

# BMJ Open Background incidence rates of health outcomes of interest for COVID-19 vaccine safety monitoring in a US population: a claims database analysis

Sonja Banga <sup>1</sup>, Alena Khromava,<sup>1</sup> Laurence Serradell,<sup>2</sup> Anne-Laure Chabanon,<sup>2</sup> Chunshen Pan,<sup>3</sup> Irisdaly Estevez,<sup>4</sup> Samantha Schilsky,<sup>4</sup> Hannah Kreisberg<sup>4</sup>

**To cite:** Banga S, Khromava A, Serradell L, *et al.* Background incidence rates of health outcomes of interest for COVID-19 vaccine safety monitoring in a US population: a claims database analysis. *BMJ Open* 2024;**14**:e083947. doi:10.1136/bmjopen-2024-083947

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<https://doi.org/10.1136/bmjopen-2024-083947>).

Received 03 January 2024  
Accepted 04 June 2024



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>Sanofi, Toronto, Ontario, Canada

<sup>2</sup>Sanofi, Lyon, France

<sup>3</sup>Sanofi, Bridgewater, New Jersey, USA

<sup>4</sup>Aetion Inc, New York, New York, USA

## Correspondence to

Dr Sonja Banga;  
[sonja.banga2@sanofi.com](mailto:sonja.banga2@sanofi.com)

## ABSTRACT

**Objective** To evaluate background incidence rates of 59 health outcomes of interest (HOI) in a diverse population, including important subpopulations, during the pre-COVID-19 era (1 January 2017–31 December 2019) and the COVID-19 era (1 March 2020–31 December 2020), before the introduction of COVID-19 vaccines.

**Design** Observational retrospective cohort study. Annual incidence rates and 95% confidence intervals (CIs) of HOIs were estimated for each population of interest, stratified by: age, sex, age and sex and seasonality.

**Data source** Optum's de-identified Clinformatics Data Mart Database (CDM).

**Participants** Individuals from the US general population and four subgroups of interest: influenza-vaccinated, paediatric (<18 years of age), elderly (≥65 years of age) and pregnant women.

**Results** During the COVID-19 era, the incidence of several cardiac conditions, coagulation disorders and acute liver injury increased across all populations assessed while the rates of some dermatological and neurological HOIs decreased relative to the pre-COVID-19 era. The incidence of acute respiratory distress syndrome (ARDS) varied considerably by subgroup: among the elderly, it decreased annually during the pre-COVID-19 era but peaked during the COVID-19 era; among pregnant women, it slightly increased annually during the pre-COVID-19 era and substantially increased during the COVID-19 era; among paediatrics, it decreased annually over the entire study. The incidence of the majority of HOIs increased with age, but were generally comparable between sexes with few exceptions. Cardiac, gastrointestinal, neurological and haematological HOIs, along with acute kidney injury and ARDS, were more common in males, whereas several immunological HOIs and chilblain-like lesions were more common in females. Pregnancy-related HOIs did not increase during the COVID-19 era, except for spontaneous abortions which increased annually over the entire study.

**Conclusion** These observations help contextualise fluctuations in background rates of adverse events noted during the COVID-19 era, and provide insight on how their use may impact safety surveillance for other vaccines.

## STRENGTH AND LIMITATIONS OF THIS STUDY

- ⇒ This study used real-world data for about 16 million eligible patients/year (from 2017 to 2020) across the USA to examine a wide range of health outcomes of interest (HOIs) that could be relevant for monitoring the safety of COVID-19 vaccines and other vaccines.
- ⇒ The study provides robust estimates of annual background incidence rates for HOIs for the general population and subgroups of interest including influenza-vaccinated, paediatrics (aged <18 years), elderly (aged ≥65 years) and pregnant women.
- ⇒ The richness of available information allowed for additional analyses by stratification for a number of factors including age, sex, age and sex (dual stratification) and seasonality.
- ⇒ The patient population was limited to those with health insurance coverage, and related elderly patients receiving Medicare.
- ⇒ Incidence rates may be overestimated or underestimated due to the potential for misclassification.

## INTRODUCTION

As of 21 April 2024, over 775 million cases of coronavirus disease 2019 (COVID-19) have been confirmed worldwide, with over 6.9 million deaths reported; of these, more than 100 million confirmed cases and over 1.1 million deaths have been reported in the US alone.<sup>1</sup> The emerging threat of COVID-19 led to the rapid development and authorisation of novel COVID-19 vaccines. However, clinical trials evaluating COVID-19 vaccine safety have been performed in highly selective populations that exclude more vulnerable populations, such as pregnant women, children and older adults, with relatively short follow-up periods (≤12 weeks),<sup>2</sup> and generally lacked power to detect rare adverse events (AEs).<sup>3</sup>

Owing to the restrictive nature of clinical trials (ie, restrictive inclusion criteria, limited sample size and follow-up period which may limit generalisability), postmarketing surveillance of AEs is



not only vital to fully understand the safety profile of COVID-19 vaccines, but also essential from a public-health perspective. In particular, the generation of robust background data for potential health outcomes of interest (HOI) is crucial for vaccine safety monitoring.<sup>4</sup> The background rate of an AE is the incidence rate of the event that would be observed in a given population in the absence of a given vaccine or any other intervention.<sup>1 2</sup> Background rates are important for contextualising rates from vaccine safety surveillance data by comparing the expected cases from the general population with the observed cases from the vaccinated population.<sup>5-7</sup> They can also inform the public on the number of specific events that will occur coincidentally among those receiving the vaccine.

Background rates can be obtained from multiple sources, such as published literature and healthcare databases.<sup>1 2</sup> Different methods for signal evaluation, such as observed versus expected (O/E) analysis allow the identification of potential safety signals.<sup>5-7</sup> Several initiatives have been formed to monitor the real-world safety of COVID-19 vaccines, with most focusing on background rates of HOIs. The vaccine COVID-19 monitoring readiness (ACCESS) project, funded by the European Medicine Agency (EMA), generated estimates of background rates for 41 HOIs to assess the impact of COVID-19 vaccines.<sup>8</sup> Similarly, studies from the Observational Health Data Sciences and Informatics (OHDSI) community project (a multistakeholder collaboration) assessed historical background rates in the general population of 15/16 HOIs potentially associated with COVID-19 vaccines across 8<sup>4</sup> and 11 countries,<sup>9</sup> respectively, with the former study also comparing these rates to those following COVID-19 infection. Additionally, the Food and Drug Administration (FDA) Biologics Effectiveness and Safety (BEST) initiative compared HOIs post-COVID-19 vaccination with historical incidence rates of HOIs pre-COVID-19.<sup>3</sup>

Despite these initiatives, data on the background rates of HOIs in time periods with and without COVID-19 circulation are limited. These data are essential to capture shifting societal and healthcare system practices during a time in which COVID-19 infections were prevalent, and when vaccines to manage these infections were not yet available. Furthermore, as the majority of studies only focused on a limited number of HOIs in the general population, data on a broader range of HOIs are lacking.

