Drug Safety (2026) 49:147–162
https://doi.org/10.1007/s40264-025-01604-y

CURRENT OPINION

Charting and Sidestepping the Pitfalls of Disproportionality Analysis

Michele Fusaroli[1] · Daniele Sartori[1] · Eugène P. van Puijenbroek[2,3] · G. Niklas Norén[1]

Received: 31 March 2025 / Accepted: 11 August 2025 / Published online: 24 September 2025

Abstract

Disproportionality analysis is used by many pharmacovigilance organizations for detecting and assessing signals of potential adverse drug reactions. However, its goal is often misunderstood and the approach misapplied, leading to erroneous conclusions due to neglected violated assumptions. In this paper we illustrate how simplistic use and interpretation of disproportionality analysis can lead to incorrect conclusions. Using VigiBase, the WHO global database of adverse event reports, and the Information Component disproportionality metric, we provide selected examples to highlight common sources of error that can introduce spurious disproportionalities or lead to missing important signals: confounding (by age, sex, indication, comedication), effect modification (by age), notoriety bias, masking, misclassification (by miscoding, incomplete or imprecise event retrieval), neglecting report utility, and violated independence assumption. Additionally, we present how sophisticated analyses may introduce new biases or amplify existing ones, such as collider bias or masking amplification. Due to its pitfalls, disproportionality analysis plays a supportive rather than decisive role in signal detection and assessment. Careful design and interpretation of disproportionality analysis, with appropriate subgrouping and clinical assessment, are essential. While subgrouping can mitigate some pitfalls, it reduces sample size and may introduce or amplify existing biases and needs to be used with care. Further development of tools to detect and mitigate biases in disproportionality analyses, and to assess their risk of bias, is needed.

Key Points

The goal of disproportionality analysis is often misunderstood and the approach misapplied, leading to erroneous conclusions.

The types of pitfalls to be aware of include confounding, effect modification, notoriety bias, masking, misclassification, neglecting report utility, violated independence assumption, and collider biases.

Careful design and interpretation of disproportionality analysis, with appropriate subgrouping and clinical assessment, are essential.

✉ Michele Fusaroli
michele.fusaroli@who-umc.org

[1] Uppsala Monitoring Centre, Uppsala, Sweden

[2] Netherlands Pharmacovigilance Centre Lareb, 's-Hertogenbosch, The Netherlands

[3] University of Groningen, Groningen, The Netherlands

1 Introduction

Disproportionality analysis is a core method in pharmacovigilance, aiding researchers and professionals in identifying and assessing signals of potential adverse reactions by analyzing vast databases of adverse event (AE) reports [1–4]. Conceptualized in the 1970s [5], and first implemented by Stricker and Tijssen in 1992 [6], disproportionality analysis was originally adopted for routine screening by a small set of pharmacovigilance stakeholders worldwide, starting in the late 1990s [7–10]. Its primary purpose is to detect and prioritize possible signals, directing clinical experts' attention for in-depth assessment. With the growing number of AE reports, disproportionality analysis is now widely used across numerous pharmacovigilance organizations [11, 12]. Disproportionality analysis can also contextualize the size of a case series during signal validation and assessment. Recently, the increasing availability of public databases (e.g., FAERS–FDA Adverse Event Reporting System) [13], the deceptive simplicity of its calculations [14], and the rise of large language models [15] have triggered an explosion of studies using this method alone to draw conclusions

about causality [16–18], which is a questionable practice at best. Even when disproportionality results are presented as inconclusive, they can still be problematic: identifying mere statistical associations between reports does not inherently advance scientific understanding [19] or improve drug safety [20].

While disproportionality analysis can offer valuable insights, it comes with risks of misinterpretation that users must be aware of. Researchers who uncritically apply disproportionality analysis may fall prey to apophenia—perceiving meaningful connections between unrelated things, specifically leading to false-positive and false-negative disproportionalities, drawing misleading conclusions [21–24]. These misinterpretations contribute to research waste [21]. Effectively, they carry the risk of selecting non-relevant drug-event combinations for the clinical review of case series in signal validation and assessment (e.g., when reporting habits, confounding or masking fully explain disproportionality but are not accounted for [25–27]), but also of missing true signals. Ultimately, poor quality disproportionality analyses harm pharmacovigilance and patient safety. Although the theoretical limitations and assumptions behind disproportionality analysis have been discussed by many [3, 24, 28–33], these discussions are often highly technical, and their practical implications may not be clear to all pharmacovigilance professionals.

2 Aim

This manuscript aims to provide an overview of possible pitfalls[1] of disproportionality analysis, and strategies to avoid or to minimize them.

Specifically, after introducing AE reports (Sect. 3.1) and disproportionality analysis (Sect. 3.2), we use real-world examples from VigiBase—the WHO global database of adverse event reports for medicines and vaccines—to illustrate how a simplistic or careless application of disproportionality analysis can yield false positive and false negative results, and mislead researchers (Sect. 4, Fig. 1). Using the same examples, we emphasize how a thoughtful, informed design can help researchers avoid common biases[2] and other pitfalls, ultimately leading to more accurate and meaningful findings. Finally, we point to some possible mitigation strategies (Sect. 5).

[1] From the Oxford English Dictionary, a pitfall is a hidden or unsuspected danger, difficulty, or opportunity for error.

[2] From the Oxford English Dictionary, a bias is a [systematic] distortion of a statistical result arising from the method of sampling, measurement, analysis, etc.

3 Material and Methods

3.1 Adverse Event Reports

Adverse event reports document suspected adverse reactions and other AEs seemingly occurring following the exposure to medicinal products (henceforth: drugs). These reports can be submitted by healthcare professionals, patients, and others, and are collected by regulatory bodies and pharmaceutical companies. They do not represent a comprehensive or unbiased sample of all AEs. Instead, they are shaped by a complex generative process (see Fig. 1), in which several stages of information capture and processing may introduce sources of systematic variability that can lead to distortions in subsequent analyses, unless properly accounted for:

Exposure and Occurrence: By design, no reports are available for individuals who did not experience any AE or were not exposed to any drug[3]. Certain AEs may occur more frequently in individuals exposed to a specific drug due to higher susceptibility in these patients independent of the drug itself (e.g., because of underlying diseases or disorders, demographic factors, common co-medications).

Recognition: Not all AEs are easy to recognize. Some require specialized tests or expert knowledge, which may not be equally available or applied to different patient groups. Some exposures, such as over-the-counter medications, traditional medications and herbals, can also be difficult to recognize—especially for healthcare professionals. Other exposures may not be recalled due to a long delay from the last administration to the event onset. Whether AEs are labeled in product information may also influence the extent to which they are recognized and suspected as reactions in the individual patient.

Reporting: Only a fraction of all AEs are reported after drug use. The likelihood of a recognized AE or exposure being reported (i.e., reporting rate) depends on factors that are relative to the event (like its severity), the drug (like its novelty), and the drug-event combination (like media attention and labeling in product information), as well as on reporters' awareness on the opportunity of reporting and motivation to do so. Also, reporting rates may vary over time, between countries, etc.

Coding: Events are generally coded using the Medical Dictionary for Regulatory Activities (MedDRA). Despite general guidelines for coding, variability in practical application can introduce redundancies and ambiguities. For example, the same event might be

[3] With the exception of a few cases reporting "no adverse event", and of some centers accepting reports on exposure to drugs not followed by an adverse event.

