Drug Safety
https://doi.org/10.1007/s40264-025-01628-4

ORIGINAL RESEARCH ARTICLE

Causal Inference Tools for Pharmacovigilance: Using Causal Graphs to Identify and Address Biases in Disproportionality Analysis

Michele Fusaroli[1,2] · Joseph Mitchell[3] · Annette Rudolph[3] · Elena Rocca[4] · Riccardo Fusaroli[5,6,7]

Received: 28 August 2024 / Accepted: 14 October 2025

Abstract

Introduction Disproportionality analysis, finding associations in the co-reporting of drugs and events, is widely used in pharmacovigilance to detect potential safety signals of adverse drug reactions. However, inherent biases and unique data features often cause disproportionality to diverge from causation, and a comprehensive framework to address these issues is lacking.

Objective We showcase how directed acyclic graphs (DAGs) can enhance disproportionality analysis-related inferences, better qualifying its limitations and catalysing its inclusion in the broader evidence landscape.

Methods We introduce a DAG-based causal framework to systematically document and address biases in disproportionality analyses (e.g., confounding, colliders, measurement and reporting biases). We illustrate its application to case studies from the Food & Drug Administration (FDA) Adverse Event Reporting System—using the Information Component as a disproportionality metric and restriction as conditioning.

Results Directed acyclic graphs facilitate the formalisation of existing knowledge and causal assumptions, optimise the design of disproportionality analysis to mitigate biases—thereby enhancing sensitivity and specificity—improve transparency, better enable the formulation of critiques, highlight limitations of disproportionality and guide follow-up studies to address residual confounding and broader evidence synthesis.

Conclusion Using DAGs to map and mitigate biases requires caution and does not allow to obtain definitive answers to causal questions. Still, it results in more reliable and knowledge-based safety signals, reducing and mapping the gap between what we find (association) and what we look for (causation). Additional research should further tailor DAGs to pharmacovigilance challenges, map the generative mechanisms of pharmacovigilance data, and better integrate disproportionality analysis results into evidence-synthesis workflows.

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

1 Unit of Pharmacology, Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy

2 Pharmacovigilance Science Section, Research Department, Uppsala Monitoring Centre, Uppsala, Sweden

3 Signal Management Section, WHO Liaison Department, Uppsala Monitoring Centre, Uppsala, Sweden

4 Department of Life Sciences and Health, Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway

5 Department of Linguistics, Cognitive Science and Semiotics, School of Communication and Culture, Aarhus University, Aarhus C, Denmark

6 The Interacting Minds Center, School of Culture and Society, Aarhus University, Aarhus, Denmark

7 Linguistic Data Consortium, University of Pennsylvania, 3600 Market Street, Suite 810, Philadelphia, PA 19104-2653, USA

Key Points

Neglecting biases results in misleading or unqualified disproportionality results.

A cautious use of directed acyclic graphs facilitates evidence synthesis, increases the predictive value of disproportionality analysis and highlights its limitations.

1 Background

1.1 Adverse Event Reports

Reports of suspected adverse drug reactions (ADRs) are collected by pharmacovigilance systems to document real-world settings and complement observational studies and

Published online: 11 December 2025

randomised clinical trials (RCTs) [1], which frequently miss events related to drug interactions, comorbidity and specific populations [2, 3]. The data are standardised according to the International Council for Harmonisation (ICH) guideline E2B, and can include administrative and identification information, suspected ADRs, patient characteristics, concomitant drugs, indications, relevant medical history including tests and procedures, and narrative case summaries [4]. The generative process of adverse event reports is complex and difficult to track—even more than that of epidemiological records. The spontaneous reporting of potential ADRs, peculiar to pharmacovigilance, is affected by many poorly understood factors. Therefore, the rates calculated on reports are not representative of occurrence rates [5, 6] and quantifying causal effects from these data alone is not possible, but it requires careful integration into the epidemiological and clinical evidence landscape.

