Study information

Title	A Non-Interventional Post-Approval Safety Study of the Pfizer-BioNTech COVID-19 mRNA Vaccine in the United States.
Protocol number	C4591009
Protocol synopsis version identifier	1.0
Date	01 April 2021
EU Post Authorization Study, PAS register number	To be registered before the start of data collection
Active substance	COVID-19 mRNA Vaccine is single-stranded, 5'-capped messenger RNA (mRNA) produced using a cell-free in vitro transcription from the corresponding DNA templates, encoding the viral spike (S) protein of SARS-CoV-2.
Medicinal product	Pfizer-BioNTech COVID-19 Vaccine
Research question and objectives	Research question: What are the incidence rates/prevalence of safety events of interest among individuals vaccinated with the Pfizer-BioNTech COVID-19 vaccine within selected United States data sources participating in the Sentinel System, overall and in subpopulations of interest (pregnant women, immunocompromised individuals, and individuals with a history of COVID-19) compared with rates of those events in individuals who have not received any vaccination for COVID-19?

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780098

	Primary objectives • To estimate the relative risk (RR) or prevalence ratio of safety events of interest following receipt of at least one dose of the Pfizer-BioNTech COVID-19 vaccine within the overall study population • To estimate the RR or prevalence ratio of safety events of interest following receipt of at least one dose of the Pfizer-BioNTech COVID-19 vaccine in pregnant women, in immunocompromised individuals, and in individuals with a history of COVID-19 Secondary objectives • To describe the proportion of individuals receiving at least one dose and a complete dose series of the Pfizer-BioNTech COVID-19 vaccine, within the overall study population, in pregnant women, in immunocompromised individuals, and in individuals with a history of COVID-19 • To describe—among individuals who receive a first dose of the Pfizer-BioNTech COVID-19 vaccine—the timing and type of second dose of COVID-19 vaccine (Pfizer-BioNTech COVID-19 vaccine or other COVID-19 vaccine), within the overall study population, in pregnant women, in immunocompromised individuals, and in individuals with a history of COVID-19 • To describe baseline characteristics (demographics and comorbidities) of individuals who receive at least one dose of the Pfizer-BioNTech COVID-19 vaccine and those who receive no COVID-19 vaccine of any type, within the overall study population, in pregnant women, in immunocompromised

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780099

	individuals, and in individuals with a history of COVID-19
Author	Alison Kawai, ScD RTI Health Solutions Waltham, Massachusetts Jeffrey Brown, PhD Department of Population Medicine Harvard Medical School & Harvard Pilgrim Health Care Institute Boston, Massachusetts Cynthia de Luise, MPH, PhD Risk Management and Safety Surveillance Research Pfizer, Inc. New York, New York

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780100

TABLE OF CONTENTS

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780101

APPENDICES

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780102

1. TITLE

A Non-Interventional Post-Approval Safety Study of the Pfizer-BioNTech COVID-19 mRNA Vaccine in the United States

2. RATIONALE AND BACKGROUND

The novel coronavirus SARS-CoV-2, the cause of COVID-19, has resulted in a global pandemic. On December 11, 2020, the Pfizer-BioNTech COVID-19 vaccine was authorized for emergency use by the Food and Drug Administration (FDA) to prevent COVID-19 in individuals aged 16 years and older in the United States (US) (FDA, 2020). Pfizer-BioNTech is submitting a Biological License Application (BLA) for marketing approval of the vaccine for the prevention of SARS-CoV-2 infection in individuals aged 12 years and older.

Post-authorization observational studies using real-world data are needed to assess the association between the Pfizer-BioNTech COVID-19 vaccine and pre-determined safety events of interest in individuals administered the vaccine in the general population and in subpopulations of interest (e.g., pregnant women, immunocompromised individuals, and individuals with a history of COVID-19). This protocol synopsis describes a proposed observational study of safety events of interest occurring in recipients of the Pfizer-BioNTech COVID-19 vaccine using claims data and electronic health records data (where available) from Data Partners participating in the Sentinel System. The safety events of interest in this study are based on those included in COVID-19 vaccine safety surveillance in the FDA's Biologics Effectiveness and Safety (BEST) System (Wong et al., 2021) and the Centers for Disease Control (CDC) Vaccine Safety Datalink (VSD) (Shimabukuro et al., 2021). Additional safety events of interest may be added as new evidence develops during the pandemic. This protocol synopsis summarizes a proposed non-interventional study, designated as a post-authorization safety study, which is anticipated as a commitment to the FDA.

3. RESEARCH QUESTION AND OBJECTIVES

Research question: What are the incidence rates/prevalence of safety events of interest among individuals vaccinated with the Pfizer-BioNTech COVID-19 vaccine within selected US data sources participating in the Sentinel System, overall and in subpopulations of interest (pregnant women, immunocompromised individuals, and individuals with a history of COVID-19), compared with rates of those events in individuals who have not received any vaccination for COVID-19?