This study aimed to evaluate the background incidence rates for 59 HOIs in the US general population and four subgroups of interest (influenza vaccinated, paediatric (<18 years of age), elderly (≥65 years of age) and pregnant women) during the pre-COVID-19 era and the COVID-19 era, prior to the introduction of COVID-19 vaccines.

## METHODS

### Data sources

Deidentified administrative health claims data for members of large commercial and Medicare Advantage health plans were obtained from the Optum's deidentified

Clinformatics Data Mart Database (CDM).<sup>10</sup> The database contains claims submitted by providers and pharmacies and includes data on healthcare costs and utilisation, including outpatient visits, inpatient stays and outpatient prescription records, for a geographically diverse population across all 50 states. This includes claims for about 67 million deidentified patients,<sup>11</sup> and provides a comprehensive, single-payer view of data on patient demographics, diagnoses using International Classification of Diseases 10th Revision, Clinical Modification (ICD-10-CM) codes, and procedures using Current Procedural Terminology (CPT)-4 revenue codes and Healthcare Common Procedure Coding System (HCPCS).<sup>12</sup>

### Study period

This study included patients from 1 January 2017 through to 31 December 2020. The study period was divided into the two time periods: pre-COVID-19 era (1 January 2017–31 December 2019) and the COVID-19 era (1 March 2020–31 December 2020). The pre-COVID-19 era included calendar year-specific patient cohorts; each encompassing patient data obtained from 1 January to 31 December of the specified year. Stratified results for all study years are presented in the online supplemental material.

### Study populations

The general study population consisted of eligible patients (commercially insured and Medicare Advantage) who received a healthcare claim for any service (inpatient, outpatient, emergency room visit, durable medical equipment, skilled nursing facility, medication or laboratory work) during each calendar year of interest. For the influenza vaccine subgroup, patients were included if they had voluntarily received an influenza vaccine between 1 September of the prior year and 28 February of the year of interest (reflecting the influenza vaccination period in the US). The pregnancy subgroup consisted of female patients aged 12–55 years with documented evidence of pregnancy within the calendar year of interest. The paediatric population consisted of patients who were aged <18 years on the date of any healthcare claim (index date) in a given calendar year; however, patients aged 17 years on index were not censored once they turned 18 years. The elderly population consisted of patients aged ≥65 years on the date of any healthcare claim (index date) in a given calendar year.

### Health outcomes of interest

A full list of HOIs assessed in this study are shown in online supplemental table S1. A total of 59 HOIs, including cardiac, dermatological, endocrine, gastrointestinal, haematological, immunological, neurological, renal and respiratory outcomes, along with pregnancy-specific HOIs evaluated in the pregnancy subgroup, were assessed. Notably, these HOIs were not selected based on adverse events (AEs) observed in preauthorisation or precensure studies of COVID-19 vaccines; rather

many were suggested by the Safety Platform for Emergency vACCines (SPEAC) and subsequently adopted by the EMA, while other HOIs were included for additional preparedness. HOI algorithms were drawn from the FDA Sentinel Initiative and the FDA Best initiatives,<sup>3</sup> the OHDSI Atlas,<sup>4</sup> Health Outcomes of Interest for Evaluation in the Post-Licensure Rapid Immunization Safety Monitoring Program,<sup>13</sup> the ACCESS project (VAC4EU)<sup>8</sup> and additional studies that used algorithms of interest.<sup>14–23</sup> Validated ICD-10 algorithms were used if available, and validated or unvalidated ICD-9 algorithms were mapped to ICD-10 codes. A full list of the ICD-10 codes used in this study is also shown in online supplemental table S1.

The date of cohort entry into each specific calendar year was the index date. Specific washout periods were applied to each health outcome to ensure separate incident cases were captured. Continuous enrolment was required during respective washout periods; for example, for a 365-day washout period, patients were required to be enrolled during all 365 days. Previously used algorithms specific to COVID-19 research (see previous paragraph) were applied for definitions of HOIs, where possible, to reduce misclassification of events. Follow-up began on study entry and patients were censored, depending on which of the following occurred first: occurrence of the HOI being evaluated, end of the calendar year, end of the study period, no enrolment recorded for greater than 31 days (inactivity), death or end of pregnancy (pregnancy subgroup only, see below). For pregnancy-related HOIs, women pregnant on 1 January of a given calendar year had an index date of 1 January, and pregnancies that began after 1 January of a given year, were indexed on the first recorded evidence of pregnancy. For non-pregnancy-related HOIs, the index date was at the first recorded evidence of pregnancy or end of the HOI-specific washout period during pregnancy (in the case of ongoing pregnancy on 1 January of any given calendar year), whichever came later. Women could re-enter this subgroup multiple times during each calendar year if they had multiple pregnancies, as long as they met the HOI-specific washout periods and pregnancy requirements per calendar year. In addition, women in this subgroup were censored on the end of the pregnancy for the assessment of non-pregnancy HOIs.

### Statistical analysis

Incidence rates per hundred thousand person-years were calculated as the number of events divided by the total number of person-years and multiplied by one-hundred thousand. If 10 or more events were observed, the CIs were calculated using the normal distribution approximation; otherwise, CIs were computed using Poisson distribution. Among the general population and the four subgroups of interest, incidence rates and 95% CIs were calculated overall and by age group (0–5, 6–11, 12–15, 16–17, 18–24, 25–39, 40–49, 50–64, 65–74 and ≥75 years), sex at birth (male, female) and seasonality (yearly quarters). Age groups were aligned with vaccine exposure

data reporting categories established by the Centers for Disease Control and Prevention (CDC) during the pandemic. Aetion Substantiate was used to conduct these analyses in the CDM and the analysis of the pregnancy subgroup was implemented in a Spark cluster using R statistical software.

### Patient and public involvement

No patients were involved in the development, design, conduct or analysis of this study.

