COVID-19 Vaccine Safety Protocol: Comparative Risk of Myocarditis or Pericarditis following COVID-19 mRNA Vaccination

Biologics Effectiveness and Safety (BEST) Initiative

Center for Biologics Evaluation and Research (CBER)
Office of Biostatistics and Epidemiology (OBE)

December 10, 2021

STUDY TEAM

FDA/CBER/OBE
Hui Lee Wong, PhD Cindy Ke Zhou, PhD Patricia Lloyd, PhD Yun Lu, PhD Joyce Obidi, PhD Richard Forshee, PhD Azadeh Shoaibi, PhD, MS, MPH Steven A. Anderson, PhD
ACUMEN, LLC & STANFORD UNIVERSITY
Mao Hu Ellen Tworkoski, MS, MPhil Kathryn Matuska Elizabeth Smith Sandia Akhtar Manzi Ngaiza, MPH Yue Wu, MS Zhiruo Wan, MS An-Chi Lo, MS, MPH Yuqin Wei, MS Yuhui Feng, MS Yoganand Chillarige, MPA John Hornberger, MD, MS, FACP
IQVIA/HEALTHCORE
Alex Secora, PhD Daniel Beachler, PhD, MHS
OPTUM
Kandace Amend, PhD, MPH Rachel P. Ogilvie, PhD, MPH John Seeger, DrPH, PharmD
CVS HEALTH
Cheryl N McMahill-Walraven PhD, MSW Smita Bhatia, MS Annemarie Kline, MS Djeneba Audrey Djibo, PhD Charlalynn Harris, PhD, MPH

VERSION CONTROL

Editors	Date	Version	Description
Acumen LLC	August 9, 2021	0.1	First draft
FDA/CBER/OBE Acumen, LLC IQVIA/HealthCore Optum	August 13, 2021	0.2	Second draft
Acumen LLC	October 19, 2021	0.3	Third draft
FDA/CBER/OBE Acumen, LLC	November 17, 2021	0.4	Fourth draft
FDA/CBER/OBE Acumen, LLC IQVIA/HealthCore Optum	December 10, 2021	0.5	Fifth draft

Table of Contents

1. Objectives

To compare the risk of myocarditis or pericarditis as a composite outcome (myocarditis or pericarditis) in specified risk windows following first dose, second dose, and any dose of Pfizer-BioNTech and Moderna mRNA COVID-19 vaccination in the 18–64-year-old population overall and stratified by age and sex groups.

2. Overview

The coronavirus disease 2019 (COVID-19) is a contagious respiratory illness caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus. While a number of COVID-19 symptoms seem to be similar to those of influenza, COVID-19 spreads more aggressively and presents more severely in some patients [1]. The first COVID-19 case was reported in China in December 2019 [2] and in the United States (US) the first case of laboratory-confirmed COVID-19 was reported to the Centers for Disease Control and Prevention (CDC) on January 22, 2020 [3]. As of October 15, 2021, a total of 254 million cases and 5.1 million deaths have been reported worldwide [4]. The highest number of cases and deaths has been reported from the US (47.4 million cases and >760,000 deaths) [4].

Currently, three COVID-19 vaccines are granted emergency use authorization or licensed by the US Food and Drug Administration, (FDA), including two mRNA vaccines developed by Moderna and Pfizer-BioNTech[5, 6]. As with all licensed vaccines, there can be limitations in the safety profile assessed during the pre-licensure clinical studies of a COVID-19 vaccine. Potential safety risks of COVID-19 vaccines may not be captured in clinical trials, particularly for rare events. Post-authorization or licensure active monitoring and reporting of COVID-19 vaccine-related adverse event of special interest (AESIs) enables better capture of rare safety events and provides timely information to support regulatory decision-making processes.

The Israel Ministry of Health identified 148 myocarditis occurring around the time of vaccination from December 2020 to May 2021, with a possible link between vaccine and cases among younger men aged 16 to 30 [7]. In the US, the Vaccine Adverse Event Reporting System (VAERS) and the Vaccine Safety Datalink (VSD) observed greater incidence rates of myocarditis or pericarditis than expected rates following mRNA COVID-19 vaccination [8]. The higher rates have also been observed in other data sources [9]. A benefit-risk analysis assessed by the CDC Advisory Committee on Immunization Practices (ACIP) concluded that the use of the mRNA vaccines prevents morbidity and mortality from COVID-19 that exceeds the expected number of myocarditis or pericarditis cases [8].

The primary objective of this study is to compare the risk of the composite outcome of myocarditis or pericarditis between the COVID-19 Moderna vaccines versus Pfizer-BioNTech vaccines among adults 18 to 64 years of age using administrative claims data sources. The 18–64-year age group was chosen since both vaccines have authorization or licensure for persons 18 years and older and adults 65 years or older are not extensively captured in the private health insurance databases used for the study.