Primary objectives

- To estimate the relative risk (RR) or prevalence ratio of safety events of interest following receipt of at least one dose of the Pfizer-BioNTech COVID-19 vaccine within the overall study population

- To estimate the RR or prevalence ratio of safety events of interest following receipt of at least one dose of the Pfizer-BioNTech COVID-19 vaccine in pregnant women, in immunocompromised individuals, and in individuals with a history of COVID-19

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780103

Secondary objectives

- To describe the proportion of individuals receiving at least one dose and a complete dose series of Pfizer-BioNTech COVID-19 vaccine, within the overall study population, in pregnant women, in immunocompromised individuals, and in individuals with a history of COVID-19

- To describe—among persons who receive a first dose of the Pfizer-BioNTech COVID-19 vaccine—the timing and type of second dose of COVID-19 vaccine (Pfizer-BioNTech COVID-19 vaccine or other COVID-19 vaccine), within the overall study population, in pregnant women, in immunocompromised individuals, and in individuals with a history of COVID-19

- To describe baseline characteristics (demographics and comorbidities) of individuals who receive at least one dose of the Pfizer-BioNTech COVID-19 vaccine and those with no record of COVID-19 vaccination of any type, within the overall study population, in pregnant women, in immunocompromised individuals, and in individuals with a history of COVID-19

4. STUDY DESIGN

This is a retrospective cohort study comparing vaccinated individuals with concurrent unexposed comparators. Vaccinated individuals will be matched to concurrent unexposed comparators (in a ratio of at least 1:2) on data source and calendar time for analysis in the overall study population, immunocompromised individuals, and individuals with a history of COVID-19. In pregnant women, those who are vaccinated will be matched to concurrent unexposed comparators (in a ratio of at least 1:2) on maternal age and pregnancy start. Propensity score methods will be used to adjust for confounding through matching or in regression analysis. The study will use data from five Data Partners that participate in the Sentinel System.

The study period will start on the date that the Pfizer-BioNTech COVID-19 vaccine was granted emergency use authorization in the US (December 11, 2020) and will end a minimum of 3 years after this date.

5. POPULATION

The source population for this study will be commercial health plan enrollees from five Data Partners that contribute claims and electronic health records data to the Sentinel System: Aetna/CVS Health, HealthCore/Anthem, HealthPartners, Humana, and Optum/UnitedHealthcare.

Individuals of all ages will be included in the descriptive analysis of Pfizer-BioNTech COVID-19 vaccine utilization.

Safety analysis is planned to be limited to individuals within the age-indicated population for the Pfizer-BioNTech COVID-19 vaccine. However, if the number of individuals receiving the vaccine outside the age-indicated range is substantial (criteria for determining this to be

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780104

defined in the statistical analysis plan [SAP]), then these individuals will be included in the safety analysis.

Individuals will be eligible for the study if they have continuous medical and pharmacy insurance coverage for at least 12 months before baseline (i.e., the start of follow-up). Women will be eligible to be included in analysis of the pregnant population if they were pregnant for at least 1 day during the study period (regardless of the timing of pregnancy start). Analysis of congenital malformations, preterm birth, and small size for gestational age will be limited to pregnancies ending in a livebirth; linked data on infants during the first year of life will be used to identify select safety events of interest. Additional study eligibility criteria are described in the data analysis section (Section 9).

6. VARIABLES

6.1. Safety events

Safety events of interest will be identified in claims and electronic health records (where available, as not all Data Partners will have access to electronic health records) using diagnosis codes, with procedure and/or pharmacy dispensing codes as appropriate. Detailed definitions will be included in the SAP.

Outcomes likely to be misclassified (to be determined based on clinical expert opinion and review of prior validation studies if available) may be validated. As needed, validation of select outcomes will be performed by clinicians without knowledge of vaccination status by review of medical records or patient profiles (i.e., chronological listings of codes in claims data).

The following safety events of interest will be assessed in the general population, immunocompromised individuals, individuals with a history of COVID-19, and pregnant women. This list comprises events being monitored in rapid cycle analysis of COVID-19 vaccines in the FDA's BEST System (Wong et al., 2021) and the CDC's VSD (Shimabukuro et al., 2021), with the addition of vaccine-associated enhanced respiratory disease:

- Acute disseminated encephalomyelitis
- Acute myocardial infarction
- Acute respiratory distress syndrome
- Anaphylaxis
- Appendicitis
- Bell's palsy
- Convulsions/seizures
- Disseminated intravascular coagulation

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780105

- Encephalomyelitis
- Guillain-Barré syndrome
- Thrombotic thrombocytopenic purpura
- Immune thrombocytopenia
- Kawasaki disease
- Multi inflammatory syndrome
- Myocarditis/pericarditis
- Narcolepsy
- Stroke, hemorrhagic
- Stroke, ischemic
- Transverse myelitis
- Deep vein thrombosis
- Vaccine-associated enhanced respiratory disease
- Venous thromboembolism
- Pulmonary embolism (subset of venous thromboembolism)

The following pregnancy outcomes will be assessed in pregnant women or their infants:

- Spontaneous abortion (spontaneous pregnancy loss before 20 completed weeks gestation)
- Stillbirth (fetal deaths at or after 20 completed weeks gestation)
- Preterm birth
- Major congenital malformations
- Small size for gestational age

Other emergent safety events of interest may be added as the understanding of the safety profile of the Pfizer-BioNTech COVID-19 vaccine evolves and feasibility of their assessment permits in the data sources.

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780106

Risk windows will be defined for safety events of interest that have a hypothesized increased risk during specific time periods following vaccination. For other safety events of interest, patients will be followed for safety events for a maximum of 1 year.

