University of Dundee

Lessons Learned on Observed-to-Expected Analysis Using Spontaneous Reports During Mass Vaccination

Gordillo-Marañón, María; Candore, Gianmario; Hedenmalm, Karin; Browne, Kate; Flynn, Robert; Piccolo, Loris

[Link to publication in Discovery Research Portal](https://...)

CURRENT OPINION

Lessons Learned on Observed-to-Expected Analysis Using Spontaneous Reports During Mass Vaccination

María Gordillo-Marañón[1,2] · Gianmario Candore[3] · Karin Hedenmalm[1] · Kate Browne[4] · Robert Flynn[1,5] · Loris Piccolo[1] · Aniello Santoro[6] · Cosimo Zaccaria[6] · Xavier Kurz[1]

Abstract

During the COVID-19 vaccination campaign, observed-to-expected analysis was used by the European Medicines Agency to contextualise data from spontaneous reports to generate real-time evidence on emerging safety concerns that may impact the benefit-risk profile of COVID-19 vaccines. Observed-to-expected analysis compares the number of cases spontaneously reported for an event of interest after vaccination ('observed') to the 'expected' number of cases anticipated to occur in the same number of individuals had they not been vaccinated. Observed-to-expected analysis is a robust methodology that relies on several assumptions that have been described in regulatory guidelines and scientific literature. The use of observed-to-expected analysis to support the safety monitoring of COVID-19 vaccines has provided valuable insights and lessons on its design and interpretability, which could prove to be beneficial in future analyses. When undertaking an observed-to-expected analysis within the context of safety monitoring, several aspects need attention. In particular, we emphasise the importance of stratified and harmonised data collection both for vaccine exposure and spontaneous reporting data, the need for alignment between coding dictionaries and the crucial role of accurate background incidence rates for adverse events of special interest. While these considerations and recommendations were determined in the context of the COVID-19 mass vaccination setting, they are generalisable in principle.

Key Points

Observed-to-expected analysis has proven beneficial in terms of generating rapid evidence during the early phases of the COVID-19 mass vaccination campaign.

Several assumptions and important aspects need to be taken into account when considering observed-to-expected analysis for the purpose of signal detection and validation.

✉ María Gordillo-Marañón
maria.maranon.16@ucl.ac.uk

1 Data Analytics and Methods Task Force, European Medicines Agency, Amsterdam, The Netherlands

2 Institute of Cardiovascular Science, Faculty of Population Health, University College London, London, UK

3 Medical Affairs and Pharmacovigilance, Pharmaceuticals, Bayer AG, Berlin, Germany

4 Stakeholders and Communication, European Medicines Agency, Amsterdam, The Netherlands

5 Medicines Monitoring Unit, University of Dundee, Dundee, UK

6 Pharmacovigilance Office, European Medicines Agency, Amsterdam, The Netherlands

1 Introduction

Safety monitoring during mass vaccination campaigns poses specific challenges, mainly due to a rapid influx of spontaneous reports within a short timeframe. Between 2021 and 2022, the European Medicines Agency (EMA) processed over 2.8 million safety reports for COVID-19 vaccines in addition to 3.6 million safety reports related to other medicinal products in the EudraVigilance (EV) database, the system for collecting, managing and analysing suspected adverse reactions to medicines authorised in the European Economic Area (EEA) [1]. These corresponded to almost 4000 reports for COVID-19 vaccines per day [1]. Besides the immense volume of data to be processed and evaluated in an unprecedented short timeframe, causality assessment was also complicated by the occurrence of coincidental adverse events in close temporal association with vaccination. Therefore, quantitative pharmacovigilance tools that contextualise data from safety reports were needed to generate real-time evidence on emerging safety signals that might have impacted the benefit-risk profile of newly authorised vaccines [2]. These tools included an observed-to-expected (O/E) analysis which compares the observed

number of cases to the expected number of cases anticipated to occur in the same number of individuals had they not been vaccinated.

The observed number of cases could be derived from different sources, including observational data such as electronic health records; however, these data have limitations which may preclude the prompt detection and characterisation of rare adverse reactions. These limitations include a time lag in terms of data availability, insufficient statistical power to detect associations with rare events and, in some healthcare databases, lack of data concerning vaccine exposure and representativeness of the study population. During the initial phases of a mass vaccination programme, spontaneous reporting systems (SRS) seem to offer advantages to promptly identify and evaluate emerging safety concerns, particularly for rare events.