3. Data Sources

The study will utilize four participating data sources within the FDA Center for Biologics Evaluation and Research (CBER) Biologics Effectiveness and Safety (BEST) Initiative including Optum pre-adjudicated, IQVIA/HealthCore, CVS Health, and Blue Health Intelligence® (BHI®)[1] administrative claims. Table 1 provides information about the data sources including the patient population size and data lag. The data lag indicates the amount of time between the service date and the data availability date for use by research teams. The data lag varies from a few days to a several months depending on the care setting in which the service is administered. It also includes the frequency in which the database is updated and prepared for research activities.

Table 1. Description of Data Sources

Administrative Claims Data Sources	Update frequency	Data Lag*	Population
Blue Health Intelligence (BHI)	Monthly	4 months for >80% of inpatient claims	National; 23–28 million enrollees annually
CVS Health	Monthly	Approximately 2.5 months from the most recent service date for 80% inpatient claim	National; over 37 million patients cumulative, and 22 million enrollees annually
IQVIA/HealthCore	Monthly	Approximately 3 months for complete data	National; ~ 76 million patients cumulative, and 25.9 million enrollees annually
OptumServe Pre-Adjudicated	Bi-weekly	Approximately 2 months for inpatient at 90% completeness and 1 month for outpatient at >70% completeness	National; ~30 million patients cumulative and 15-16 million enrollees annually

** Data lag can vary by outcome and the care setting in which the outcome is diagnosed; we will produce outcome- and care setting-specific delay profiles.*

[1] Blue Health Intelligence® (BHI®) is a trade name of Health Intelligence Company, LLC, an independent licensee of the Blue Cross Blue Shield Association.

The Optum data includes enrollment, prescription drug and pre-adjudicated hospital and physician health insurance claims. The pre-adjudicated claims database includes claims for privately insured and Medicare Advantage enrollees. Hospital and physician claims undergo initial processing on a daily basis from a large number of providers across the U.S. who accept patients with health insurance. Optum has established an ongoing weekly update schedule to incorporate newly processed claims into the pre-adjudicated claims database. This data source was utilized to reduce the delay between the occurrence of healthcare services and their presence in the database. The pre-adjudicated claims have an approximately two-month delay for 90% completeness for inpatient claims and over 70% completeness at one-month for outpatient claims.

HealthCore, Inc. is a wholly-owned, research subsidiary of Anthem, Inc., a holding company owning several large US health plans. HealthCore has a commercially insured US population database, the HealthCore Integrated Research Environment (HIRE), with longitudinal data on health plan enrollees. This proprietary research environment combines medical and pharmacy claims, and laboratory results, drawn from nearly 76 million unique individuals with medical coverage, with approximately 59 million also having pharmacy coverage dating back to 2006. The data currently contain over 23 million members in 2020 with a data lag of 3 months for complete data and is 1–3 months for pharmacy dispensings and early settled outpatient claims.

CVS Health Clinical Trial Services LLC transforms health plan enrollment, demographic, and clinical claims data containing longitudinal health information for all individuals enrolled in Aetna commercial and Medicare Advantage health plans from January 2018 through March 2021 into a patient-centered, comprehensive Common Data Model (CDM). The CDM contains over 37 million individuals in total and on average about 22 million individuals annually.

BHI data provide longitudinal enrollment, demographic, and claims information from Blue Cross and Blue Shield commercial health insurance plans in the US for the last ten years. Data available for this study were limited to a cohort of all enrollees who received a biologic product, had a pregnancy outcome, or were born after October 1, 2015, contributing to approximately 23–28 million enrollees annually.

3.1 Common Data Model

This study will be conducted using four databases in the BEST Initiative. The four databases will transform their data into the Observational Medical Outcomes Partnership (OMOP) CDM version 5.3.1[10]. If a database is not in the OMOP CDM framework, study-specific standard analytical files (SAFs) that mimic the structure of the data extract from OMOP will be used.

4. Methods

The study will use an active, comparative cohort design, comparing the rate of myocarditis or pericarditis as composite outcome in post-vaccination risk-windows for people vaccinated with the Pfizer-BioNTech and Moderna COVID-19 vaccines. A cohort design was selected for rapid comparison of safety risk between two vaccines, without the need to wait for accrual of a post-vaccination control window required in a self-controlled design.

The analysis will be conducted at the vaccine dose level which means each individual can contribute up to two observations (i.e., the unit of analysis is a vaccine dose and not a patient). Poisson regression models will be used to estimate incidence rate ratio and risk difference between mRNA vaccines.

4.1 Study Population

The study population will include individuals who received an mRNA COVID-19 vaccine dose (Pfizer-BioNTech or Moderna) and were 18–64 years old at time of vaccination of first observed dose. To identify post-vaccination myocarditis or pericarditis incident events, post-vaccination risk windows and pre-vaccination clean windows are pre-specified in Table 2. Vaccinated individuals aged 12–17 will be excluded for comparison since only the Pfizer-BioNTech vaccine is authorized for use in this age group.