6.2. Vaccine exposures

Exposures to the Pfizer-BioNTech COVID-19 vaccine will be identified in claims and electronic health records data via pharmacy dispensing and/or procedure codes. It is anticipated that some vaccinations administered outside traditional medical care settings may not be captured in claims or electronic health records data. Therefore, sensitivity analysis will incorporate study designs that are less susceptible to misclassification of unexposed status (Section 9.4). For analysis of pregnancy outcomes, except for congenital malformations, exposures occurring anytime during pregnancy (excluding those that occur after the at-risk period for the outcome where applicable, e.g., in analysis of preterm birth, a woman vaccinated after 37-weeks gestation would not be considered exposed) will be considered. For congenital malformations, only pregnancies with vaccinations occurring during the exposure window (i.e., first trimester) will be considered exposed.

6.3. Covariates

The following potential confounders will be identified:

- Demographics: age, sex, and race/ethnicity (if feasible). Data on race/ethnicity are anticipated to be incomplete, and the feasibility of including this variable in the analyses will be assessed before study start

- Comorbidities, identified in claims or electronic health records data (where available, as not all Data Partners will have access to diagnosis, procedure, and dispensing codes from electronic health records) in the 12 months before the index date: history of anaphylaxis, history of allergies, diabetes (type 1 and type 2), hypertension, cardiovascular disease, cerebrovascular diseases, chronic respiratory disease, chronic kidney disease, chronic liver disease, cancer, autoimmune disorders, influenza and other respiratory infections. Comorbidities will be identified via diagnosis codes (with procedure and/or pharmacy dispensing codes as appropriate).

- Medications and non–COVID-19 vaccinations in the 12 months before the index date (including vaccines administered concomitantly with the Pfizer-BioNTech COVID-19 vaccine), identified in claims or electronic health records data via procedure and/or pharmacy dispensing codes

- Health care utilization in the 12 months before the index date: number of hospitalizations; number of emergency department visits, cancer screening(s); skilled nursing facility, nursing home, or extended care facility stay; other preventive health care services, as appropriate

Immunocompromising conditions (to be considered as a potential confounder in analysis of the overall study population and to identify cohorts for safety and descriptive analysis in

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780107

immunocompromised individuals) will be identified using diagnosis, procedure, and dispensing codes for immunodeficiencies, immunosuppressant medication use, human immunodeficiency virus and other immunosuppressing conditions, and receipt of organ or bone marrow transplants.

Pregnancy status (to be considered as a potential confounder in the analysis of the overall study population and to identify cohorts for safety and descriptive analysis in pregnant women) will be identified in claims or electronic health records data. An algorithm that incorporates diagnosis and/or procedure codes will identify the final pregnancy outcome as well as the start and end dates for each pregnancy episode.

History of COVID-19 before the index date (to identify cohorts for safety and descriptive analysis in individuals with prior history of COVID-19) will be identified in claims or electronic health records data via diagnosis codes.

7. DATA SOURCES

This study will use data from five Data Partners, including data from four national insurers (CVS Health/Aetna, HealthCore/Anthem, Humana, and Optum/UnitedHealthcare) and one regional insurer (HealthPartners). Each Data Partner is a participant in the FDA Sentinel System. These data sources capture longitudinal medical care information on outpatient medication dispensings, vaccine administrations, and inpatient and outpatient diagnoses and procedures. The data sources also capture member demographic and health plan enrollment information. Each Data Partner can request access to full-text medical records for outcome validation for at least a subset of participants. All Data Partners are able to link to external data sources such as state immunization registries and can collect additional information via surveys in at least a subset of members. As part of their participation in Sentinel, three Data Partners (CVS Health, HealthCore, and Optum) maintain a mother-infant linkage table to support studies of medication exposures during pregnancy. As all Data Partners contribute data to the Sentinel System, this study will leverage Sentinel data and distributed querying infrastructure, including quality-checked and curated data formatted to the Sentinel Common Data Model (SCDM) and the publicly available Sentinel analytic tools (Curtis et al., 2012; Sentinel, 2018).

The Sentinel System is an active surveillance system that uses routine querying and analytical tools to evaluate electronic health care data from a distributed data network for monitoring the safety of regulated medical products in the US, established under the Sentinel Initiative (Behrman et al., 2011; Platt et al., 2018). The average enrollment length for patients across data sources in Sentinel is similar to that in other claims databases of members with medical and pharmacy coverage; approximately 25% of patients have over 3 years of enrollment, and patients with chronic conditions such as diabetes and older members typically have longer than average enrollment periods within these databases.

The Data Partners use the SCDM (Curtis et al., 2012; Sentinel, 2018) for standardization of demographic and clinical data elements. Publicly available routine analytical tools (i.e., reusable, modular Statistical Analysis System [SAS] programs) designed to be executed

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780108

against the SCDM permit rapid and standardized queries across partners, including descriptive analyses and complex methodologies (e.g., comparative analyses).