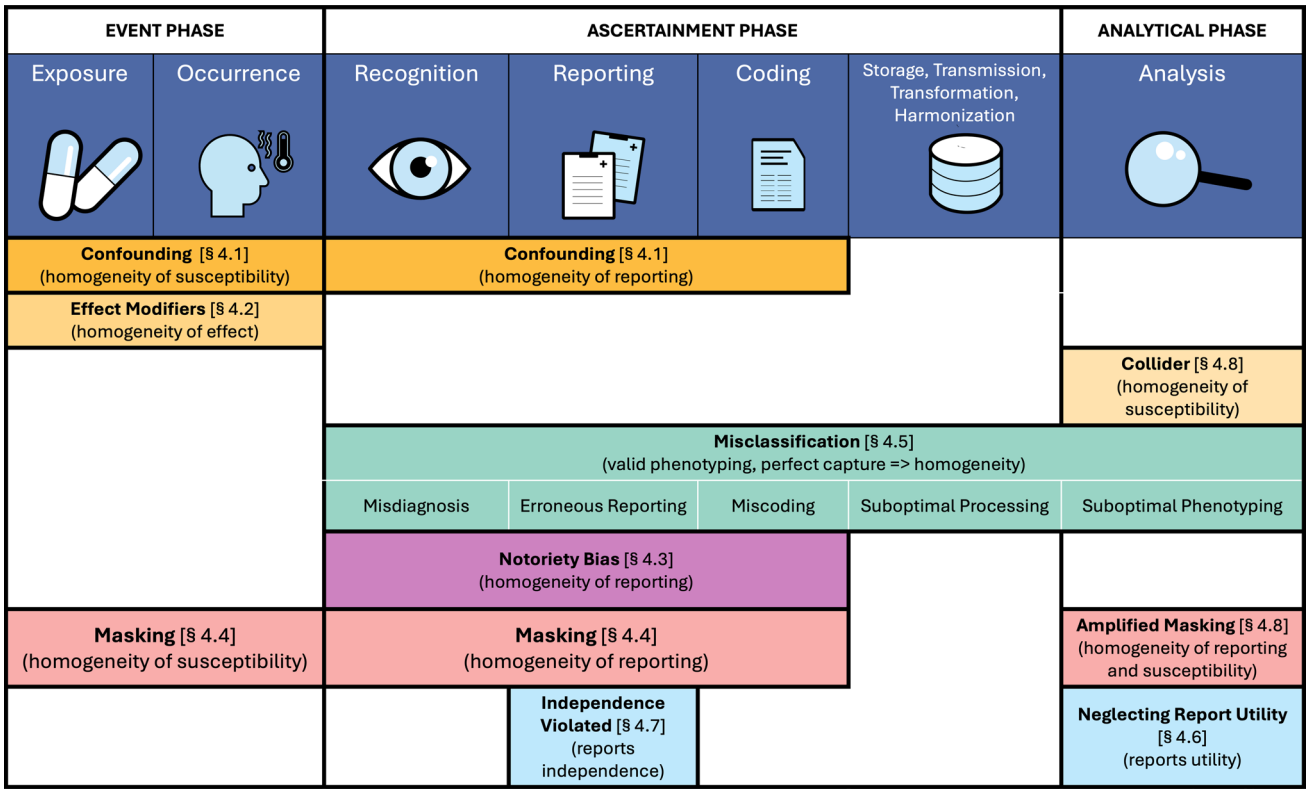

Fig. 1 Conceptual overview of pitfalls. Pitfalls are systematized according to the stage of the generative and analytical process. Squared brackets denote the sections of the manuscript presenting the specific pitfalls. The assumptions that when violated result in the pitfalls are presented in round brackets (see also Sect. 5)

recorded with one or multiple (more or less specific) terms. Similar considerations apply to drugs, which are often reported as free text and need to be coded to a standardized lexicon [34].

Data Transmission, Storage, Transformation, and Standardization: Transmission between clinical investigators, pharmaceutical companies and regulatory authorities does not always occur for every report, nor is the procedure of transmission the same. For example, the extent to which solicited reports are shared with VigiBase can vary between member organizations, and this variability is often difficult to account for [35, 36]. The choice of the storage format can result in data loss or auto-compilation of partial fields [13]. Procedures such as duplicate detection and removal, pre-computation of time to onset, and flagging of certain types of reports (e.g., those related to pregnancy [37, 38]) vary over time and between organizations and can affect data quality and completeness.

Analyses should seek to mitigate any distortion due to systematic variability. Because signal detection relies on general designs believed to be broadly effective, it will often be relevant to design and execute additional analyses for the assessment of a detected signal, bearing in mind all potential distortions.

3.2 Disproportionality Analysis and Its Fragility

Disproportionality analysis is a quantitative method used to support the identification of potential adverse reactions by determining whether a drug-event pair is reported more frequently than expected by chance, according to some model and assumptions [1, 11]. Specifically, disproportionality analysis compares the proportion of an AE of interest between a foreground (reports recording the exposure to the drug of interest) and a background (reports recording any other drugs but not the one of interest) (Fig. 2). Since disproportionality analysis uses AE reports, the results can be influenced by biases from the data generative process and the specific population represented in the analyzed database [39]. These biases emerge from either the nature of AE reports or the specific assumptions on which the disproportionality analytical approach is built.

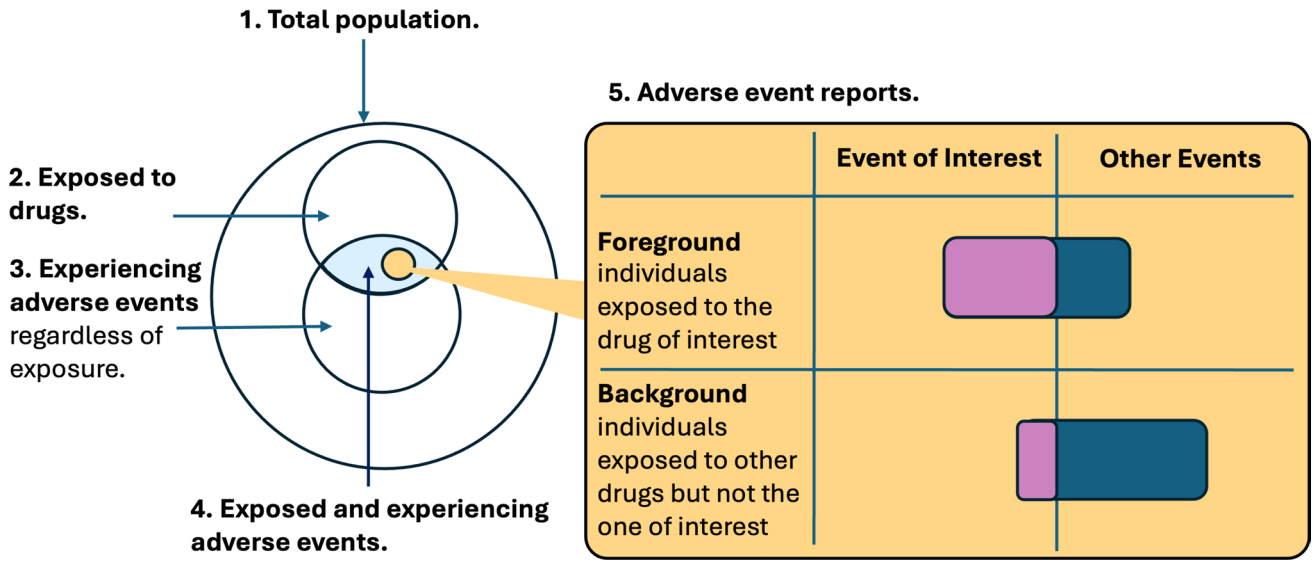

Fig. 2 Definition of foreground and background for disproportionality analysis, within the dataset of AE reports. The stacked bars within the contingency table show the different proportions of reports with the event of interest in the foreground and the background

4 Examples of Disproportionality Analysis Pitfalls

This section illustrates examples of how crude disproportionality analysis can lead to misleading conclusions and how more nuanced approaches can help mitigate (but sometimes exacerbate) these pitfalls. In these analyses, we use VigiBase (accessed December 31st, 2024, excluding suspected duplicates [40] and non-domestic reports[4]: 39,908,928 reports), and we use the Information Component (IC) as a disproportionality metric, along with its 95% credibility interval to identify statistical signals of disproportionate reporting ($IC_{025} > 0$) [41]. Any disproportionality suggests an association between a drug and an AE, which could be either causal or purely statistical. The shrinkage applied by the IC pulls the point estimate towards 0 and reduces the risk of false positives but does not formally control the familywise error or false discovery rates. The use of confidence intervals (CIs) provides further protection against spurious disproportionalities, although it only accounts for random variability and not systematic biases. Each of these strategies may of course come at the cost of missing some weaker signals. Examples are provided together with the observed and expected counts and forest plots showing the IC. Analyses were performed using the DiAna R package version 2.1.0 [42].

[4] Reports submitted from a member organization different from the country in which the adverse event occurred.

4.1 Confounding

Confounding occurs when an external factor influences both drug exposure and the outcome risk (or their reporting), resulting in a different susceptibility or reporting in exposed and non-exposed. Common confounders include age, sex, indication, and country.

4.1.1 Confounding by Age: Growth Retardation with Poliomyelitis Vaccine

A crude, database-wide disproportionality analysis suggests an association between the poliomyelitis vaccine and growth retardation (36 observed reports vs 11 expected) (Fig. 3). However, patients most at risk for growth retardation are also the typical recipients of the vaccine. This can introduce confounding by age. To obtain a less biased estimate, the foreground and background must be comparably susceptible to the outcome, and we have therefore to account for age in our analysis. Stratifying by age, we find approximately as many reports as expected in each age group (e.g., 34 reports vs 31 expected in the 28 days to 2 years age group, with a non-significant credibility interval), or even fewer (e.g., 0 reports vs 33 expected in the 2 to 11 years age group), thus dismissing a plausibly spurious disproportionality.

4.1.2 Confounding by Indication: Ischemic Stroke with Insulin

A crude analysis highlights a disproportionality between ischemic stroke and insulin (167 observed vs 114 expected; Fig. 4). However, the underlying reason for prescribing