To be included in the analysis, individuals are required to have continuous enrollment in a medical insurance plan in the 365 days prior to vaccination and on the date of vaccination (day 0). Subjects who experienced myocarditis or pericarditis during the pre-vaccination clean window will be excluded from the analysis.

4.2 Study Period

The study period will be from December 18th, 2020 (the date of Emergency Use Authorization for the Moderna COVID-19 vaccine) through the latest available data cut date for each data partner. Pfizer-BioNTech doses from December 11th to December 17th, 2020 will not contribute to person time or events, but will be used to assign dose ordering for later administrations.

4.3 Exposure

The exposure is receipt of COVID-19 Pfizer-BioNTech or Moderna vaccine, as identified by vaccine product codes such as Current Procedural Terminology (CPT)/Healthcare Common Procedure Coding System (HCPCS) codes or National Drug Codes (NDCs) in the professional, outpatient institutional, inpatient, or prescription drug care settings. The list of valid codes will be continuously reviewed and updated if new codes are added.

Dose number (Dose 1, Dose 2) will be assigned by administration code if available. If an administration code is not available (for example, for doses identified by NDCs), the temporal ordering of doses of the same vaccine will be used to determine dose ordering. 'All doses' of either vaccine will consist of all administrations identified as Dose 1 or Dose 2, and potential doses beyond Dose 2 will not be assessed.

4.4 Index Date and Follow-up

Each vaccination administration date will serve as an index date (day 0). Follow-up will begin on day 1 after vaccination and continue until the end of the risk windows detailed in Table 2. Follow-up time will be censored based on disenrollment, study end date, or the occurrence of a subsequent dose so that overlapping risk-window time will not be double counted.

In the dose-specific analyses, Dose 1 and Dose 2 time are defined respectively as exposed time following the first administration of a vaccine and following the second administration of the same vaccine, among persons who may have any number of observed doses for the vaccine.

4.4 Outcomes

The composite outcome of myocarditis or pericarditis will be defined as in the "COVID-19 Vaccine Safety Surveillance: Active Monitoring Master Protocol" [11].

Table 2. Myocarditis/Pericarditis Definitions

Analysis	Age Group of Interest	Claim Setting	Clean Window	Risk Window
Primary	18–64	IP, OP/PB*	365 days	1–7 days [8]
Sensitivity	18–64	IP, OP/PB*	365 days	1–21 days
Sensitivity	18–64	IP, OP/PB*	365 days	1–42 days [12]

**All facility and professional claims in the inpatient or outpatient setting will be used to define the outcome, excluding claims that cover only lab services.*

4.4.1 Clean and Risk Windows

The clean window is defined as an interval used to define incident myocarditis or pericarditis where an individual enters the study cohort only if myocarditis or pericarditis did not occur during that interval. The risk window is defined as an interval during which occurrence of myocarditis or pericarditis will be included in the analyses and the risk window for the primary analysis was informed by data from VAERS[9]. In the case that a risk window following vaccination is censored by a following dose, the remaining events or person-time will not be double-counted.

4.5 Covariates

The following covariates will be summarized descriptively and may be used in the inferential models:

- Demographics at time of each vaccination dose
  - Age (18–25, 26–35, 36–45, 46–55, and 56–64 years)
  - Sex (Male or Female)
  - US Department of Health and Human Services (HHS) Region
  - Urban/Rural residency status
  - Calendar time at time of vaccination measured from December 18, 2020 in weeks
- Claims-based covariates relative to each vaccination dose
  - Any COVID-19 diagnosis prior to or on the date of vaccination. COVID-19 events beyond 1 year prior to vaccination will be used, but the enrollment requirement will not be extended. The U07.1 diagnosis code will be used to identify COVID-19. Although other diagnoses codes were used to identify COVID-19 prior to April 1st, 2020, the performance of these codes were uncertain, and they were used for a relatively short time. Therefore, these codes will not be used.

4.6 Descriptive Analyses

Counts of eligible vaccination and incident events for each vaccine brand will be calculated. Stratified counts by dose number, age group, and sex will be calculated. Distribution of covariates by vaccine brand will also be calculated.

Based on the descriptive analysis, the feasibility of including certain age groups or covariates will be assessed. An example table representing the proposed descriptive statistics can be found below in Table 3.

Table 3. Example table of descriptive statistics

Stratification or Statistic	Value
Risk Window	1-7, 1-21, 1-42 day risk window
Product	Pfizer-BioNTech, Moderna
Dose	1, 2, or All
Age Group	18–25, 26–35, 36–45, 46–55, and 56–64 years
Sex	Male or Female
HHS Region	HHS Regions 1-10
COVID-19 Diagnosis Prior to Vaccination	Yes or No
Urban/Rural Residency	Urban or Rural
Total Number of Eligible Vaccinations	Number of eligible vaccinations
Total Number of Events	Number of events

4.7 Statistical Analysis

Poisson regression models will be estimated to assess the incidence rate ratio and risk difference between Moderna and Pfizer-BioNTech (reference). Model-based incidence rates post Moderna and Pfizer-BioNTech vaccination will also be estimated. The analysis will be conducted separately within each data source, and a meta-analysis will be conducted pooling results across all four data sources.