1.2 Disproportionality Analysis and Signal Management

Nevertheless, as millions of reports are collected annually, the generation of hypotheses concerning ADRs (*signal detection*) makes increasing use of fast quantitative analytical techniques beyond traditional and slower qualitative analyses (case-by-case and case series review [7–9]). Disproportionality analysis (DPA), on which this work focuses, detects signals of disproportionate reporting when the exposure to a drug (D) is disproportionally associated with the occurrence of an adverse event (E), as derived from reports, compared to what would be expected by chance if they were reported independently [1, 10]. In other words, the primary goal of disproportionality analysis is to identify drug-event combinations that are plausible ADRs. This relies on important assumptions that must be carefully assessed and should not have the pretence of quantifying the causal effect. The emerging working hypotheses, when *validated* through clinical and pharmacological reasoning, may justify further resource allocation for comprehensive *signal assessment,* encompassing all available evidence, and *evaluation* through tailored trials and epidemiological studies.

1.3 The Importance of Avoiding Overstatements

Associations inferred from adverse event reports might be biased by data generating mechanisms in common with epidemiology (e.g., confounding and measurement error), but also by mechanisms specific to the field: different reporting habits, duplicates, incomplete and unverified data [11], all mean that statistical association alone does not imply causation. Despite this, exaggerating the meaning of associations or neglecting measurement error and reporting (i.e., spin) is common in pharmacovigilance [12]. To avoid unwarranted claims, more conservative practices confine DPA to finding associations and refrain from discussing causality [13]. Aside from occasional—often poorly justified—sensitivity analyses, biases are typically relegated to the "limitations" sections, and causal implications are shallowly avoided by adopting terms like "association", "link", "increased reporting", "correlation", and "pattern".

1.4 The Harm Related to Understatements

Paradoxically, avoiding causal discussions might foster a comfort zone where the impact of biases and the conditions under which causality could be inferred remain under-explored. Pharmacovigilance aims to minimise medicine-related harms and should not be satisfied with merely identifying associations. The ease of conducting DPA generates thousands of otherwise unqualified signals, which can be misinterpreted by researchers and the public, especially if biases are only superficially acknowledged and not properly discussed [14, 15]. Even regulators can be easily overwhelmed by the amount of disproportionality-generated signals, despite existing prioritisation strategies [16, 17]. Nonetheless, reports and DPA can still provide important—however preliminary—evidence of medicine-related problems. Dismissing them and expecting—even in principle—causality to be only inferred through RCTs or pharmacoepidemiological studies [18] is problematic, since timely warnings might have important implications.

Recent work by Scosyrev and colleagues algebraically formalises the strong, and often unrealistic, assumptions required for DPA metrics to equal a causal risk ratio, concluding that most signals are likely non-causal [5]. Rather than seeing this as a reason to abandon causal inquiry, we take it as a call for more structured reasoning. If these assumptions are consistently violated, it is imperative to use tools that can: a) formally map why they are violated (e.g., confounding, selection bias [6]), b) guide analyses that attempt to mitigate these issues, and c) catalyse better efforts to synthesise evidence from multiple sources to better identify the effects. As Bradford Hill emphasised: *"All scientific work is incomplete—whether it be observational or experimental. All scientific work is liable to be upset or modified by advancing knowledge. That does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time"* [19]. In such scenarios where both overstatements and understatements can lead to problematic conclusions, it is essential to use tools that qualify causal claims by explicitly connecting them to underlying assumptions and data patterns [20].

1.5 Causal Inference: Qualifying Statements on Underlying Assumptions

Directed acyclic graphs (DAGs) have emerged across disciplines like computer science, epidemiology, economics, and psychology to provide a structured framework for addressing causal questions [21–24]. In principle, their utility lies in offering a true representation of causal relationships in the world. However, as human knowledge of these relationships is generally limited, DAGs are crucial to visually represent the assumed relationships between variables based on the researcher's current knowledge or beliefs. Thus, DAGs make assumptions explicit, identify potential biases, and coherently guide analytical decisions. Even after a study has been performed, DAGs can highlight relevant missing information, catalyse further evidence synthesis to fill these gaps, and enable easier communication and collective scientific development.