Specific information in the SCDM includes, but is not limited to, the following types of data:

- Enrollment data, including one record per covered individual per unique enrollment span.
- Individuals are assigned a unique identifier by their insurer that is linkable to all other data in the SCDM. Each record in the enrollment file indicates the patient identifier, enrollment start and end dates, and whether the patient was enrolled in medical coverage, pharmacy coverage, or both during that range.
- Demographic data, including birth date, sex, race/ethnicity, and ZIP code of their most recently recorded primary residence.
- Outpatient pharmacy dispensing data, including the date of each vaccination or prescription dispensing, the National Drug Code, NDC identifier associated with the dispensed product, the nominal days' supply, and the number of individual units (e.g., pills, tablets, vials) dispensed. Products purchased over the counter or at some cash-only retail locations selling prescription drug products (e.g., through the Walmart Prescription Program) are not consistently captured.
- Medical encounter data, including the health care provider most responsible for the encounter as well as the facility at which the encounter occurred and its ZIP code. Admission and discharge dates (if applicable) are also included in addition to the encounter type (i.e., an ambulatory visit, Emergency Department visit, inpatient hospital stay, non-acute inpatient stay, or otherwise unspecified ambulatory visit). Discharge disposition (i.e., alive, expired, or unknown) as well as discharge status (i.e., to where a patient was discharged) are also included for inpatient hospital stays and non-acute inpatient stays.
- Diagnosis data, including the date of diagnosis, its associated encounter identifier, admission date, provider identifier, and encounter type. Diagnoses are recorded with ICD-9-CM, *International Classification of Diseases, 9th Revision, Clinical Modification* and ICD-10-CM, *International Classification of Diseases, 10th Revision, Clinical Modification* codes. For inpatient hospital and non-acute inpatient stay encounters, the SCDM includes both principal and non-principal discharge diagnosis data.
- Procedure data, including the procedure date (e.g., date of vaccination), its associated encounter identifier, admission date, provider identifier, and encounter type, are coded as ICD-9-CM and ICD-10-Procedure Coding System procedure codes; *Current Procedural Terminology*, CPT categories II, III, or IV codes; revenue codes and Healthcare Common Procedure Coding System (HCPCS) levels II and III codes.

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780109

A brief description of each individual data source is below.

7.1. CVS Health, Aetna

Aetna, a CVS Health company, is one of the nation's leading health care benefits companies, serving 38 million people. Aetna became an FDA Sentinel Data Partner in 2008. Aetna's SCDM captures longitudinal information on dispensed prescriptions, inpatient and outpatient diagnoses, and inpatient and outpatient treatments and procedures.

7.2. HealthCore

HealthCore, Inc., a participant in Sentinel since 2008, is a significant contributor to the Sentinel Collaboration and Sentinel Distributed Database.

As of February 2021, there were 79 million unique individuals with medical coverage and approximately 60 million with medical and pharmacy coverage available for research.

7.3. HealthPartners

HealthPartners is the largest consumer-governed nonprofit health care organization in the US, providing care, insurance coverage, research, and education to its members. HealthPartners serves more than 1.8 million medical and dental health plan members and more than 1.2 million patients. HealthPartners Institute has been a member of the Sentinel Initiative since 2008.

7.4. Humana

Humana/Comprehensive Health Insights (CHI) is a health economics and outcomes research subsidiary of Humana Pharmacy Solutions, which focuses on treatment effectiveness, drug safety, adherence, medical and pharmacy benefit design, disease management programs, and other health care services based on the Humana health plan member population. Humana/CHI has been an active collaborator and Data Partner in the FDA Sentinel System. Humana databases represent geographic coverage for the entire US population and represent over 27 million lives.

7.5. Optum Research Database

The Optum Research Database (ORD) is a proprietary research database that contains eligibility data and medical claims and includes health plan members who are geographically diverse across the US and comprise approximately 3% to 4% of the US population. Optum has curated and quality-checked data formatted to the SCDM available for use and is a longtime participant in the Sentinel System.

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780110

8. STUDY SIZE

Assuming that the true RR = 1 and a matching ratio of 1:2, the table below presents the probability that the upper limit of the 95% confidence interval (CI) for the observed RR will be below 1.5, 2.0, 2.5, and 3.0 for study sizes ranging from 500,000 to 20,000,000 vaccinated individuals (1,000,000 doses to 40,000,000 doses, under the assumption that each person will receive 2 doses). These estimates are presented to cover a range of safety events of interest with respect to rareness, based on background rates in the general population.

Safety event of interest	**Estimated background rate per 100,000 person-years (Black et al., 2021)**	**Number of individuals vaccinated**	**Probability that the upper confidence limit of RR will be below the following thresholds[a] :**			
			1.5	**2.0**	**2.5**	**3.0**
Guillain-Barré syndrome	1.68	500,000	0.07	0.12	0.18	0.24
		1,000,000	0.10	0.20	0.31	0.42
		2,500,000	0.18	0.42	0.64	0.80
		5,000,000	0.31	0.70	0.91	0.98
		10,000,000	0.54	0.94	1.00	1.00
		20,000,000	0.83	1.00	1.00	1.00
Bell's palsy	25.2	500,000	0.31	0.70	0.91	0.98
		1,000,000	0.54	0.94	1.00	1.00
		2,500,000	0.90	1.00	1.00	1.00
		5,000,000	1.00	1.00	1.00	1.00
		10,000,000	1.00	1.00	1.00	1.00
		20,000,000	1.00	1.00	1.00	1.00

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780111

Safety event of interest	Estimated background rate per 100,000 person-years (Black et al., 2021)	Number of individuals vaccinated	Probability that the upper confidence limit of RR will be below the following thresholds[a]:			
			1.5	2.0	2.5	3.0
Myocardial infarction	208	500,000	0.99	1.00	1.00	1.00
		1,000,000	1.00	1.00	1.00	1.00
		2,500,000	1.00	1.00	1.00	1.00
		5,000,000	1.00	1.00	1.00	1.00
		10,000,000	1.00	1.00	1.00	1.00
		20,000,000	1.00	1.00	1.00	1.00

RR = relative risk.

a. Estimates in this table assume a risk window duration of 42 days for Guillain-Barré syndrome, and 28 days for Bell's palsy and myocardial infarction.