The primary analysis will use myocarditis or pericarditis in the 1–7 days post-vaccination for all doses, consisting of either first or second doses of mRNA COVID-19 vaccine. The sensitivity analyses will assess the 1–21 and 1–42 days post-vaccination risk windows.

In addition, secondary analyses will assess comparative risk post Dose 1 and Dose 2 specifically. If all-dose outcome counts stratified by dose remain sufficiently large in age and sex strata (i.e, at least five events in each stratum), then the dose-specific models will include all age, sex, and vaccine interaction effects specified in the all-dose model[13]. However, if there are insufficient outcome counts in dose-specific analyses, then the model selection process will be repeated for the dose-specific analyses.

4.7.1 Poisson Regression with Age Group and Sex

Multiple Poisson regression models including treatment status (Pfizer-BioNTech (reference) vs. Moderna), age groups (18–25, 26–35, 36–45, 46–55, and 56–64 years), sex, and interactions between these variables will be estimated. A final model will be selected by applying likelihood ratio tests (LRTs). Robust variance estimation will be used to account for correlated observations in the all-dose analysis.

Depending on outcome counts assessed from descriptive statistics, the models assessed may range from a fully interacted model to a simple model including only main-effect terms:

- Main-Effects Model (Model I)

    - $\log(E(Y|X)) = \alpha + \beta' x_{main} + \log(time_{adj})$
    - Degrees of Freedom (DF) = Moderna vaccine (1) + age groups (4) + sex (1) = 6
- Two-way Interactions Model (Model II)
    - $\log(E(Y|X)) = \alpha + \beta' x_{main} + \theta' x_{age:sex,vaccine::sex,vaccine:age} + \log(time_{adj})$
    - DF = main effects (6) + age:sex (4) + vaccine:sex (1) + vaccine:age (4) = 15
    - Variants of this model including only one of the treatment-related two-way interactions will be assessed:
        - Model IIa: age:sex and vaccine:sex only
        - Model IIb: age:sex and vaccine:age only
- Fully Interacted Model (Model III)
    - $\log(E(Y|X)) = \alpha + \beta' x_{main} + \theta' x_{age:sex,vaccine:sex,vaccine:age} + \gamma x_{age:sex:vaccine} + \log(time_{adj})$
    - DF = main effects (6) + two-way interactions (9) + age:sex:vaccine (4) = 19

Where Y is the outcome count associated with a strata identified by vector of predictors and corresponding interaction terms X = [I(Moderna vaccine), I(Age:26-35), …, I(Age:56-64), I(Male), …], x_{main}, $x_{age*sex,vaccine*sex,vaccine*age}$, and $x_{age:sex:vaccine}$ are subsets of X corresponding to main effects, two-way interactions, and three-way interactions, and β, θ, and γ are coefficient vectors corresponding to the main effects, two-way interactions, and three-way interactions specifically. For example, $\exp(\beta_1)$ corresponding to I(Moderna vaccine) in Model I can be interpreted as the rate ratio of Moderna vaccine vs. Pfizer-BioNTech vaccine adjusting for other covariates.

The final model used for inference will be selected by applying the following procedure:

- Step 1: The fully interacted model (Model III) will be compared against the two-way interactions model (Model II)
    - If the LRT is statistically significant, Model III will be selected as the final model
    - If not, then proceed to Step 2
- Step 2: The two-way interactions model (Model II) will be compared against the main-effects model (Model I)
    - If the LRT is statistically significant, proceed to Step 3
    - If not, then Model I will be selected as the final model
- Step 3: Model II will be compared against the partial two-way models (Model IIa, Model IIb)
    - If the LRT is significant for both comparisons, Model II will be selected as the final model
    - If the LRT is significant for one comparison and not the other, then the partial model (Model IIa, Model IIb) with the non-significant comparison will selected as the final model
    - If the LRT is not significant for either comparison, Model II will be selected as the final model due to lack of evidence for selecting a particular partial model

If some models cannot be successfully estimated due to lack of events, the selection procedure will start from the following less-complex model. In addition to the above, the Akaike Information Criterion (AIC) will be calculated for each model and compared.

4.7.2 Poisson Regression with Age Group, Sex, and Other Covariates

Poisson regression models including treatment status, age group, sex, and other covariates will be estimated. The same model estimation and selection procedure, as explained above, will be repeated except with the addition of covariate main effects to provide adjustment for potential confounding. Categorical covariates must have at least one event in each category in order to be included in the model.

Adjusted incidence rates within age and sex groups developed from models with other covariates will be assessed using the average within that age and sex group. For example, among males 18–25 years old, the proportion of observations with rural residency status, proportion with COVID-19 prior to vaccination, and mean week of vaccination will be used to assess the incidence rate within the group if these covariates are included within the model.