Crucially, the use of DAGs in DPA cannot per se quantify causal effects nor provide definitive answers to causal questions from adverse event reports alone. The DAGs do not constitute a new statistical estimator requiring validation, but rather an epistemological framework for structuring, developing and communicating the assumptions that underlie any disproportionality analysis. DAGs are tools to enhance the rigor of causal reasoning by bridging data and theory, allowing a pragmatic approach to research that explicitly builds on previous evidence, accounts for uncertainty, and can be more easily criticised and further developed by successive studies. In other words, the goal of such causal inference framework is to make analytic reasoning transparent and criticisable, moving the field towards a more rigorous and cumulative scientific process of critical epistemic iteration, in which models are iteratively proposed, criticised and built upon [25].

1.6 Aim

In this paper, we propose integrating DAGs within DPAs to tackle the challenges of providing qualified causal hypotheses, anchored in clearly articulated assumptions and sources of uncertainty. Qualitative studies already use causal inference techniques to address these challenges in individual case assessment [7] and case-series, either by excluding alternative causes or by following Bradford Hill's viewpoints [26, 27]. On the contrary, explicit implementation of causal inference techniques in DPA is rare (e.g., to address confounding by indication and concomitance [11]). Causal inference techniques can help pharmacovigilance better fulfil its goal by clarifying the conditions under which an identified disproportionality could support a potential causal relationship between a drug and an event. They highlight, the information that is still missing, the potential sensitivity to biases, the inability to correct for unmapped confounders, the need for improved information capture, and the important aspects that should be accounted for when planning the integration of further evidence sources, including designing evaluation studies [13, 14]. This approach prioritises the reliability of identifying potential causation rather than estimating effect magnitude (or influence [28]), given the unique features of reports. We here provide a tutorial demonstrating the use and utility of DAGs, discussing their current benefits and limitations for DPA, providing indications for deepening one's understanding on how to use causal inference, and suggesting future research directions.

2 Materials and Methods

To showcase the integration of causal inference with DPA, we introduce DAGs as a formalised tool to address both simple and complex causal scenarios. To maintain a clear pedagogical flow, we first (Sect. 3.1) isolate the causal inquiry from the effects of sampling variability and structural biases (confounders, colliders, measurement errors, and reporting biases), which are then gradually addressed in Sect. 3.2 and exemplified in Sect. 3.3. The DAG examples provided are thus initially simplistic representations of potential effects for pedagogical purposes and should not be taken as generalisable DPAs. They increase in complexity throughout the article, up to the complex example of Sect. 3.4. These examples from the Food and Drug Administration Adverse Event Reporting System (FAERS) crucially demonstrate how biases can be tackled beyond mere acknowledgment. Even though causal models may not fully reflect reality in all its complexity (i.e., may not be completely accurate), they can still capture important aspects of the clinical reality and of the data generation process and therefore be useful to address causal questions [29].

We used DAGs to explicitly represent expected biases and identify the most suitable study design to address them (i.e., using D-separation criteria; see Sect. 3.2). We calculated association as disproportionality using both a crude study design (with the entire FAERS as background) and a more nuanced design (selecting the background to reduce variability in variables that introduce distortion). We compared the results using forest plots to illustrate how DAGs can help predict and address biases in pharmacovigilance, given an appropriate causal model.

As a disproportionality metric, we used the Information Component (IC), often conceptualised as a Bayesian measure, with a prior centred on 0, which deals with chance associations due to limited sampling by conservatively shrinking associations towards the null, albeit at the cost of distorting the effect size estimate [30] (see Supplementary Table S1). A signal of disproportionate reporting is identified by the IC

when the lower limit of its 95% credibility interval $IC_{025} > 0$.[1] Consistent with current pharmacovigilance practices, we interpret disproportionality results as either significant or non-significant regardless of effect size.

Analyses were performed using the DiAna R package version 2.1.0: an open-access toolkit for DIsproportionality ANAlysis and other pharmacovigilance investigations in FAERS [31], considering data from January 2004 to March 2023. Directed acyclic graphs were drawn using functions made available through the PVdagger R package version 1.0.0: visualisation of causal structures in pharmacovigilance data using DAGs [31]. A step-by-step tutorial, including the R code and its description, is provided in the supplementary materials for full transparency and reproducibility.