## RESULTS

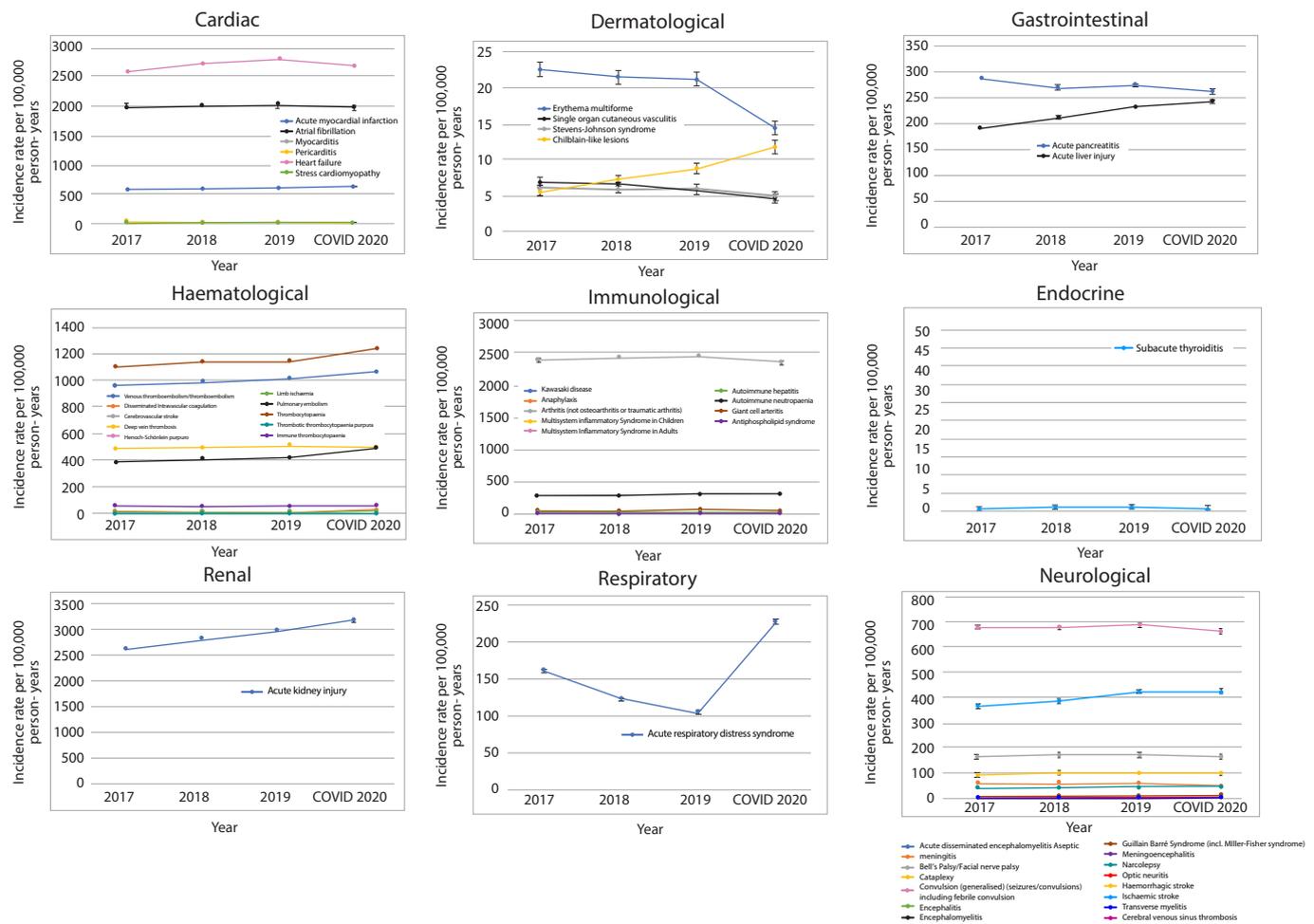
### Population characteristics

The number of patients with any healthcare service claim who met eligibility criteria identified for pre-COVID-19 and COVID-19 eras were as follows: 16 108 837 patients in 2017; 16 449 405 patients in 2018; 16 609 816 patients in 2019; and 15 463 486 patients in the 2020 COVID-19 era (1 March 2020–31 December 2020). We considered the period January–February 2020 as part of the pre-COVID-19 era and excluded these data from the rest of the 2020 data (11 014 663 patients excluded). The demographics for the general population and each subgroup (except for the pregnancy subgroup) for each year assessed are shown in online supplemental table S2. The mean age in the general population ranged from 45.5 (SD 23.8) years in 2017 to 49.0 (23.8) years in 2020. In contrast, the mean age in the paediatric and elderly subgroups remained relatively consistent between the pre-COVID-19 and COVID-19 era. There was a slightly larger population of females than males in most groups, except for the paediatric subgroup, which had an equal distribution across sex. The majority of patients included in the influenza vaccinated and paediatric subgroups, and general population were commercial payers, whereas the majority of patients in the elderly subgroup received Medicare.

### HOI incidence rates pre-COVID-19 era versus COVID-19 era

#### General population

Rates of HOIs varied between the pre-COVID-19 and COVID-19 eras for each HOI category in the general population (figure 1). Among cardiac HOIs, rates of heart failure were the highest throughout the study. All dermatological HOIs decreased in the COVID-19 era compared with the pre-COVID-19 era, except for incidence of chilblain-like lesions, which was higher during the COVID-19 era. Among gastrointestinal HOIs, the incidence of acute pancreatitis slightly decreased throughout the study period, whereas the incidence of acute liver injury increased and was highest during the COVID-19 era. The incidence rates of haematological HOIs, including venous thromboembolism (VTE), disseminated intravascular coagulation (DIC), pulmonary embolism and thrombocytopenia, were generally lower in the pre-COVID-19 era than in the COVID-19 era. For immunological HOIs, the incidence of antiphospholipid



**Figure 1** Health outcome of interest incidence stratified by year for the general population.

syndrome was low throughout the study period and decreased further in the COVID-19 era. Conversely, incidence of autoimmune neutropaenia remained relatively stable throughout the study period. Generally, the incidence of neurological HOIs was lower in the COVID-19 era than in the pre-COVID-19 era, with the exception of cerebral venous sinus thrombosis, and haemorrhagic and ischaemic stroke, which were lowest in 2017. Notably, convulsions had the highest incidence rate compared with all other neurological HOIs. Incidence rates of acute kidney injury and acute respiratory distress syndrome (ARDS) were higher in the COVID-19 era than in the pre-COVID-19 era.

### Paediatric population

Among paediatric patients, incidence rates of dermatological and gastrointestinal HOIs, along with acute kidney injury, were similar to those observed in the general population (figure 2); however, in contrast with the general population, incidence of ARDS declined over the study period. Additionally, cardiac (except for myocarditis), haematological, immunological and neurological HOIs remained stable throughout the study period. Overall, the incidence of seizures, convulsions and thrombocytopaenia was higher than all other HOIs.