4.7.3 Person-Time Adjustment for Observation Delay

Person-time attributed to each vaccination will be adjusted for observation delay ($time_{adj}$). The adjustment will be performed by segmenting study time for each vaccination by weeks of delay relative to the data cut date, and then multiplying the time by an estimated proportion of 'data completeness' obtained from historical data. Within each stratum defined by vaccine, age, sex, and other covariates, the adjusted time will be calculated as follows:

$$time_{adj} = \sum_{t=1}^{s} \sum_{i=1}^{n_t} \sum_{w=0}^{T_i} l_{stiw} P(s,t,w)$$

where:

- s represents the study time period (e.g., study week) at which the analysis is planned (e.g., s = 15)
- t in 1 … s.
- n_t is the number of subjects vaccinated during time period t. Ineligible doses due to lack of enrollment or having an AESI in the cleaning period will be excluded.
- *i* in 1 … n_t.
- T_i represents the exposed weeks at risk following a dose (i.e., within the AESI-specific risk window) for subject *i*.
- w in 0 … T_i. Study weeks in which AESIs occur are represented by t + w, with t + 0 being the same study week as the vaccine dose, t + 1 the next week, etc.
- l_{stiw} in 0 … 7 is the number of exposed days following a vaccine dose in study week w post vaccine dose for a patient identified by *i, t* in the group based on data at observation week s.
    - If the dose is administered on day 5 of study week *t*, then the first 1–2 days of 1–42 days post-vaccination risk window would occur in w = 0 post dose (with $l_{sti(w=0)}$ = 2). The next 3–9 days would occur in week w = 1 post dose (with $l_{sti(w=1)}$ = 7), etc.

- The occurrence of a second dose may contribute additional exposed time but overlapping time will not be double counted. For example, if the risk window length is 42 days, a second dose 22 days after the first dose will result in an overall risk period of 1 to (42 + 42 – 21 overlap) = 63 days. In the case when the gap between doses is larger than the risk interval post first dose, the follow-up after the first dose would be censored at the end of the planned risk window and restarted at the second dose.

- *P(s,t,w)* is the proportion of AESIs occurring in study week *t + w* that would be observed by week *s*. This adjustment factor adjusts for the observation delay due to the use of partially accrued data.
  - Cumulative proportion of data completeness will be estimated from myocarditis or pericarditis events in 2019 data.
  - To account for variability in the observation delay adjustment for short delays (0–1 weeks), only events and person time 2 or more weeks delayed relative to the data cut date will be included in the analysis.

4.7.4 Meta-Analysis of Poisson Regression Results

Given a common study protocol and a standard analytical package will be used across commercial claims databases that cover populations with similar demographics, a meta-analysis will be performed to pool results to gain higher precision and statistical power using both random-effects and fixed-effect models. The pooled model-based incidence rates, covariate adjusted, and unadjusted relative risk of myocarditis or pericarditis will be presented overall and by age group and sex.

Below, $\hat{\theta}_k$ is used to represent the estimator (log of rate ratio or log of incidence rate) estimated from each data source, where k = 1,2,3,4 indicates the data sources — Optum, IQVIA/HealthCore, CVS Health and BHI, respectively. The goal is to estimate the pooled result:

$$\hat{\theta} = \frac{\sum_{k=1}^{K} \hat{\theta}_k \omega_k}{\sum_{k=1}^{K} \omega_k}$$

The same notation applies to estimate the pooled effect with or without covariate adjustment, within different risk windows, or for subgroup analysis.

4.7.4.1 Additional Poisson Regression Models for Meta-Analysis

Given the low observed event counts and the variability of counts between data partners, the regression models identified by the model-selection process described in Section 4.7.1 may not be mutually consistent between data partners. For example, the model selected for one data partner may include vaccine and sex interactions, whereas another data partner may not include any vaccine interactions, which may complicate comparisons across data partners. Therefore, to facilitate the meta-analysis, a set of simplified, common models will be fit within each of the individual data sources. In addition to the models described in Section 4.7.1, we will also fit the following models:

- Models stratified by sex (Model 4):
    - Variants of this model with or without a treatment-related two-way interaction will be assessed, as the larger number of cases observed in males may allow for estimation of age interaction effects:
        - Model 4a: vaccine brand, age and vaccine:age, with or without other covariates
        - Model 4b: vaccine brand and age only, with or without other covariates
- Overall models (Model 5)
    - Variants of this model with or without a treatment-related two-way interaction will be assessed:
        - Model 5a: vaccine brand, age, sex and vaccine:age, with or without other covariates
        - Model 5b: vaccine brand, age, and sex only

The models will include adjustment with additional covariates if there are at least two myocarditis or pericarditis events within each level of the covariate (e.g., COVID-19 prior to vaccination will only be included if there are at least two cases of no COVID-19 prior to vaccination and at least two cases of COVID-19 prior to vaccination among patients with outcomes in the data source).