9. DATA ANALYSIS

All analyses will be conducted separately within each data source. Pooled analysis of RR and prevalence ratio estimates from all data sources will be conducted using meta-analysis techniques or other appropriate analytic techniques.

Detailed methodology for summary and statistical analyses of data collected in this study will be documented in the SAP, which will be dated, filed, and maintained by the sponsor. The SAP may modify the plans outlined in the protocol; any major modifications of primary endpoint definitions or their analyses will be reflected in a protocol amendment.

9.1. Descriptive analysis

Descriptive analysis will be conducted in order to report on utilization of the Pfizer-BioNTech COVID-19 vaccine during the overall study period and during the study period, stratified in 12-week increments. The proportion of individuals receiving at least one dose and a complete dose series of the Pfizer-BioNTech COVID-19 vaccine will be estimated within the overall study population, in pregnant women, in immunocompromised individuals, and in individuals with a history of COVID-19. Among patients who receive a first dose of the Pfizer-BioNTech COVID-19 vaccine, the proportion of patients will be reported by type

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780112

of second dose of COVID-19 vaccine and time between the two doses will be described using summary measures.

9.2. Comparative analysis

9.2.1. Overall approach

Because the Pfizer-BioNTech COVID-19 vaccine is currently recommended in a series of two doses given 3 weeks apart, separate exposed cohorts will be formed for each dose. The dose 1 and dose 2 cohorts will be matched to concurrent unexposed comparators (ratio of at least 1:2) on calendar time for safety comparative analyses in the general population, immunocompromised individuals, and individuals with a history of COVID-19; and on maternal age and pregnancy start for safety comparative analysis of pregnant women. Unexposed comparators may contribute to the exposed cohorts if they subsequently receive the Pfizer-BioNTech COVID-19 vaccine during the study period. Separate analyses will be conducted for each safety event. Individuals with safety events of interest in a pre-specified washout period before the index date will be excluded to ensure that incident events are identified during the study period. Confounding bias will be addressed with propensity scores through matching or in regression analysis. Data from the dose 1 and dose 2 cohorts and their matched comparators will be combined, if appropriate, to obtain incidence rates/proportions and RR/incidence proportions of safety events following receipt of at least one dose of the Pfizer-BioNTech COVID-19 vaccine.

9.2.2. Analyses in overall study population, immunocompromised individuals, and individuals with history of COVID-19

For analysis of safety events in the overall study population, individuals receiving each vaccine dose will be matched to unexposed concurrent comparators within Data Partner on time period-specific propensity scores within 1-month periods of calendar time. The index date (i.e., start of follow-up) in the vaccinated cohorts will be the date of vaccination.

In unexposed comparators, the index date selected will be within close temporal proximity (e.g., within the same calendar month) to the date of vaccination in the corresponding exposed persons. Individuals will be eligible to be selected as unexposed comparators if they have no record of COVID-19 vaccination before the potential index date. Estimating propensity scores and matching within narrow time intervals accounts for the changing predictors of being vaccinated over time, seasonality (i.e., temporal trends) of SARS-CoV-2 and other respiratory infections, and changes in health care utilization over time. Follow-up in both groups will begin on or the day after the index date (depending on the safety event of interest) and end at the earliest of the following: end of the study period, end of data availability, disenrollment from the health plan, death, occurrence of the safety event of interest, end of the duration of the outcome-specific risk window (or 1 year for outcomes without a known risk window), or receipt of a dose of the Pfizer-BioNTech COVID-19 vaccine or any other COVID-19 vaccine. If dose 2 is the Pfizer-BioNTech COVID-19 vaccine, individuals will stop follow-up in the dose 1 cohort and will start follow-up in the dose 2 cohort. Immunocompromised individuals and individuals with a history of COVID-19 before the index date will be drawn and analyzed separately from the general population using the same approach.

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780113

9.2.3. Analyses in the pregnant population

An algorithmic approach will be used to identify pregnancies in women of reproductive age, as well as pregnancy start and end dates. Each vaccine dose during pregnancy will be matched to unexposed concurrent comparator pregnancies (matching ratio of at least 1:2) within Data Partners on maternal age and estimated pregnancy start date (i.e., date of last menstrual period [LMP]).

For analysis of all outcomes in pregnant women, except for congenital malformations, the index date (i.e., start of follow-up) in the vaccinated cohorts (dose 1 and dose 2 cohorts) will be the date of vaccination. The index date in unexposed comparators will be assigned to the equivalent of the gestational age at vaccination of the corresponding exposed comparators.

For analysis of non-pregnancy–related outcomes, pregnancies will be eligible to be unexposed comparators if they have no record of COVID-19 vaccination before the potential index date. For analysis of pregnancy outcomes, except congenital malformations, pregnancies will be eligible to be included in the unexposed comparator group if they have no record of COVID-19 vaccination any time during pregnancy before the potential index date or within the 42 days before pregnancy start; women who have received any COVID-19 vaccine more than 42 days before their pregnancy begins will be eligible to be included in analyses of these outcomes. Matching on pregnancy start date adjusts for seasonality (i.e., temporal trends) of SARS-CoV-2 and other respiratory infections as well as changes in health care utilization over time. Matching exposed and unexposed pregnancies on gestational age at the start of follow-up avoids bias due to the changing underlying risk of pregnancy outcomes over the course of pregnancy.

