COVID-19 Vaccine-Mediated Enhanced Disease and Vaccine Effectiveness in the Vaccine Safety Datalink

VSD Study #1341

PROTOCOL VERSION: 1.2

DATE: 27 April 2021

KEY STUDY STAFF

Marshfield Clinic Research Institute (MCRI)	
Principal Investigator	**Edward Belongia, MD**
Co-Investigators	**Thomas Boyce, MD, MPH** **James Donahue, DVM, PhD** **Dave McClure, PhD**
Biostatistician	**Burney Kieke Jr., MS**
Project Managers	**Kayla Hanson, MPH** **Hannah Berger, MPH**
Data Manager	**Erica Scotty, MS**
Project Administrator	**Bobbi Bradley, MPH, MBA**

Vaccine Safety Datalink Collaborating Site Investigators and Project Managers	
Denver Health	**Joshua Williams, MD** **Jonathan Block, MEd, PMP**
Harvard Pilgrim	**Katherine Yih, PhD, MD** **Jessica LeBlanc, MPH**
HealthPartners	**Malini DeSilva, MD, MPH** **Leslie Kuckler, MPH**
Kaiser Permanente Colorado	**Matt Daley, MD** **Jo Ann Shoup, PhD**
Kaiser Permanente Northern California (AIM 2)	**Nicky Klein, MD, PhD** **Kristin Goddard, MPH**
Kaiser Permanente Northwest	**Stephanie Irving, MHS** **Tia Kauffman, MPH**
Kaiser Permanente Southern California (AIM 2)	**Bruno Lewin, MD** **Cheryl Carlson, MPH** **Denison Ryan, MPH**
Kaiser Permanente Washington	**Lisa Jackson, MD, MPH** **Erika Kiniry, MPH**

Centers for Disease Control and Prevention (CDC) Investigators	
Immunization Safety Office	**Jonathan Duffy, MD, MPH** **Michael McNeil, MD, MPH**
COVID-19 VE Team	**TBN**

PROTOCOL CHANGE HISTORY

Version/Date	Summary of Change
1.0/26 March 2021	Original protocol
1.1/21 April 2021	Minor revisions to eligibility criteria, exposure classification, and covariates
1.2/27 April 2021	Minor revisions to data lag and Aim 2 endpoints

PROTOCOL SYNOPSIS

Title: COVID-19 VMED and Vaccine Effectiveness in the VSD	
Short name: COVID-19 VMED/VE	
Study Rationale	Vaccine-mediated enhanced disease (VMED) can result from immune complex formation and Th2 biased cellular immune response when vaccinated individuals become infected with the target pathogen. Animal studies of candidate vaccines against SARS-CoV-1 have raised concerns about potential VMED in recipients of COVID-19 vaccines. Challenge studies in nonhuman primates after COVID-19 vaccination did not demonstrate VMED, and human clinical trials of mRNA vaccines have found no evidence of VMED. Despite these reassuring findings, long-term monitoring is needed since declining antibody titers could alter VMED risk over time. An observational study of VMED can simultaneously estimate vaccine effectiveness (VE) for hospitalized COVID-19, an important endpoint to assess the impact of vaccination on disease burden. This study in the Vaccine Safety Datalink (VSD) will estimate VE against COVID-19 hospitalization and provide data to assess the risk of VMED in vaccine recipients.
Study Aims	1. Estimate overall and product-specific VE against COVID-19 hospitalization and severe respiratory illness in a cohort of VSD enrollees. This aim will generate a signal for VMED if VE <0 with a 95% confidence interval that excludes 0. 2. Compare severity and outcomes in vaccinated and unvaccinated patients hospitalized for acute COVID-19 illness. To identify VMED, we will test the hypothesis that COVID-19 vaccination is associated with COVID-19-related ICU admission or increasing severity based on the WHO Clinical Progression Scale.
Study Design	**Aim 1.** Prospective cohort study of VSD enrollees who are age-eligible for COVID-19 vaccination based on Emergency Use Authorization. **Aim 2.** Cross-sectional analysis of severity in fully vaccinated and unvaccinated patients hospitalized with COVID-19.
Study Duration	Follow-up will begin on February 1, 2021, approximately when states began prioritizing COVID-19 vaccines to persons aged ≥65 years old. Follow-up will continue for approximately 24 months, until Q1 of 2023.
Eligibility Criteria	**Aim 1**. For the cohort analysis, VSD dynamic data files (DDF) will identify eligible individuals at each VSD site, initially based on age group. Eligible individuals must have at least 12 months of continuous VSD enrollment at the start of follow-up to ensure ascertainment of relevant medical conditions and prior COVID-19 hospitalization. Individuals with a prior inpatient diagnosis of COVID-19 will be excluded.

Title: COVID-19 VMED and Vaccine Effectiveness in the VSD	
Short name: COVID-19 VMED/VE	
	Aim 2. For the cross-sectional severity analysis, COVID-19 hospitalized individuals will be sampled each month. Fully vaccinated patients who are hospitalized for COVID-19 will be frequency matched to unvaccinated COVID-19 patients based on VSD site and age group. Individuals who are partially vaccinated at the time of hospitalization will be excluded.
Classification of Vaccination Status	**Aim 1**. For two dose vaccines, person-time will be categorized into 3 groups: • Unvaccinated (all unvaccinated and vaccinated person-time until <14 days after dose 1); • Partially vaccinated (14 days after dose 1 until <14 days after dose 2); • Fully vaccinated (≥14 days after receipt of dose 2). For one dose vaccines, person-time will be categorized into 2 group: • Unvaccinated (all unvaccinated and vaccinated person-time until <14 days after dose 1); • Fully vaccinated (≥14 days after receipt of dose 1) Individuals will accrue fully or partially vaccinated person-time on the first day of follow-up if doses were received at least 14 days prior to the calendar date when study follow-up period begins. A 14-day window for development of immunity will be used for Pfizer/BioNTech, Moderna, and Janssen vaccines. Vaccine manufacturer, dates of administration, and dose number will be extracted from VSD data files. The completeness of VSD COVID-19 vaccine records will be validated by medical record abstraction in Aim 2. **Aim 2**. Vaccination status will be assessed as fully vaccinated or unvaccinated at the time of COVID-19 hospitalization. Individuals will be classified as fully vaccinated if they completed the 2 dose series (Pfizer and Moderna) at least 14 days before hospital admission, or if they received a single dose vaccine (Janssen) at least 14 days before hospital admission. The unvaccinated group will be restricted to individuals who received 0 doses of COVID-19 vaccine before hospital admission. Partially vaccinated individuals will be excluded from the Aim 2 analysis.
Endpoints	**Aim 1**. The primary endpoint is hospitalization with an ICD-10 diagnosis code for COVID-19. These include U07.1 (COVID-19 infection or disease), J12.82 (pneumonia due to COVID-19), or M35.81 (multisystem inflammatory syndrome due to COVID-19 [MIS-C]). We will not require laboratory confirmation of SARS-CoV-2 infection since diagnostic tests may be performed prior to hospital admission and not available in VSD data files. The secondary Aim 1 endpoint is respiratory failure or acute respiratory distress syndrome (ARDS) due to COVID-19, identified by the combination of U07.1, J12.82, or M35.81 and a diagnosis code for either ARDS (J80) or acute respiratory failure (J96.0*) during the same hospitalization.