As with the models described in Section 4.7.1, the goodness-of-fit between Model 4a vs. 5b and 5a vs. 5b will be assessed using LRTs and AIC.

4.7.4.2 Random-Effects Meta-Analysis

We will use random-effects meta-analysis to account for the between-study heterogeneity across multiple data sources.

The random-effects model takes the form:

$$\hat{\theta}_k = \mu + \zeta_k + \epsilon_k$$

Where μ is the global true effect of interest, ζ_k is a data source specific random error term and ϵ_k is data source specific sampling error term. The pooled effect can be estimated by the inverse-variance method

$$\hat{\theta} = \frac{\sum_{k=1}^{K} \hat{\theta}_k \omega_k^*}{\sum_{k=1}^{K} \omega_k^*}$$

$$\omega_k^* = \frac{1}{s_k^2 + \tau^2}$$

where s_k^2 represents variance of θ_k estimated for each study, and τ^2 is the variance of the distribution of true effect sizes.

The Paule-Mandel method will be used to estimate τ^2 as suggested by Veroniki (2016) [14]. Bakbergenuly (2020) additionally found that the Paule-Mandel estimator is well-suited for when the

number of studies is small [15]. As a sensitivity analysis, we will use the DerSimonian-Laird and restricted maximum likelihood (RMLE) estimators for τ^2 in order to assess the variation in the estimation of τ^2 due to the small number of studies.

4.7.4.3 Fixed-Effect Meta-Analysis

To address concerns that the random-effects method may not perform well when the number of studies is small, we will additionally run a fixed-effect meta-analysis. The fixed-effect model takes the form:

$$\hat{\theta}_k = \mu + \epsilon_k$$

Where μ is the global true effect of interest, and ϵ_k is data source specific sampling error term. The pooled effect can be estimated by the inverse-variance method

$$\hat{\theta} = \frac{\sum_{k=1}^{K} \hat{\theta}_k \omega_k}{\sum_{k=1}^{K} \omega_k}$$

$$\omega_k = \frac{1}{s_k^2}$$

where s_k^2 represents variance of $\hat{\theta}_k$ estimated for each study.

4.7.4.4 Evaluating Between-Study Heterogeneity

This section describes the methods we will use to evaluate between-study heterogeneity.

Forest Plots

Forest plots will be generated to visualize the variation of rate ratios and 95% CIs across data sources.

Cochran's Q

Cochran's *Q* is defined as a weighted sum of squares (WSS).

$$Q = \sum_{k=1}^{K} \omega_k \left(\hat{\theta}_k - \hat{\theta}\right)^2$$

We will use the value of *Q* to check if there is excess variation in our data. If there is no between-study heterogeneity, Q will approximately follow a chi-square distribution with K-1 degrees of freedom. As *Q* may be sensitive to the number of studies assessed, an additional complementary statistic will be assessed.

Higgins & Thompson's I^2 Statistic

We will also calculate Higgins & Thompson's I^2 statistic [16], which is defined as the percentage of variability in the effect sizes that is not caused by sampling error:

$$I^2 = \frac{Q - (K-1)}{Q}$$

- I^2 = 25%: low heterogeneity
- I^2 = 50%: moderate heterogeneity
- I^2 = 75%: substantial heterogeneity

The 95% CI for I^2 will also be calculated.

4.7.4.6 Meta-Analysis Specifications

We will conduct the following meta-analyses to combine estimates from the same model results across different data sources:

Table 4. Estimands and Specifications for Meta-Analysis

Models Meta-Analyzed	Age group (years)	Sex	Estimand	Method
Model 4a	18–25, 26–35, 36–45, 45–55, 56–64 years	Female	RR (Pfizer-BioNTech as reference)	Fixed-effect; Random-effects (Paule-Mandel, DerSimonian-Laird, RMLE)
Model 4b	Overall	Female	RR (Pfizer-BioNTech as reference)	Fixed-effect; Random-effects (Paule-Mandel, DerSimonian-Laird, RMLE)
Model 4a/b	18–25, 26–35, 36–45, 45–55, 56–64 years	Female	IR (Moderna), IR (Pfizer-BioNTech)	Random-effects (Paule-Mandel)
Model 4a	18–25, 26–35, 36–45, 45–55, 56–64 years	Male	RR (Pfizer-BioNTech as reference)	Fixed-effect; Random-effects (Paule-Mandel, DerSimonian-Laird, RMLE)
Model 4b	Overall	Male	RR (Pfizer-BioNTech as reference)	Fixed-effect; Random-effects (Paule-Mandel, DerSimonian-Laird, RMLE)

Models Meta-Analyzed	Age group (years)	Sex	Estimand	Method
Model 4a/b	18–25, 26–35, 36–45, 45–55, 56–64 years	Male	IR (Moderna), IR (Pfizer-BioNTech)	Random-effects (Paule-Mandel)
Model 5a	18–25, 26–35, 36–45, 45–55, 56–64 years	Overall	RR (Pfizer-BioNTech as reference)	Fixed-effect; Random-effects (Paule-Mandel, DerSimonian-Laird, RMLE)
Model 5b	Overall	Overall	RR (Pfizer-BioNTech as reference)	Fixed-effect; Random-effects (Paule-Mandel, DerSimonian-Laird, RMLE)
Model 5a/b	18–25, 26–35, 36–45, 45–55, 56–64 years	Overall	IR (Moderna), IR (Pfizer-BioNTech)	Random-effects (Paule-Mandel)

RR=relative risk; IR=incidence rate.