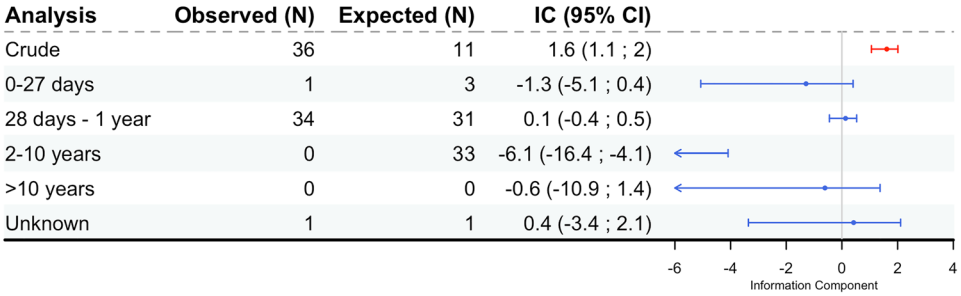

Fig. 3 Positive confounding by age: poliomyelitis vaccine and growth retardation

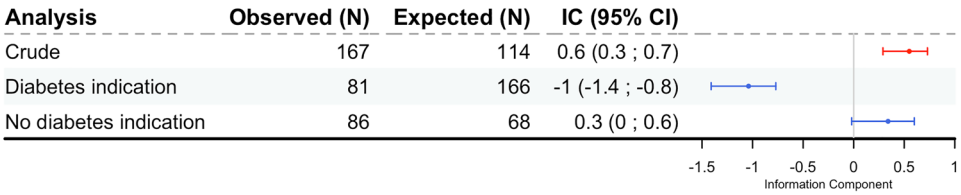

Fig. 4 Positive confounding by indication: ischemic stroke with insulin

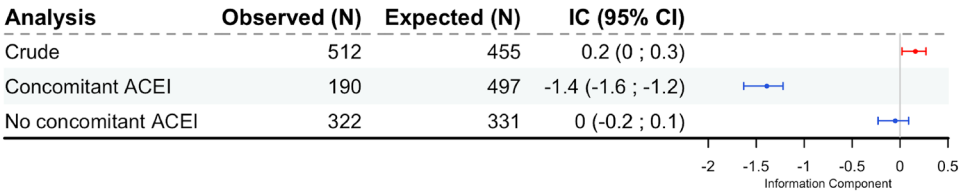

Fig. 5 Positive confounding by comedication: cough with hydrochlorothiazide

insulin, diabetes mellitus, is a risk factor for ischemic stroke, suggesting a possibility of confounding by indication. When looking within the stratum reporting a diabetes indication, there are fewer reports in the foreground than we would expect based on the background (i.e., 81 observed vs 166 expected when restricting to the therapeutic indication "hyperglycaemia/new onset diabetes mellitus Standardized MedDRA Queries [SMQ]"). That the association disappears (is even reversed, with fewer reports than expected) when accounting for diabetes, indicates that the confounding by indication fully explains the association. In the stratum without a diabetes indication, the IC is still positive although the lower limit falls just below 0, but this is likely residually confounded due to misclassification (not all reports list an indication for treatment, and we know insulin is exclusively used in diabetes[5]).

4.1.3 Confounding by Comedication: Cough with Hydrochlorothiazide

Analyzing cough with hydrochlorothiazide shows a disproportionality (512 observed vs 455 expected; Fig. 5). However, hydrochlorothiazide is frequently co-prescribed with angiotensin converting enzyme inhibitors (ACEIs), which are known causes of cough. Here, confounding by comedication could frame hydrochlorothiazide as "culprit", even though hydrochlorothiazide is not believed to have a causal role in inducing cough (i.e., it is just an "innocent bystander"). When excluding reports with ACEIs from both foreground and background, the disproportionality is not observed (322 observed vs 331 expected)[6].

4.1.4 Confounding by Sex: Breast Cancer with Finasteride

No disproportionality is observed between finasteride and breast cancer in a crude analysis of the entire database (59 reports vs 47 expected, with non-significant interval; Fig. 6). However, finasteride is predominantly prescribed to men, who have a lower baseline risk of breast cancer, suggesting

[5] See also Sect. 4.5 (and Section 3.3.3 of [33] for a presentation of how misclassification by incompleteness can invalidate the reliability of subgrouping in achieving homogeneity of susceptibility between foreground and background.

[6] Hydrochlorothiazide is also indicated for heart failure, which can itself cause coughing complaints, also introducing a potential confounding by indication.

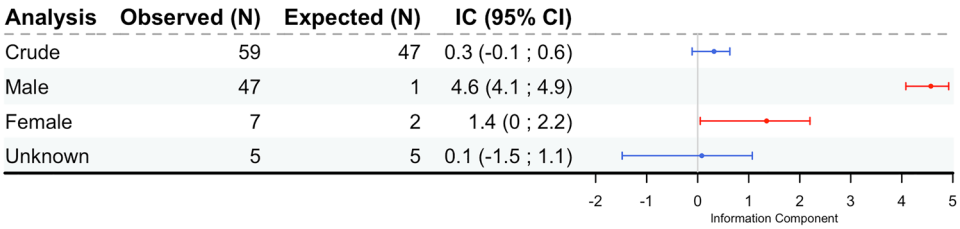

Fig. 6 Negative confounding by sex: breast cancer with finasteride

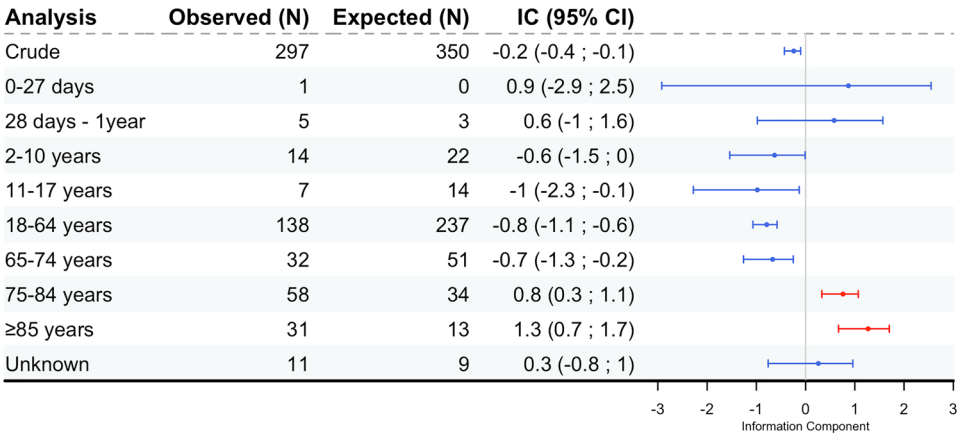

Fig. 7 Age-specific effect: hepatitis with ceftriaxone

the possibility of confounding by sex. Stratifying by sex reveals a strong disproportionality in males (47 observed vs 1 expected), and a weaker but significant disproportionality in females (7 observed vs 2 expected). The higher susceptibility in the background (which includes more women) relative to the foreground results in negative confounding that is sufficient to obscure the disproportionality. In fact, male breast cancer is included in the product information of finasteride based on an assessment of case reports [43], but registry studies have given conflicting results, and the association could also result from further confounding beyond that generated by the systematic difference in sex distribution (e.g., related with different occupation, radiation exposure, physical inactivity) [44].

4.2 Effect Modifiers

Effect modifiers influence how different subsets of the population respond to a drug. Sex, age, pharmacokinetic, and pharmacodynamic factors (e.g., liver/kidney injury, obesity) can be significant effect modifiers. Neglecting effect modification can lead to missing signals or misinterpreting adverse effects in certain subpopulations as applying to the entire exposed population.

4.2.1 Age-Specific Effect: Hepatitis with Ceftriaxone

Crude analysis of ceftriaxone and hepatitis shows slightly fewer reports than one would expect (297 reports vs 350 expected; Fig. 7). However, a feasibility study of signal detection for risk groups in VigiBase identified a signal of hepatitis due to ceftriaxone in the elderly based on disproportional reporting in this subgroup [45, 46]. After stratifying by age, disproportionality emerges (58 reports vs 34 expected for the 75 to 86 years age group, and 31 reports vs 13 expected for the > 85 years age group). Indeed, hepatitis is now labeled for ceftriaxone. As there is no known hepatotoxic effect in younger patients without severe hepatic impairment, searching for a causal effect on the entire exposed population may dilute the effect and miss it.

4.3 Notoriety Bias

Notoriety bias may result when an AE has a greater chance of being recognized and reported (and coded using a specific term) because the drug is known or thought to cause it [47].

4.3.1 Notoriety Bias: Autism with MMR Vaccine

A sharp increase in the reporting of the MMR (measles, mumps, rubella) vaccine with autism occurred in 1999 (42 reports vs 0.4 expected), while up to 1998 no significant disproportionality could be seen (1 report vs 0.03 expected; Fig. 8). This can be explained by the now retracted publication of fraudulent findings claiming a link between the two [48, 49], and the subsequent heightened media and societal attention. The increased reporting also led to a ripple effect

of the disproportionality of autism with other vaccines, despite the lack of any causal relation [47, 49].

4.4 Masking

Disproportionality analysis considers the overall number of reports including the drug of interest and the overall proportion of reports including the AE, in order to compute an expected value for the number of reports including the two together. Litigation, regulatory action, or general publicity may lead to very large numbers of reports on a specific drug—AE combinations. This may inflate the expected values for other combinations involving the same drug or AEs and suppress disproportionality—a phenomenon referred to as masking, cloaking, or competition bias. It may lead to missed or delayed detection of essential safety signals. A related but reverse effect is the so-called background dilution [25, 33] where large numbers of reports unrelated to both the drug and AE of interest decrease their expected value, inflating disproportionality.

4.4.1 Masking: Rhabdomyolysis with Venlafaxine

Crude analysis identifies a disproportionality between rhabdomyolysis and venlafaxine (274 vs 71 expected, with IC = 1.9 [1.7; 2.1]). When looking at the time trend of the IC, the disproportionality is only visible since 2010 (81 reports vs 23 expected), notwithstanding venlafaxine was already approved by FDA in the 1993. In fact, we know statins are an established cause of rhabdomyolysis [50] and constitute a large proportion of rhabdomyolysis reports in the database (43% today), improperly increasing the number of expected reports and masking the signal. Removing reports of statins with rhabdomyolysis from the database (unmasking) increases the disproportionality at the end of 2024 even further (255 vs 36 expected, with IC = 2.8 [2.6; 2.9]) and allows the disproportionality to emerge already in 2003 (14 vs 6 expected, compared to 15 vs 23 expected in the crude analysis) (Fig. 9). This example shows how masking can delay the detection of a signal. The dilution of the masking with time could be due to reduced reporting of expected reactions and an attenuated notoriety bias when cerivastatin was removed from the market.