Confounding will be adjusted for in regression analysis with the use of propensity scores. Follow-up for non-pregnancy–related outcomes in pregnant women and for pregnancy outcomes (except congenital malformations, small size for gestational age, and preterm birth) will begin on the index date or the day after the index date (depending on the safety event of interest) and end at the earliest of the following: end of the study period, end of data availability, disenrollment from the health plan, death, occurrence of the safety event of interest, end of the duration of the outcome-specific risk window (or 1 year for outcomes without a known risk window), receipt of the Pfizer-BioNTech COVID-19 vaccine or any other COVID-19 vaccine, or end of pregnancy (for analysis of pregnancy outcomes). If dose 2 is the Pfizer-BioNTech COVID-19 vaccine, individuals will stop follow-up in the dose 1 cohort and will continue follow-up in the dose 2 cohort. Small size for gestational age and preterm birth will be identified at birth or shortly after birth in the mother or infant.

For analysis of congenital malformations, each exposed pregnancy will be matched to unexposed comparator pregnancies (matching ratio of at least 1:2) within Data Partners on maternal age and estimated pregnancy start date. The index date (i.e., cohort entry date and baseline period for defining covariates) in both the exposed and unexposed cohorts will be pregnancy start. The prevalence of congenital malformations in infants born to mothers with exposure to at least one dose of Pfizer-BioNTech COVID-19 vaccine during the exposure window (e.g., first trimester of pregnancy) will be compared with those born to mothers with no exposure to any COVID-19 vaccine during the exposure window. Propensity scores will

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780114

be used to analytically adjust for confounding. The time period for identifying congenital malformations in the infant will start at birth and will end at age 1 year, disenrollment from the health plan, diagnosis of the outcome, death, end of data availability, or end of the study period, whichever is earliest.

9.3. Further details on comparative analysis

For the cohort design, the distribution of demographics, comorbidities, and other potential confounders will be reported and compared between the matched exposed and unexposed cohorts. Balance in the matched cohorts will be assessed using standardized differences or other suitable methods.

Before comparative analysis of the matched cohorts, propensity scores will first be estimated by Data Partners separately in the overall study population, pregnant women, immunocompromised individuals, and individuals with a history of COVID-19, as the probability of being vaccinated vs. the probability of being in the comparator cohort at the index date using logistic regression and baseline variables; propensity scores will be calculated in distinct time periods (e.g., within the same calendar month) to account for changes in predictors of vaccination over time.

Incidence rates (for all outcomes except pregnancy-related outcomes) or incidence/prevalence proportions (for pregnancy-related outcomes) and 95% CIs of each of the safety events of interest will be estimated separately for the matched exposed and unexposed cohorts.

For comparative analyses of all outcomes except pregnancy-related outcomes, Cox models or Poisson regression will be used to estimate hazard ratios or incidence rate ratios and 95% CIs within the propensity score-matched cohorts. For comparative analysis of pregnancy outcomes, logistic regression will be used to estimate prevalence or incidence proportion ratios and 95% CIs. Comparative analysis in pregnant women will analytically adjust for propensity score. Risk/prevalence or incidence rate differences (depending on the safety event of interest) and 95% CIs will be estimated for all safety events of interest. Because each dose of vaccine within the same person will be considered a separate observation but combined in analysis, the estimation of variance will account for the correlation between dose 1 and dose 2 using appropriate statistical methods.

9.4. Sensitivity analyses

9.4.1. Self-controlled risk interval design

Sensitivity analyses incorporating a SCRI design will be implemented for acute outcomes with known risk periods in the overall study population, pregnant women (non-pregnancy–related outcomes only), immunocompromised individuals, and individuals with a history of COVID-19. Pregnancy-related outcomes will not be analyzed with the SCRI design. Only vaccinated individuals will be included in the SCRI analysis; the rate of a specific safety event in a post-vaccination risk interval will be compared with the rate in a control interval within the same person. The control interval will be the same length as the risk interval and will comprise person-time after the dose 2 risk interval; these intervals will be outcome

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780115

specific and specified in more detail in the SAP. A washout period between the risk and control intervals may be incorporated for safety events for which the risk interval is not well known. Because of the self-controlled nature of the design, bias of RR estimates arising from differences in the distribution of time-constant confounding factors between vaccinated and unvaccinated individuals is avoided with the SCRI design. Furthermore, as the design only includes vaccinated individuals, it avoids the potential for misclassification of unexposed status due to incomplete capture of COVID-19 vaccinations in claims or electronic health record data.

9.4.2. Cohort design with historical comparators

If feasible, sensitivity analysis of non-acute safety outcomes in the general population and in pregnant women will incorporate historical comparator cohorts of influenza vaccinees from a time period before the introduction of COVID-19 vaccines. The use of a historical unexposed comparator avoids the potential for misclassification of unexposed status due to incomplete capture of COVID-19 vaccinations in claims or electronic health record data. The feasibility of this analysis will depend on the absence of trends in coding for each safety event of interest over time in the historical comparator period and the study period.