3 DAGs Framework for DPA

3.1 DAGs to Explore Simple Scenarios

Directed acyclic graphs are a graphical representation of the assumed causal relations between variables relevant to the study. Directed acyclic graphs represent variables as nodes and causal connections as directed edges (i.e., arrows), providing a structured framework for predicting which associations we would observe if the DAGs were accurately reflecting reality. Arrows indicate generic causal relations, which could take any mathematical form (e.g., $y = 3x$; or $y = 2 + e^x$). Often researchers do not yet know the specific mathematical relation, beyond that there is a relation. Nevertheless, DAGs by definition enforce the key principle of acyclicity, that is, that:

- The connection between two links can only be in one direction (e.g. if x causes y, y cannot cause x).
- There is no cyclical path (i.e., a path starting in one variable and ending back in the same variable by following the direction of the connections).

If DAGs respect this principle, even in absence of further knowledge about the specific mathematical form of the connections, they can be used to inform scientific inferences, e.g. identifying confounders and colliders. If more is known, scientific inferences can be further informed. We start with the simple case of acyclicity, but, for instance, in Sect. 3.2, we showcase that if we know or assume that links identify monotonic homogenous relationships, then we can talk about signed DAGs, which have useful properties for identifying the direction and impact of biases.

Consider the hypothesis that D is a direct cause of E (Fig. 1). $D \overset{?}{\rightarrow} E$ formalises our inquiry, contrasting the hypothetical scenarios where D causes E ($D \rightarrow E$) and where D does not cause E ($D\ E$). Under the assumption of infinite sample sizes and no distortion from external variables or measurement error, statistical findings are assumed to be "faithful": causal dependence reliably manifests as statistical associations [32]. Under these assumptions, lack of statistical association suggests lack of causal relationship (panel A), while statistical association suggests causal relationship. Even in this over-simplified scenario, statistical association cannot discriminate between direct causality (panel B, where drug D causes event E) and reverse causality (panel C, where condition E causes drug D administration), which can only be discerned by domain-knowledge, considering factors like on-label or off-label drug use and the possibility of paradoxical reactions [33].

3.2 DAGs to Explore Complex Scenarios

Once introduced the simple scenarios of Sect. 3.1, we can gradually show how DAGs help navigate more realistic causal scenarios involving multiple interconnected factors and complex relationships [11]. In particular, we will rely on the following features from traditional DAGs (see Fig. 2, left panel, and Sect. 3.3 for concrete examples):

- *Front-door paths:* $D \rightarrow E$ and $D \rightarrow M \rightarrow E$ (indirect), with M as a mediator of D's effect on E.
- *Back-door path:* $D \leftarrow C \rightarrow E$ identifies C as a common cause of both D and E (i.e., a confounder).
- *Collider*: $D \rightarrow F \leftarrow E$ identifies F as a common effect of D and E.
- *Independent risk factor*: $J \rightarrow E$ identifies J as a competing exposure.

Front-door paths, regardless of the number of mediators, correspond to the DAG in Fig. 1B, indicating that the event is a reaction to the drug. However, even excluding reverse causality, a reaction is not the only possible explanation for a statistical association between D and E. Back-door paths allow the flow of information, as variability in the confounder affects both the exposure and the outcome, potentially leading to statistical dependence even without direct causal relationship. In contrast, colliders only affect inference if included in the statistical model, and competing exposures never affect the flow of information.

In complex scenarios with multiple elements, we can predict which statistical associations should be present if the DAG was accurate using D-separation criteria [23]: a statistical association indicates a causal association only when

[1] Note that according to Bayesian conventions and IC literature [30], we discuss credibility and not confidence intervals, that is, the range of estimated values compatible with the observed data given the model assumptions.

