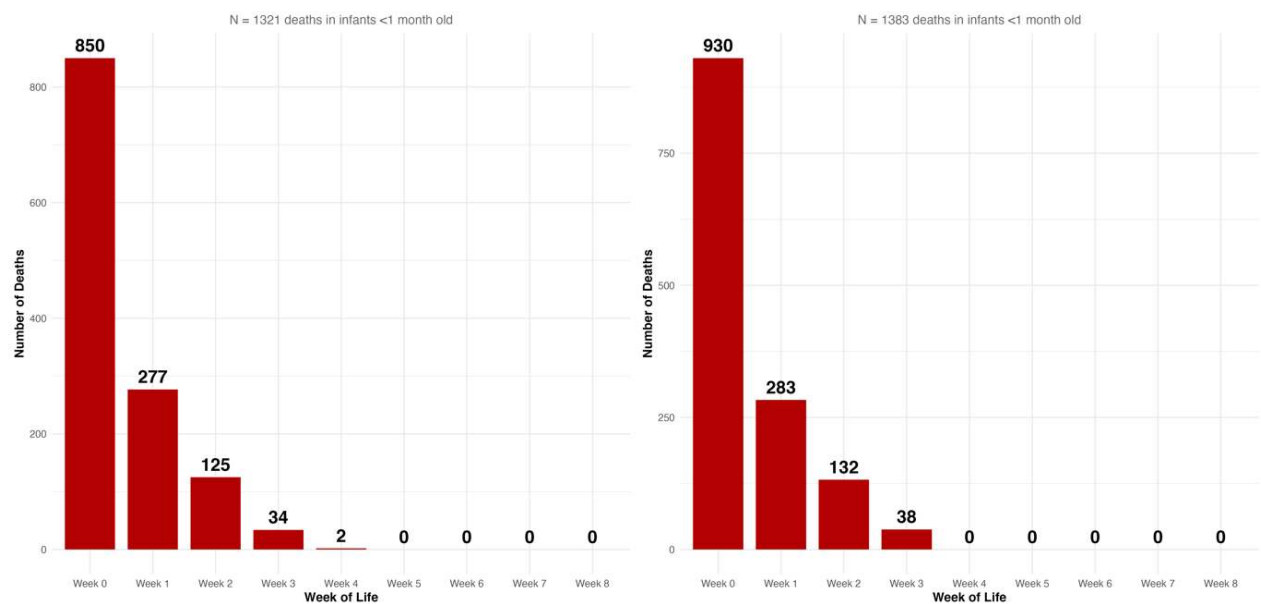
Normalized Myocarditis Reporting Rate by Age Group (VAERS Data through September 2023)



HepB Death Reports in Infants (< 2 Years): Analysis: Jessica Rose



Figure 7. VAERS newborn data (age restricted to 1 year or less) showing comparison of counts when using AGE_YRS (left) vs. CAGE_MO (right) variables