4.7.5 Presentation of Results

For each analysis, the estimates of rate ratio will be presented using tables (Table 5) and graphs (e.g., a forest plot).

Table 5. Table Shell for Effect Estimates

Dose	Sex	Age Group* (years)	Incidence Rate	Rate Ratio (95% CI)	Risk Diff. per 1M Doses (95% CI)
All Doses	Male	18–25	-	-	-
All Doses	Male	…	-	-	-
All Doses	Male	56–64	-	-	-
All Doses	Female	18–25	-	-	-
All Doses	Female	…	-	-	-
All Doses	Female	56–64	-	-	-

*Age categories may change if groups are combined, etc.

The risk difference/attributable risk per 1 million doses will be calculated by subtracting the estimated number of events post 1 million Pfizer-BioNTech vaccine doses from the estimated number of events post 1 million Moderna vaccine doses. The number of cases post vaccination will be calculated as: 1 – exp(–IR * PT), (*IR*: model-based incidence rate for each vaccination brand, *PT*: person time at risk for each vaccination brand). The 95% CI of the risk difference will also be reported[17].

The parameter estimates, variance/covariance matrix, and other model outputs or diagnostics will be reported for verification purposes.

4.8 Medical Record Review

As further evaluation, we will conduct a chart review of representative events to confirm or rule out outcomes and assess outcome onset date in relation to vaccination date.

- We will apply a Brighton case definition algorithm to classify cases by different levels of probability/plausibility[18].
- Cases from the 1-21 day window will be requested to account for the possibility that cases following chart review are reclassified into the 1-7 day window based on symptom onset date
- Events with an onset date prior to vaccination will be excluded. For those with onset after vaccination, we will assess the onset pattern by weeks/days instead of relying on admission date from the electronic data.

5. Limitations

This study has several limitations. Comparisons of the vaccines may be biased by confounding with minimal adjustment in the statistical models due to the modest sample size. The Poisson regression model will include age groups and sex with adjustment for a limited number of potential confounders such as urban/rural residency status or prior COVID-19, when sample size permits. Other potential confounding, such as previous cardiac conditions, medication use, or other health characteristics will not be assessed or adjusted in this analysis.

Statistical power is limited due to the low incidence of myocarditis or pericarditis events. In addition, while bias due to incomplete observation of events may be reduced due to the observation delay adjustment, this approach may produce biased results if historical observation delay does not reflect current data processing patterns per vaccine brands.

Finally, as with all observational studies using data not primarily collected for research purposes, administrative claims may have some degree of outcome and exposure misclassification, which might potentially affect the estimates. However, as the study only includes vaccinated individuals, exposure misclassification may have a limited impact on estimates provided it is similar between the two cohorts. In addition, prior reports of myocarditis or pericarditis events in younger males may be misclassified as incident events despite a 365-day cleaning window resulting in biased estimates.

6. References

1. *Coronavirus Disease 2019 (COVID-19)*. 2020 5/13/2020 11/8/2020]; Available from: https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html.
2. Wang, C., et al., *A novel coronavirus outbreak of global health concern.* The Lancet, 2020. **395**(10223): p. 470-473.
3. Stokes, E.K., et al., *Coronavirus disease 2019 case surveillance—United States, January 22–may 30, 2020.* 2020. **69**(24): p. 759.