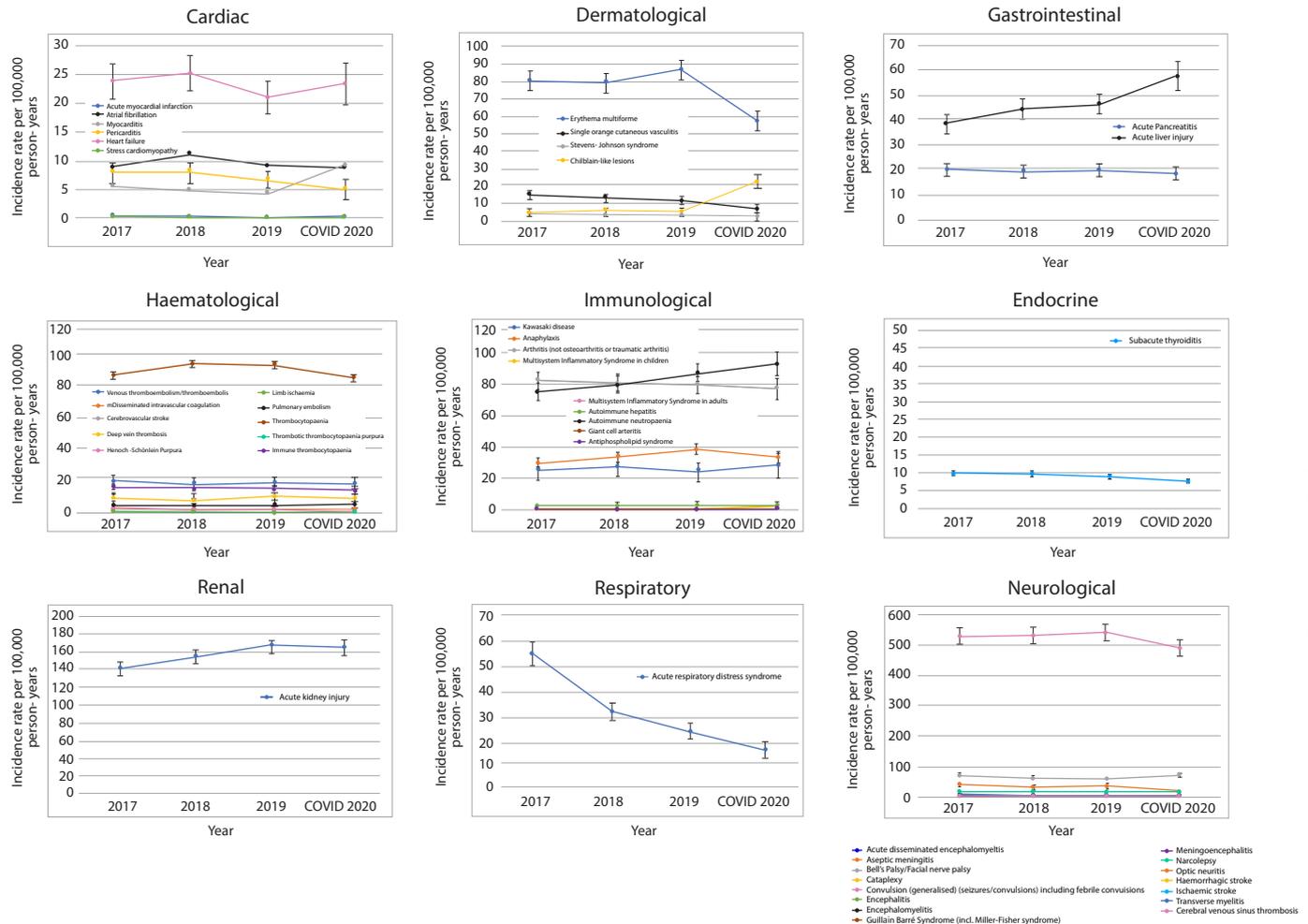
### Elderly population

In the elderly subgroup, the majority of cardiac, neurological, haematological and immunological HOIs remained stable throughout the study; however, a few differences were observed compared with the general population (figure 3). The incidence of several neurological HOIs (aseptic meningitis, Bell's palsy, convulsion and ischaemic and haemorrhagic stroke), immunological HOIs (arthritis, giant cell arteritis and antiphospholipid syndrome) and acute kidney injury were higher in the pre-COVID-19 era than the COVID-19 era. Additionally, the incidence of atrial fibrillation and acute pancreatitis decreased over the study period, whereas rates of acute liver injury increased steadily from 2017 to 2019 then remained relatively stable through 2020. As expected, incidence of ARDS was highest in the COVID-19 era.

### Pregnancy population

In the pregnancy subgroup, incidence of all non-pregnancy-specific HOIs followed similar trends in general to those observed with the general population during pre-COVID-19 and COVID-19 eras (figure 4).

There was a general trend towards increased incidence of pregnancy-related HOIs from 2017 to 2018



**Figure 2** Health outcome of interest incidence stratified by year for the paediatric subgroup.

(except for spontaneous abortions and stillbirths), which then remained relatively stable through 2019 (except full term births), before subsequently decreasing in the COVID-19 era back to similar or lower incidence rates to those observed in 2017 (figure 5). The exceptions were spontaneous abortions which slightly increased annually throughout the study period and stillbirths which remained relatively unchanged.

### Influenza-vaccinated population

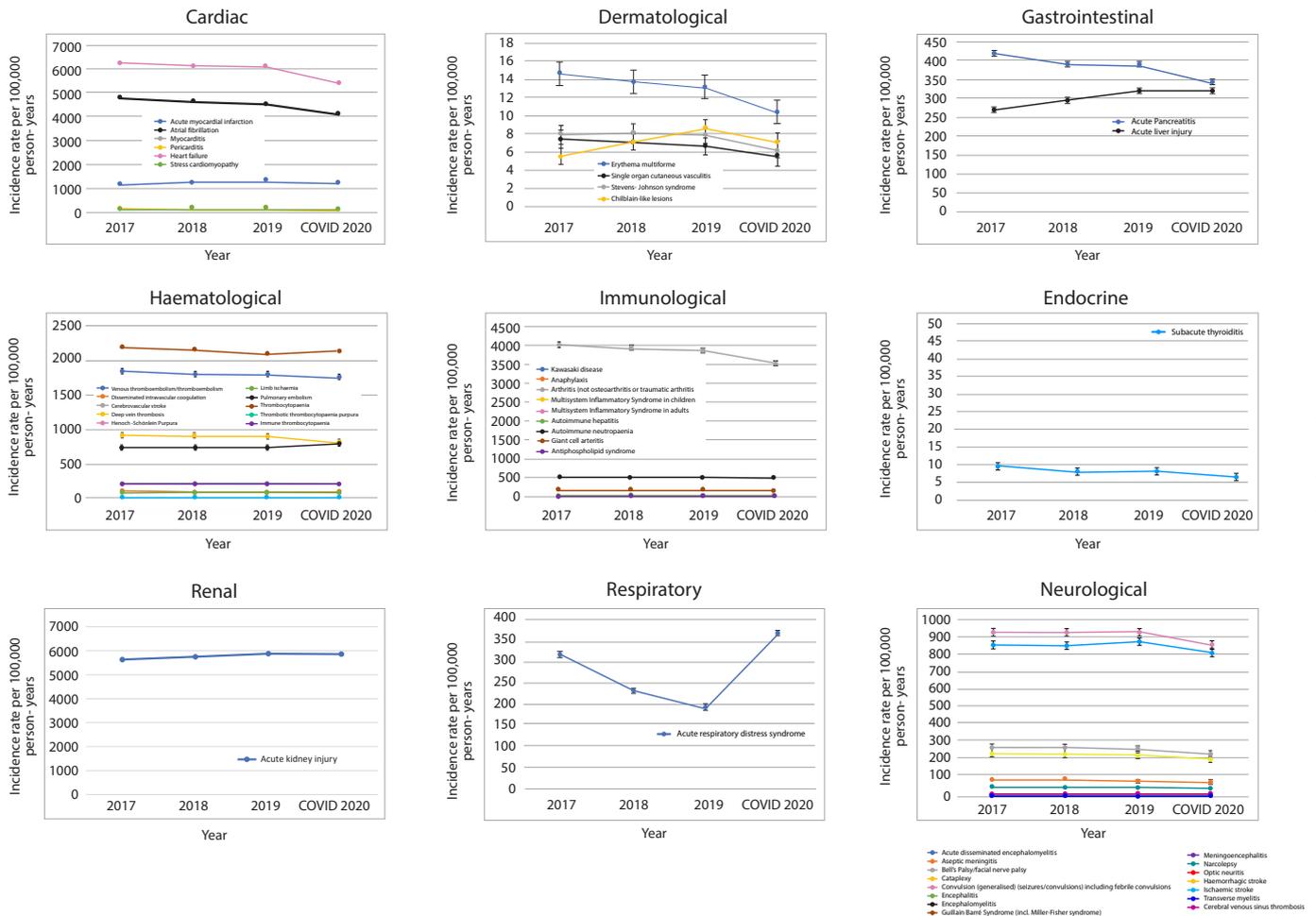
Among influenza-vaccinated patients, incidence of cardiac, dermatological and gastrointestinal HOIs, along with the incidence of acute kidney injury and ARDS were comparable with the general population, though slightly elevated (online supplemental figure S1). Incidence rates of several HOIs were higher than with the general population including heart failure, atrial fibrillation, arthritis, convulsions (except during the COVID-19 era), erythema multiforme, acute kidney injury and ARDS.