Title: COVID-19 VMED and Vaccine Effectiveness in the VSD	
Short name: COVID-19 VMED/VE	
	Aim 2. The co-primary severity endpoints are 1) admission to intensive care unit (ICU) and 2) WHO Clinical Progression Scale. The latter is an ordinal scale based on setting and need for respiratory or cardiovascular support. For hospitalized patients, scores range from 4 (no oxygen therapy) to 9 (mechanical ventilation with vasopressors, dialysis, or extracorporeal membrane oxygenation (ECMO). The score is 10 for COVID-19 death. Endpoints for Aim 2 will be manually abstracted from medical records. The abstraction will also be used to validate the ICD-10 codes (J80 and J96.0*) that define severe COVID-19 respiratory illness in Aim 1. Medical records will be reviewed (when available) from up to 50 patients per month across all participating VSD sites. Vaccinated hospitalized patients will be frequency matched to unvaccinated hospitalized patients in strata defined by age group, VSD site, and month of admission. Sampling of COVID-19 hospitalizations for abstraction will occur throughout the 24-month study period.
Analysis	For both Aims, data will be extracted from VSD files with a 12 week lag to ensure nearly complete capture of COVID-19 vaccinations and hospital admissions. **Aim 1**. The specific analytic approach will be described in a separate statistical analysis plan. A Cox proportional hazards regression analysis will generate hazard ratios for COVID-19 hospitalization and severe COVID-19 (ARDS or respiratory failure). VE will be defined as 100*(1-HR). We will construct separate models for each endpoint and product. Covariates will include age, sex, VSD site, race/ethnicity, receipt of other vaccines, and specific comorbid conditions associated with increased COVID-19 severity. We will assess changes in VE with increasing time after vaccination. We will separately estimate VE for fully vaccinated and partially vaccinated person-time relative to unvaccinated person-time. If feasible, several secondary analyses will be conducted. These include estimating VE for the following: 1) strata defined by age group (<50 years, 50-64 years, ≥65 years) and sex; 2) persons with a prior COVID-19 diagnosis; 3) pregnant women; 4) mixed-product recipients (e.g., 1 dose each of Pfizer/BioNTech and Moderna). We will perform a sensitivity analysis using a 7-day window after each vaccine dose (rather than 14 days) as the threshold for development of immunity. We may estimate VE in children and adolescents if one or more vaccines are authorized for use in persons under 16 years of age. We will conduct two interim analyses at approximately 2-4 and 12 months after initiating data extraction. The specific timing of interim analyses will be determined in consultation with CDC. The final analysis will occur after approximately 24 months of follow-up. **Aim 2**. The specific analytic approach will be described in a separate statistical analysis plan. For each severity endpoint, doubly robust estimation of potential outcomes will be used to examine the association with COVID-19 vaccination

Title: COVID-19 VMED and Vaccine Effectiveness in the VSD	
Short name: COVID-19 VMED/VE	
	status. The primary model will include hospitalized patients who are either unvaccinated or fully vaccinated with any FDA-authorized COVID-19 vaccine. If feasible, separate models will assess specific products. Covariates may include any of those described in Aim 1 and any sourced from medical record abstraction. One interim analysis is planned after approximately 600 chart abstractions (50% of planned total) have been completed. A significant association between COVID-19 vaccination and severity will be interpreted as evidence supporting (but not proving) VMED.

TABLE OF CONTENTS

1 Background

COVID-19 has created once-in-a-lifetime public health crisis. Nearly 30 million COVID-19 cases have been diagnosed in the US, and over 541,000 people have died from this disease (covid.cdc.gov/covid-data-tracker) Vaccine development has occurred at a record pace, and three highly efficacious vaccines have received Emergency Use Authorization (EUA) by the Food and Drug Administration (FDA). Two vaccines use an mRNA platform with a lipid nanoparticle (LNP) delivery system. The modified mRNA codes for the stabilized prefusion form of SARS-CoV-2 spike protein. The Moderna vaccine (mRNA-1273) is authorized for use in persons aged ≥18 years; it is administered as a 100 µg intramuscular (IM) dose at 0 and 28 days. The Pfizer/BioNTech vaccine (BNT162b2) is authorized for use in persons aged ≥16 years. It is administered as a 30 µg IM dose at 0 and 21 days. The Janssen vaccine uses a replication incompetent human adenovirus (Ad26) vector to deliver a transgene for stabilized spike protein. It is administered as a single IM dose of 5×10^{10} viral particles. Additional vaccines are in late phase clinical development, including a two dose chimp adenovirus vector vaccine (AstraZeneca) and a two dose recombinant spike protein nanoparticle vaccine (Novavax). One or more of these vaccines may receive EUA in the coming months.

One particular safety concern for all COVID-19 vaccines is *vaccine-mediated enhanced disease* (VMED), also known as vaccine-associated enhanced disease. VMED was first reported in the 1960s in children who received an investigational formalin-inactivated respiratory syncytial virus (RSV) vaccine [1]. The vaccine generated an antibody response, but a paradoxical effect was observed with more severe RSV illness in vaccine recipients. Eighty percent of RSV-infected vaccine recipients were hospitalized compared to only 8% of RSV-infected children in the control group. Similar evidence of enhanced disease was observed in children who received an inactivated measles vaccine [2].