The key requirements and statistical methodology for O/E analysis have been described in European regulatory guidelines [3, 4]. The main principles, assumptions and considerations have been discussed by Mahaux et al. [5]. Despite the challenges and limitations inherent to the methodology that are detailed below, O/E analysis played a vital role in the early stages of the COVID-19 vaccination campaign, before other data sources were available. In this article, we discuss updated pharmacoepidemiological considerations for O/E analysis based on our experience from the safety monitoring of COVID-19 vaccines. Within those analyses, the 'observed' is determined from SRS while the 'expected' is ascertained based on historical background rates derived from healthcare databases. Note that other methodologies used for signal detection (e.g. disproportionality methods [6]) are out of scope and are not discussed in the following sections. Whilst aspects described in this paper are specific to the context of mass vaccination, the lessons learned and recommendations are in principle generalisable, although other considerations may be needed for other medicinal products.

2 O/E Analysis Performed by EMA

Observed-to-expected analysis became an established tool used by EMA to support the safety monitoring of COVID-19 vaccines [2], facilitating further contextualisation of spontaneous safety data to support regulatory decision making and enriching the risk communication to public health bodies, patients and healthcare professionals. O/E analysis supports safety assessments in two specific contexts:

- Signal detection: facilitating the prioritisation of adverse events of special interest (AESIs) in the routine review phase. O/E analysis was used to monitor certain

AESIs proposed by the Brighton Collaboration within the Safety Platform for Emergency vACcines (SPEAC) [7], such as Guillain–Barré syndrome and transverse myelitis.
- Signal validation: generating evidence while investigating a potential safety signal. O/E analysis also contributed to the assessment of safety signals, such as thrombosis with thrombocytopenia syndrome (TTS) associated with Vaxzevria (previously COVID-19 Vaccine AstraZeneca) [8], and myocarditis/pericarditis with COVID-19 mRNA-based vaccines [9–11].

To perform O/E analysis, the following key elements were defined by EMA:

2.1 Selection of the Risk Period

The risk period is considered as the time period after the vaccination during which there is clinical and biological plausibility for the adverse event to occur. Risk periods for individual adverse events were determined based on published case definitions, biological plausibility or the time to onset (TTO) distributions of the cases reported to EV following COVID-19 vaccination. As detailed in the next sections, the risk period is important because it is used to estimate both the 'observed' and the 'expected' number of cases.

2.2 Estimation of the Observed Number of Cases

The methodology used by EMA to estimate the number of 'observed' cases was based on the number of individuals in the EEA who experienced a specific AESI or other adverse event following immunisation, referred hereafter as simply an 'adverse event', and were reported to EV. The set of 'observed' cases included reports retrieved from EV using dedicated Medical Dictionary for Regulatory Affairs (MedDRA) search criteria for specific adverse event(s) that occurred within the risk period.

2.3 Estimation of the Expected Number of Cases

The number of 'expected' cases was estimated based on EEA vaccine exposure data from the European Centre for Disease Prevention and Control (ECDC) website [12], the risk period and incidence rates of the adverse event estimated from populations prior to COVID-19 disease (i.e. background incidence rates), as described by Mahaux et al. [5].

ECDC published updated national exposure data, stratified by vaccine brand, dose series and specific age groups, on a weekly basis. Data were further processed by EMA to represent total vaccine doses by vaccine brand, sex and age groups of interest, using age and sex distribution data requested to Member States at different points during the COVID-19 vaccination campaign. The cut-off date for the data on COVID-19 vaccination was set to 2 weeks before the date used for the extraction of the observed cases to allow for the adverse event to be reported to EV. The total person-time at risk was then calculated using the processed exposure data and the risk period as described by Mahaux et al. [5], taking into account vaccinations with multiple doses and whether the risk periods overlap.

Background incidence rates for adverse events from the ACCESS project (vACCine covid-19 monitoring readinESS) [13, 14], calculated using longitudinal healthcare databases, were primarily used by EMA to conduct O/E analysis. For adverse events that were not available via the ACCESS project or requiring further refinement (e.g. adjustments of event definitions), background rates were sourced from other EMA-funded studies (e.g. thromboembolic events [15, 16]), or ascertained from healthcare databases directly accessible to EMA [17].

2.4 Sensitivity Analyses

Sensitivity analyses were performed for O/E analysis to adjust for potential under-reporting to EV, incomplete data within spontaneous reports regarding TTO of specific adverse events by using different risk periods and uncertainty around the risk periods. To account for regional heterogeneity between the incidence of specific adverse events, a sensitivity analysis was also performed using different healthcare data sources. Sensitivity analyses were used to consider uncertainty around case definitions (see Sect. 3.3).

3 Considerations on Feasibility, Design and Interpretation of O/E Analysis

As part of EMA's safety monitoring for COVID-19 vaccines, O/E analysis complemented data from other sources, such as clinical, biological and pathophysiological information, and was treated as an initial approach to facilitate prioritisation and to generate additional evidence while investigating a potential safety signal. We outline below some considerations on assessing the feasibility of the analyses and on the design and interpretation of the results based on the experience during the COVID-19 vaccination campaign, which are also summarised in Table 1.