4.5 Misclassification

Misclassification leads to inaccurate measurements of key study variables (e.g., due to incorrect or incomplete coding), potentially resulting in a distortion, systematic or not, in the measure of disproportionality (i.e., information bias).

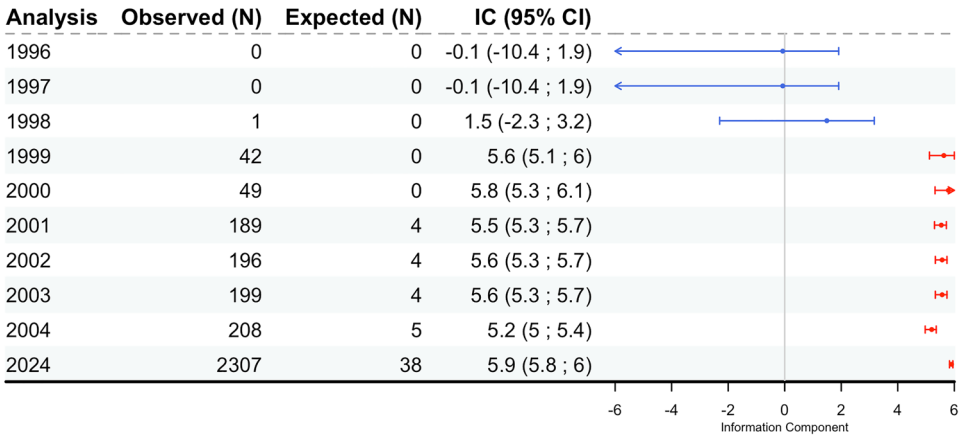

Fig. 8 Notoriety bias: autism with MMR vaccine

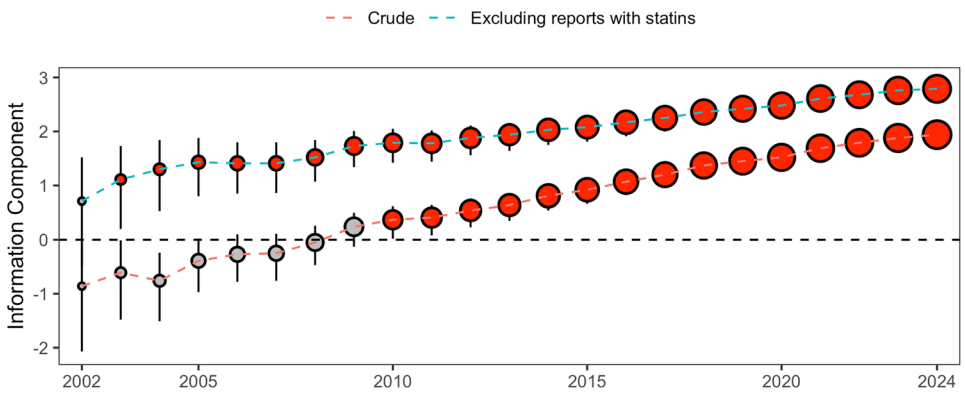

Fig. 9 Masking: rhabdomyolysis with venlafaxine. The size of the points is proportional to the count of observed

Misclassification can occur at various stages: recognition (e.g., misdiagnosis or failure to diagnose), reporting and coding (e.g., coding using false friends, when done by non-native English speakers, as in the case of mistaking delirium for delusion due to similar spelling in Italian [51]), processing (e.g., data loss and improper harmonization), and analysis (e.g., suboptimal phenotyping, as in the case of missing reports of oculo-muco-cutaneous syndrome because only symptoms are recorded). Beyond outcomes, misclassification can also affect exposures (e.g., reports from blinded clinical trials can report both drug and placebo without being necessarily exposed to both), drug-event relationships (e.g., the drug was taken before the AE or for treating the AE, or the AE resulted from drug inefficacy, without a suspicion of ADR), or covariates (impeding the achievement of homogeneity when subgrouping, resulting in residual bias [52]).

4.5.1 Miscoding Indication as AE: Hereditary Angioedema with Lanadelumab

A strong disproportionality is observed for hereditary angioedema with lanadelumab (1707 reports vs 0.6 expected; Fig. 10). However, lanadelumab is exclusively prescribed for hereditary angioedema, and the disproportionality can be explained by reverse causality. This misclassification can be due to precautionary reporting (i.e., reporting any event that occurs in connection with drug exposure, even when there is no suspicion that the drug caused the event), including that from patient support programs [27]. It can also result from inappropriate coding of the indication as an AE. The term hereditary itself should raise concerns, regardless of the drug of interest. However, drugs may worsen a hereditary condition and trigger symptomatic flares. Clinical knowledge is usually sufficient to identify reverse causality. In contrast, reverse causality can be very difficult to identify with purely statistical methods.

4.5.2 Incomplete Event Retrieval: Infections with Tisagenlecleucel

Investigating 'respiratory tract infection' with tisagenlecleucel shows no disproportionality (3 reports vs 2 expected, with an IC credibility interval spanning 0; Fig. 11). However, any analysis based on MedDRA must account for the many different ways in which similar AEs may be reported and coded. Expanding the search to the HLT level to include additional respiratory infection terms reveals a disproportionality (109 reports vs 83 expected, including 57 "pneumonia" cases and 28 "sinusitis" cases). A disproportionality is also seen at the SOC level, including all infections and infestations (659 reports vs 304 expected). In fact, tisagenlecleucel can result in immunosuppression and increase the risk for infections, particularly serious ones, including respiratory tract infections. Relying solely on one MedDRA preferred term without considering the different diagnoses and codes related to respiratory tract infections can miss the disproportionality through fragmentation of the relevant case series.

4.5.3 Imprecise Event Retrieval: Peripheral Neuropathies with Efalizumab

Investigating peripheral neuropathies with efalizumab, no disproportionality emerges for the PT nor for the HLGT for "neuropathy peripheral" (1 vs 9 and 23 vs 16, respectively, with non-significant interval; Fig. 12). However, both of these case definitions are very broad. Considering the more specific terms "demyelinating polyneuropathy" and "Guillain Barré" a disproportionality emerges (10 vs 2 and 3 vs 0.1, respectively), as has already been presented by Hill et al [53]. Although Sect. 4.5.2 showed the potential appropriateness of using higher levels of the MedDRA hierarchy, here it would be insufficient to detecting a signal of disproportionate reporting. The HLGT peripheral neuropathies introduces related but unspecific terms in the background.

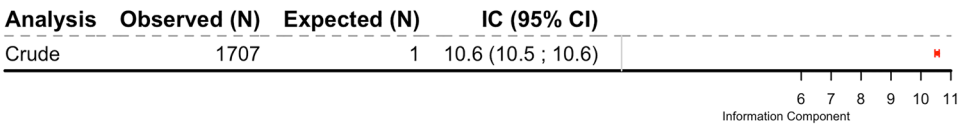

Fig. 10 Miscoding indication as adverse event: hereditary angioedema with lanadelumab

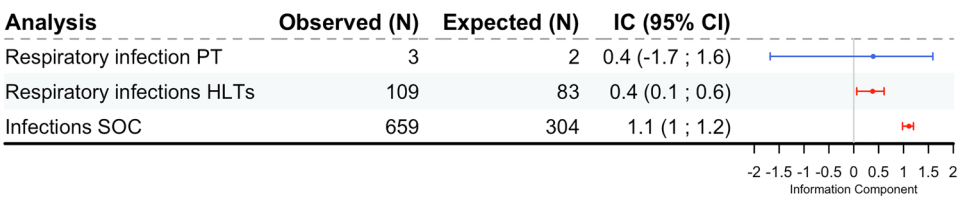

Fig. 11 Incomplete event retrieval: respiratory infections with tisagenlecleucel

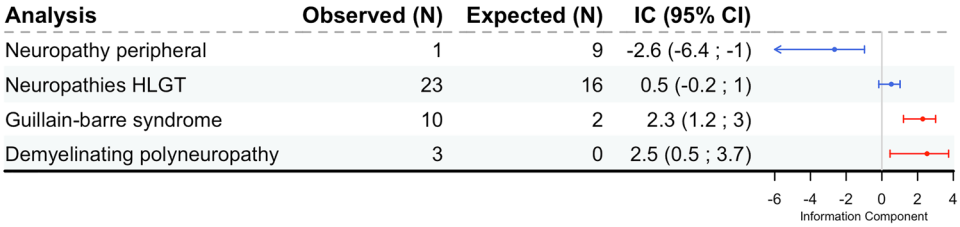

Fig. 12 Imprecise event retrieval: peripheral neuropathies with efalizumab

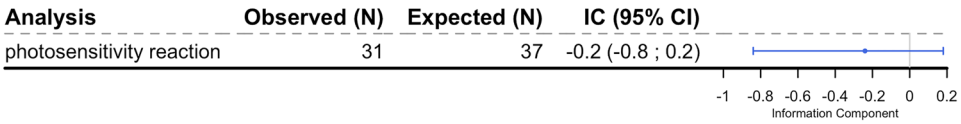

Fig. 13 High pharmacovigilance utility: photosensitivity with montelukast

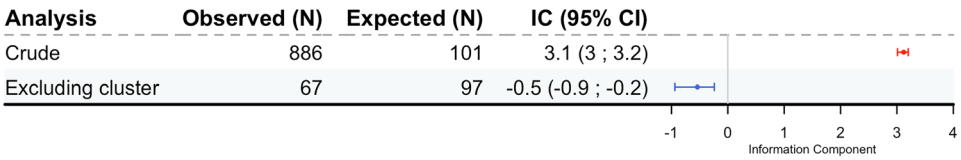

Fig. 14 Violated independence: anhedonia with docetaxel

4.6 Neglecting Report Utility

Distinct reports can have different utility based on their completeness and relevance for causality assessment. Reports may differ in reporters' suspicion levels (see precautionary reports [27]), clinical utility [54, 55] for causal assessment, and completeness [56, 57]. However, this is not accounted for in disproportionality analysis, which attributes the same weight to all reports. Some reports can be strong enough to identify a signal in the absence of disproportionality [58, 59], while others clearly do not support a signal even if there is disproportionality.

