Global Vaccine Data Network®

Global Vaccine Data Network™

Background rates of adverse events of special interest following COVID-19 vaccination

Study protocol

Version 1.2
20 October 2022

TERMS OF USE

This work is subject to the *Terms of Use* set out at https://www.globalvaccinedatanetwork.org/aboutus/terms-use. Please read these carefully, it is very important that you are aware of the limits of the permitted uses and your obligations.

This work is made available for the purpose of academic research and other non-commercial purposes. If you wish to use this work for commercial purposes, contact us at gvdn@auckland.ac.nz for licensing fees and terms.

Copies of GVDN protocols must include the copyright notice.

Research proposals, protocols, reports, or publications using and/or derived from a Protocol(s) must include the following attribution: This work was based on the original *Global Vaccine Data Network (GVDN) Background rates of adverse events of special interest following COVID-19 vaccination study protocol* with the following citation: Global Vaccine Data Network. Background rates of adverse events of special interest following COVID-19 vaccination study protocol. Version 1.2. 2022 [updated 2022 October; cited YYYY month DD]. Available from: https://www.globalvaccinedatanetwork.org/Data-Dashboards.

THE GLOBAL VACCINE DATA NETWORK

The Global Vaccine Data Network™ (GVDN®) constitutes a multinational network of sites conducting globally coordinated active surveillance epidemiologic studies of the safety of vaccines, including COVID-19 vaccines. The GVDN network currently consists of 22 partners across 18 countries and is expanding. The GVDN is supported by the Global Coordinating Centre (GCC), hosted by UniServices at University of Auckland, Waipapa Taumata Rau in New Zealand. Through international collaboration with capacity for data linkage, it is now possible to have a large enough population to conduct robust analyses of rare events following vaccination.

GLOBAL **CO**VID **V**ACCINE **S**AFETY (GCoVS) PROJECT

Through UniServices, the GVDN was awarded a federal grant from the CDC/HHS to implement, host and manage a project titled "Assessing the safety of COVID-19 vaccines across large and diverse populations using the 17-country Global Vaccine Data Network Consortium", which is referred to as the **G**lobal **Co**VID **V**accine **S**afety (GCoVS) project.

FUNDING

This project is supported by the Centers for Disease Control and Prevention (CDC) of the U.S. Department of Health and Human Services (HHS) as part of a financial assistance award totalling US$5,643,515 with 100% percentage funded by CDC/HHS. The contents are those of the author and do not necessarily represent the official views of, nor an endorsement, by CDC/HHS, or the U.S. Government. For more information, please visit cdc.gov.

PROTOCOL DEVELOPMENT

This protocol was developed as one component of the GCoVS project, by the Background Rates and Observed vs. Expected Work Group. Members of the Work Group were associates of the GCoVS project partner sites who volunteered their time and expertise to develop a protocol suitable for use by multiple international sites, to harmonise collected data that could be amalgamated to increase study power by the GVDN. Expertise and administrative support were provided by the GVDN Global Coordinating Centre team, who are primarily based in Auckland, New Zealand.

Work group lead

Anders Hviid[1,2]

Work group members

Nelson Aguirre Duarte[3]
Karin Batty[4]
Steve Black[3,5]
Hannah Chisholm[3]
Hazel Clothier[6,7,8]
Heather Gidding[9,10,11]
Petteri Hovi[12]
Yannan Jiang[3]
Janine Paynter[3]
Helen Petousis-Harris[3]
Anastasia Phillips[9]
John Sluyter[3]
Thuan Vo[12,13]
Daniel Walsh[3]

Affiliations

1. Statens Serum Institut, Copenhagen, Denmark
2. University of Copenhagen, Copenhagen, Denmark
3. University of Auckland | Waipapa Taumata Rau, Auckland, New Zealand
4. Auckland UniServices Limited at University of Auckland | Waipapa Taumata Rau, Auckland, New Zealand
5. University of Cincinnati and Children's Hospital, Ohio, USA
6. Victorian Department of Health, Melbourne, Australia
7. SAFEVIC, Murdoch Childrens Research Institute, Melbourne, Australia
8. Centre for Health Analytics, Melbourne Children's Campus, Melbourne, Australia
9. National Centre for Immunisation Research and Surveillance of Vaccine Preventable Diseases, Sydney, Australia
10. Kolling Institute, Northern Sydney Local Health District, St Leonards, Australia
11. University of Sydney Northern Clinical School, Sydney Australia
12. Finnish Institute for Health and Welfare, Helsinki and Oulu, Finland
13. Tampere University, Tampere, Finland