Decades later, experimental vaccines against severe acute respiratory syndrome coronavirus 1 (SARS-CoV-1) were associated with more severe disease after challenge in animal models [3]. Two components of the adaptive immune response appear to mediate VMED in animals. First, immune complex formation with complement deposition led to activation of inflammatory cytokines in the lungs. Second, the vaccine triggered a T helper type 2 (Th2) biased immune response of CD4+ T cells. The Th2-type cytokines include interleukin 4 (IL4), IL5, and IL13, which are associated with immunoglobulin E (IgE) and allergic eosinophilic responses. Our current understanding is that two immune markers may indicate higher risk of VMED: elevated ratio of binding antibodies to neutralizing antibodies, and Th2 biased CD4+ response. The pathogenesis of VMED appears to be distinct from *antibody-dependent enhancement* (ADE), a phenomenon that has been associated with a dengue vaccine [4].

The risk of VMED after COVID-19 vaccination is theoretical. Limited challenge studies in nonhuman primates after SARS-CoV-2 vaccination have not demonstrated evidence of VMED, but follow-up time was short [5]. A trial of the Moderna mRNA vaccine in humans did not observe a Th2 biased response or a high level of binding (nonneutralizing) antibodies [6, 7]. The Pfizer/BioNTech phase 2/3 trial did not identify any clinical or immunologic findings that suggest VMED [8]. Similarly, the Janssen Ad26 vaccine generated neutralizing antibodies that increased up to 57 days after vaccination, and the T-cell response was Th1-biased [9]. All three vaccines generate high levels of neutralizing antibodies and there has been a Th1 biased response rather than Th2. In FDA EUA submissions, all three vaccines were reported to have 85% to 100% efficacy against severe COVID-19. All COVID-19 hospitalizations and deaths occurred in placebo recipients, providing further evidence that VMED does not occur soon after vaccination.

Despite these reassuring findings, long-term monitoring is needed since declining antibody titers could alter VMED risk over time. Studies to assess VMED risk can also be used to estimate vaccine effectiveness (VE) against serious COVID-19 illness, since VMED and VE are measuring risk in opposite directions. This protocol describes

an observational study in the Vaccine Safety Datalink (VSD) that will estimate VE against COVID-19 hospitalization and monitor for evidence of VMED during the study period.

2 Study Aims

1. Estimate overall and product-specific VE against COVID-19 hospitalization and severe respiratory illness in a cohort of VSD enrollees. This aim will generate a signal for VMED if VE <0 with a 95% confidence interval that excludes 0.
2. Compare COVID-19 severity and outcomes in fully vaccinated and unvaccinated patients hospitalized for acute COVID-19 illness. To identify VMED, we will test the hypothesis that COVID-19 vaccination is associated with 1) COVID-19-related ICU admission; and 2) increasing severity based on WHO Clinical Progression Scale after adjustment for potential confounders in the potential outcomes framework.

3 Study Design and Population

The VSD is a collaboration between the Immunization Safety Office at the CDC and nine integrated healthcare systems in the US (https://www.cdc.gov/vaccinesafety/ensuringsafety/monitoring/vsd/index.html). Healthcare systems contribute data on their members and patients, creating a large population of individuals for whom near complete immunization records and healthcare claims data are available. The annual VSD population is approximately 12 million people, or 3% of the US population. The VSD population includes individuals across the age spectrum; about 20% are children and 16% are 65 years or older.

All VSD sites will participate in this study including Denver Health (Denver, CO), Harvard Pilgrim (Boston, MA), HealthPartners (Minneapolis, MN), Kaiser Permanente Colorado (Denver, CO), Kaiser Permanente Northern California (Oakland, CA), Kaiser Permanente Northwest (Portland, OR), Kaiser Permanente Southern California (Los Angeles, CA), Kaiser Permanente Washington (Seattle, WA), and Marshfield Clinic (Marshfield, WI). All sites will contribute data except Harvard Pilgrim. Additionally, Kaiser Permanente Northern California and Kaiser Permanente Southern California will only contribute data for Aim 2.

We will address **Aim 1** with a prospective cohort study of VSD enrollees who are age-eligible for COVID-19 vaccination based on FDA EUA criteria and ACIP recommendation. We will address **Aim 2** with a cross-sectional analysis of severity in fully vaccinated and unvaccinated patients who are hospitalized with a diagnosis of COVID-19.

Follow-up time will begin to accrue on February 1, 2021, approximately when states began prioritizing COVID-19 vaccines for individuals aged ≥65 years old Vaccinated and unvaccinated person-time will accrue for approximately 24 months, until Q1 of 2023.

3.1. Inclusion and Exclusion Criteria

Aim 1. For the cohort analysis, VSD dynamic data files (DDF) will identify individuals ≥16 years of age at participating VSD sites. The age criteria may change over time if COVID-19 vaccines are authorized and recommended for use in children. Eligible individuals must have at least 12 months of continuous VSD enrollment at the start of follow-up; enrollment gaps <32 days will be ignored. Individuals with a prior inpatient diagnosis of COVID-19 will be excluded. Prior COVID-19 hospitalizations will be identified by ICD-10 codes B97.29 (1/1/2020-4/1/2020), U07.1, J12.82, or M35.81. Person-time follow-up will begin on February 1, 2021 or when an individual meets the study eligibility criteria. Person-time will be terminated at the time of COVID-19 hospital admission, death, disenrollment from VSD, or end of the study period.

Aim 2. For the cross-sectional severity analysis, VSD members aged ≥16 years that were hospitalized for COVID-19 after February 1, 2021 will be identified from DDF at participating sites. Individuals will be excluded if have

<12 months of continuous enrollment prior to hospitalization (enrollment gaps <32 days will be ignored) or had a prior inpatient diagnosis of COVID-19. Individuals will be sampled in strata defined by vaccination status at the time of hospitalization (fully vaccinated, unvaccinated), VSD site, month of admission, and age group. Individuals who are partially vaccinated at the time of hospital admission will be excluded from **Aim 2**.

3.2. COVID-19 Vaccination Status

Exposure to COVID-19 vaccines will be identified by CVX codes in VSD data files (**Table 1**). This list will be updated as more vaccine products become available in the US. We will ascertain vaccination date, product, manufacturer, and dose number for each exposure. Most VSD sites maintain linkage with state immunization information systems (IIS). We therefore expect to capture nearly all COVID-19 vaccines delivered outside the VSD sites. We will validate this as part of the medical record abstractions in Aim 2.