4.6.1 Reports with High Pharmacovigilance Utility: Photosensitivity with Montelukast

Montelukast and photosensitivity are not disproportionally reported overall (31 reports vs 37 expected), or in any of the subgroups that we have considered; nor does a disproportionality between the two appear to be subject to masking by a specific event[7] (Fig. 13). However, this was identified as a signal by the UMC in 2009 based on a strong underlying case series [60]: the 16 cases available at the time came from 8 different countries, and 15 reported montelukast as the solely suspected drug; in 4 cases, photosensitivity resolved upon de-challenge (montelukast discontinuation) and in 2 of these the event recurred upon rechallenge (montelukast reintroduction). A small number of reports with high clinical utility, as in this case, can often offer more meaningful insights for causality assessment than even a knowledge-driven disproportionality analysis. This is because scientific understanding evolves iteratively, with reasonable assumptions potentially shifting over time. Patterns that are not currently observable may emerge in the future. Therefore, signals should never be dismissed solely due to the absence of disproportionality.

4.7 Violated Independence

One common assumption in statistical analyses is that individual data points are independent of each other. In AE reports, dependence can arise in cases such as duplicate entries or clusters of reports generated during a study or legal proceedings.

4.7.1 Clusters of Reports from Court Litigations: Anhedonia with Docetaxel

Docetaxel and anhedonia show a strong disproportionality in the crude analysis (886 observed vs 101 expected; Fig. 14). However, most of these reports were submitted between 2018 and 2019, and 95% of them record all of the following AEs: anhedonia, alopecia, anxiety, discomfort, emotional distress, injury, pain. This is a very unusual combination only reported six other times in all VigiBase. It is highly likely that these are dependent reports, plausibly produced in a cluster due to a single court litigation [61]. After removing

[7] While it is possible that many events commonly reported with montelukast – particularly those of a neuropsychiatric nature because of notoriety and solicited reporting – are contributing to a form of diffuse masking, especially difficult to detect and address.

all the reports in the database co-reporting these seven AEs, the disproportionality disappears (67 vs 97 expected).

4.8 Biases Introduced by Sophisticated Analyses

Finally, sophisticated analyses can introduce or amplify biases if not carefully applied [62, 63]. Restricting to a specific subpopulation with a less-representative background can amplify masking, and restricting on effects shared by exposure and outcome can introduce collider biases.

4.8.1 Amplification of Masking: Anemia with Tisagenlecleucel

Investigating anemia with tisagenlecleucel, a CAR-T cell therapy used to treat acute lymphoblastic leukemia, shows a significant disproportionality in the crude analysis (70 reports vs 23 expected; Fig. 15). However, leukemia can itself cause anemia, suggesting a potential confounding by indication. Restricting to reports recording leukemias as an indication for treatment, the disproportionality disappears (33 reports vs 36 expected). While this could suggest that the disproportionality is spurious, the background restriction also amplifies masking. In fact, the background after restriction has a clear over-representation of drugs (13% ibrutinib, 10% imatinib, 10% venetoclax) associated with anemia (both as treatment and cause). As a consequence, the background includes more reports on anemia, and, even in the presence of a potential causal effect, the disproportionality with tisagenlecleucel is masked. The analysis based on reports that do not list leukemia as the indication indicates disproportional reporting (37 reports vs 14 expected), but this analysis—while avoiding masking amplification—likely suffers from residual confounding due to misclassification (many reports do not list an indication for treatment).

4.8.2 Introduction of a Collider Bias: Cardiac Failure with Pembrolizumab

Crude analysis shows disproportionality for cardiac failure with pembrolizumab (359 vs 152 expected; Fig. 16). To account for different reporting practices, one may be tempted to restrict the analysis to serious reports, which eliminates the disproportionality (261 vs 291). The cardiotoxicity of pembrolizumab is well established [64]. The fact that a plausible signal disappears when restricting to serious reports is due to a collider bias because both the event of interest and the drug of interest (or its indication for use, in this case cancer) increase the likelihood of the report being classified as serious. Intuitively, by restricting to serious reports, we retain a higher percentage of reports on pembrolizumab—compared to other drugs—where the AEs are not inherently serious, because patients using pembrolizumab are frailer. This leads to a higher relative frequency of cardiac failure in the background after restricting to serious reports, ultimately reducing the IC.

5 Assumptions of Disproportionality Analysis and Strategies to Mitigate Pitfalls

This section presents a conceptual overview of the key assumptions of disproportionality analysis, and mitigation strategies to address the presented pitfalls.

5.1 Assumptions Underlying Disproportionality Analysis

For disproportionality analysis to be effective, the proportion of the events in the foreground and background should correlate with the risk of the event in the exposed population and an equally susceptible reference group, respectively.

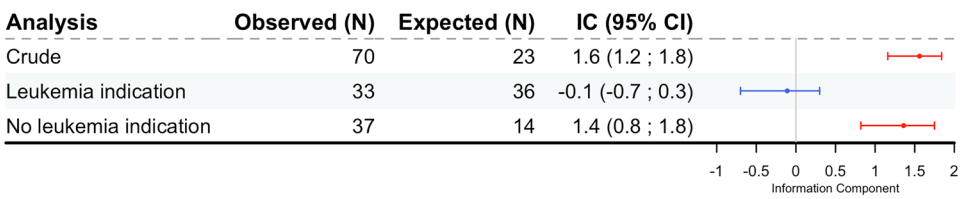

Fig. 15 Amplification of masking: anemia with tisagenlecleucel

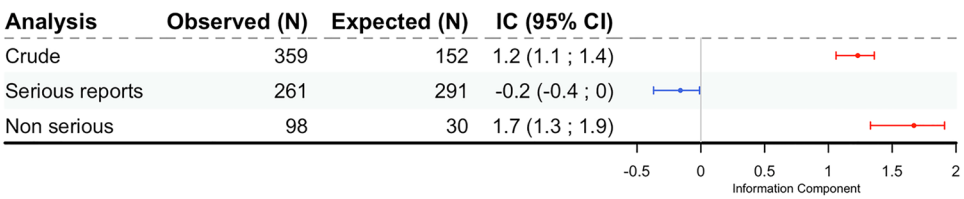

Fig. 16 Collider bias introduced by restricting to only serious reports: cardiac failure with pembrolizumab

These assumptions can be violated at different stages of the generative process of the AE reports (Fig. 1), and, while frequently overlooked, rarely hold true. For an overview of the different assumptions, refer to Table 1.

All assumptions should be explicitly stated at the design stage and critically evaluated during analysis and interpretation. The evaluation should consider differences between datasets of AE reports, as these may affect the degree to which assumptions are violated and the ability to capture and correct for important factors. Realistically, all disproportionality analyses violate at least one of the assumptions, but when violations are minor, disproportionality analysis can still be useful, when properly contextualized. Similarly, while more sophisticated epidemiological methods have proven to be powerful in structured real-world data, their application in disproportionality analysis must be approached with caution. This is because the foundational assumptions of these methods often do not hold in AE reports databases. In traditional epidemiology, multivariate adjustment is generally preferred over subgrouping, as the latter can reduce statistical power, complicate interpretation, and limit generalizability. However, in the context of AE reports, subgrouping may often be more appropriate. This is because it avoids relying on assumptions that are frequently majorly violated in these data sources—particularly regarding the completeness and accuracy of covariate information. For instance, we can place greater confidence in the presence of an indication when it is explicitly recorded in a report than in its absence when it is not mentioned. A disproportionality analysis based on reports without the indication of interest may suffer from residual confounding due to misclassification, and this bias would typically carry over to an adjusted analysis.

Finally, while disproportionality analyses do rely on strong assumptions, they were developed in part to address challenges posed by other approaches—particularly those that combine numerator and denominator data from different sources, combinations that can introduce additional layers of bias and distortion [65].

5.2 Mitigation Strategies

To address these pitfalls, one may account for heterogeneity in susceptibility, effect, and reporting using subgrouping, stratification, regression [8, 33, 66–68], devise queries to reduce the impact of misclassification (e.g., based on MedDRA hierarchy, Standardized MedDRA Queries, knowledge engineering techniques, or clinical knowledge [53]), and integrate disproportionality with a case-by-case assessment.

If pitfalls are addressed appropriately, disproportionality analysis may be used beyond signal detection, in signal assessment. So, mitigation strategies for disproportionality analysis will vary depending on its intended use and should consider the general characteristics of the database.

5.2.1 For Signal Detection

For signal detection, mitigation strategies may prioritize sensitivity and timeliness to avoid missing important signals. However, one must also account for the potential opportunity costs associated with overly sensitive thresholds that may trigger excessive false positives, resource-intensive follow-up, and alert fatigue [12].