HISTORY OF PROTOCOL VERSIONS

Version number	Date	Summary of changes
0.1	28 September 2021	Protocol created
0.2–0.9	10 October 2021–12 January 2022	Protocol drafted by work group lead and members
1.0	14 February 2022	Protocol finalised by work group lead and members for review by GCoVS sites
1.1	28 February 2022	Minor amendments made to protocol wording in response to feedback from GCoVS sites
1.2	20 October 2022	Protocol formatted for publication and open access from the GVDN website

TABLE OF CONTENTS

LIST OF TABLES

ABBREVIATIONS

Abbreviation	Term
AESI	adverse event of special interest
CDC	Centers for Disease Control and Prevention
CI	confidence interval
COVID-19	coronavirus disease due to SARS-CoV-2
ED	emergency department
GCC	Global Coordinating Centre for the Global Vaccine Data Network
GCoVS	Global COVID-19 Vaccine Safety project
GVDN	Global Vaccine Data Network
HHS	U.S. Department of Health and Human Services
HREC	Human Research Ethics Committee
ICD-10	International Classification of Diseases 10th revision
ID	identification number
IEC	Independent Ethics Committee
IRB	Institutional Review Board
PYRS	person-years
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
vs.	versus

PROTOCOL SYNOPSIS

Title

Background rates of adverse events of special interest following COVID-19 vaccination

Background

The estimation of background rates provide important context for regulatory and public health agencies to quickly assess emerging safety signals. Comparison of background rates and post-vaccination rates is a rapid and useful tool for the surveillance of vaccine adverse events of special interest (AESI). In the context of a global collaboration, estimation of background rates is feasible for many countries since only outcome and population estimates as denominator are required.

Aim and objective

To summarise global background rates of adverse events of special interest and collate data on background rates from different sites (countries) and population subgroups (age and sex) on a set of consistently defined AESI outcomes.

Study design

This is an observational retrospective study designed to estimate the baseline incidence of selected AESIs that could be potentially associated with COVID-19 vaccination.

Population

Patients presenting to site healthcare facilities (hospital inpatient, outpatient, emergency department, and primary care) during the period of the study.

Study period

Data collection is between 01 January 2015 to 31 December 2020 to reflect the situation before the availability of COVID-19 vaccines.

Outcome event

An outcome event is any one of 13 AESI defined by harmonised ICD-10 codes occurring during the study period within an individual where no previous outcome events have occurred within a washout duration of 365 days.

Analyses

The cohort will be described in terms of proportion of the population, sex ratio, and age structure. Background rates for each age group, sex, and period combination will be calculated as the expected count per 100,000 person-years. Exact confidence intervals on the lower and upper count bounds will be constructed using the Poisson distribution. The exact 95% confidence interval will be calculated by dividing the lower and upper count bounds by person-years. Sites that are prohibited to share table cells with low counts will suppress the cell values as appropriate.

1. BACKGROUND

Before the unprecedented global rollout of COVID-19 vaccines, a list of adverse events of special interest (AESI) was developed based on the pathophysiology of SARS-CoV-2 infection and what was known about vaccine safety issues in general. Post vaccination rollout, further events have been added in response to the safety signals of thrombosis with thrombocytopenia syndrome and myocarditis, respectively.

The estimation of background and post-vaccination rates is a rapid and useful tool for the surveillance of vaccine AESI. In the context of a global collaboration, estimation of background rates is feasible for many countries since only outcome and population estimates as denominator are required. Background rates provide important context for regulatory and public health agencies to quickly assess emerging safety signals. Countries with access to immunisation registers can also provide post-vaccination rates, which allows for observed versus expected comparisons of AESI. Such comparisons have the potential to investigate early safety concerns, inform vaccination policies and can be conducted rapidly; well before a more sophisticated analysis can be planned and carried out.

One highly relevant example of this approach was the thrombosis with thrombocytopenia signal, which prompted the suspension of the use of Oxford/AstraZeneca COVID-19 vaccine on 11 March 2021 in Denmark and Norway. Immediately, a collaboration between Denmark and Norway was formed to provide observed vs. expected comparisons for a range of thrombotic events based on nationwide register data. The results showed an increased risk of serious thrombotic events primarily in the form of cerebral venous sinus thrombosis following vaccination with the adenoviral vector vaccine, corresponding to one case per ~40,000 vaccinations.[a] On March 25, the vaccine was removed from the Danish programme. Norway similarly removed the vaccine from the national programme on May 12. Additional studies have confirmed this vaccine risk.