Table 1. CVX Codes for COVID-19 Vaccines

Vaccine Product	CVX Code	CVX Code Description	Vaccination Schedule
Moderna COVID-19 Vaccine	207	SARS-CoV-2 (COVID-19) vaccine, mRNA, spike protein, LNP, preservative free, 100 mcg/0.5mL dose	2 doses separated by at least 28 days
Pfizer/BioNTech COVID-19 Vaccine	208	SARS-CoV-2 (COVID-19) vaccine, mRNA, spike protein, LNP, preservative free, 30 mcg/0.3mL dose	2 doses separated by at least 21 days
Janssen COVID-19 Vaccine	212	SARS-CoV-2 (COVID-19) vaccine, vector non-replicating, recombinant spike protein-Ad26, preservative free, 0.5 mL	1 dose
N/A	213	SARS-CoV-2 (COVID-19) vaccine, unspecified	N/A

Aim 1. For the cohort analysis, person-time will be classified as unvaccinated, partially vaccinated, or fully vaccinated based on an individual's exposure to COVID-19 vaccine(s). For two dose vaccines, partially vaccinated person-time will begin 14 days after receipt of dose 1. Fully vaccinated person time will begin 14 days after dose 2. For one dose vaccines, fully vaccinated person-time will begin 14 days after receipt of dose 1. The 14-day window for development of immunity will be used for Pfizer/BioNTech, Moderna, and Janssen vaccines. For each mRNA vaccine, we will perform a sensitivity analysis using a 7-day window after each dose for development of immunity.

Individuals who receive dose 1 and dose 2 from different manufacturers will be included in the overall analysis but excluded from product-specific analyses. Second doses will be invalid and excluded from analyses if administered more than four days prior to the recommended date (i.e., before day 17 for a 21 day dosing interval). Second doses administered after the recommended date will be valid regardless of timing.

Aim 2. For the cross-sectional analysis, vaccination status will be assessed at the time of COVID-19 hospitalization and classified as unvaccinated or fully vaccinated. Persons will be classified as fully vaccinated if they have received both doses of a two dose vaccine and it has been ≥14 days since receipt of dose two, or if they have received a one dose vaccine and it has been ≥14 days since vaccination. Persons will be classified as unvaccinated if they have received 0 doses of vaccine before hospital admission. Partially vaccinated individuals will be excluded for the Aim 2 analysis.

3.3. Outcome Measures

Aim 1. The primary outcome is hospitalization with an ICD-10 diagnosis code for COVID-19. The following ICD-10 codes will be used to identify COVID-19 cases:

- U07.1 (COVID-19 infection or disease);
- J12.82 (pneumonia due to COVID-19);
- M35.81 (multisystem inflammatory syndrome due to COVID-19).

CDC coding guidelines indicate that U07.1 should be used only for persons with a confirmed diagnosis of COVID-19. CDC guidelines indicate that both U07.1 and J12.82 should be assigned for cases of COVID-19 pneumonia (www.cdc.gov/nchs/data/icd/Announcement-New-ICD-code-for-coronavirus-19-508.pdf). Since adherence to coding guidelines is unknown, we will accept any of these three codes assigned in the inpatient setting as evidence of COVID-19 hospitalization. We will not require laboratory confirmation of SARS-CoV-2 infection because diagnostic tests may be performed prior to hospital admission and might not be available in VSD data files.

The secondary endpoint is acute respiratory failure or acute respiratory distress syndrome (ARDS) due to COVID-19. This outcome will be defined as:

1. Inpatient diagnosis code for COVID-19 (U07.1, J12.82, or M35.81)
   AND
2. Inpatient diagnosis code for ARDS (J80) or acute respiratory failure (J96.0*) during the same hospitalization.

Aim 2. The co-primary severity endpoints are 1) admission to intensive care unit (ICU); and 2) increasing score on WHO Clinical Progression Scale (**Table 2**). The latter is an ordinal scale based on setting and need for respiratory or cardiovascular support [10]. For hospitalized patients, scores range from 4 (no oxygen therapy) to 9 (mechanical ventilation with vasopressors, dialysis, or extracorporeal membrane oxygenation (ECMO). The score is 10 for COVID-19 death.

Table 2. WHO Clinical Progression Scale for COVID-19

Patient State	Descriptor	Score
Uninfected	Uninfected; no viral RNA detected	0
Ambulatory; mild disease	Asymptomatic; viral RNA detected	1
	Symptomatic; independent	2
	Symptomatic; assistance needed	3
Hospitalized; moderate disease	Hospitalized; no oxygen therapy	4
	Hospitalized; oxygen by mask or nasal prongs	5
Hospitalized; severe disease	Hospitalized; oxygen by non-invasive ventilation or high flow	6
	Intubation and mechanical ventilation; SpO2/FiO2 ≥200	7
	Mechanical ventilation; SpO2/FiO2 ≥150 but <200 or vasopressors	8
	Mechanical ventilation; SpO2/FiO2 <150 and vasopressors, dialysis, or ECMO	9
Death	Death	10

Abbreviations: ECMO=extracorporeal membrane oxygenation; FiO2=fraction of inspired oxygen; SpO2=oxygen saturation.

3.3.1. Medical Record Abstraction for Aim 2

Outcome measures for Aim 2 will be ascertained by manual abstraction of medical records. Medical records will be reviewed (when available) from up to 1200 hospitalized patients over 24 months.

Patients will be selected for medical record abstraction based on COVID-19 vaccination status at the time of hospital admission. We anticipate fewer vaccinated vs unvaccinated patients based on high vaccine efficacy against severe COVID-19 in phase 3 trials. The sampling strategy and volume may vary over time based on the number of eligible hospitalizations. We will initially sample all fully vaccinated patients who are hospitalized with

a diagnosis of COVID-19 in a given month. For the same month, we will sample unvaccinated patients with frequency matching on age group (<50 years, 50-64 years, ≥65 years) and VSD site. The ratio of vaccinated to unvaccinated patients may vary over time due to the limited capacity for medical record abstraction (approximately 50 records per month across all sites). We will abstract all COVID-19 hospitalizations in a given month if the total is ≤50. Sampling of COVID-19 hospitalizations and abstractions will occur throughout the 24 month study period. One interim analysis is planned after approximately half (n=600) of the abstractions are completed.