References

1. U.S. Department of Health and Human Services. (n.d.). Vaccine Adverse Event Reporting System (VAERS). <https://vaers.hhs.gov>
2. U.S. Department of Health and Human Services. (2020, November). VAERS data use guide. https://vaers.hhs.gov/docs/VAERSDataUse-Guide_November2020.pdf
3. National Childhood Vaccine Injury Act of 1986, H.R. 5546, 99th Cong. (1986). <https://www.congress.gov/bill/99th-congress/house-bill/5546>
4. Centers for Disease Control and Prevention. (2021, January). Vaccine Adverse Event Reporting System (VAERS) standard operating procedures for COVID-19. <https://icandecide.org/wp-content/uploads/2023/02/VAERS-SOP-29-JAN-2021.pdf>
5. National Institute on Aging. (2018). NIA adverse event and serious adverse event guidelines. <https://www.nia.nih.gov/sites/default/files/2018-09/nia-ae-and-sae-guidelines-2018.pdf>
6. World Health Organization. (2010). Minimum requirements for a functional pharmacovigilance system. <https://docslib.org/doc/11315194/minimum-requirements-for-a-functional-pharmacovigilance-system>
7. World Health Organization. (2021). Regulation and prequalification: What is pharmacovigilance? <https://www.who.int/teams/regulation-prequalification/regulation-and-safety/pharmacovigilance>
8. Block, J. (2023). Is the US's Vaccine Adverse Event Reporting System broken? *BMJ*, 383, Article p2582. <https://doi.org/10.1136/bmj.p2582>
9. Rose, J. (2021). Critical appraisal of VAERS pharmacovigilance: Is the U.S. Vaccine Adverse Events Reporting System (VAERS) a functioning pharmacovigilance system? *Science, Public Health Policy and the Law*, 3, 65–89. <https://publichealthpolicyjournal.com/critical-appraisal-of-vaers-pharmacovigilance-is-the-u-s-vaccine-adverse-events-reporting-system-vaers-a-functioning-pharmacovigilance-system/>
10. National Vaccine Information Center. (n.d.). VAERS Wayback Machine. <https://medalerts.org/vaersdb/wayback/>
11. Varricchio, F., Iskander, J., Destefano, F., Ball, R., Pless, R., Braun, M. M., & Chen, R. T. (2004). Understanding vaccine safety information from the Vaccine Adverse Event Reporting System. *The Pediatric Infectious Disease Journal*, 23(4), 287–294. <https://doi.org/10.1097/00006454-200404000-00002>
12. Singleton, J. A., Lloyd, J. C., Mootrey, G. T., Salive, M. E., & Chen, R. T. (1999). An overview of the vaccine adverse event reporting system (VAERS) as a surveillance system. *Vaccine*, 17(22), 2908–2917. [https://doi.org/10.1016/S0264-410X\(99\)00132-2](https://doi.org/10.1016/S0264-410X(99)00132-2)
13. Geier, D. A., & Geier, M. R. (2004). A review of the Vaccine Adverse Event Reporting System database. *Expert Opinion on Pharmacotherapy*, 5(3), 691–698. <https://doi.org/10.1517/14656566.5.3.691>
14. Zhou, W., Pool, V., Iskander, J. K., English-Bullard, R., Ball, R., Wise, R. P., Haber, P., Pless, R. P., Mootrey, G., Ellenberg, S. S., Braun, M. M., & Chen, R. T. (2003). Surveillance for safety after immunization: Vaccine Adverse Event Reporting System (VAERS)—United States, 1991–2001. *MMWR Surveillance Summaries*, 52(1), 1–24. (Erratum published 2003, *MMWR Morbidity and Mortality Weekly Report*, 52(6), 113). <https://www.cdc.gov/mmwr/preview/mmwrhtml/ss5201a1.htm>
15. Iskander, J. K., Miller, E. R., & Chen, R. T. (2004). The role of the Vaccine Adverse Event Reporting System (VAERS) in monitoring vaccine safety. *Pediatric Annals*, 33(9), 599–606. <https://doi.org/10.3928/0090-4481-20040901-11>
16. Varricchio, F., Iskander, J., Destefano, F., Ball, R., Pless, R., Braun, M. M., & Chen, R. T. (2004). Understanding vaccine safety information from the Vaccine Adverse Event Reporting System. *The Pediatric Infectious Disease Journal*, 23(4), 287–294.
17. Chen, R. T., Rastogi, S. C., Mullen, J. R., Hayes, S. W., Cochi, S. L., Donlon, J. A., & Wassilak, S. G. (1994). The Vaccine Adverse Event Reporting System (VAERS). *Vaccine*, 12(6), 542–550. [https://doi.org/10.1016/0264-410X\(94\)90315-8](https://doi.org/10.1016/0264-410X(94)90315-8)
18. Shimabukuro, T. T., Nguyen, M., Martin, D., & DeStefano, F. (2015). Safety monitoring in the Vaccine Adverse Event Reporting System (VAERS). *Vaccine*, 33(36), 4398–4405. <https://doi.org/10.1016/j.vaccine.2015.07.035>