2. AIM AND OBJECTIVE

2.1 Aim

The study aims to summarise global background rates of adverse events of special interest.

2.2 Objective

Collate data on background rates from different sites (countries) and population subgroups (age and sex) on a set of consistently defined AESI outcomes.

3. METHODS

3.1 Study design

This is an observational retrospective study designed to estimate the baseline incidence of selected AESIs that could be potentially associated with COVID-19 vaccination.

3.2 Participant selection

Participants are patients presenting to site healthcare facilities during the period of the study, refer below for the description of 'patient types'.

3.2.1 Patient types

Patient types include hospital inpatients, outpatients, emergency department (ED) patients, and primary care patients. Definitions will vary between countries. In some countries separate datasets exist for emergency and outpatient departments. In countries without clearly defined patient types, contact duration (if available) can be used as a proxy for patient types at discretion of the site lead(s). As an example, a contact duration of 24 hours or longer can be used as a proxy for inpatients.

a. Pottegard A, Lund LC, Karlstad O, Dahl J, Andersen M, Hallas J, et al. Arterial events, venous thromboembolism, thrombocytopenia, and bleeding after vaccination with Oxford-AstraZeneca ChAdOx1-S in Denmark and Norway: Population based cohort study. BMJ. 2021;373:n1114.

3.2.2 Age group intervals

The age group interval could be 5-years (preferred interval), 10-years, or 20-years depending on the rarity of the outcome events at participating sites, refer to Table 1. Please also note that it is possible to combine different age groupings, e.g., 0–19 years, 20–24 years, 25–29 years, 30–34 years, 35–39 years, 40–49 years, 50–59 years, 60–79 years, and 80+.

Table 1. Age group intervals

5-year age group (preferred)	10-year age group	20-year age group
0–4	0–9	0–19
5–9		
10–14	10–19	
15–19		
20–24	20–29	20–39
25–29		
30–34	30–39	
35–39		
40–44	40–49	40–59
45–49		
50–54	50–59	
55–59		
60–64	60–69	60–79
65–69		
70–74	70–79	
75–79		
80+	80+	80+

3.3 Study period

Data collection should reflect the situation (AESI rates) before the availability of COVID-19 vaccines. It is known that vaccine availability in each country was variable. For the purposes of the study the period for the data collection is between 01 January 2015 to 31 December 2020.

3.4 Study variables

3.4.1 Outcomes

The outcomes are defined in Table 2. Each AESI was defined by harmonised ICD-10 codes. Some AESIs are defined by one ICD-10 code, others are defined by more than one, in this case any one of the ICD-10 codes listed constitutes an AESI case. The ICD-10 codes should be included/searched in primary and/or secondary diagnoses (may also be termed “associated” or “related”). Please see Table 2 for a list of the ICD-10 codes accompanied by their text.

3.4.2 Demographic variables

Individual-specific data will be collected at the site level (individual register per row). Requested variables are unique local/site identification number (ID), patient’s date of birth, and gender.

3.4.3 Event details

We will be collecting the admission date, discharge date and patient type (ED presentation, hospital inpatient, outpatient clinic, or primary care consultation). These are the event definitions for the study.

Table 2. Study outcome measures

Category	AESI	ICD-10 Code
Neurological conditions		
	Guillain-Barré syndrome	G61.0
	Transverse myelitis	G37.3
	Facial palsy	G51.0
	Acute disseminated encephalomyelitis	G04.0
	Febrile seizures	R56.0
	Generalised seizures	G40.0–G40.9, G41.0, R56.8
Haematological conditions		
	Thrombocytopenia	D69.5, D69.6
	Idiopathic thrombocytopenia	D69.3, D69.4
	Pulmonary embolism	I26.0, I26.9
	Cerebral venous sinus thrombosis	I63.6, I67.6
	Splanchnic vein thrombosis	I81, I82.0, I82.3
Cardiovascular conditions		
	Myocarditis	I40.1, I40.8, I40.9, I51.4
	Pericarditis	I30.0, I30.8, I30.9

3.4.3.1 Outcome event

An outcome event is defined by a relevant diagnosis and the corresponding admission date in the outcome source dataset. Depending on the outcome source dataset, an outcome event could e.g., constitute a hospital contact, an inpatient hospitalisation, or an ED visit. Typically, a recording in an outcome source dataset will consist of a diagnosis, and admission and discharge dates. The source dataset may contain multiple event rows relating to the same hospitalisation e.g., if a patient is transferred between departments however this is effectively the same hospitalisation. The process to be followed here is to combine these events and their diagnoses when there is **<24 hours (one day)** between them and the **earliest start date** for these bundled events is then used.