Medical record abstraction will include but is not limited to demographic characteristics, smoking status, residential setting, symptom onset date, hospital admission and discharge date, reason for admission, discharge disposition, symptoms, vital signs and oxygen saturation on day of admission, pre-existing chronic conditions, ICU admission, noninvasive and invasive respiratory support, and pulmonary and non-pulmonary complications. Specific COVID-19 therapies will be abstracted. Dates and results of diagnostic tests for SARS-CoV-2 (PCR or antigen test) will be recorded until discharge or death.

COVID-19 vaccination status and dates will be independently abstracted from medical records for validation of electronic VSD data files. Abstracted medical record data on respiratory disease severity will be used to validate the ICD-10 codes that define severe COVID-19 respiratory illness in Aim 1 (J80 and J96.0*). For this validation, any patient with a score of 7-10 on the WHO Clinical Progression Scale will be defined as having severe respiratory disease. We will assess the sensitivity, specificity and predictive value of the ICD-10 codes among the hospitalized patients with abstracted data.

MCRI will develop and pilot a standardized chart review instrument. A detailed instruction manual will be developed and distributed to ensure consistent data collection across sites; training sessions for abstractors may also be conducted. Each site will receive a line list of cases to review that will contain at a minimum the VSD ID number and date of COVID-19 hospitalization. Chart review data will be directly entered into an online REDCap database hosted by MCRI. Completed abstractions will be reviewed and adjudicated by investigators at MCRI.

4 Analytic Approach

4.1. Aim 1

A separate statistical analysis plan will be developed to describe analytic methods in detail. Briefly, we will estimate hazard ratios (HR) for COVID-19 hospitalization and severe COVID-19 (ARDS or respiratory failure) using Cox proportional hazards models. VE will be defined as 100*(1-HR). We will construct separate models for each endpoint and product. An overall model will include all FDA-authorized COVID-19 vaccines. Covariates will include age, sex, race/ethnicity, VSD site, receipt of other vaccines, and specific comorbid conditions. We anticipate conducting two interim analyses at approximately 2-4 and 12 months after the start of data extraction at participating VSD sites. This specific timing and criteria for interim analyses will be determined in consultation with CDC. The final analysis will occur after approximately 24 months of follow-up. Data will be extracted from VSD files with a 12 week lag to ensure nearly complete capture of COVID-19 vaccinations and hospital admissions.

The time scale for the regression models will be calendar time and time zero will be February 1, 2021. Analyses will utilize the counting process formulation of the Cox model which readily accommodates key elements of this study such as left truncation and time-dependent covariates [11].

CDC has identified several pre-existing chronic disease categories associated with increased risk for severe COVID-19 illness or death (https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html). In addition, the CDC list includes a behavioral risk factor (smoking) and pregnancy. We will ascertain the presence of the chronic disease categories on the CDC list based on ICD-10 codes. Note, we

will not ascertain HIV status as a separate variable, rather include HIV in the immunocompromised state category. An individual will be classified as having a particular condition if a corresponding ICD-10 code is present in the inpatient, outpatient, or emergency department setting during the preceding 12 months before hospital admission. The prevalence of these conditions is much greater than the incidence, and these will be fixed covariates assigned at the time of entry into the cohort. Each of the various chronic disease categories will be included in the models *a priori*.

Body mass index (BMI) from height and weight measurements at each VSD site may be used to enhance the classification of overweight, obesity and severe obesity. Smoking is not included since it is not a pre-existing disease and smoking status is not collected in a consistent manner across VSD sites.

If feasible, we will categorize pregnancy status of women as a time-varying covariate using output from the Dynamic Pregnancy Algorithm (DPA). An exploratory analysis would estimate the hazard ratio for COVID-19 hospitalization in fully vaccinated vs unvaccinated pregnant women with adjustment for potential confounders.

Frailty is a likely risk factor for severe COVID-19, but frailty cannot be directly identified using electronic health record data. We anticipate that demographic and chronic disease variables in the model will reduce the potential for residual confounding due to frailty. Many of these variables are included in existing frailty measures (such as Charlson Comorbidity Index). We therefore do not intend to include a separate frailty index in each model.

We will also extract information on past receipt of influenza, pneumococcal and herpes zoster vaccines and include as covariates in an effort to control for the healthy vaccinee effect. We will look back 24 months for receipt of ≥1 influenza vaccine, back to age 50 for receipt of ≥1 herpes zoster vaccine (zoster vaccine live or recombinant zoster vaccine), and back to age 65 for receipt of ≥1 pneumococcal vaccine (pneumococcal conjugate or pneumococcal polysaccharide).

We will assess changes in VE with increasing time after vaccination by creating a composite exposure variable that reflects time since vaccination among the vaccinated. In addition, pre-existing natural immunity to COVID-19 may modify the association between vaccination and the **Aim 1** endpoints. We will evaluate this potential phenomenon by including an interaction term between the composite exposure variable and pre-existing natural immunity to COVID-19. Person-time will be classified as 'prior immunity' if there was an outpatient diagnosis of COVID-19, identified by ICD-10 code U07.1. J12.82, or M35.81, or a positive COVID-19 RT-PCR or rapid antigen test at least 28 days before the calendar date when follow-up time begins at each site. After this date, we will classify person-time as 'prior immunity' beginning 28 days after an outpatient COVID-19 diagnosis or positive lab test.

A recent topic of interest in the VE literature is the effect of depleting susceptibles on estimates of the waning of VE over time. We will review this literature and address this issue in the statistical analysis plan if feasible [12-15].

We will perform secondary analyses of VE in strata defined by age group (<50 years, 50-64 years, ≥65 years), and sex. If feasible, we will estimate VE among persons who received a COVID-19 diagnosis prior to vaccination, pregnant women, and mixed-product recipients (e.g., 1 dose each of Pfizer/BioNTech and Moderna). We will perform a sensitivity analysis using a 7-day window after each mRNA vaccine dose (rather than 14 days) as the threshold for development of immunity. We may estimate VE in children and adolescents if one or more vaccines are authorized for use in persons under 16 years of age.