3.1 Consideration 1: Types of Adverse Events

O/E analysis using SRS and historical background incidence rates from healthcare databases is more appropriate for certain types of adverse events depending on their characteristics and the feasibility of ascertaining a reliable background incidence rate. For example, some events are known vaccination-induced adverse events for which a causal association can more readily be established based on a review of the reported cases, such as adverse events indicative of reactogenicity. O/E analysis is less useful for such events. The type of adverse event, how it is reported to SRS and how the event is captured in healthcare databases informs the decision to perform O/E analysis as described below.

3.1.1 Reporting of the Event to SRS

Certain adverse events occurring in vaccinated patients, particularly those that are rare, acute, serious or associated with vaccines in general, for example, Guillain–Barre syndrome, will most often be reported to EV [18–20]. However, as this assumption cannot be presumed, it is still recommended to undertake sensitivity analyses to evaluate the impact of different levels of under-reporting. Alternatively, adverse events with high background incidence rates irrespective of vaccination (e.g. myocardial infarction) or with a long latency (e.g. cancer) may not be suspected by the patient or the physician as being related to the vaccine, except in certain cases when the event occurred in close temporal association with vaccination. The level of under-reporting associated with such adverse events is generally expected to be high. While a sensitivity analysis can be used to adjust the O/E analysis for under-reporting (methodology and illustrations described in [5]), there is uncertainty inherent in estimating the level of under-reporting. Interpretation of O/E analysis of these events in the context of vaccination may be difficult.

3.1.2 Representation of the Event in Healthcare Databases

Background incidence rates are often estimated from electronic health records or claims databases, which are based on healthcare encounters. In these circumstances, O/E analysis is appropriate for events that result in patients seeking healthcare and lead to a clear diagnosis that is recorded within the respective database. It is less appropriate for events reported as symptoms that are transient or for which a patient may not seek healthcare, (e.g. vomiting and dizziness) or that could be caused by a range of different conditions (e.g. vertigo). Background incidence rates are generally not readily available for certain adverse events that are

Table 1 Summary table of considerations and recommendations for O/E analysis

Consideration	Recommendation
Consideration 1 Types of adverse events suitable for O/E analysis	The usefulness of performing O/E analysis should be evaluated for vaccine-induced adverse events for which a causal association can be reliably assessed based on a review of the reported cases Sensitivity analyses adjusting for different levels of under-reporting in SRS associated with different type of adverse event should be performed The feasibility assessment of O/E analysis should evaluate if the event of interest is adequately captured in electronic health records or claims databases so background incidence rates can be calculated. Proxy background rates can be considered when data is limited
Consideration 2 Risk period for observed cases	Risk periods should be estimated based on non-clinical data, clinical guidelines and case definitions provided by the Brighton Collaboration [21] Descriptive analyses of the 'observed' cases, including the time to onset distributions, should be conducted to inform sensitivity analyses
Consideration 3 Alignment of coding dictionaries	Expert review is needed to ensure alignment between the codes used to identify the 'observed' cases from SRS and the codes used to estimate the number of 'expected' cases Sensitivity analyses using narrow and broad endpoint definitions should be used when differences are found between the coding dictionaries used to describe the events in the spontaneous reporting system and in healthcare databases To assess if a mapping between the relevant codes already exists, for example, through ongoing international standardisation initiatives
Consideration 4 Adverse events caused by the disease	When the potential adverse events associated with the vaccines may also be caused by the disease itself, the results should be interpreted cautiously and sensitivity analyses to adjust for changes in the background incidence rates should be considered
Consideration 5 Accuracy of the expected number of cases	The selection of the background incidence rate should take into account the characteristics of the database such as geographic region, coverage, latency, accessibility and availability of variables that may influence the susceptibility to an adverse event. Sensitivity analyses with different background incidence rates should be considered When possible, data sources that combine primary care data with specialist and hospital data should be considered over unlinked databases as they are likely to be more representative
Consideration 6 Heterologous vaccinations, multiple dose schedules and virus variants	For vaccinations with multiple doses, the time at risk should be adjusted if the time between the administration of the doses is known across the population Data on vaccine exposure and reported to SRS should be as granular and harmonised as possible to allow for stratified analyses, for example by dose number (e.g. first vs second) Adapted versions of already authorised vaccines should be considered independent vaccines, and O/E analysis performed for either the original or the adapted vaccine, if possible

O/E observed-to-expected, *SRS* spontaneous reporting systems

extremely rare or represent new clinical entities (e.g. TTS). In this situation, proxy background rates (i.e. background rates from other categories considered to present the same risk of occurrence and similar aetiology) could be used. For example, background incidence rates of cerebral venous sinus thrombosis were used to conduct O/E analysis that facilitated evaluation of the signal of TTS [8].