Among the haematological HOIs, the incidence rates of both thrombocytopenia and pulmonary embolism were higher in the COVID-19 era compared with the pre-COVID-19 era. The majority of neurological HOIs remained relatively stable throughout the study period;

however, incidences of convulsions decreased slightly in the COVID-19 era.

### HOIs incidence stratified by age

The age-stratified incidence rates for all HOIs observed in the general population for 2019 are shown in online supplemental figure S2. Overall, the incidence of cardiac, gastrointestinal, neurological HOIs (except for acute disseminated encephalomyelitis and encephalomyelitis), along with acute kidney injury and ARDS, increased with age. Incidence of dermatological HOIs varied between age groups, with incidence of erythema multiforme highest in patients aged 0–5 years and Stevens Johnson syndrome highest in patients aged ≥75 years. The incidence of several immunological HOIs also varied between age groups, with the incidence of anaphylaxis decreasing with age, but the incidence of arthritis, antiphospholipid syndrome, giant cell arteritis, autoimmune hepatitis and neutropaenia increasing with age. Kawasaki disease was only observed in patients aged 1–5 years and therefore no age trend could be established. Age-stratified incidence rates for all HOIs observed in the general population for 2017, 2018, 2019 and 2020 are shown in online supplemental tables S3–S6.



**Figure 3** Health outcome of interest incidence stratified by year for the elderly subgroup (age  $\geq 65$  years).

Similar trends were observed in the influenza-vaccinated (online supplemental figure S3 and tables S7–S10) and pregnancy subgroup (online supplementary Figure S4, and S5 (pregnancy subgroup) and tables S11–S14).

### HOIs incidence stratified by sex

The sex-stratified incidence rates for all HOIs observed in the general population for 2019 are shown in online supplemental table S15. The majority of HOIs were comparable across sexes; however, a few sex-specific trends were observed. The majority of cardiac, gastrointestinal, neurological and haematological HOIs, along with acute kidney injury and ARDS, were more commonly reported in males than females. Among immunological HOIs, incidence of arthritis was consistently higher among males, whereas incidence of autoimmune hepatitis, autoimmune neutropaenia, subacute thyroiditis and giant cell arteritis were consistently higher among females. Notably, incidence of Kawasaki disease was generally higher in males than females. For dermatological HOIs, chilblain-like lesions occurred at higher rates in females than males.

Similar trends were observed in the influenza vaccinated, paediatric and elderly subgroups in 2019, with a few exceptions. In the paediatric subgroup, the incidence of arthritis was lower among males and incidence

of autoimmune neutropaenia was lower among females (online supplemental table S15). In the elderly subgroup, incidence of autoimmune neutropaenia was comparable across sexes and the incidence of ARDS was higher in males than females (online supplemental table S15).

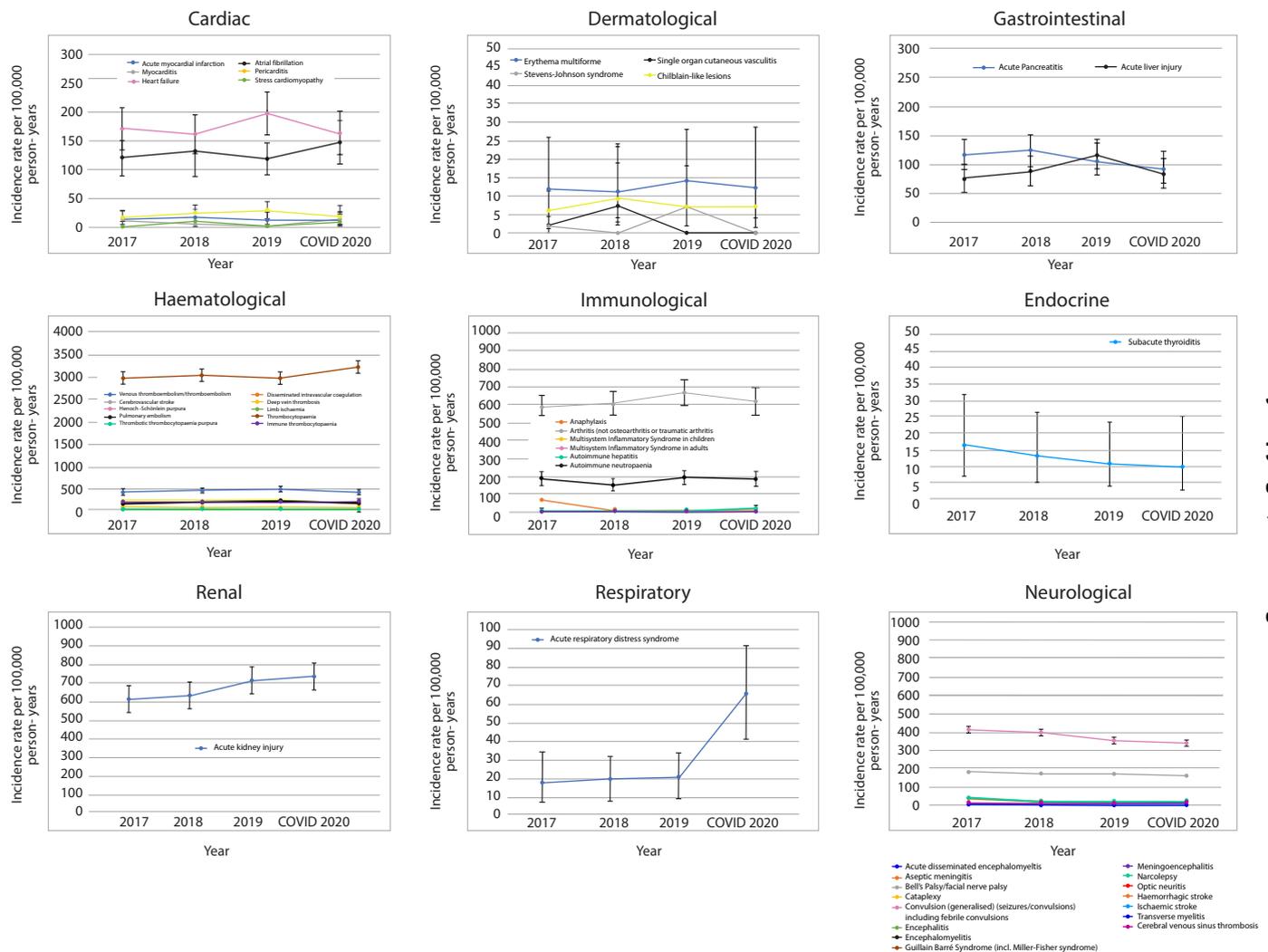
Sex-stratified incidence rates for all HOIs observed in the general population and all subgroups for 2017, 2018 and 2020 are shown in online supplemental tables S16–S18. Additionally, dual-stratified (sex and age) incidence rates for some HOIs observed in the general population for 2017–2020 are shown in online supplemental tables S19–S22.