The figure below shows the number of hospitalizations needed to detect hazard ratios (HR) from 0.05 to 6.0 for α=0.05 (type I error rate) and power=0.80 (type II error rate 0.20) for a two-sided test between vaccinated and unvaccinated at various levels of vaccination (proportion exposed) in the cohort. HR below 1.0 indicates vaccine effectiveness for preventing COVID-19 hospitalization. VE can be estimated as 100*(1-HR). HR above 1.0 is

consistent with VMED. Note that for Cox regression models, the sample size is the number of non-censored events (hospitalizations). With anticipated COVID-19 hospitalization on the order of hundreds per week in the VSD data, very large and very small HRs could be detected relatively soon. Calculations were performed in SAS/STAT 15.1 "Proc Power CoxReg" with 0.20 correlation of vaccination status with other covariates and proportion vaccinated ranging from 0.05 to 0.50 exposed in the population (note that results are symmetric for proportions vaccinated >0.50).

Under these scenarios, fewer than 100 hospitalizations are required to detect a strong VMED signal with HR >4.0 (**Figure 1A**) or high vaccine effectiveness with HR <0.25 (**Figure 1B**).

Figure 1A. Aim 1 Sample Size Estimates for Detecting an Increased Hazard Ratio for Hospitalization (VMED signal; VE <0) among Vaccinated Individuals (80% power, α=0.05, 2-sided test, R^2=0.20)

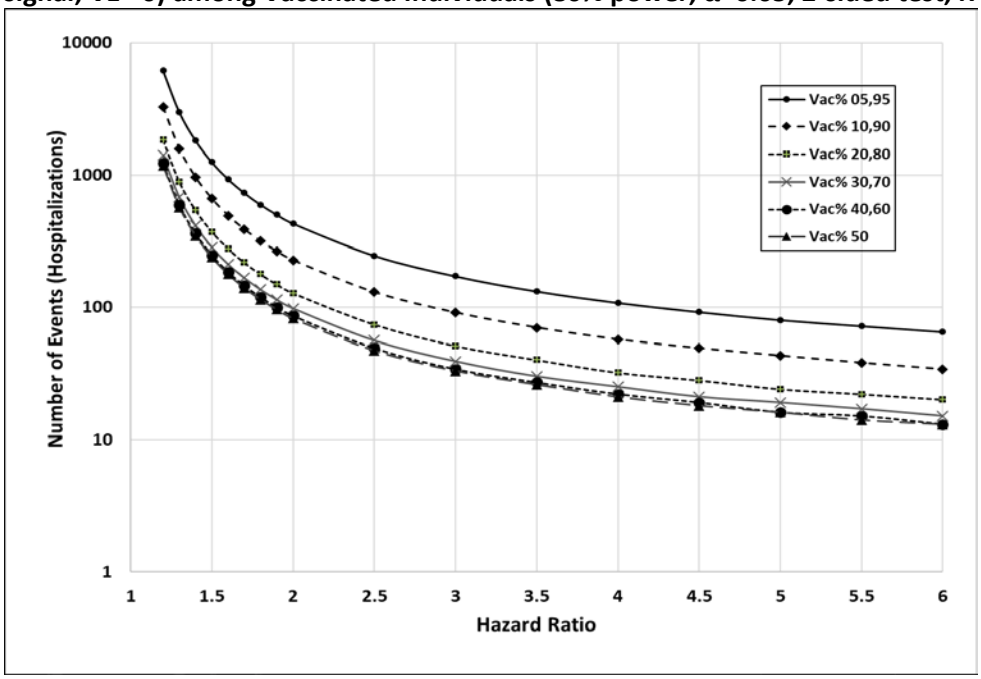
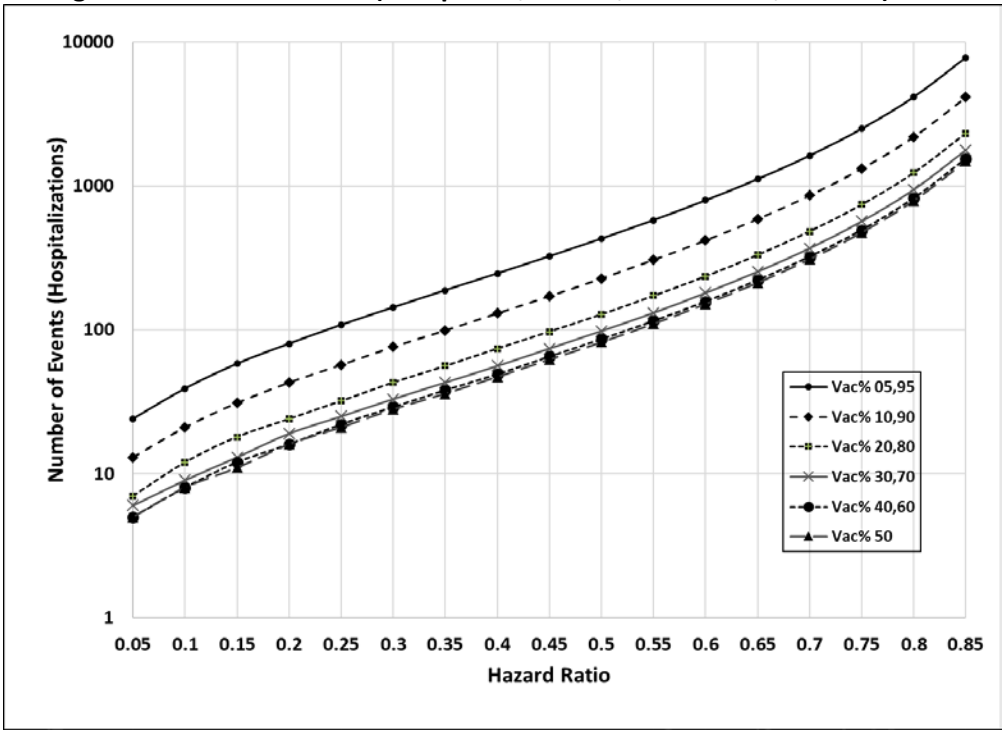


Figure 1B. Aim 1 Sample Size Estimates for Detecting a Reduced Hazard Ratio for Hospitalization (VE >0) among Vaccinated Individuals (80% power, α=0.05, 2-sided test, R^2=0.20)



4.2. Aim 2

Up to 600 fully vaccinated hospitalized patients will be frequency matched to unvaccinated patients with stratification by month of admission, age group (<50 years, 50-64 years, ≥65 years), and VSD site. The ratio of vaccinated to unvaccinated may vary over time due to the limited capacity for medical record abstractions each month (50 abstractions per month across all sites). The total number of medical record abstractions is estimated to be approximately 1200. If VMED occurs, evidence is likely to emerge during the latter half of the study period. One interim analysis is planned after approximately 600 chart abstractions (50% of planned total) have been completed. The specific timing for the interim analysis will be determined in consultation with CDC. Data will be extracted from VSD files with a 12 week lag to ensure nearly complete capture of COVID-19 vaccinations and hospital admissions.