Possible strategies include: stratification by common confounders, effect modifiers (e.g., sex, age), and reporting

Table 1 Assumptions underlying disproportionality analysis

Assumption	Description
Valid phenotyping	(a) *Only* outcomes, exposures, and relationships* relevant to the research question are retrieved (b) *All* outcomes, exposures, and relationships relevant to the research question are retrieved
Homogeneity	(a) Homogeneity of *susceptibility* to the event between foreground and background (b) Homogeneity of drug *effect* across individuals in the foreground (c) Homogeneity of *reporting* (and recognition and coding) across strata and drug-event combinations**
Reports independence	(a) Different reports document *different events* (b) Reports are *produced independently*
Reports utility	(a) Reports have the same degree of *reporter suspicion* (b) Reports have the same degree of *assessor suspicion* (c) *Perfect capture* of relevant variables allows correct classification (d) Reports *support the causal relation* with the same strength

*When the drug and event are each in scope with the research inquiry, the report can still refer to an out-of-scope relationship (e.g., incompatible temporal relationship)

**Non-systematic variations in reporting at the level of specific drugs or adverse events should cancel out with disproportionality analysis

practices (e.g., country, reporter type), followed by the inspection of subgroup-specific estimates [66, 67]. However, subgrouping can reduce sample size (risk of missing signals because of lack of sufficient evidence) and may introduce collider biases or amplify masking. Also, performing multiple variations of the same analysis in parallel increases the risk of spurious disproportionalities and requires corrections. For example, beyond looking at 95% CIs and shrinkage used in regular IC analysis, one may use a 99% CI or any other method in order to control the false discovery rate [66, 69].

Ensuring proper capture of variables and running the analysis on multiple queries, potentially also using the MedDRA hierarchy, can address misclassifications [70, 71]. Unsupervised machine learning approaches may help identify case series of interest based on the co-reported events [72–75].

Filtering out precautionary reports can be very difficult, because there are no methods to clearly flag them at the stage of reporting. However, when successfully implemented, these methods help reduce spurious disproportionalities—such as those related to reverse causality introduced by miscoding [27]. Predictive models for disproportionality analysis that account for other aspects of strength of evidence [76–78] can identify signals that would not be detected using disproportionality alone. Finally, once an unexpected disproportionality is identified, case-by-case review can help refine the study design by accounting for drug event-specific pitfalls that may not have been apparent from the broader design choices used in hypothesis-free, untargeted signal detection. For example, individual case review may identify a violation of independence when many cases are submitted simultaneously by a single source, like a poison control center. Extremely high disproportionality scores can sometimes suggest a misclassification of the indication for use as an AE, but they can also emerge from highly specific or well-known reactions.

5.2.2 For Signal Assessment

Generally, mitigation strategies for signal assessment should seek to increase specificity. However, if disproportionality is used to assess signals detected in other types of evidence, one should also consider negative biases, which could reduce sensitivity.

In signal assessment, a general characterization of the foreground relative to the background can be useful a priori to better inform the analysis and interpretation. Such characterization, together with the assumptions derived from a priori knowledge, should drive the subgrouping on expected confounders, effect modifiers, and reporting practices. To account for potential notoriety bias, one might use time series analyses of disproportionality in relation to regulatory warnings, court litigations, and other factors that may impact the reporting rate. Sudden shifts of disproportionality score over time can be indicative of either emerging issues or the impact of reporting biases. To account for masking and background dilution, one should exclude competing drugs or events and large inputs of unrelated reports, when relevant. Subgrouping on variables influenced by prescriptions and outcomes should generally be avoided, as this may introduce collider bias (e.g., not only outcome seriousness, but also other drugs and events). Inspecting the background after restrictions helps detect and adjust over-representation of specific drugs and events, which may amplify masking.

Unmeasured confounders can limit the effectiveness of pooling or subgrouping (e.g., many reports without smoking history could still relate to smokers), leading to residual confounding in the stratum without a specific covariate. Subgrouping can also reduce sensitivity, by decreasing the population size enough to lose true signals and generalizability. Also, apparent confounding by co-medication (e.g., rhabdomyolysis with azithromycin co-reported with statins) does not rule out drug-drug interactions [79].

To address misclassification, it is important to apply clinical judgment in selecting queries, assessing retrieved cases individually to ensure relevance, and using sensitivity analyses when uncertain. As there is no universally applicable choice of coded terms or sets thereof, any choice of terms should be commensurate with the specific inquiry at hand and requires extensive familiarity with (inter)national reporting practices, medical terminologies, and coding practices and conventions.

Moreover, disproportionality analysis usually does not account for time to onset (TTO), dose-response relationship, or reversibility, which can contextualize findings [80].

Deduplication can help avoid some false positives [40], and the same methods sometimes identify clusters of reports that violate the assumption of report independence [40, 81].

6 Discussion

The potential pitfalls in disproportionality analysis are wide ranging and can have substantial impact. Some of them apply to any analysis of AE reports. We hope to have demonstrated the challenge related with pitfalls and to have highlighted the assumptions underpinning disproportionality analysis, the conditions in which they can be violated, and strategies to mitigate the risk of false positives and false negatives.

Our examples demonstrate that subgroup analysis can help uncover and avoid many of the pitfalls inherent in disproportionality analysis, although careless subgrouping may itself introduce or exacerbate bias. These findings support a

more nuanced approach—one that relies on prior knowledge to guide subgrouping—challenging the view that the nature of AE reports and disproportionality analysis precludes more sophisticated analytical methods [20].

The use of disproportionality analysis should be justified and its ability to address the question of interest explicitly assessed. The latter engenders proper design, based on existing causal knowledge, and minimizes the gap between the potential and actual value of disproportionality analysis. Similarly, assumptions should be assessed and transparently reported, and their violations addressed, as we have shown to be at least partly possible. Failure to do either, renders disproportionality analysis vulnerable to deliberate malpractice, or oversights, leading to poor quality studies, potentially missing important signals, generating unjustified alerts, and dissipating clinical review resources on spurious disproportionalities. With growing use by scientists who may not have a full appreciation of the underlying data and the strengths and limitations of disproportionality analysis, there is an increased risk of succumbing to the pitfalls that may have a consequential impact on public health and patient safety.

Even when following best practices for disproportionality analysis, and applying the most appropriate strategies to address biases, there remains a large gap between disproportionality and causation. So, associational and causal terms should be used accordingly [19]. Even more fundamentally, proportions of AEs in the database do not reflect the true frequency of events in the general population, and disproportionality analysis does not correspond to relative risk, even after adjustments like those proposed in this article. Therefore, disproportionality should not be interpreted as a measure of effect, but rather as a mere indicator of the potential presence or absence of a causal relation. The assumption-based methods presented above should be viewed merely as strategies to improve the predictive validity of this indicator. Moreover, disproportionality should never be used in isolation for establishing causality, but as one component of a broader evidence puzzle. A case review is crucial, and other factors supporting the existence of the signal must be considered.

This work is not a complete inventory of all possible pitfalls and is anecdotal by design. It does not discuss how common and consequential different pitfalls are. The pitfalls illustrated here generally apply to a wide variety of databases of AE reports, but there may still be some variability. For example, some pitfalls related to geographical variability may be less evident in national databases, while databases of pharmaceutical companies with a limited sample of drugs may be more susceptible to masking. Moreover, when applying disproportionality analysis to other data sources one has to consider that further data source-specific pitfalls may become relevant [82–84]. Beyond affecting the reliability of disproportionality analysis, several of these pitfalls also apply to any other analysis of AE reports. For example, if the case definition is too narrow, this will result in incomplete case series and impact any subsequent analyses.

Future research may focus on developing a systematic framework to identify confounders, avoid colliders, and address measurement errors and differential reporting. Such a framework should be built on a more comprehensive understanding of the generative process of AE reports, than the one we used. Studies that systematically assess the frequency, detectability, and public health consequences of the presented biases—especially when signals are either falsely amplified or obscured—would significantly advance the field.

This manuscript, together with other efforts [28, 29, 33, 85], may contribute to the formalization of biases in disproportionality analysis and AE reports in a common catalogue of bias, alongside other observational and experimental biases (e.g., https://catalogofbias.org/), and in the development of a risk of bias tool. Such a tool could allow for better qualified and more useful disproportionality analysis.

7 Conclusion

This work highlights the complexity of disproportionality analysis in pharmacovigilance, emphasizing the risks of conflating disproportionality with causality. While disproportionality analysis is valuable for signal detection and assessment, it operates on several assumptions that must be carefully considered and addressed to avoid misleading conclusions. Future research may seek to develop a systematic framework for incorporating knowledge about the biases affecting AE reports into the design of pharmacovigilance analyses.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40264-025-01604-y.

Acknowledgments The authors are indebted to the members of the WHO Programme for International Drug Monitoring who contribute reports to VigiBase. However, the opinions and conclusions of this study are not necessarily those of the various member organisations nor of the WHO. The views expressed are those of the authors and not necessarily those of the Uppsala Monitoring Centre or Lareb. The authors would like to thank Judith Brand (who helped in the conceptualization of the study and provided constructive feedback on drafts of the manuscript), Oskar Gauffin (who helped in the conceptualization of the study) for their contributions, Boukje Raemaekers, and Peter Hjelmström (who provided helpful comments on the final draft of the paper). MedDRA® trademark is registered by ICH. MedDRA® the Medical Dictionary for Regulatory Activities terminology is the international medical terminology developed under the auspices of the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). This was an invited paper from Drug Safety.