19. Finsterer, J. (2022). Vaccine Adverse Event Reporting System could miss or misinterpret neurological side effects of COVID-19 vaccinations. *Annals of Neurology*, 92(1), 157–158. <https://doi.org/10.1002/ana.26369>
20. Evans, S. J., Waller, P. C., & Davis, S. (2001). Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiology and Drug Safety*, 10(6), 483–486. <https://doi.org/10.1002/pds.677>
21. Rothman, K. J., Lanes, S., & Sacks, S. T. (2004). The reporting odds ratio and its advantages over the proportional reporting ratio. *Pharmacoepidemiology and Drug Safety*, 13(8), 519–523. <https://doi.org/10.1002/pds.1001>
22. Shimonovich, M., Pearce, A., Thomson, H., Keyes, K., & Katikireddi, S. V. (2021). Assessing causality in epidemiology: Revisiting Bradford Hill to incorporate developments in causal thinking. *European Journal of Epidemiology*, 36(9), 873–887. <https://doi.org/10.1007/s10654-020-00703-7>
23. Fedak, K. M., Bernal, A., Capshaw, Z. A., & Gross, S. (2015). Applying the Bradford Hill criteria in the 21st century: How data integration has changed causal inference in molecular epidemiology. *Emerging Themes in Epidemiology*, 12, Article 14. <https://doi.org/10.1186/s12982-015-0037-4>
24. Lazarus, R. (n.d.). Electronic support for public health–Vaccine Adverse Event Reporting System (ESP:VAERS) (Grant Final Report, Grant ID: R18 HS 017045). Agency for Healthcare Research and Quality. <https://digital.ahrq.gov/ahrq-funded-projects/electronic-support-public-health-vaccine-adverse-event-reporting-system>
25. McLachlan, S., Osman, M., Dube, K., & et al. (2021). Analysis of COVID-19 vaccine death reports from the Vaccine Adverse Events Reporting System (VAERS) database: Interim results and analysis. *ResearchGate*. <https://doi.org/10.13140/RG.2.2.26987.26402>
26. Medical Dictionary for Regulatory Activities. (n.d.). MedDRA. <https://www.meddra.org/>
27. World Health Organization. (n.d.). Causality assessment of an adverse event following immunization (AEFI). <https://iris.who.int/server/api/core/bitstreams/f6816ba8-3c56-4a80-b98a-f2ac1c1cac85/content>
28. Delage, G. (2000). Rotavirus vaccine withdrawal in the United States; the role of postmarketing surveillance. *The Canadian Journal of Infectious Diseases*, 11(1), 10–12. <https://doi.org/10.1155/2000/414396>
29. Isgut, M., Gloster, L., Choi, K., Venugopalan, J., & Wang, M. D. (2023). Systematic review of advanced AI methods for improving healthcare data quality in post COVID-19 era. *IEEE Reviews in Biomedical Engineering*, 16, 53–69. <https://doi.org/10.1109/RBME.2022.3216531>
30. Acceldata. (n.d.). How AI is transforming data quality management. <https://www.acceldata.io/blog/how-ai-is-transforming-data-quality-management>
31. Tsai, T., Lee, J. J., Phillips, R., & Lin, S. (2025). Data transformation to advance AI/ML research and implementation in primary care. *Annals of Family Medicine*, 23(4), 363–367. <https://doi.org/10.1370/afm.240459>
32. Cima, G., Console, M., Lenzerini, M., & Poggi, A. (2023). A review of data abstraction. *Frontiers in Artificial Intelligence*, 6, Article 1085754. <https://doi.org/10.3389/frai.2023.1085754>
33. U.S. Food and Drug Administration. (n.d.). FDA Adverse Event Reporting System (FAERS). <https://open.fda.gov/data/faers/>
34. Liu, S., Yuan, Z., Rao, S., Li, W., Wu, T., Deng, S., Zhong, Y., Lin, J., Guo, W., & Yuan, S. (2025). Adverse drug reactions related to methotrexate: A real-world pharmacovigilance study using the FAERS database from 2004 to 2024. *Frontiers in Immunology*, 16, Article 1586361. <https://doi.org/10.3389/fimmu.2025.1586361>
35. Noguchi, Y., Aoyama, K., Kubo, S., Tachi, T., & Teramachi, H. (2020). Improved detection criteria for detecting drug-drug interaction signals using the proportional reporting ratio. *Pharmaceuticals*, 14(1), Article 4. <https://doi.org/10.3390/ph14010004>
36. Lane, S., Yeomans, A., & Shakir, S. (2022). Systematic review of spontaneous reports of myocarditis and pericarditis in transplant recipients and immunocompromised patients following COVID-19 mRNA vaccination. *BMJ Open*, 12(7), Article e060425. <https://doi.org/10.1136/bmjopen-2021-060425>
37. Manus. (n.d.). Manus 1.5 release. <https://manus.im/blog/manus-1.5-release>
38. Moderna. (n.d.). Moderna COVID-19 vaccine: Package insert / prescribing information. *Drugs.com*. <https://www.drugs.com/pro/moderna-covid-19-vaccine.html>
39. Centers for Disease Control and Prevention. (n.d.). COVID-19 vaccination demographics in the United States.