3.2 Consideration 2: Definition of the Risk Window

The risk period is considered as the time period after the vaccination during which there is clinical and biological plausibility for the adverse event to occur. This can be estimated based on non-clinical data, clinical guidelines and case definitions provided by the Brighton Collaboration [21]. For example, cases of Guillain–Barre syndrome following vaccination are generally expected to occur within 2–42 days [22]. Reported cases of this adverse event with a longer TTO should therefore be carefully ascertained before

inclusion in the 'observed'. In some circumstances, the TTO of a vaccine-induced event may differ from the TTO for an adverse event due to other biological causes [23]. For example, a study by Nguyen et al. [24] showed that the TTO for vaccine-associated myocarditis was significantly shorter (10 days, 95% confidence interval 6–12 days) than the TTO for immunotherapy myocarditis (33 days, 95% confidence interval 20–88 days). One approach to take this into consideration is to examine the distribution of the TTO of the reported cases for the event of interest and evaluate if the majority of the cases fall within the expected TTO. Then, a sensitivity analysis using a risk window during which the majority of the events with information about the TTO have been reported can be defined. This descriptive approach prior to the O/E analysis has also proven useful to characterise adverse events that are defined by a broad MedDRA code (e.g. Standardised MedDRA Queries [25]) encompassing multiple clinical entities, each of which may have a different TTO. By exploring the distribution of the TTO, one can

infer which of the events is likely to be the one reported for a suspected adverse drug reaction. However, it is important to exercise caution for events with a variable or a latent TTO and therefore, the use of these types of analyses should be limited to inform sensitivity analyses.

3.3 Consideration 3: Alignment of Coding Dictionaries

A significant limitation of O/E analysis conducted using SRS and longitudinal healthcare databases is the potential for misalignment because of differences in respective coding dictionaries used to ascertain both the 'expected' and 'observed'. For instance, the algorithms we used to ascertain data from SRS were based on medical dictionaries for regulatory activities (e.g. MedDRA) while the strategies to ascertain background incidence rates were based on clinical vocabularies for use in healthcare databases (e.g. International Classification of Diseases, Tenth Revision [26], SNOMED CT [27]).

Given that there are differences between coding dictionaries, alignment of case definitions for adverse events between respective dictionaries is essential to ensure the validity of the comparison between 'expected' and 'observed'. The codes that best describe a clinical entity may be too generic or too broad in scope and therefore not aligned with the codes used to capture the adverse event within SRS. In such cases, the use of both narrow (i.e. the list of codes includes only terms specific to the adverse event) and broad (i.e. the list includes the terms of the narrow definition and those potentially linked to the adverse event) definitions in both the calculation of the 'expected' and 'observed' cases within the context of sensitivity analyses can be informative. In addition, the set of clinical terms used to define the event should include conditions potentially caused by the vaccine and exclude terms that imply other clear aetiologies. For example, spontaneous safety reports of cytomegalovirus myocarditis or malarial myocarditis have a clearly identified aetiology other than the vaccine, hence, in general terms, they would not be considered relevant in an analysis aimed at investigating a possible association between a vaccine and myocarditis. By analogy, such terms should not be used as part of the definition of the adverse event when calculating background incidence rates for myocarditis. Furthermore, undertaking manual expert review is crucial to refine the case definition for adverse events, increase specificity and ensure alignment between the codes used to identify the 'observed' cases and the codes used to estimate the number of 'expected' cases.

When conducting O/E analysis for COVID-19 vaccines, we observed that diagnosis codes used in healthcare databases and SRS can lack specificity. For example, a potential signal for vasculitis [28], a heterogeneous group of diseases,

was evaluated by EMA in late 2021 based on literature and case review [29]. While an O/E analysis was performed to generate evidence on a possible association, it proved inconclusive because of a misalignment between coding dictionaries used to ascertain the expected and observed, which resulted in a high background incidence rate. Despite the lack of evidence from O/E analysis, there was sufficient evidence from the qualitative case review to include "small vessel vasculitis with cutaneous manifestations" in Section 4.8 of the SmPC of the Jcovden vaccine (previously COVID-19 Vaccine Janssen) [30]. Because of the lack of granularity of the International Classification of Diseases, Tenth Revision codes used to identify the expected cases, it was not possible to restrict the case definition to those patients with a form of vasculitis with similar characteristics to those of the case reports from EV.