9.4.3. Alternative risk intervals

For events for which the risk intervals are not well known (to be defined in the SAP), descriptive analyses of the timing of events relative to vaccination will be conducted. If temporal clusters of increased risk following vaccination are identified, sensitivity analyses will be conducted using alternative risk intervals.

9.5. Monitoring and interim analysis

Before the final safety analysis, results from the monitoring analysis will be reported in Q3 2022. The monitoring reports will describe the number and proportion of total individuals who have received the Pfizer-BioNTech COVID-19 vaccine in the overall study population, in immunocompromised individuals, and in individuals with a history of COVID-19. The completeness of exposure data and the number of vaccinated individuals outside the age indication will also be assessed in monitoring analyses.

Results from an interim analysis will be reported in Q3 2023. The results from the interim analysis will include those reported from the monitoring analysis. Additionally, the distribution of characteristics in the vaccinated and unvaccinated individuals and incidence rates of each safety event of interest (overall, not by exposure status) will be reported in the overall study general population, individuals with a history of COVID-19, and immunocompromised individuals. Notably, pregnant women will only be analyzed as a separate population in the final analysis to allow sufficient time for data on pregnancies to accrue, as pregnancies will be identified by their outcomes (e.g., livebirth, stillbirth, spontaneous abortion) and the data lag is approximately 6 to 9 months for most Data Partners.

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780116

10. STRENGTHS AND LIMITATIONS

A major strength of this study is that it will include a very sizeable source population in the US, as the participating Data Partners together collect data on more than 100 million individuals. The use of secondary data will enable the efficient assessment of several safety events of interest identified by the CDC's VSD and the FDA's BEST Initiative, in addition to pregnancy-related safety events of interest in pregnant women, while using robust study design and analytic approaches to adjust for potential confounding. Moreover, the secondary use of administrative data collected as part of routine medical care will avoid recall bias; this approach also avoids selection bias that might occur in primary data collection studies, as a patient's inclusion in this study is not voluntary.

Nevertheless, this study is subject to limitations arising from the use of secondary data and the selected study designs. Limitations related to the data sources include the potential for lack of recording in claims and electronic health records of COVID-19 vaccines administered without reimbursement from health insurers. This situation may lead to misclassification of exposed individuals as "unexposed" comparators, which will underestimate vaccine coverage rates and bias comparative risk estimates for the cohort design with concurrent unexposed comparators. The completeness of exposure data will be assessed in monitoring analyses before the end of the study; if the data appear to be substantially incomplete, then the primary study design may be reconsidered (for example, the SCRI and the cohort design with historical unexposed comparators may be designated as the primary study designs; and/or linkage to immunization registries may be considered). Additionally, the use of claims data and electronic health records data may lead to some misclassification of outcomes (e.g., false positives and false negatives). Some events, such as spontaneous abortion, will be incompletely captured in existing databases. Conversely, there have been limited validation studies of ICD-10-CM based algorithms for safety events of interest, and the accuracy of algorithms for many safety events of interest are unknown. When possible, validated algorithms will be used, and outcomes that are likely to be misclassified (based on prior validation studies and clinical expert input) may be validated through review of medical records or claims profiles, depending on the safety event of interest.

A study design-related limitation of both the cohort and SCRI designs is that any uncertainty regarding risk periods will lead to misclassification and attenuation of risk estimates. Sensitivity analyses with alternative risk intervals will be considered for outcomes for which the risk interval is not well known.

A limitation of the cohort design is the potential for residual or unmeasured confounding because it is unlikely that the data sources will have information on all potential confounders. To address potential confounding, the SCRI, which automatically adjusts for time-invariant confounders, will be used as a secondary approach where feasible. However, the SCRI is not well suited to study outcomes with gradual onset, long latency, or risk periods that are not well known. The SCRI may also be subject to bias for outcomes that affect exposure probability.

A limitation specific to the cohort design with concurrent unexposed comparators is that unvaccinated individuals may become exposed to the COVID-19 vaccine at any time during

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780117

the study; if this situation occurs frequently, the amount of unexposed person-time in the unexposed comparator group will be substantially reduced, which will limit the precision of RR estimates. Forming two separate exposed cohorts by dose number and matching unexposed to exposed at the time of each vaccine will minimize the loss of unexposed person-time due to receipt of vaccine in these individuals between the first and second doses. Additionally, the sensitivity analyses with the historical unexposed comparator cohort and the SCRI will not be subject to this limitation.

11. MILESTONES

Below is a proposed schedule of milestones, subject to discussion with the FDA.

Milestone	Planned date	Description of milestone
Registration in the EU PAS register	TBD	To be registered before the start of data collection.
Start of data collection, estimated[a]	Q2 2022	Start of data collection is the planned date for starting data extraction for the purposes of the study analysis.
Monitoring Analysis 1 Report[b,c]	Q3 2022	Vaccine counts and proportions of individuals in the databases who were vaccinated, within the overall study population, in immunocompromised individuals, and in individuals with a history of COVID-19.
Interim Study Report	Q3 2023	Vaccine counts and proportions of individuals in the databases who were vaccinated, within the overall study population, in immunocompromised individuals, and in individuals with a history of COVID-19. Distribution of characteristics among exposed and unexposed individuals within the overall study population, in immunocompromised individuals, and in individuals with a history of COVID-19. Incidence rates of safety events of interest, overall, without regard to exposure status in the overall study population, in immunocompromised individuals, and in individuals with a history of COVID-19.
End of data collection	Q1 2025	End of data collection is the planned date on which the analytical data set will first be completely available. The analytical data set is the minimum set of data required to perform the statistical analysis for the study objectives.