### HOIs incidence stratified by seasonality

Across all populations analysed, trends in annual incidence of HOIs by calendar year were consistent throughout seasons; however, the incidence of all HOIs was higher between January and March than other time periods, particularly for cardiac conditions (online supplemental tables S23–S27).

### DISCUSSION

This US-based retrospective cohort study estimated the yearly background incidence rates of 59 HOIs, such as

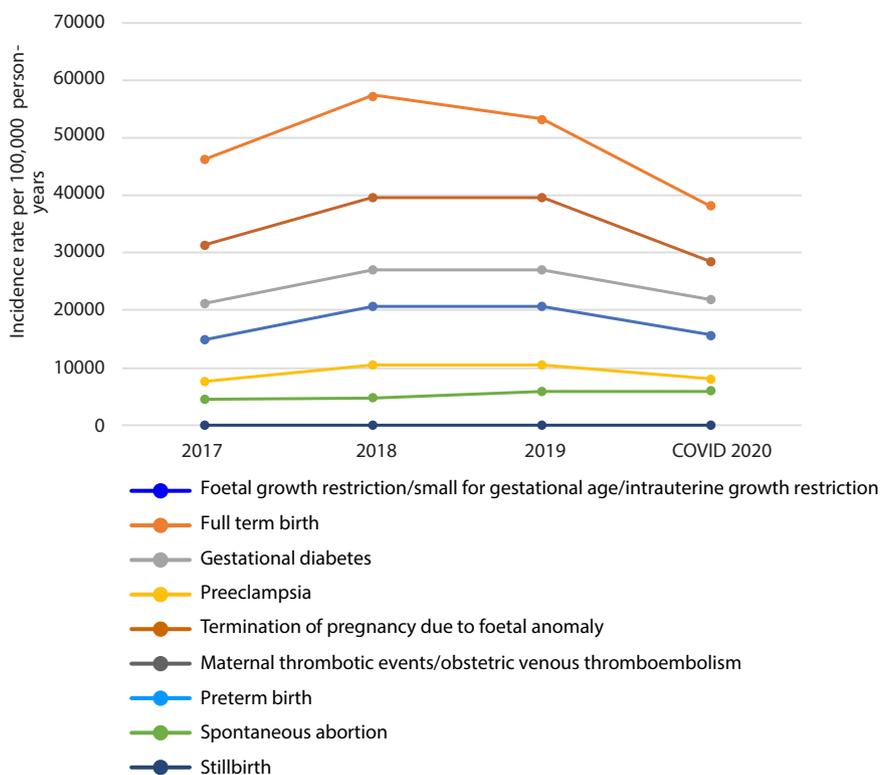


**Figure 4** Health outcome of interest incidence stratified by year for the pregnancy subgroup.

thrombotic thrombocytopenic syndrome, and multi-system inflammatory syndrome in children and adults, in roughly 16 million patients per calendar year, during two different time periods relative to the COVID-19 pandemic. Overall, variations in the incidence rates of HOIs by age and sex strata and among four key subgroups of interest in the pre-COVID-19 versus COVID-19 era were noted. As expected, higher rates of cardiac conditions, including acute myocardial infarction and myocarditis, and coagulation disorders, such as VTE, DIC, pulmonary embolism and thrombocytopenia, were noted during the COVID-19 era compared with the pre-COVID-19 era. These findings are in line with previous literature and are likely directly related to COVID-19 infection, which induces a prothrombotic state.<sup>24</sup> The higher incidence of all health outcomes between January and March, particularly for cardiac conditions reflects the seasonal variability observed during the winter months.<sup>25</sup> Similarly, incidence of ARDS, a common complication of COVID-19 infection, nearly doubled during the COVID-19 era.<sup>26</sup> Furthermore, incidence of acute liver injury was found to be highest during the COVID-19 era, which may be, in part, related to the increased prevalence of substance abuse

during the pandemic<sup>27</sup>; however, increased rates of acute liver injury could also be directly attributed to COVID-19 infection.<sup>28 29</sup> In contrast, the incidence of some dermatological and neurological conditions, acute pancreatitis and subacute thyroiditis, decreased slightly during the COVID-19 era. A similar pattern was also observed among the elderly subgroup and the influenza-vaccinated subgroup, though rates were generally higher compared with the general population. During this period, before the introduction of COVID-19 vaccines, there were a number of lockdowns, which likely had an impact on population health through disruption of health services and changes in health-seeking behaviours, particularly for certain subgroups who were classified as being high risk, which may reflect a shift in care-seeking practices during the period when COVID-19 was highly prevalent.<sup>3 30 31</sup>

The higher rates observed for several HOIs such as heart failure, atrial fibrillation, arthritis, convulsions (except during the COVID-19 era), erythema multiforme, acute kidney injury and ARDS in the influenza-vaccinated patient subgroup compared with the general population may be attributable to recommendations to target individuals at high risk of these events (ie, those with



**Figure 5** Pregnancy-specific health outcomes of interest stratified by year for the pregnancy subgroup.

comorbidities and underlying conditions) for influenza vaccination.<sup>32</sup> Of note, accumulating evidence suggests that influenza vaccination decreases the risk of cardiovascular and respiratory adverse outcomes (or exacerbation of underlying conditions) and related mortality in the general population as well as those with pre-existing conditions.<sup>33–38</sup> Among pregnant women, the incidence of many non-pregnancy-related HOIs were similar to the general population. Most pregnancy-related HOIs remained low during the COVID-19 era compared with prior years, possibly due to changes in health-seeking behaviours. For example, the incidence of acute pancreatitis, deep vein thrombosis and aseptic meningitis were lowest during the COVID-19 era. Additionally, the incidence of full-term birth and termination of pregnancy due to fetal anomaly were lowest during the COVID-19 era, which may reflect fewer planned pregnancies due to fear of COVID-19. Interestingly, there was a decline in the rates of pregnancy and births during 2020 as reported in the CDC's National Vital Statistics System.<sup>39</sup> Overall, the results from this study align with the findings from previous literature examining rates of select COVID-19 HOIs between 2017 and 2020, such as the ACCESS<sup>8</sup> and OHDSI projects<sup>4</sup>; however, some variations were noted. The differences observed in this study compared with other similar studies could be due in part to variations in study design, data sources and geographical regions. Further, it is important to note that background rates of HOIs in this study were assessed on an active user population (ie, a population that may be more likely to

seek medical care and thus be more similar with respect to healthcare-seeking behaviours to those who would elect to get a COVID-19 vaccination). In order to be well prepared for potential safety signals, we elected to examine a large number of HOIs, including the AEs of special interest suggested by SPEAC.<sup>3 4 8</sup>