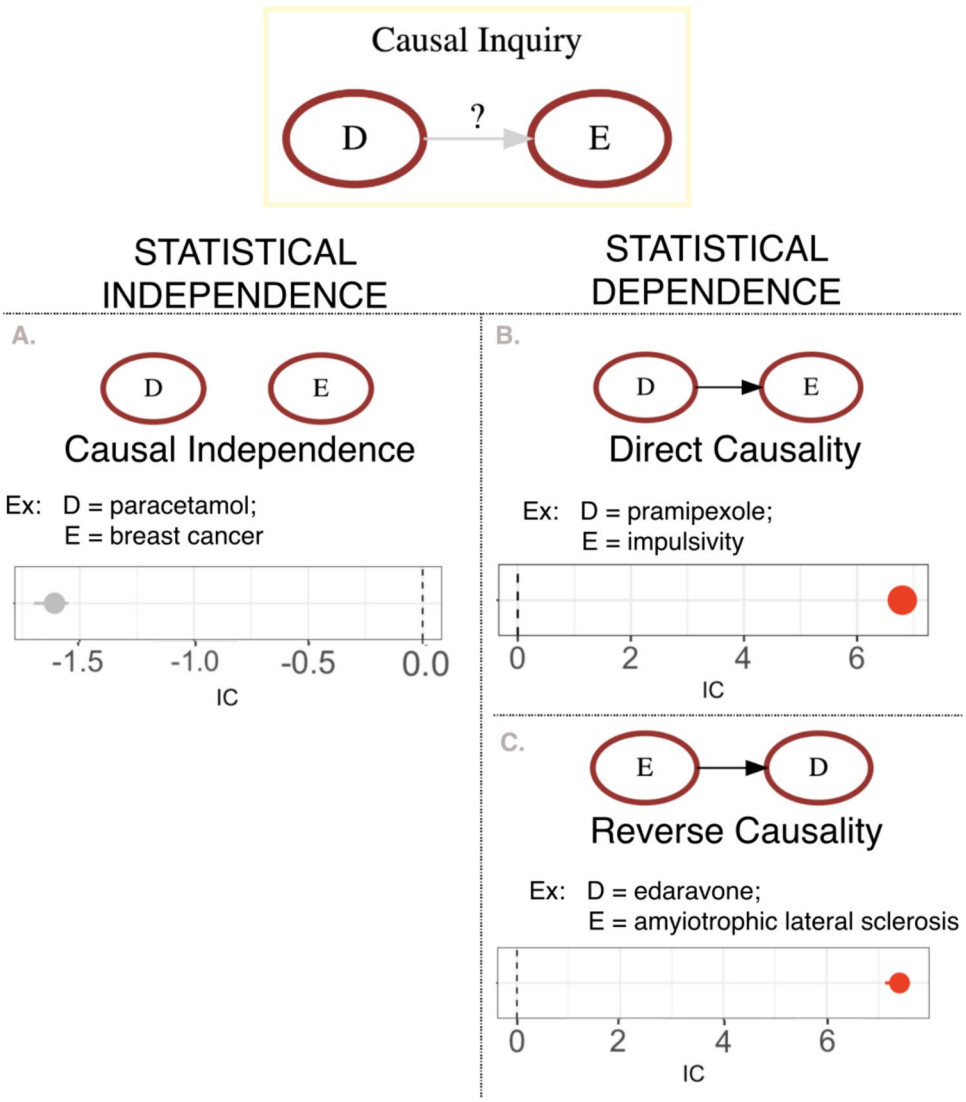


Fig. 1 Statistical and causal dependence: our inquiry (grey arrow with a question mark) concerns the existence of a direct causal relationship between drug D and event E. Assuming, for the sake of the example, a simplified scenario with infinite sample size and no systematic bias, there are only two possible scenarios: statistical independence suggesting causal independence (panel **A**) or statistical dependence suggesting causal dependence (panels **B** and **C**). Disproportionality analysis (DPA) helps to discriminate between these two scenarios but cannot distinguish between (**B**) direct and (**C**) reverse causality on its own. Forest plots display the information component (IC) on the *X*-axis, with a vertical dashed line indicating the threshold for statistical significance. The result of the DPA is shown as a dot (placed at the IC point estimate) with a 95% credibility interval. An association exists when the entire credibility interval is above the threshold (red vs grey). Note that these introductory examples still ignore reporting biases and therefore should not be considered generalisable approaches to DPA analyses