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780118

Milestone	Planned date	Description of milestone
Final Study Report	Q3 2025[d]	Descriptive analysis of vaccine utilization in the overall study population, in immunocompromised individuals, in individuals with a history of COVID-19, and in pregnant women. Comparative safety analysis in the overall study population, in immunocompromised individuals, in pregnant women, and in individuals with a history of COVID-19.

BLA = Biological License Application; PAS = post-authorization study; TBD = to be determined.

a. If BLA approved in Q3 2021.

b. Only includes Data Partners with a data lag of < 6 months.

c. Monitoring counts will not incorporate enrollment or any other study eligibility criteria.

d. Report may be delayed to Q4 2025, depending on the extent of validation and/or the need for external linkages.

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780119

12. REFERENCES

Behrman RE, Benner JS, Brown JS, et al. Developing the Sentinel System--a national resource for evidence development. N Engl J Med. 2011;3646:498-99. doi: 10.1056/NEJMp1014427 [published Online First: 2011/01/14].

Black SB, Law B, Chen RT, Dekker CL, Sturkenboom M, Huang WT, et al. The critical role background rates of possible adverse events in the assessment of COVID-19 vaccine safety. Vaccine. 2021. doi: https://doi.org/10.1016/j.vaccine.2021.03.016.

Curtis LH, Weiner MG, Boudreau DM, O'Cooper W, Daniel G, Nair Vet al. Design considerations, architecture, and use of the Mini-Sentinel distributed data system. Pharmacoepidemiol Drug Saf. 2012;2(Suppl 1):23-31. doi:10.1002/pds.2336 [published Online First: 2012/01/25].

Platt R, Brown JS, Robb M, McClellan M, Ball R, Nguyen MD, et al. The FDA Sentinel Initiative—an evolving national resource. N Engl J Med. 2018;37922:2091-93.

Sentinel [Internet]. Silver Spring, MD. Food and Drug Administration, FDA c2010. 2018. Distributed Database and Common Data Model. Available at: https://www.sentinelinitiative.org/sentinel/data/distributed-database-common-data-model. Accessed March 11, 2021.

Shimabukuro, T. Advisory Committee on Immunization Practices, ACIP: COVID-19 Vaccine Safety Update. 2021. Available at: https://www.cdc.gov/vaccines/acip/meetings/downloads/slides-2021-02/28-03-01/05-covid-Shimabukuro.pdf. Accessed March 5, 2021.

U.S. Food and Drug Administration (FDA). FDA Takes Key Action in Fight Against COVID-19 By Issuing Emergency Use Authorization for First COVID-19 Vaccine. 2020. Available at: https://www.fda.gov/news-events/press-announcements/fda-takes-key-action-fight-against-covid-19-issuing-emergency-use-authorization-first-covid-19. Accessed March 5, 2021.

Wong HL, Zhou CK, Thompson D, Dimova R, Clarke T, Forshee R, et al. COVID-19 Vaccine Safety Surveillance: Active Monitoring Master Protocol. 2021. Available at: https://www.bestinitiative.org/wp-content/uploads/2021/02/C19-Vaccine-Safety-Protocol-2021.pdf. Accessed March 5, 2021.

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780120

Appendix 1. Investigator list

Principal Investigators and Contributors to the Protocol Synopsis

Name, degrees	Job title	Affiliation
Cynthia de Luise, MPH, PhD	Senior Director, Epidemiology	Pfizer
Alison Kawai, ScD	Senior Research Epidemiologist	RTI Health Solutions
Jeffrey Brown, PhD	Associate Professor in the Department of Population Medicine	Harvard Medical School & Harvard Pilgrim Health Care Institute
J Bradley Layton, PhD	Senior Research Epidemiologist	RTI Health Solutions
Catherine Panozzo, PhD	Assistant Professor in the Department of Population Medicine	Harvard Medical School & Harvard Pilgrim Health Care Institute
Alicia Gilsenan, PhD, FISPE	Senior Director and Head, Epidemiology	RTI Health Solutions
Brian Calingaert, MS	Director, Epidemiology Analysis	RTI Health Solutions
Catherine Johannes, PhD	Senior Director, Epidemiology	RTI Health Solutions

Data Partner Coordinating Investigators

Name, degrees	Job title	Affiliation
Cheryl N McMahill-Walraven, MSW, PhD	Director, Safety & Collaboration	CVS Health
Aziza Jamal-Allial, PhD	Senior Epidemiologist	HealthCore
Kevin Haynes, PharmD, MSCE	Principal Scientist	HealthCore
Pamala A. Pawloski, PharmD., BCOP, FCCP	Senior Research Investigator	HealthPartners Institute
Vinit Nair, BPharm, MS, RPh	Principal & Director , Government Research & Consortiums	Humana
Florence Wang, ScD	Vice President, Epidemiology	Optum
Jessica Franklin, PhD	Principal Consultant, Epidemiology	Optum

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780121

Name, degrees	Job title	Affiliation

Note: Data Partner Coordinating investigators have reviewed and contributed to this protocol.

090177e196b004bd\Approved\Approved On: 13-Apr-2021 20:49 (GMT)

FDA-CBER-2021-5683-0780122