Funding Not applicable.

Declarations

Conflict of Interest The authors declare no conflict of interest specific for this research. GNN and EvP are Editorial Board members of Drug Safety. GNN and EvP were not involved in the selection of peer reviewers for the manuscript nor any of the subsequent editorial decisions.

Ethics Approval Not applicable. This study did not use personal data.

Consent to Participate Not applicable. This study did not involve human participants.

Consent for Publication Not applicable.

Availability of Data and Material The data that support the findings of this study are not publicly available. Access to the data is restricted based on the conditions for access within the WHO Programme for International Drug Monitoring. Subject to these conditions, data are available from the authors on reasonable request. For further inquiries, please contact Uppsala Monitoring Centre via https://who-umc.org/contact-information/.

Code Availability The code used for this study is publicly available as a supplementary material and uses functions from the DiAna package [42]. Analyses were performed using R (version 4.4.2)

Author Contributions MF, DS, GNN participated in the conceptualization and design of the study. MF performed the data analysis and the visualization and wrote the original draft. MF, DS, EvP, GNN contributed to the interpretation of results, and to the review and editing of the draft. All the authors read and approved the final version.

References

1. Bate A, Evans SJW. Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiol Drug Saf. 2009;18:427–36.
2. Council for International Organizations of Medical Sciences, editor. Practical aspects of signal detection in pharmacovigilance: report of CIOMS Working Group VIII. Geneva: CIOMS; 2010.
3. Hauben M, Madigan D, Gerrits CM, Walsh L, Van Puijenbroek EP. The role of data mining in pharmacovigilance. Expert Opin Drug Saf. 2005;4:929–48.
4. Raschi E, Moretti U, Salvo F, Pariente A, Antonazzo IC, Ponti FD, et al. Evolving roles of spontaneous reporting systems to assess and monitor drug safety. Pharmacovigilance. 2018. https://doi.org/10.5772/intechopen.79986.
5. Finney DJ. Systematic signalling of adverse reactions to drugs. Methods Inf Med. 1974;13:01–10.
6. Stricker BH, Tijssen JG. Serum sickness-like reactions to cefaclor. J Clin Epidemiol. 1992;45:1177–84.
7. Bate A, Lindquist M, Edwards IR, Olsson S, Orre R, Lansner A, et al. A bayesian neural network method for adverse drug reaction signal generation. Eur J Clin Pharmacol. 1998;54:315–21.
8. DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. Am Stat. 1999;53:177–90.
9. Evans SJW, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. Pharmacoepidemiol Drug Saf. 2001;10:483–6.
10. van Puijenbroek E, Diemont W, van Grootheest K. Application of quantitative signal detection in the Dutch spontaneous reporting system for adverse drug reactions. Drug Saf. 2003;26:293–301.
11. Fusaroli M, Raschi E, Poluzzi E, Hauben M. The evolving role of disproportionality analysis in pharmacovigilance. Expert Opin Drug Saf. 2024;23:981–94.
12. Wisniewski AFZ, Bate A, Bousquet C, Brueckner A, Candore G, Juhlin K, et al. Good signal detection practices: evidence from IMI PROTECT. Drug Saf. 2016;39:469–90.
13. Giunchi V, Fusaroli M, Hauben M, Raschi E, Poluzzi E. Challenges and opportunities in accessing and analysing FAERS data: a call towards a collaborative approach. Drug Saf. 2023;46:921–6.
14. Poluzzi E, Raschi E, Piccinni C, Ponti FD. Data mining techniques in pharmacovigilance: analysis of the publicly accessible FDA adverse event reporting system (AERS). In: Data mining applications in engineering and medicine. 2012. p. 265–302.
15. Kendall G, Teixeira da Silva JA. Risks of abuse of large language models, like ChatGPT, in scientific publishing: authorship, predatory publishing, and paper mills. Learn Publ. 2024;37:55–62.
16. Loke YK, Mattishent K, Navaneetharaja N. New adverse drug reaction signals from 2017 to 2021—genuine alerts or false alarms? Pharmacy. 2024;12:33.
17. Sartori D, Aronson JK, Norén GN, Onakpoya IJ. Signals of adverse drug reactions communicated by pharmacovigilance stakeholders: a scoping review of the global literature. Drug Saf. 2023;46:109–20.
18. Wang L, Feng W, Duan J, Liang J. Pharmacovigilance bibliometrics: visualizing thematic development in the category of pharmacology and pharmacy in web of science. Front Pharmacol. 2021;12:731757.
19. Hernán MA. The C-Word: scientific euphemisms do not improve causal inference from observational data. Am J Public Health. 2018;108:616–9.
20. Hammad TA, Naylor M, Ely DM, Davies S. Exploring the complexities of disproportionality analysis in pharmacovigilance: reflections on the READUS-PV guideline and a call to action. Front Pharmacol. 2025;16:1573353.
21. de Boer A. When to publish measures of disproportionality derived from spontaneous reporting databases? Br J Clin Pharmacol. 2011;72:909–11.
22. Gagne JJ. Finding meaningful patterns in adverse drug event reports. JAMA Intern Med. 2014;174:1934–5.
23. Greenblatt DJ. The pharmacovigilance syndrome. J Clin Psychopharmacol. 2015;35:361–3.
24. Hauben M, Reich L, Van Puijenbroek EP, Gerrits CM, Patadia VK. Data mining in pharmacovigilance: lessons from phantom ships. Eur J Clin Pharmacol. 2006;62:967–70.
25. Adjaï F, Fournier D, Dolladille C, Lebrun-Vignes B, Bihan K. Impact of the vaccination against SARS-CoV-2 campaign on disproportionality indicator from the WHO pharmacovigilance