A separate statistical analysis plan will be developed to describe the modelling approach for **Aim 2**. In general, causal inference using the potential outcomes framework will guide the analysis [16]. Doubly robust estimation will be used for either endpoint outcome [17]. The WHO Clinical Progression Scale may be collapsed to two categories. The collapsed categories would include moderate disease (score 4-5) and severe disease/death (score 6-10), with the former serving as the referent group in the regression analyses. Analysis will be in two stages. In the first stage, inverse probability weights will be constructed from a vaccination status propensity score model. In the second stage, the inverse probability weights will be combined with a regression adjusted outcome model to produce double robust estimation. Covariates for the models may include any covariates described in Aim 1 such as age, sex, race/ethnicity, VSD site, receipt of other vaccines, specific comorbid conditions, and any covariates sourced from the medical record abstractions.

The power table below shows the minimum detectible odds ratios using Fisher's exact conditional test with a 2-sided alpha of 0.05 and a percent of severe outcomes among the unvaccinated of 5, 10, and 20% (**Table 3**).

Table 3. Aim 2 Power Estimates for Risk of ICU Admission

Risk among Unvaccinated	Odds Ratio	No. of Abstracted Hospital Admissions	Power
5%	1.5	1200	0.343
10%	1.5	1200	0.592
20%	1.5	1200	0.828
5%	2.0	1200	0.837
10%	2.0	700	0.858
20%	2.0	400	0.832

5 Human Subjects and Confidentiality

Institutional Review Board (IRB) approval may be required at some participating VSD sites, although we anticipate the protocol will fall under the category of public health surveillance rather than human subjects research at CDC. MCRI will work with participating sites at the beginning of the project to determine which, if any, sites need to submit project-specific documents, including the protocol and chart review forms, for local IRB review.

The privacy and confidentiality of all subjects under surveillance will be strictly protected according to standard VSD procedures. The VSD project is covered by an Assurance of Confidentiality. CDC has obtained an Assurance of Confidentiality under Section 308(d) of the Public Health Service Act (42 U.S.C. 242 m(d)), which provides that this data can only be used for the purpose for which it is obtained, unless such institution or individual has consented to that disclosure. Pursuant to this, all CDC and VSD site project personnel have signed a nondisclosure statement.

Whenever possible, only aggregate data will be transferred to MCRI for analysis. When individual-level data are needed, variables will be restricted to those needed for the analysis and will not contain any direct identifiers. The data will include indirect identifiers such as date of birth, date of vaccination, and date of diagnosis, which are considered protected health information (PHI) under HIPAA. Transfer of data will occur via the secure VSD Hub; pre-existing agreements between VSD sites permit the transfer of limited datasets using this method.

Individual medical records will only be reviewed for abstraction of data as specified in the study protocol. Chart reviews will be conducted at individual sites by designated personnel. Direct identifiers required to link the automated data with a medical record will not be transferred to MCRI. Chart review data will be directly entered into an online REDCap database hosted by MCRI. Each individual user will receive a unique username and password. Data access groups will be configured to ensure that users only have access to data on individuals from their site, with the exception of MCRI, which will have access to all data captured in the database. If portions of the medical record are requested, non-essential patient identifiers will be redacted prior to upload into REDCap. REDCap databases are secure and HIPAA compliant and REDCap has previously been used for chart abstraction studies in the VSD. If needed, technical information will be provided to participating sites so that a security review or risk assessment can be conducted.

This project does not involve intervention or interaction with human subjects and is an analysis of existing data collected for non-research purposes. Risks are minimal and are limited to the inadvertent disclosure of PHI.

6 Data Management and Quality

Automated data will be extracted from existing, standardized VSD dynamic data files (DDF). Specific data files will include: CONSTANT, COVLRSLT, COVLTEST, DEATH, ENROLL, INPT, OUTPT, and VACCINE as well as CURRPREG, the output from the DPA. Sites have agreed to update the DEATH file on a monthly basis for another ongoing study; all other files are updated on a weekly basis. Additionally, an ancillary height/weight (HTWT) file will be requested from all participating sites at the time of each analysis. The unique VSD identification number will be used to link information between data files. Variables of interest are broken down by data file below:

- CONSTANT: VSD ID, site, date of birth, sex, race/ethnicity
- COVLRSLT: VSD ID, COVID-19 test result
- COVLTEST: VSD ID, specimen collection date, COVID-19 test type
- CURRPREG: VSD ID, pregnancy start date, pregnancy end date, pregnancy outcome, number and dates of pregnancy indicators
- DEATH: VSD ID, date of death
- ENROLL: VSD ID, enrollment dates
- HTWT: VSD ID, height and weight measurements, measurement dates
- INPT: VSD ID, date of admission, length of stay, ICD-10 diagnosis code, discharge disposition
- OUTPT: VSD ID, date of visit, department, ICD-10 diagnosis code, discharge disposition
- VACCINE: VSD ID, date of vaccine administration, vaccine product, vaccine manufacturer

For analytic data extractions, we will impose a 12-week data lag to allow for near-complete accrual of exposures and outcomes. Monthly monitoring programs may be run without a data lag.

SAS programs will be developed, tested, and distributed to all participating sites that will extract necessary information from each site's data files. Programs will be run separately at each site and study-specific datasets will be uploaded to the VSD Hub for secure transfer to MCRI. Datasets from each site will be downloaded and combined for analysis at MCRI.

7 Challenges and Limitations

This is an observational study utilizing electronic and clinical health records, and multiple factors may influence the accuracy and precision of risk estimates. In early 2021, COVID-19 activity was declining across most of the United States. The number of hospitalizations in VSD will be influenced by overall COVID-19 activity in the coming months, and the recent identification of highly transmissible variants (B.1.1.7 and B.1.351) may affect future disease activity. Increasing circulation of the B.1.351 in the US may cause confounding over time since some vaccines generate reduced neutralizing antibody titers and/or reduced protected against symptomatic illness caused by this strain.