3.4.3.2 Incident case (new onset case)

This is an outcome event occurring during the study period within an individual where no previous outcome events have occurred within a washout duration. Consequently, an individual may contribute multiple incident cases as long as they are separated in time by at least the washout duration (**365 days**). Chronic conditions can only occur once in an individual, but all the study outcomes in this protocol (refer to section 3.4.3.1) can reoccur.

3.5 Data preparation

3.5.1 Definitions

3.5.1.1 Washout duration

Defined as a lag period of one year after an outcome event. The same type of event should it occur within one year is not included as an outcome event.

3.5.1.2 Ascertainment period

Period where data needs to be available to identify incident cases in the study period taking into account washout periods: 2014–2020 (corresponding to a maximum washout duration of one year for the 2015–2020 study period).

A separate dataset will be prepared for each AESI, as defined by harmonised ICD-10 codes and data source in the case of e.g., hospitals and EDs, and organised in long format, i.e., each row has information on one outcome event for one individual.

In some data sources, one hospitalisation can comprise many consecutive event recordings (if the data has not been previously cleaned into hospitalisation courses or episodes that take this into account using the one-day separation criterion described in section 3.4.3.1, the washout duration restriction will ensure that consecutive recordings are not counted as new events).

All outcome events occurring in the ascertainment period are initially included before the wash-out criterion is applied – see below. This is to ensure that an outcome event occurring in the beginning of the study period, e.g., January 2015, is not counted as an incident case if outcome events have occurred before the start of the study period and within the wash-out duration, e.g., in July 2014. Each individual can be represented in multiple rows if the event has occurred multiple times during the ascertainment period.

3.5.2 Identifying incident outcomes – incident cases dataset

For each ID:

a) Combine consecutive recordings within individuals where previous EVENT_END equals subsequent EVENT_START.
b) For each outcome event, look back 365 days (the washout period duration). If there is another event in this period (the washout period), then the outcome event is not counted as an incident case. This should be done at the same time for all possible outcome events.
c) Remove outcome events occurring before the study period start. The resulting dataset is the incident cases dataset.

3.5.3 Source population datasets

3.5.3.1 Scenario I: When individual-level information on the source (the source of the outcomes) population is available for linkage

a) Creation of source population dataset

Rows: A unique individual.
Columns: ID, SEX, DOB, EXIT_DATE

ID
Individual-level identifier, often pseudo-anonymised. String, e.g., ID0014984.

SEX
M, F, O where M=male, F=female, O= other gender or missing.

DOB
Date of birth, two-digit day, three-letter abbreviation of the month, four-digit year, e.g., 04JUL2022.

EXIT_DATE
Date of potential exit from follow-up due to e.g., death, emigration etc. Two-digit day, three-letter abbreviation of the month, four-digit year, e.g., 04JUL2022.

b) Linkage of incident outcomes to source population

The population dataset is merged with the incident cases dataset (LEFT_JOIN), such that for each ID in the population dataset, incident cases from the outcome datasets are linked, generating a row for each outcome – if there are no incident events for an ID, a row is generated with missing variables for the outcome variables. The merged dataset will have the following columns: ID, SEX, DOB, EXIT_DATE, AESI, EVENT_START, EVENT_END, PATIENT_TYPE.

c) Construction of aggregated outcome data

This dataset is then converted into an aggregated dataset with the columns; AESI, PATIENT_TYPE, AGE, SEX, PERIOD, COUNT, PYRS.

PATIENT_TYPE
1 = emergency department, 2 = hospital inpatient, 3 = hospital outpatient, 4 = primary care, 5 = all of hospital inpatient, outpatient, emergency department, primary care, 99 = missing.

AGE
The pre-defined age group intervals (5-year, 10-year, or 20-year). Where possible, the most narrow age intervals should be used.