Ongoing initiatives that aim to address the discrepancies between codes used in the clinical and pharmacovigilance setting exist. For example, the WEB-RADR2 project aims to develop mappings between SNOMED CT and MedDRA to support interoperability between the terminologies [31]. SNOMED CT is a multi-lingual clinical terminology in use globally that has been recommended by the European Health Data Space Project as a standard terminology to improve interoperability between data sources [32].

3.4 Consideration 4: Adverse Events Caused by the Disease

In the context of a pandemic mass vaccination campaign, background incidence rates for O/E analysis are calculated based on pre-pandemic data due to the practicalities associated with obtaining adequate information in a short period of time after the beginning of the pandemic and vaccine deployment. However, some adverse events associated with the vaccines may also be caused by the disease itself, such as myocarditis and COVID-19 disease [33, 34]. This should be taken into account when comparing the numbers of 'observed' and 'expected' cases, by undertaking sensitivity analyses, adjusting the case definition or using incidence rates based on the pandemic period. For example, using incidence rates calculated in the early pandemic period (e.g. 2020) when the SARS-CoV-2 infection rate was high for O/E analysis for COVID-19 vaccines.

3.5 Consideration 5: Criteria for Determining the Expected Number of Cases

The calculation of the 'expected' number of cases is based on background incidence rates of the adverse event, vaccine exposure data and the risk period. The incidence rates used to estimate the 'expected' number of cases should

be calculated using a population similar to the vaccinated population, not only in terms of age and sex but also other variables that may influence the susceptibility to an adverse event. Therefore, caution is needed when selecting the most appropriate data source for the background incidence rates, taking into account the characteristics of the database such as geographic region, coverage, latency, accessibility and availability of variables that may influence the susceptibility to an adverse event. In an ideal scenario, background incidence rates would be derived from the same setting as the 'observed'; however, at a practical level this may not be feasible, particularly during the early stages of a mass vaccination campaign. Similarly, vaccine exposure in the target population stratified by brand, dose, age and sex should be accessible and in line with the timelines used to identify the 'observed' cases. Therefore, using the most appropriate background incidence rate and vaccine exposure data is crucial; however, several considerations should be taken into account when determining the most suitable data source.

3.5.1 Relevance of the Healthcare Settings Available for the Estimation of Background Incidence Rates

Background incidence rates are estimated from electronic healthcare records from different settings, including secondary care data (specialist and hospital data), primary care data or administrative claims data. In principle, data sources that combine primary care data with secondary care data would allow greater confidence that the ascertainment of events is more complete, captures a wider population and more encounters with the healthcare system, and is therefore likely to be representative. However, the decision between background incidence rates calculated in primary versus secondary care should be guided by the type of adverse event, given that some would be expected to be diagnosed in, for example, a secondary care setting, and not well captured in primary care databases.

3.5.2 Differences Between Countries and Over Time

Background incidence rates are estimated from electronic healthcare records with country-specific clinical and coding practices. These, together with population-specific characteristics, can lead to differences between the background incidence rates, with some countries showing a higher incidence of certain adverse events than others. Several studies have shown that such differences in the underlying populations lead to substantial heterogeneity between data sources [15, 35, 36]. In addition, variations can be found within databases across different time periods. Therefore, it is recommended to estimate the background incidence rates at different timepoints and then evaluate the consistency across

time periods taking into account potential factors that may explain these variations (e.g. seasonal effects). Sensitivity analyses using different background incidence rates indicative of distinct scenarios should be considered to evaluate the presence of bias. For example, Mahaux et al. illustrate the impact on the O/E analysis conclusions using different background incidence rates [5].

3.5.3 Stratification of the Background Incidence Rates

When possible, background incidence rates should be stratified by age and sex to allow O/E analysis in certain subsets of the population. For example, O/E analysis for the adverse event of TTS performed for Vaxzevria found a greater association in young vaccinees compared to other age groups [8]. Similarly, a greater association in young male compared to female individuals or other age groups was found for myocarditis and pericarditis and the mRNA vaccines Comirnaty and Spikevax [9, 11]. Without the stratification of background incidence rates by age and sex, the O/E analysis would have not suggested the association [9]. Beyond age and sex, other variables may affect susceptibility to an event and such potentially relevant risk factors or confounders should be taken into account. For instance, mass vaccination campaigns target initially individuals at highest risk (e.g. immunocompromised individuals) followed by additional groups deemed to benefit from vaccination. Therefore, when at-risk groups are prioritised for vaccination, the calculation of background incidence rates should be tailored and not be based on the general population [3].