This study had several strengths that warrant discussion. Our study assessed 59 HOIs compared with 15, 17 and 41 in the OHDSI Atlas, FDA Best initiatives and ACCESS studies, respectively.<sup>3 4 8</sup> Incidence rates of several key HOIs were dual stratified by age and sex, compared with these other studies which only assessed by age and sex, separately.<sup>3 4 8</sup> In this large population, the dual stratification allowed for more precise estimation of rates, particularly for those HOIs affected by age and sex, such as myocarditis. To our knowledge, this study examined the widest range of HOIs in the general population that could be relevant for safety monitoring of COVID-19 vaccines, and future vaccines, and included an important subgroup of interest not previously assessed, pregnant women. Moreover, this study examined a subgroup of patients who received an influenza vaccine, allowing for better understanding of healthcare-seeking behaviour among those more willing to use preventative medicine, and may help inform future vaccination programmes.

There are a few limitations in this study. The study population, consisting of only health insurance-covered patients and elderly Medicare recipients who actively used the healthcare system, may not fully represent the entire US population as they may have more comorbidities than

those without insurance coverage or those who do not actively engage with the healthcare system. Conversely, it could be a ‘healthier’ population who may be more likely to take preventative health measures. However, due to the detailed information on interactions for millions of patients and a similar age and sex distribution to the general US population, we did not consider this to greatly impact the results.<sup>40–43</sup> The data source used in this study also had inherent limitations that could lead to either overestimation or underestimation of incidence. For example, administrative claims data carries the potential for misclassification bias. As claims data are primarily collected for reimbursement purposes rather than research, they are prone to incomplete or inaccurate coding of diagnoses, potentially leading to misclassification or under-reporting of AEs/outcomes of interest, and general misclassification of exposure. Furthermore, the claims-based algorithms used to identify events may underestimate incidence due to their low sensitivity/positive predictive value. Conversely, the use of algorithms from prior research, which lack validation, could have potentially led to overestimation or underestimation of incidence rates. However, the use of broad definitions for outcomes should allow rare events to be captured.

Our findings also need to be considered in the context of changes in healthcare-seeking behaviour during the COVID-19 era.<sup>44 45</sup> Given the changes in healthcare practices and usage during the COVID-19 era, there is a need to exercise caution when using data from 2020.

## Conclusions

Accurate background rates for HOIs are essential for contextualisation while monitoring real-world safety of COVID-19 vaccines and other vaccines. The findings of this study are generally comparable with other similar studies assessing incidence rates of HOIs pre-COVID-19 and during the COVID-19 era using US-based and international data, such as OHDSI, ACCESS and the BEST initiative. However, this study assessed a wide range of HOIs (59) and included pregnant women, a subgroup of interest not previously assessed. Overall, the background rates of HOIs in this study provide further insight and should help to better contextualise the fluctuations in background rates of AEs related to changes in societal practices, healthcare practices and healthcare utilisation in response to the global COVID-19 pandemic, and how they may impact vaccine safety surveillance and/or clinical safety research of COVID-19 vaccines and other vaccines.

**Acknowledgements** Medical writing and editorial assistance were provided by Nichola Cruickshanks PhD, inScience Communications, Springer Healthcare Ltd, UK. Funding for this assistance was provided by Sanofi.

**Contributors** Concept and/or design of the study: SB, AK, LS, A-LC, CP, IE, SS, HK. Acquisition of data: the datasets used and/or analysed during the current study were obtained using standard contracts and data use agreements. Analysis of data: IE, SS, HK. Data interpretation: SB, AK, LS, A-LC, CP, IE, SS, HK. SB, guarantor.

**Funding** This study was funded by Sanofi. Sanofi participated in the conceptualization of the study and had a role in the statistical analyses, writing of this report, and in the decision to publish.

**Competing interests** SB, AK, LS, A-LC and CP are employees of Sanofi and may hold shares and/or stock options in the company. IE, SS and HK are employees of Aetion and hold stock options in the company.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this research.

**Patient consent for publication** Not applicable.

**Ethics approval** The requirement for ethics approval for this research was waived by the Western Institutional Review Board-Copernicus Group. This study involved de-identified data that cannot be linked back to specific individuals in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** No data are available. The datasets used and/or analysed during the current study were obtained using standard contracts and data use agreements. The deidentified Clinformatics Data Mart (CDM) administrative claims dataset are proprietary to Optum and, therefore, cannot be broadly disclosed or made publicly available at this time. The disclosure of these data to third parties would require a data use agreement with Optum.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## ORCID iD

Sonja Banga <http://orcid.org/0000-0003-3581-9318>

## REFERENCES

- World Health Organization. WHO Coronavirus (COVID-19) Dashboard, 2023. Available: <https://covid19.who.int>
- Black SB, Law B, Chen RT, *et al*. The critical role of background rates of possible adverse events in the assessment of COVID-19 vaccine safety. *Vaccine* 2021;39:2712–8.
- Moll K, Lufkin B, Fingar KR, *et al*. Background rates of adverse events of special interest for COVID-19 vaccine safety monitoring in the United States, 2019–2020. *Vaccine* 2023;41:333–53.
- Li X, Ostropolets A, Makadia R, *et al*. Characterising the background incidence rates of adverse events of special interest for COVID-19 vaccines in eight countries: multinational network cohort study. *BMJ* 2021;373:n1435.
- Caster O, Juhlin K, Watson S, *et al*. Improved statistical signal detection in Pharmacovigilance by combining multiple strength-of-evidence aspects in vigiRank. *Drug Saf* 2014;37:617–28.
- EMA. Guideline on good pharmacovigilance practices (GVP), 2013. Available: [https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-product-population-specific-considerations-i-vaccines\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-product-population-specific-considerations-i-vaccines_en.pdf)
- Mahaux O, Bauchau V, Van Holle L. Pharmacoepidemiological considerations in observed-to-expected analyses for vaccines. *Pharmacoepidemiol Drug Saf* 2016;25:215–22.
- Willame C, Dodd C, Durán CE, *et al*. Background rates of 41 adverse events of special interest for COVID-19 vaccines in 10 European Healthcare databases - an ACCESS cohort study. *Vaccine* 2023;41:251–62.
- Voss EA, Shoaibi A, Yin Hui Lai L, *et al*. Contextualising adverse events of special interest to Characterise the baseline incidence rates in 24 million patients with COVID-19 across 26 databases: a multinational retrospective cohort study. *EClinicalMedicine* 2023;58:101932.
- Ingrasciotta Y, Jin Y, Foti SS, *et al*. Real-world patient characteristics and use of disease-modifying anti-rheumatic drugs in patients