SEX
M, F, O (O = other genders or missing).

PERIOD
Calendar years in the study period, 2015, 2016, 2017, 2018, 2019, 2020.

COUNT
The number of outcome incident cases in the age-, sex-, period- group specified by the other columns.

PYRS (person-years)
The cumulative amount of follow-up (in years) in the source population in the age-, sex-, period- group specified by the other columns.

This type of aggregated dataset is often generated using survival analysis tools in R or SAS that split a dataset with individuals in rows according to different time-periods. The resulting dataset will contain multiple rows for each ID, one for each combination of AGE, SEX and PERIOD that the person has contributed follow-up time in (also accounting for when that person enters the study and exits the study – study entry is 2015.01.01 and study exit is 2020.12.31 or date of loss from source data due to, e.g., death or emigration – EXIT_DATE). All cases and all follow-up for each combination of AGE, SEX and PERIOD can then be summed over all individuals creating the final aggregated dataset.

3.5.3.2 Scenario II: Aggregated information on the source (the source of the outcomes) population is available

If an aggregated dataset with the columns; AGE, SEX, PERIOD, POPULATION SIZE, can be obtained from demographic information, then we can aggregate the incident case dataset according to AGE, SEX and PERIOD and merge these two datasets on "AGE, SEX, PERIOD" resulting in a dataset with the following columns: one, AESI, PATIENT_TYPE, AGE, SEX, PERIOD, COUNT (from the incident case dataset), POPULATION SIZE. We then replace POPULATION SIZE with PYRS = POPULATION SIZE * one year (if PERIOD is one year, e.g., 2015).

3.5.3.3 Other possible scenarios

a) No linkage is possible to determine the incident cases of the outcome (i.e., we don't know what events pertain to the same person in each dataset as records within the one dataset-such as hospitalisations-are not able to be identified and thus combined). This should be 'do-able' by all sites.
b) Sites may only be able to contribute some datasets (e.g., only hospitalisations) either as a) linked or b) stand alone.

3.5.4 Adverse events of special interest (AESI) codes

Table 3 provides the code to identify each AESI in the dataset.

Table 3. AESI dataset codes

Code	AESI
	Neurological conditions
NE_GBS	Guillain-Barré syndrome
NE_TRM	Transverse myelitis
NE_BP	Facial palsy
NE_ADM	ADEM
NE_FSZ	Febrile seizures
NE_GSZ	Generalised seizures
	Haematological conditions
HM_THR	Thrombocytopenia
HM_ITC	Idiopathic thrombocytopenia
HM_PEM	Pulmonary embolism
HM_CER	Cerebral venous sinus thrombosis
HM_SVT	Splanchnic vein thrombosis
	Cardiovascular conditions
CV_MYO	Myocarditis
CV_PER	Pericarditis

3.6 Data analysis

3.6.1 Descriptive analysis

Briefly describe the cohort in terms of proportion of the population, sex ratio, and age structure.

3.6.2 Statistical analysis

3.6.2.1 Construction of background rates

Background rates for each age group, sex, and period combination can now be calculated as the expected count per 100,000 person-years. Exact confidence intervals on the lower and upper COUNT bounds can be constructed using the Poisson distribution (refer to https://epid.blogspot.com/2012/08/how-to-calculate-confidence-interval-of.html). The exact 95% confidence interval (CI) can then be calculated dividing the lower and upper COUNT bounds by PYRS.

3.6.2.2 Combining rates between sites

Ideally, the COUNTS and PYRS are **aggregated** across sites, and then the combined rate is calculated with exact 95% CIs at the coordinating site. Sites that are prohibited to share table cells with low counts will suppress the cell values as appropriate.

4. COMPLIANCE

4.1 Ethics approval and local authorisations

The study will be conducted in full conformance with local authority requirements. Ethical approval will be sought from the local Human Research Ethics Committee (HREC)/Institutional Review Board (IRB)/ Independent Ethics Committee (IEC) as required.

4.2 Changes to the protocol

Written protocol amendments will be documented as changes to the protocol, which will be summarised in the table of protocol amendments at the beginning of this publication. Major changes will almost always need to be approved by the appropriate HREC/IRB/IEC if such review is required at the site. In such circumstances, the adjustment will not be enacted until it has received approval. The investigator will file minor protocol adjustments, such as administrative changes, at each participating site and submit them to the appropriate local HREC/IRB/IEC as required.