3.5.4 Accessibility to Vaccine Exposure Data

Ideally, the populations used to calculate both background incidence rates and data for vaccine exposure should be as similar as possible. For example, national data regarding vaccine exposure should be combined with country-specific background incidence rates to calculate the 'expected'. Those are then compared to the 'observed', which are also identified at national level. As for any scenario, when O/E analysis is performed across countries (e.g. EEA level), vaccine exposure data should be stratified following the same principles as with the background incidence rates. In addition, when the vaccination campaign involves different vaccine brands, information on the number of vaccinees by vaccine brand is required, particularly if the mechanism of action of the vaccines differs across types.

3.6 Consideration 6: Heterologous Vaccinations, Multiple Dose Schedules and Virus Variants

One of the assumptions inherent in O/E analysis is that the number of vaccinated individuals over time, ideally stratified

by age group and sex, is available. For vaccinations with multiple doses, Mahaux et al. [5] described how to calculate the cumulative time at risk, provided that the time between dose administration is known and constant across the population. Such adjustment of the time at risk is particularly relevant when the risk period is longer than the average time window between the scheduled vaccine doses. In this scenario, the risk periods overlap, i.e. the administration of the second dose has taken place when the risk period following the first administration is not yet over, thus the doses cannot be considered independently as doing so would overestimate the total time at risk by double counting [5].

To investigate if there is an effect associated with exposure to multiple vaccine doses, information by dose number is needed for both the expected and observed cases. During the COVID-19 vaccination campaign, dose information by country at EEA level was accessible through data published by ECDC, which made it possible to stratify the number of expected cases by dose. However, the reporting of the dose for cases submitted to EV was not standardised and in many cases this information was missing or not available in the structured fields of the individual case safety reports (ICSRs) submitted to the database. A standardised structured method for reporting dose information in ICSR forms should be enforced to facilitate analyses by dose number during future vaccination campaigns.

The use of heterologous vaccination within national vaccination campaigns, including 'mix and match' combinations of different COVID-19 vaccine platforms, posed another challenge in terms of conducting O/E analysis. For example, after confirmation of the risk of TTS in association with Vaxzevria (COVID-19 Vaccine AstraZeneca at the time), a proportion of individuals that received a first dose of this vaccine subsequently received an mRNA vaccine for subsequent doses. However, the safety profile of heterologous COVID-19 vaccine schedules could not be evaluated by O/E analysis given vaccine exposure data published by ECDC did not include dose data regarding heterologous vaccination and data concerning the dose series for such cases was often incomplete within ICSRs submitted to EV.

The effectiveness of vaccines is also impacted by emerging variants of circulating viruses that escape vaccine-induced immune responses, necessitating the development of adapted vaccines. In terms of safety monitoring, adapted vaccines should be considered independent vaccines. Consequently, O/E analysis should be performed for either the original or the adapted vaccine, if feasible. A lack of granularity regarding the type of vaccine within ICSRs precluded the ability to distinguish between originally authorised and adapted COVID-19 vaccines, which posed a significant limitation for O/E analysis.

4 Conclusions

The application of O/E analysis during the safety monitoring of COVID-19 vaccines has demonstrated its usefulness to generate rapid evidence to support decision making during the early phases of mass vaccination campaigns. However, we have summarised important aspects that need to be taken into account when considering this methodology for the purpose of signal detection and validation. We highlight the following lessons learned from our experience performing O/E analysis within the context of a mass vaccination setting:

- Data collected on vaccine exposure and reported to SRS should be as granular and harmonised as possible to allow for stratified analyses. As a minimum requirement, data on vaccine brand, dose, age and sex should be ascertained.
- The rapid influx of spontaneous reports within a short timeframe hinders the manual review of spontaneous case reports. Initiatives that aim to standardise codes used between respective coding dictionaries for clinical and pharmacovigilance resources should be reinforced to promote alignment between the codes used to identify the 'observed' cases and those used to estimate the number of 'expected' cases.
- Accurate background incidence rates are essential to estimate the number of 'expected' cases in the absence of an unvaccinated population. Several studies have shown differences in background incidence rates between healthcare settings, countries and time periods. Expert review is warranted to inform the selection of background incidence rates for the main and sensitivity analyses.
- Adverse events caused by the disease, heterologous vaccinations, multiple dose schedules and virus variants pose additional challenges to the design and interpretability of O/E analysis. In such circumstances, the feasibility of O/E analysis should be carefully evaluated. When performed, the assumptions should be clearly described and the limitations acknowledged.

In conclusion, the considerations described in this article should inform the design of O/E analysis and provide a framework to assess its strengths and limitations when contextualising the results during future mass vaccination campaigns.

supported and continue to support the safety monitoring of the COVID-19 vaccines.