This study relies primarily on COVID-19 diagnosis codes to identify cases rather than RT-PCR laboratory results. This limitation is due to incomplete access to SARS-CoV-2 test results for VSD enrollees who may have been tested through a commercial or public health laboratory. Due to limited supply, vaccine distribution is being prioritized for specific groups, but recommendations vary in different states. Prioritization of risk groups based on occupation or exposure risk may be a source of residual confounding since VSD files do not include this information.

Phase 3 clinical trials have provided strong evidence that VMED is not occurring during the first few months after vaccination. VMED is a theoretical risk beyond the first few months if neutralizing antibody titers decline, but the study is not powered to detect an elevated risk of VMED during the latter half of the observation period.

8 Site Responsibilities

Investigators at participating sites and the CDC will contribute to review of the draft protocol and data collection forms. They will provide scientific input on the analytic approach along with interpretation and publication of results. Sites will be responsible for the following tasks:

- Provide the MCRI study team with documentation of IRB and data transfer approval, if required;
- Review and approve SAS programs to generate study-specific datasets;
- Generate an updated ancillary HTWT file when requested; and
- Conduct local medical record abstractions to ascertain severity endpoints and relevant clinical data.

9 Timeline

Timeframe	Project Activity
April 2021	Finalize protocol and send to participating sites for regulatory approval
April-May 2021	Develop chart review form and manual
April-May 2021	Develop data extraction program
June 2021	Begin monthly data extraction and chart reviews
Q2-3 2021	Expected interim analysis #1 for AIM 1 (determined in consultation with CDC)
Q2-3 2022	Expected interim analysis #2 for AIM 1 (determined in consultation with CDC)
Q1 2023	Final analysis
Q2 2023	Submit manuscript to CDC clearance
Ongoing	Provide project updates on VSD calls and the VSD website

10 References

[1] Kim HW, Canchola JG, Brandt CD, Pyles G, Chanock RM, Jensen K, et al. Respiratory syncytial virus disease in infants despite prior administration of antigenic inactivated vaccine. Am J Epidemiol. 1969;89:422-34. https://doi.org/10.1093/oxfordjournals.aje.a120955

[2] Fulginiti VA, Eller JJ, Downie AW, Kempe CH. Altered reactivity to measles virus. Atypical measles in children previously immunized with inactivated measles virus vaccines. JAMA. 1967;202:1075-80. https://doi.org/10.1001/jama.202.12.1075

[3] Lambert PH, Ambrosino DM, Andersen SR, Baric RS, Black SB, Chen RT, et al. Consensus summary report for CEPI/BC March 12-13, 2020 meeting: Assessment of risk of disease enhancement with COVID-19 vaccines. Vaccine. 2020;38:4783-91. https://doi.org/10.1016/j.vaccine.2020.05.064

[4] Shukla R, Ramasamy V, Shanmugam RK, Ahuja R, Khanna N. Antibody-Dependent Enhancement: A Challenge for Developing a Safe Dengue Vaccine. Front Cell Infect Microbiol. 2020;10:572681. https://doi.org/10.3389/fcimb.2020.572681

[5] Corbett KS, Flynn B, Foulds KE, Francica JR, Boyoglu-Barnum S, Werner AP, et al. Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in Nonhuman Primates. N Engl J Med. 2020;383:1544-55. https://doi.org/10.1056/NEJMoa2024671

[6] Jackson LA, Anderson EJ, Rouphael NG, Roberts PC, Makhene M, Coler RN, et al. An mRNA Vaccine against SARS-CoV-2 - Preliminary Report. N Engl J Med. 2020. https://doi.org/10.1056/NEJMoa2022483

[7] Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, et al. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. N Engl J Med. 2021;384:403-16. https://doi.org/10.1056/NEJMoa2035389

[8] Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. N Engl J Med. 2020;383:2603-15. https://doi.org/10.1056/NEJMoa2034577

[9] Stephenson KE, Le Gars M, Sadoff J, de Groot AM, Heerwegh D, Truyers C, et al. Immunogenicity of the Ad26.COV2.S Vaccine for COVID-19. JAMA. 2021. https://doi.org/10.1001/jama.2021.3645

[10] WHO Working Group on the Clinical Characterisation and Management of COVID-19 Infection. A minimal common outcome measure set for COVID-19 clinical research. Lancet Infect Dis. 2020;20:e192-e7. https://doi.org/10.1016/S1473-3099(20)30483-7

[11] Therneau TM, Grambusch PM. Modeling Survival Data: Extending the Cox Model. New York: Springer-Verlag; 2000.

[12] Lipsitch M, Goldstein E, Ray GT, Fireman B. Depletion-of-susceptibles bias in influenza vaccine waning studies: how to ensure robust results. Epidemiol Infect. 2019;147:e306. https://doi.org/10.1017/S0950268819001961

[13] Ray GT, Lewis N, Klein NP, Daley MF, Lipsitch M, Fireman B. Depletion-of-susceptibles Bias in Analyses of Intra-season Waning of Influenza Vaccine Effectiveness. Clin Infect Dis. 2020;70:1484-6. https://doi.org/10.1093/cid/ciz706

[14] Fireman B, Gruber S, Zhang Z, Wellman R, Nelson JC, Franklin J, et al. Consequences of Depletion of Susceptibles for Hazard Ratio Estimators Based on Propensity Scores. Epidemiology. 2020;31:806-14. https://doi.org/10.1097/EDE.0000000000001246

[15] Tokars JI, Patel MM, Foppa IM, Reed C, Fry AM, Ferdinands JM. Waning of Measured Influenza Vaccine Effectiveness Over Time: The Potential Contribution of Leaky Vaccine Effect. Clin Infect Dis. 2020;71:e633-e41. https://doi.org/10.1093/cid/ciaa340

[16] Rubin DB. Causal Inference Using Potential Outcomes. J Am Stat Assoc. 2005;100:322-31. https://doi.org/10.1198/016214504000001880

[17] Funk MJ, Westreich D, Wiesen C, Sturmer T, Brookhart MA, Davidian M. Doubly robust estimation of causal effects. Am J Epidemiol. 2011;173:761-7. https://doi.org/10.1093/aje/kwq439