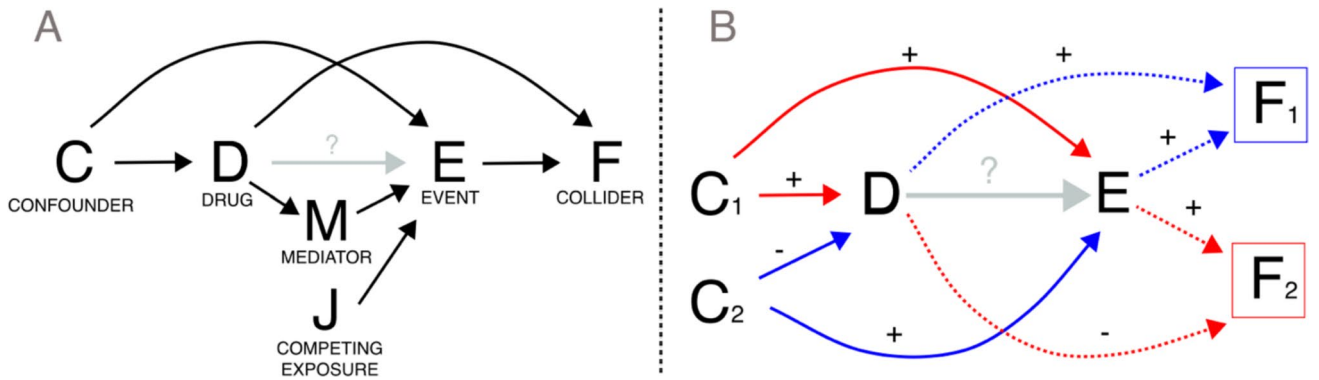
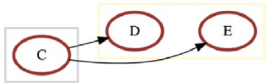
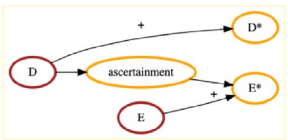
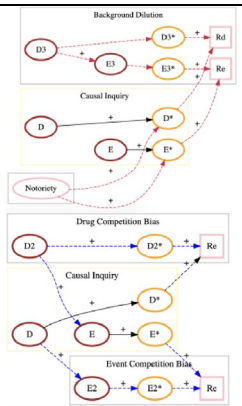


Fig. 2 Complex directed acyclic graphs (DAGs). In the left panel (**A**), we show DAG components. In the right panel (**B**), we show four possible scenarios involving two different bias directions: red paths (C_1: positive confounder, F_2: positive collider) potentially result in false positives; blue paths (C_2: negative confounder, F_1: negative collider) potentially result in false negatives. Enclosing the variable in a square means it has been conditioned on

Table 1 Mechanisms potentially underlying association. The table shows the different mechanisms that can underlie a signal of disproportionate reporting, formalised using directed acyclic graphs (DAGs)

Mechanism		Explanation	Prototypical DAG	Examples
Direct Causality		D causes E.		**Adverse drug reaction**
Reverse Causality		E causes the prescription of D.		**Exposure driven by the event** (approved, off-label, or improper use).
Systematic Biases	Confounding Biases (Section 3.3.1)	D and E have a common cause C.		**Confounding** by: • indication/contraindication • sex/age • comorbidity/coprescription
	Collider Biases (Section 3.3.2)	D and E have a common effect F which is conditioned on.		**Collider bias** by: • prophylaxis • outcome
	Measurement Error (Section 3.3.3)	Taking the drug D affects the chance of detecting and reporting E, or vice versa.		**Measurement errors** for exposure and outcome. **Noise** from confounder and collider measures. **Composite measures** **Ascertainment bias** (by drug or event)
	Reporting Biases (Section 3.3.4)	Any occurrence, to be reported, needs at least one suspected drug (Rd) and one event (Re). Moreover, external factors (such as reporter education, motivation, and resources) can influence whether an occurrence is reported or not. This selection, related to the peculiar generative process of adverse event reports, potentially introduces colliders and confounders as we include other drugs and events in the causal models.		**Positive reporting biases** • ***notoriety bias***: >> reports with both D and E (DE*) • ***background dilution***: >> reports with other drugs and events (nDnE*) **Negative reporting biases** • ***event competition bias***: >> reports with same drug but other events (DnE*) • ***drug competition bias***: >> reports with same event but other drugs (nDE*)
Chance association		A random sample can by chance show an association. Particularly with few cases.		

all back-door paths are blocked (suppressing the information flow), and colliders are not accidentally opened [34, 35], ensuring that information flows only through front-door paths. To change the status of a path (close a back-door or open a collider) we apply conditioning procedures. Randomisation of exposure would be the gold standard for closing all known and unknown back-door paths. However, this is not feasible with adverse event reports, which are observational in nature. In the analysis of reports we can achieve conditioning through background restriction (limiting analysis to a subset that is homogeneous for the value of the confounder) and covariate adjustment (including an interaction term between exposure and confounder).