Declarations

References

1. EU Member States and the European Medicines Agency (EMA). Report on pharmacovigilance tasks 2019–2022. 2023. https://www.ema.europa.eu/en/documents/report/report-pharmacovigilance-tasks-eu-member-states-european-medicines-agency-ema-2019-2022_en.pdf. Accessed 19 Mar 2023.

2. Durand J, Dogné J-M, Cohet C, Browne K, Gordillo-Marañón M, Piccolo L, et al. Safety monitoring of COVID-19 vaccines: perspective from the European Medicines Agency. Clin Pharmacol Ther. 2023;113:1223–34.

3. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP): product- or population-specific considerations. I. Vaccines for prophylaxis against infectious diseases. P I. 2013. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-product-or-population-specific-considerations-i-vaccines-prophylaxis-against-infectious-diseases_en.pdf. Accessed 19 Mar 2023.

4. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Guide on methodological standards in pharmacoepidemiology (Revision 10). EMA/95098/2010. http://www.encepp.eu/standards_and_guidance. Accessed 19 Mar 2023.

5. Mahaux O, Bauchau V, Van Holle L. Pharmacoepidemiological considerations in observed-to-expected analyses for vaccines. Pharmacoepidemiol Drug Saf. 2016;25:215–22.

6. European Medicines Agency. Screening for adverse reactions in EudraVigilance. 2016. https://www.ema.europa.eu/en/documents/other/screening-adverse-reactions-eudravigilance_en.pdf. Accessed 19 Mar 2023.

7. Safety Platform for Emergency vACcines. SO2-D2.5.2.1: AESI case definition companion guide for 1st tier AESI. 2021. https://brightoncollaboration.us/wp-content/uploads/2021/03/SPEAC_D2.5.2.1-GBS-Case-Definition-Companion-Guide_V1.0_format12062-1.pdf. Accessed 25 May 2023.

8. European Medicines Agency. Signal assessment report on embolic and thrombotic events (SMQ) with COVID-19 vaccine (ChAdOx1-S [recombinant]): Vaxzevria (previously COVID-19 Vaccine AstraZeneca) (other viral vaccines). 2021. https://www.ema.europa.eu/en/documents/prac-recommendation/signal-assessment-report-embolic-thrombotic-events-smq-covid-19-vaccine-chadox1-s-recombinant_en.pdf. Accessed 25 May 2023.

9. European Medicines Agency. Updated signal assessment report on myocarditis, pericarditis with tozinameran (COVID-19 mRNA vaccine (nucleoside-modified): COMIRNATY). 2021 https://www.ema.europa.eu/en/documents/prac-recommendation/updated-signal-assessment-report-myocarditis-pericarditis-tozinameran-covid-19-mrna-vaccine_en.pdf. Accessed 25 May 2023.

10. European Medicines Agency. Signal assessment report on myocarditis, pericarditis with tozinameran (COVID-19 mRNA vaccine (nucleoside-modified): COMIRNATY). 2021. https://www.ema.europa.eu/en/documents/prac-recommendation/signal-assessment-report-myocarditis-pericarditis-tozinameran-covid-19-mrna-vaccine_en.pdf. Accessed 25 May 2023.

11. European Medicines Agency. Signal assessment report on myocarditis and pericarditis with Spikevax: COVID-19 mRNA vaccine (nucleoside-modified). 2021. https://www.ema.europa.eu/en/documents/prac-recommendation/signal-assessment-report-myocarditis-pericarditis-spikevax-previously-covid-19-vaccine-moderna-covid_en.pdf. Accessed 25 May 2023.

12. European Centre for Disease Prevention and Control. COVID-19 vaccine tracker. https://vaccinetracker.ecdc.europa.eu/public/extensions/COVID-19/vaccine-tracker.html#distribution-tab. Accessed 25 May 2023.

13. COVID-19 vaccine monitoring: VAC4EU. https://vac4eu.org/covid-19-vaccine-monitoring/. Accessed 25 May 2023.

14. Willame C, Dodd C, Gini R, Durán C, Thomsen R, Wang L, et al. Background rates of Adverse Events of Special Interest for monitoring COVID-19 vaccines. Zenodo; 2021 Aug. https://zenodo.org/record/5255870. Accessed 19 Mar 2024.

15. Li X, Ostropolets A, Makadia R, Shoaibi A, Rao G, Sena AG, 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.

16. Burn E, Li X, Kostka K, Morgan Stewart H, Reich C, Seager S, et al. Background rates of five thrombosis with thrombocytopenia syndromes of special interest for COVID-19 vaccine safety surveillance: incidence between 2017 and 2019 and patient profiles

from 25.4 million people in six European countries. Hematology. 2022;31(5):495–510.