https://data.cdc.gov/Vaccinations/COVID-19-Vaccination-Demographics-in-the-United-St/km4m-vcsb/data_preview

40. Schmeling, M., Manniche, V., & Hansen, P. R. (2025). Safety of BNT162b2 mRNA COVID-19 vaccine batches: A nationwide cohort study. *Pharmacoepidemiology and Drug Safety*, 34(11). <https://doi.org/10.1002/pds.70270>
41. Albert, E., Aurigemma, G., Saucedo, J., & Gerson, D. S. (2021). Myocarditis following COVID-19 vaccination. *Radiology Case Reports*, 16(8), 2142–2145. <https://doi.org/10.1016/j.radcr.2021.05.033>
42. Montgomery, J., Ryan, M., Engler, R., Hoffman, D., McClenathan, B., Collins, L., Loran, D., Hrnecir, D., Herring, K., Platzer, M., Adams, N., Sanou, A., & Cooper, L. T., Jr. (2021). Myocarditis following immunization with mRNA COVID-19 vaccines in members of the US military. *JAMA Cardiology*. Advance online publication. <https://doi.org/10.1001/jamacardio.2021.2833>
43. Vogel, G., & Couzin-Frankel, J. (2021, June 6). Israel reports link between rare cases of heart inflammation and COVID-19 vaccination in young men. *Science*. <https://www.sciencemag.org/news/2021/06/israel-reports-link-between-rare-cases-heart-inflammation-and-covid-19-vaccination>
44. Mansanguan, S., Charunwatthana, P., Piyaphanee, W., Dechkhajorn, W., Poolcharoen, A., & Mansanguan, C. (2022). Cardiovascular manifestation of the BNT162b2 mRNA COVID-19 vaccine in adolescents. *Tropical Medicine and Infectious Disease*, 7(8), Article 196. <https://doi.org/10.3390/tropicalmed7080196>
45. Buerger, N., Lopez-Ayala, P., Hirsiger, J. R., Mueller, P., Median, D., Glarner, N., Rumora, K., Herrmann, T., Koechlin, L., Haaf, P., Rentsch, K., Battegay, M., Banderet, F., Berger, C. T., & Mueller, C. (2023). Sex-specific differences in myocardial injury incidence after COVID-19 mRNA-1273 booster vaccination. *European Journal of Heart Failure*. Advance online publication. <https://doi.org/10.1002/ehfj.2978>
46. Hulscher N, Hodkinson R, Makis W, McCullough PA. Autopsy findings in cases of fatal COVID-19 vaccine-induced myocarditis. *ESC Heart Fail*. 2025;12(5):3212–3225. doi:10.1002/ehf2.14680. <https://onlinelibrary.wiley.com/doi/10.1002/ehf2.14680>
47. Bouchaala, A., Nguadi, J., Benhlima, A., Arfaoui, M., Elhamzaoui, H., & Alilou, M. (2023). Post-vaccine COVID-19 acute myocarditis: Case reports and literature review. *The Pan African Medical Journal*, 44, Article 192. <https://doi.org/10.11604/pamj.2023.44.192.35425>