The D-separation criteria follow the direction of causality, independently of whether the cause increases or decreases the effect. However, if we assume that causal links follow a monotonic function (where an increase in the exposure consistently influences the outcome in the same direction) and effect homogeneity (no difference in effect directions across subpopulations), we can represent signed DAGs, where arrows indicate direct (+) or inverse (−) proportionality (see Fig. 2, right panel) [36]. These assumptions, while non-trivial and requiring cautious assessment, add valuable information to DAGs, expanding the range of possible inferences regarding the direction of biases and the expected impact of conditioning on different variables.

In particular, when dealing with confounders, we can expect that back-door paths with an even number of—or zero—negative signs create positive confounding (C_1), which in a lack of true causal relation could potentially lead to a false positive (red path, affecting specificity). Conversely, an odd number of negative signs creates negative confounding (C_2), which in the case of a true causal relation could potentially lead to a false negative (blue path, affecting sensitivity). The expectation is reversed for colliders: an even number of negative signs indicates negative collider bias, while an odd number indicates positive collider bias.

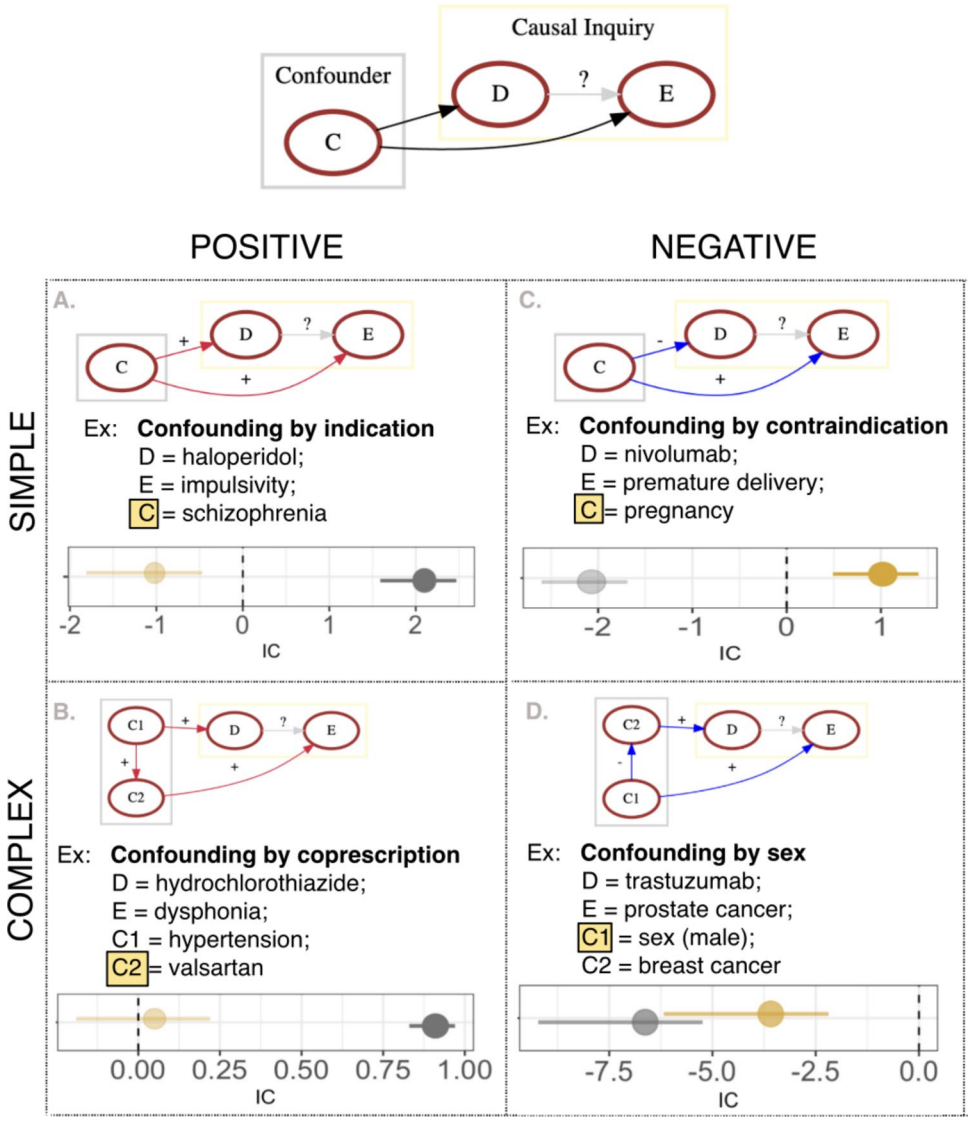


Fig. 3 Confounding biases. Confounding can introduce either a positive (red) or negative (blue) distortion, and can result from more or less complex paths. Prototypical structures are shown here together with examples showing the effect of restricting on reports recording the confounder (gold) compared to the crude analysis on the entire FAERS (gray). With positive confounding, conditioning decreases the information component (IC). With negative confounding, conditioning increases the IC. Note that these exemplary biases—established and named in the literature–are not exhaustive, but share analogous causal structures in the graphs. In all these cases, different subset populations are differently exposed to the drug and differently susceptible to the event. Therefore, conditioning on the subset population is necessary to account for the bias. The yellow square marks the variable used for conditioning. Note that these introductory examples still ignore reporting biases and therefore should not be considered generalisable approaches to DPA analyses

As the underlying assumptions do not always hold, these predictions require caution.