17. European Medicines Agency. Real-world evidence framework to support EU regulatory decision-making: report on the experience gained with regulator-led studies from September 2021 to February 2023. 2023. https://www.ema.europa.eu/en/documents/report/real-world-evidence-framework-support-eu-regulatory-decision-making-report-experience-gained_en.pdf. Accessed 9 Aug 2023.

18. Hazell L, Shakir SAW. Under-reporting of adverse drug reactions. Drug Saf. 2006;29:385–96.

19. Marques J, Ribeiro-Vaz I, Pereira AC, Polónia J. A survey of spontaneous reporting of adverse drug reactions in 10 years of activity in a pharmacovigilance centre in Portugal. Int J Pharm Pract. 2014;22:275–82.

20. Sandberg A, Salminen V, Heinonen S, Sivén M. Under-reporting of adverse drug reactions in Finland and healthcare professionals' perspectives on how to improve reporting. Healthcare (Basel). 2022;10:1015.

21. Safety Platform for Emergency vACcines. D2.3 priority list of adverse events of special interest: COVID-19. https://brightoncollaboration.us/wp-content/uploads/2021/11/SPEAC_D2.3_V2.0_COVID-19_20200525_public.pdf. Accessed 25 May 2023.

22. Safety Platform for Emergency vACcines (SPEAC). SO2-D2.5.2.1: AESI case definition companion guide for 1st tier AESI Guillain Barré and Miller Fisher syndromes. 2021. https://brightoncollaboration.us/wp-content/uploads/2021/03/SPEAC_D2.5.2.1-GBS-Case-Definition-Companion-Guide_V1.0_format12062-1.pdf. Accessed 19 May 2024.

23. Maignen F, Hauben M, Tsintis P. Modelling the time to onset of adverse reactions with parametric survival distributions. Drug Saf. 2010;33:417–34.

24. Nguyen LS, Cooper LT, Kerneis M, Funck-Brentano C, Silvain J, Brechot N, et al. Systematic analysis of drug-associated myocarditis reported in the World Health Organization pharmacovigilance database. Nat Commun. 2022;13:25.

25. Mozzicato P. Standardised MedDRA queries: their role in signal detection. Drug Saf. 2007;30:617–9.

26. International Classification of Diseases (ICD). https://www.who.int/standards/classifications/classification-of-diseases. Accessed 11 Jul 2023.

27. SNOMED International. https://www.snomed.org. Accessed 11 Jul 2023.

28. Jatwani S, Goyal A. Vasculitis. StatPearls. Treasure Island (FL): StatPearls Publishing; 2022 http://www.ncbi.nlm.nih.gov/books/NBK545186/. Accessed 13 Jan 2023.

29. Lareb. Vasculitis and administration of COVID-19 vaccines. 2021. www.lareb.nl/media/x30nhooz/signal_2021_vasculitis-and-administration-of-covid-19-vaccines_gws.pdf. Accessed 19 May 2024.

30. European Medicines Agency. Meeting highlights from the Pharmacovigilance Risk Assessment Committee (PRAC) 24 - 27 October 2022. European Medicines Agency. 2022 https://www.ema.europa.eu/en/news/meeting-highlights-pharmacovigilance-risk-assessment-committee-prac-24-27-october-2022. Accessed 28 Mar 2023.

31. SNOMED CT – MedDRA Mappings Updated|MedDRA. https://www.meddra.org/news-and-events/news/snomed-ct-meddra-mappings-updated. Accessed 25 Jan 2024.

32. Towards European Health Data Space (TEHDAS). Recommendations to enhance interoperability within HealthData@EU. 2022 https://tehdas.eu/app/uploads/2022/12/tehdas-recommendations-to-enhance-interoperability-within-healthdata-at-eu.pdf. Accessed 25 May 2023.

33. Priyadarshni S, Westra J, Kuo Y-F, Baillargeon JG, Khalife W, Raji M. COVID-19 infection and incidence of myocarditis: a multi-site population-based propensity score-matched analysis. Cureus. 2022;14: e21879.

34. Keller K, Sagoschen I, Konstantinides S, Gori T, Münzel T, Hobohm L. Incidence and risk factors of myocarditis in hospitalized patients with COVID-19. J Med Virol. 2023;95: e28646.

35. Ostropolets A, Li X, Makadia R, Rao G, Rijnbeek PR, Duarte-Salles T, et al. Factors influencing background incidence rate calculation: systematic empirical evaluation across an international network of observational databases. Front Pharmacol. 2022;13: 814198.

36. Russek M, Quinten C, de Jong VMT, Cohet C, Kurz X. Assessing heterogeneity of electronic health-care databases: a case study of background incidence rates of venous thromboembolism. Pharmacoepidemiol Drug Saf. 2023;32:1032–48.