48. Das, B. B., Moskowitz, W. B., Taylor, M. B., & Palmer, A. (2021). Myocarditis and pericarditis following mRNA COVID-19 vaccination: What do we know so far? *Children*, 8(7), Article 607. <https://doi.org/10.3390/children8070607>
49. Rose, J., Hulscher, N., & McCullough, P. A. (2024). Determinants of COVID-19 vaccine-induced myocarditis. *Therapeutic Advances in Drug Safety*, 15. <https://doi.org/10.1177/20420986241226566>
50. Liao, Y. F., Tseng, W. C., Wang, J. K., Chen, Y. S., Chen, C. A., Lin, M. T., Lu, C. W., Wu, M. H., & Chiu, S. N. (2023). Management of cardiovascular symptoms after Pfizer-BioNTech COVID-19 vaccine in teenagers in the emergency department. *Journal of the Formosan Medical Association*, 122(8), 699–706. <https://doi.org/10.1016/j.jfma.2022.12.004>
51. Parry, P. I., Lefringhausen, A., Turni, C., Neil, C. J., Cosford, R., Hudson, N. J., & Gillespie, J. (2023). 'Spikeopathy': COVID-19 spike protein is pathogenic, from both virus and vaccine mRNA. *Biomedicines*, 11(8), Article 2287. <https://doi.org/10.3390/biomedicines11082287>
52. Schaler, L., & Wingfield, M. (2022). COVID-19 vaccine - can it affect fertility? *Irish Journal of Medical Science*, 191(5), 2185–2187. <https://doi.org/10.1007/s11845-021-02807-9>
53. Zhu, H., Wang, X., Zhang, F., Zhu, Y., Du, M. R., Tao, Z. W., Sun, C., Ma, H. T., Li, Y. D., Liang, G. Q., Ren, J., & Feng, B. L. (2022). Evaluation of inactivated COVID-19 vaccine on semen parameters in reproductive-age males: A retrospective cohort study. *Asian Journal of Andrology*, 24(5), 441–444. <https://doi.org/10.4103/aja202225>
54. Edele Santos, D., Colonetti, T., Rodrigues Ugioni, M. L., Rech, P., Marcelino Baptista, M., Medeiros, L. R., Grande, A. J., & Rosa, M. I. (2023). Effects of COVID-19 or vaccines for SARS-CoV-2 on sperm parameters: A systematic review and meta-analysis. *Journal of Reproductive Immunology*, 160, Article 104140. <https://doi.org/10.1016/j.jri.2023.104140>
55. Abd, Z. H., Muter, S. A., Saeed, R. A. M., & Ammar, O. (2022). Effects of Covid-19 vaccination on different semen parameters. *Basic and Clinical Andrology*, 32(1), Article 13.

- <https://doi.org/10.1186/s12610-022-00163-x>
56. Centers for Disease Control and Prevention. (n.d.). COVID-19 vaccination demographics in the United States.

https://data.cdc.gov/Vaccinations/COVID-19-Vaccination-Demographics-in-the-United-St/km4m-vcsb/about_data

This page is intentionally left blank