- with rheumatoid arthritis: a cross-national study. *Clin Rheumatol* 2023;42:1047–59.
- 11 Indiana University Bloomington. Optum Clinformatics data Mart. 2023. Available: <https://ssrc.indiana.edu/data/optum.html>
  - 12 Optum. Clinformatics data Mart (CDM). 2017. Available: [https://www.optum.com/content/dam/optum/resources/productSheets/Clinformatics\\_for\\_Data\\_Mart.pdf](https://www.optum.com/content/dam/optum/resources/productSheets/Clinformatics_for_Data_Mart.pdf)
  - 13 Lieu TA, Nguyen MD, Ball R, et al. Health outcomes of interest for evaluation in the post-Licensure rapid immunization safety monitoring program. *Vaccine* 2012;30:2824–30.
  - 14 Riley M, Howard J, Dale K, et al. Validating Notifications of pregnancy Terminations for birth defects before 20 weeks gestation. *Health Inf Manag* 2002;30:1–8.
  - 15 Hwang JJ, Shin SH, Kim YJ, et al. Epidemiology of Antiphospholipid syndrome in Korea: a nationwide population-based study. *J Korean Med Sci* 2020;35:e35.
  - 16 Ostropolets A, Li X, Makadia R, et al. Factors influencing background incidence rate calculation: systematic empirical evaluation across an international network of observational databases. *Front Pharmacol* 2022;13:814198.
  - 17 Kawai AT. Mini-sentinel/CBER protocol vaccines administered in pregnancy and birth outcomes: A project to develop data infrastructure and design framework, 2017. Available: <https://sentinelinitiative.org/sites/default/files/vaccines-blood-biologics/assessments/Influenza-Vaccines-and-Birth-Outcomes-Revised-Protocol.pdf>
  - 18 FDA-Sentinel. COVID-19 History Master Protocol, 2021. Available: [https://www.sentinelinitiative.org/sites/default/files/Methods/COVID-19-Natural\\_History\\_Protocol\\_v3.0.pdf](https://www.sentinelinitiative.org/sites/default/files/Methods/COVID-19-Natural_History_Protocol_v3.0.pdf)
  - 19 Gubernot D, Jazwa A, Niu M, et al. Population-based background incidence rates of medical conditions for use in safety assessment of COVID-19 vaccines. *Vaccine* 2021;39:3666–77.
  - 20 Jinjuvadia K, Kwan W, Fontana RJ. Searching for a needle in a haystack: use of ICD-9-CM codes in drug-induced liver injury. *Am J Gastroenterol* 2007;102:2437–43.
  - 21 Pollissard L, Shah A, Puneekar RS, et al. Burden of illness among Medicare and non-Medicare US populations with acquired thrombotic Thrombocytopenic purpura. *J Med Econ* 2021;24:706–16.
  - 22 Fung KW, Richesson R, Smerek M, et al. Preparing for the ICD-10-CM transition: automated methods for translating ICD codes in clinical phenotype definitions. *EGEMS (Wash DC)* 2016;4:1211.
  - 23 Andrade SE, Harrold LR, Tjia J, et al. A systematic review of validated methods for identifying cerebrovascular accident or transient ischemic attack using administrative data. *Pharmacoepidemiol Drug Saf* 2012;21 Suppl 1:100–28.
  - 24 García-Ortega A, de la Rosa D, Oscullo G, et al. Coagulation disorders and thromboembolic disease in COVID-19: review of current evidence in search of a better approach. *J Thorac Dis* 2021;13:1239–55.
  - 25 Seasonal variability of disease. *JAMA* 2023;329:598.
  - 26 Hendrickson KW, Peltan ID, Brown SM. The epidemiology of acute respiratory distress syndrome before and after Coronavirus disease 2019. *Crit Care Clin* 2021;37:703–16.
  - 27 Marano G, Traversi G, Gaetani E, et al. Alcohol use disorder and liver injury related to the COVID-19 pandemic. *World J Hepatol* 2022;14:1875–83.
  - 28 Phippos MM, Barraza LH, LaSota ED, et al. Acute liver injury in COVID-19: prevalence and association with clinical outcomes in a large U.S. *Hepatology* 2020;72:807–17.
  - 29 Iheanacho CO, Enechukwu OH. COVID-19-associated liver injury, role of drug therapy and management: a review. *Egypt Liver J* 2022;12:66.
  - 30 Whaley CM, Pera MF, Cantor J, et al. Changes in health services use among commercially insured US populations during the COVID-19 pandemic. *JAMA Netw Open* 2020;3:e2024984.
  - 31 Moynihan R, Sanders S, Michaleff ZA, et al. Impact of COVID-19 pandemic on utilisation of Healthcare services: a systematic review. *BMJ Open* 2021;11:e045343.
  - 32 World Health Organization. Vaccines against influenza: WHO position paper – may 2022. *Wkly Epidemiol Rec* 2022;97:185–208.
  - 33 Cheng Y, Cao X, Cao Z, et al. Effects of influenza vaccination on the risk of cardiovascular and respiratory diseases and all-cause mortality. *Ageing Res Rev* 2020;62:S1568-1637(20)30259-2.
  - 34 Ngwudike CJ, Villalobos A. Correlation between cardiovascular protection and influenza vaccination. *Curr Cardiol Rep* 2023;25:571–6.
  - 35 Clar C, Oseni Z, Flowers N, et al. Influenza vaccines for preventing cardiovascular disease. *Cochrane Database Syst Rev* 2015;2015:CD005050.
  - 36 Vasileiou E, Sheikh A, Butler C, et al. Effectiveness of influenza vaccines in asthma: A systematic review and meta-analysis. *Clin Infect Dis* 2017;65:1388–95.
  - 37 Kopsaftis Z, Wood-Baker R, Poole P. n.d. Influenza vaccine for chronic obstructive pulmonary disease (COPD). *Cochrane Database Syst Rev* 2018;CD002733.
  - 38 Buchy P, Badur S. Who and when to vaccinate against influenza. *Int J Infect Dis* 2020;93:375–87.
  - 39 National Vital Statistics System. Vital Statistics rapid release. 2021. Available: <https://www.cdc.gov/nchs/data/vsrr/vsrr012-508.pdf>
  - 40 United States Census Bureau. Age and Sex Composition in the United States, 2018. Available: <https://www.census.gov/data/tables/2018/demo/age-and-sex/2018-age-sex-composition.html>
  - 41 United States Census Bureau. Age and Sex Composition in the United States, 2017. Available: <https://www.census.gov/data/tables/2017/demo/age-and-sex/2017-age-sex-composition.html>
  - 42 United States Census Bureau. Age and Sex Composition in the United States, 2019. Available: <https://www.census.gov/data/tables/2019/demo/age-and-sex/2019-age-sex-composition.html>
  - 43 United States Census Bureau. Age and Sex Composition in the United States, 2020. Available: <https://www.census.gov/data/tables/2020/demo/age-and-sex/2020-age-sex-composition.html>
  - 44 Vander Woude CA, King EJ, Hirschtick JL, et al. Differential care-seeking behaviors during the beginning of the covid-19 pandemic in michigan: a population-based cross-sectional study. *BMC Public Health* 2023;23:2101.
  - 45 Adjemian J, Hartnett KP, Kite-Powell A, et al. Update: COVID-19 pandemic-associated changes in emergency Department visits - United States. *MMWR Morb Mortal Wkly Rep* 2020;70:552–6.