4. *COVID-19 Dashboard*. 2020 [cited 2021 11/17/2021]; Available from: https://coronavirus.jhu.edu/map.html
5. *Moderna COVID-19 Vaccine*. 11/5/2021 [cited 2021 11/17/2021]; Available from: https://www.fda.gov/emergency-preparedness-and-response/coronavirus-disease-2019-covid-19/moderna-covid-19-vaccine.
6. *Comirnaty and Pfizer-BioNTech COVID-19*. 11/11/2021 [cited 2021 11/17/2021]; Available from: https://www.fda.gov/emergency-preparedness-and-response/coronavirus-disease-2019-covid-19/comirnaty-and-pfizer-biontech-covid-19-vaccine.
7. *Surveillance of Myocarditis (Inflammation of the Heart Muscle) Cases Between December 2020 and May 2021 (Including)*. 2021, Israel Ministry of Health: https://www.gov.il/en/departments/news/01062021-03.
8. Gargano, J.W., et al., *Use of mRNA COVID-19 vaccine after reports of myocarditis among vaccine recipients: update from the Advisory Committee on Immunization Practices—United States, June 2021*. 2021. **70**(27): p. 977.
9. Diaz, G.A., et al. *Myocarditis and Pericarditis After Vaccination for COVID-19*. JAMA 2021 August 4, 2021; August 10, 2021:[Available from: https://doi.org/10.1001/jama.2021.13443.
10. *OMOP CDM v5.3*. [cited 2021 11/17/2021]; Available from: https://ohdsi.github.io/CommonDataModel/cdm53.html.
11. US Food and Drug Administration, C.f.B.E.a.R., Office of Biostatistics and Epidemiology. *COVID-19 Vaccine Safety Surveillance: Active Monitoring Master Protocol*. 2021 February 10, 2021; Available from: https://www.bestinitiative.org/wp-content/uploads/2021/02/C19-Vaccine-Safety-Protocol-2021.pdf.
12. Kuntz, J., et al., *Myocarditis and pericarditis are rare following live viral vaccinations in adults.* Vaccine, 2018. **36**(12): p. 1524-1527.
13. Agresti, A., *Categorical data analysis*. Vol. 482. 2003: John Wiley & Sons.
14. Veroniki, A.A., et al., *Methods to estimate the between-study variance and its uncertainty in meta-analysis.* Res Synth Methods, 2016. **7**(1): p. 55-79.
15. Bakbergenuly, I., D.C. Hoaglin, and E. Kulinskaya, *Methods for estimating between-study variance and overall effect in meta-analysis of odds ratios.* Research Synthesis Methods, 2020. **11**(3): p. 426-442.
16. Higgins, J.P.T. and S.G. Thompson, *Quantifying heterogeneity in a meta-analysis.* Statistics in Medicine, 2002. **21**(11): p. 1539-1558.
17. Krinsky, I., A.L.J.T.R.o.E. Robb, and Statistics, *On approximating the statistical properties of elasticities.* 1986: p. 715-719.
18. *Myocarditis/Pericarditis Case Definition*. 2021 11/1/2021]; Available from: https://brightoncollaboration.us/myocarditis-case-definition-update/.

Appendices

Appendix A: Code Lists

Table A1. ICD-10 CM Diagnosis Codes for Myocarditis or pericarditis*

Code	Description
B33.22	Viral myocarditis
B33.23	Viral pericarditis
I30.0	Acute nonspecific idiopathic pericarditis
I30.1	Infective pericarditis
I30.8	Other forms of acute pericarditis
I30.9	Acute pericarditis, unspecified
I32	Pericarditis in diseases classified elsewhere
I41	Myocarditis in diseases classified elsewhere
I40.0	Infective myocarditis
I40.1	Isolated myocarditis
I40.8	Other acute myocarditis
I40.9	Acute myocarditis, unspecified
I51.4	Myocarditis, unspecified

* Myocarditis or pericarditis events were identified in inpatient and outpatient (outpatient facility and professional claims) care settings

Appendix B: Sample R Code for Meta-Analysis

This section provides sample R code for fixed-effect model, random-effects model and heterogeneity analysis described in section 4.7.4.

```
library(meta)

# Prepare dataset
df.sample <- data.frame (
    data_source = c('DP1', 'DP2', 'DP3', 'DP4'),
    logrrest = c(-0.05, 0.08, 0.12, -0.15),
    logrrse = c(0.25, 0.3, 0.21, 0.37)
)

# Run meta-analysis
m.gen <- metagen(
    TE = logrrest,
    seTE = logrrse,
    studlab = data_source,
    data = df.sample,
```

```
    sm = 'RR',
    method.tau = 'PM'
)

# Output meta-analysis result
m.gen
##           RR           95%-CI %W(fixed) %W(random)
## DP1   0.9512 [0.5828; 1.5527]      28.0       28.0
## DP2   1.0833 [0.6017; 1.9503]      19.5       19.5
## DP3   1.1275 [0.7471; 1.7016]      39.7       39.7
## DP4   0.8607 [0.4168; 1.7775]      12.8       12.8
##
## Number of studies combined: k = 4
##
##                          RR           95%-CI     z p-value
## Fixed effect model   1.0305 [0.7950; 1.3357] -0.23  0.8205
## Random effects model 1.0305 [0.7950; 1.3357] -0.23  0.8205
##
## Quantifying heterogeneity:
##   tau^2 = 0 [0.0000; 0.1205]; tau = 0 [0.0000; 0.3471]
##   I^2 = 0.0% [0.0%; 16.6%]; H = 1.00 [1.00; 1.09]
##
## Test of heterogeneity:
##   Q  d.f. p-value
## 0.55    3  0.9077
##
## Details on meta-analytical method:
## - Inverse variance method
## - Paule-Mandel estimator for tau^2
## - Q-profile method for confidence interval of tau^2 and tau

# Generate forest plot
forest.meta(
    m.gen,
    sortvar = TE,
    predict = TRUE,
    leftlabs = c("Data Source","LogRR","SE")
     )
```

Sample forest plot from R