database: a competition bias study from case/non-case analysis. Therapie. 2024;S0040–5957(24):00039–48.
26. Grundmark B, Holmberg L, Garmo H, Zethelius B. Reducing the noise in signal detection of adverse drug reactions by standardizing the background: a pilot study on analyses of proportional reporting ratios-by-therapeutic area. Eur J Clin Pharmacol. 2014;70:627–35.
27. Klein K, Scholl JHG, De Bruin ML, van Puijenbroek EP, Leufkens HGM, Stolk P. When more is less: an exploratory study of the precautionary reporting bias and its impact on safety signal detection. Clin Pharmacol Ther. 2018;103:296–303.
28. Cutroneo PM, Sartori D, Tuccori M, Crisafulli S, Battini V, Carnovale C, et al. Conducting and interpreting disproportionality analyses derived from spontaneous reporting systems. Front Drug Saf Regul. 2023. https://doi.org/10.3389/fdsfr.2023.1323057.
29. Faillie J-L. Case-non-case studies: principle, methods, bias and interpretation. Therapies. 2019;74:225–32.
30. Hauben M, Maignen F. Does serious consequential masking exist? An update. Pharmacoepidemiol Drug Saf. 2017;26:727–9.
31. Scosyrev E, Behr S, Jain D, Ponnuru A, Michel C. Disproportionality analysis and causal inference in drug safety. Pharm Med. 2025;39:97–107.
32. Van Holle L, Bauchau V. The upper bound to the relative reporting ratio-a measure of the impact of the violation of hidden assumptions underlying some disproportionality methods used in signal detection. Pharmacoepidemiol Drug Saf. 2014;23:787–94.
33. Fusaroli M, Mitchell J, Rudolph A, Rocca E, Fusaroli R. Causal inference tools for pharmacovigilance: using causal graphs to identify and address biases in disproportionality analysis. OSF; 2025. https://osf.io/h5w9u_v3.
34. Lagerlund O, Strese S, Fladvad M, Lindquist M. WHODrug: a global, validated and updated dictionary for medicinal information. Ther Innov Regul Sci. 2020;54:1116–22.
35. Jokinen JD, Walley RJ, Colopy MW, Hilzinger TS, Verdru P. Pooling different safety data sources: impact of combining solicited and spontaneous reports on signal detection in pharmacovigilance. Drug Saf. 2019;42:1191–8.
36. Harinstein L, Kalra D, Kortepeter CM, Muñoz MA, Wu E, Dal Pan GJ. Evaluation of postmarketing reports from industry-sponsored programs in drug safety surveillance. Drug Saf. 2019;42:649–55.
37. Sandberg L, Vidlin SH, K-Pápai L, Savage R, Raemaekers BC, Taavola-Gustafsson H, et al. Uncovering pregnancy exposures in pharmacovigilance case report databases: a comprehensive evaluation of the vigibase pregnancy algorithm. Drug Saf. 2025. https://doi.org/10.1007/s40264-025-01559-0.
38. Zaccaria C, Piccolo L, Gordillo-Marañón M, Touraille G, de Vries C. Identification of pregnancy adverse drug reactions in pharmacovigilance reporting systems: a novel algorithm developed in EudraVigilance. Drug Saf. 2024;47:1127–36.
39. Candore G, Juhlin K, Manlik K, Thakrar B, Quarcoo N, Seabroke S, et al. Comparison of statistical signal detection methods within and across spontaneous reporting databases. Drug Saf. 2015;38:577–87.
40. Norén GN, Orre R, Bate A, Edwards IR. Duplicate detection in adverse drug reaction surveillance. Data Min Knowl Discov. 2007;14:305–28.
41. Norén GN, Hopstadius J, Bate A. Shrinkage observed-to-expected ratios for robust and transparent large-scale pattern discovery. Stat Methods Med Res. 2013;22:57–69.
42. Fusaroli M, Giunchi V. DiAna. An open access toolkit for DIsproportionality ANAlysis and other pharmacovigilance investigations in the FAERS. 2024. https://fusarolimichele.github.io/DiAna_package/.
43. Finasteride: potential risk of male breast cancer. GOV.UK. https://www.gov.uk/drug-safety-update/finasteride-potential-risk-of-male-breast-cancer.
44. Kjærulff TM, Ersbøll AK, Pukkala E, Bolin K, Green A, Emneus M, et al. Characteristics of finasteride users in comparison with nonusers: a Nordic nationwide study based on individual-level data from Denmark, Finland, and Sweden. Pharmacoepidemiol Drug Saf. 2020;29:453–60.
45. Boyd I. Ceftriaxone and hepatitis in patients 75 years and older. WHO Pharm Newsl. 2018;6:24–30.
46. Sandberg L, Taavola H, Aoki Y, Chandler R, Norén GN. Risk factor considerations in statistical signal detection: using subgroup disproportionality to uncover risk groups for adverse drug reactions in VigiBase. Drug Saf. 2020;43:999–1009.
47. Pariente A, Gregoire F, Fourrier-Reglat A, Haramburu F, Moore N. Impact of safety alerts on measures of disproportionality in spontaneous reporting databases the notoriety bias. Drug Saf. 2007;30:891–8.
48. Wakefield A, Murch S, Anthony A, Linnell J, Casson D, Malik M, et al. RETRACTED: Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. Lancet. 1998;351:637–41.
49. BMJ editorial board. Wakefield's article linking MMR vaccine and autism was fraudulent | The BMJ. 2011. https://www.bmj.com/content/342/bmj.c7452.
50. Collins R, Reith C, Emberson J, Armitage J, Baigent C, Blackwell L, et al. Interpretation of the evidence for the efficacy and safety of statin therapy. Lancet. 2016;388:2532–61.
51. Garmann T. From safety reports to safety data. Exploring variability in coding of reported adverse drug reactions into standardized medical language. Oslo Metropolitan University; 2024. https://oda.oslomet.no/oda-xmlui/handle/11250/3162187.
52. Hauben M, Patadia VK, Goldsmith D. What counts in data mining? Drug Saf. 2006;29:827–32.
53. Hill R, Hopstadius J, Lerch M, Norén GN. An attempt to expedite signal detection by grouping related adverse reaction terms. Drug Saf. 2012;35:1194–5.
54. Muñoz MA, Dal Pan GJ, Wei Y-JJ, Delcher C, Xiao H, Kortepeter CM, et al. Towards automating adverse event review: a prediction model for case report utility. Drug Saf. 2020;43:329–38.
55. Gosselt HR, Bazelmans EA, Lieber T, van Hunsel FPAM, Härmark L. Development of a multivariate prediction model to identify individual case safety reports which require clinical review. Pharmacoepidemiol Drug Saf. 2022;31:1300–7.
56. Oosterhuis I, Rolfes L, Ekhart C, Muller-Hansma A, Härmark L. First experiences with a tool to measure the level of clinical information present in adverse drug reaction reports. Expert Opin Drug Saf. 2018;17:111–5.
57. Bergvall T, Norén GN, Lindquist M. VigiGrade: a tool to identify well-documented individual case reports and highlight systematic data quality issues. Drug Saf. 2014;37:65–77.
58. Berbain T, Pariente A, Miremont-Salamé G, Grandvuillemin A, Micallef J, Chouchana L, et al. Contribution of causality assessment for an automated detection of safety signals: an example using the French pharmacovigilance database. Drug Saf. 2020;43:243–53.
59. Edwards IR, Lindquist M, Wiholm BE, Napke E. Quality criteria for early signals of possible adverse drug reactions. Lancet. 1990;336:156–8.
60. Uppsala Monitoring Centre. Montelukast and Photosensitivity. WHO Signal. 2009;8–9.
61. Muñoz MA, Dal Pan GJ. The impact of litigation-associated reports on signal identification in the US FDA's adverse event reporting system. Drug Saf. 2019;42:1199–201.

62. Gravel CA, Douros A. Considerations on the use of different comparators in pharmacovigilance: a methodological review. Br J Clin Pharmacol. 2023;89:2671–6.
63. Caster O, Norén GN, Madigan D, Bate A. Large-scale regression-based pattern discovery: the example of screening the WHO global drug safety database. Stat Anal Data Min ASA Data Sci J. 2010;3:197–208.
64. Nielsen DL, Juhl CB, Nielsen OH, Chen IM, Herrmann J. Immune checkpoint inhibitor-induced cardiotoxicity: a systematic review and meta-analysis. JAMA Oncol. 2024;10:1390–9.
65. Moore N, Hall G, Sturkenboom M, Mann R, Lagnaoui R, Begaud B. Biases affecting the proportional reporting ratio (PPR) in spontaneous reports pharmacovigilance databases: the example of sertindole. Pharmacoepidemiol Drug Saf. 2003;12:271–81.
66. Hopstadius J, Norén GN. Robust discovery of local patterns: subsets and stratification in adverse drug reaction surveillance. In: Proceedings of 2nd ACM SIGHIT international health informatics symposium. New York, NY, USA: Association for Computing Machinery; 2012. p. 265–74. https://doi.org/10.1145/2110363.2110395.
67. Seabroke S, Candore G, Juhlin K, Quarcoo N, Wisniewski A, Arani R, et al. Performance of stratified and subgrouped disproportionality analyses in spontaneous databases. Drug Saf. 2016;39:355–64.
68. Woo EJ, Ball R, Burwen DR, Braun MM. Effects of stratification on data mining in the US Vaccine Adverse Event Reporting System (VAERS). Drug Saf. 2008;31:667–74.
69. Ahmed I, Dalmasso C, Haramburu F, Thiessard F, Broët P, Tubert-Bitter P. False discovery rate estimation for frequentist pharmacovigilance signal detection methods. Biometrics. 2010;66:301–9.
70. Park G, Jung I. A tree-based scan statistic for zero-inflated count data in post-market drug safety surveillance. Sci Rep. 2022;12:16299.
71. Russo M, Wang SV. An open-source implementation of tree-based scan statistics. Pharmacoepidemiol Drug Saf. 2024;33:e5765.
72. Fusaroli M, Polizzi S, Menestrina L, Giunchi V, Pellegrini L, Raschi E, et al. Unveiling the burden of drug-induced impulsivity: a network analysis of the FDA adverse event reporting system. Drug Saf. 2024;47:1275–92.
73. Norén GN, Meldau E-L, Chandler RE. Consensus clustering for case series identification and adverse event profiles in pharmacovigilance. Artif Intell Med. 2021;122:102199.
74. Gattepaille L. Using the WHO database of spontaneous reports to build joint vector representations of drugs and adverse drug reactions, a promising avenue for pharmacovigilance. 2019. p. 1–6.
75. Erlanson N, China JF, Taavola H, Norén GN. Clinical relatedness and stability of vigiVec semantic vector representations of adverse events and drugs in pharmacovigilance. Drug Saf. 2025;48:401–13.
76. Caster O, Juhlin K, Watson S, Norén GN. Improved statistical signal detection in pharmacovigilance by combining multiple strength-of-evidence aspects in vigiRank. Drug Saf. 2014;37:617–28.
77. Van Holle L, Bauchau V. Use of logistic regression to combine two causality criteria for signal detection in vaccine spontaneous report data. Drug Saf. 2014;37:1047–57.
78. Scholl JHG, van Hunsel FPAM, Hak E, van Puijenbroek EP. A prediction model-based algorithm for computer-assisted database screening of adverse drug reactions in the Netherlands. Pharmacoepidemiol Drug Saf. 2018;27:199–205.
79. Strandell J, Bate A, Hägg S, Edwards IR. Rhabdomyolysis a result of azithromycin and statins: an unrecognized interaction. Br J Clin Pharmacol. 2009;68:427–34.
80. Aiello V, Fusaroli M, Raschi E, Palazzini M, Hu L, Barbuto S, et al. Pulmonary embolism in a patient with ADPKD treated with Tolvaptan: from the clinical experience to the analysis of the Food and Drug Administration Adverse Event Reporting System registry. Kidney Int Rep. 2021;6:2472–7.
81. Tregunno PM, Fink DB, Fernandez-Fernandez C, Lázaro-Bengoa E, Norén GN. Performance of probabilistic method to detect duplicate individual case safety reports. Drug Saf. 2014;37:249–58.
82. van Stekelenborg J, Ellenius J, Maskell S, Bergvall T, Caster O, Dasgupta N, et al. Recommendations for the use of social media in pharmacovigilance: lessons from IMI WEB-RADR. Drug Saf. 2019;42:1393–407.
83. Harpaz R, DuMouchel W, Schuemie M, Bodenreider O, Friedman C, Horvitz E, et al. Toward multimodal signal detection of adverse drug reactions. J Biomed Inform. 2017;76:41–9.
84. Caster O, Dietrich J, Kürzinger M-L, Lerch M, Maskell S, Norén GN, et al. Assessment of the utility of social media for broad-ranging statistical signal detection in pharmacovigilance: results from the WEB-RADR project. Drug Saf. 2018;41:1355–69.
85. Fusaroli M, Salvo F, Begaud B, AlShammari TM, Bate A, Battini V, et al. The reporting of a disproportionality analysis for drug safety signal detection using individual case safety reports in pharmacovigilance (READUS-PV): explanation and elaboration. Drug Saf. 2024;47:585–99.