Using these basic building blocks, Table 1 provides an overview of the mechanisms that could underlie an observed statistical dependence, based on common pharmacovigilance biases, formalised using a DAG framework. In the next section, we demonstrate, through concrete examples, how this DAG framework enables a more systematic accounting of biases using consistent D-separation criteria.

3.3 Formalising Different Kinds of Bias and Examples

3.3.1 Confounding Bias

Many biases commonly observed in pharmacovigilance can be reframed as confounding biases, such as confounding by indication (Fig. 3, panel A), co-prescription (panel B), contraindication (panel C), and sex (panel D). Confounding biases are open back-door paths through which information flows even in the absence of causal dependencies. To close the back-door, we must condition on any nodes along the back-door path or, if not possible, on an effect of the confounder as a proxy (a surrogate confounder, which shares some of its variance).[2]

The flow of information introduced by a confounder can positively distort the association (e.g., individuals exposed to D might have been so due to C, which also makes them susceptible to E). This positive distortion can result in false positives, if of sufficient magnitude and in the lack of a causal effect.

Consider a case of indication bias (panel A). Even if, based on current knowledge, haloperidol does not cause

[2] We could condition on a cause of the confounder as well, which is a common cause and therefore a confounder on its own.

impulsivity, haloperidol and impulsivity are positively associated in FAERS. Schizophrenia is both an indication for haloperidol and a cause of impulsivity, thus introducing positive confounding. Restricting to reports recording schizophrenia among indications (closing the back-door), the association disappears (gold), suggesting a lack of causation[3] and allowing us to discard a spurious signal of disproportionate reporting.

Similar considerations apply to confounding by co-prescription (or, potentially, comorbidity), even if the causal diagram is slightly more complex (panel B). When examining the potential role of hydrochlorothiazide in inducing dysphonia, we find a statistical association. Nonetheless, its co-prescription with valsartan (a known cause of cough-related dysphonia), for hypertension, introduces positive confounding. Restricting to reports mentioning valsartan, the association vanished (gold), suggesting that hydrochlorothiazide-related dysphonia may be solely due to co-prescription with valsartan.

The flow of information can also negatively distort the association (e.g., C, while causing the exposure to D, reduces the susceptibility to E), potentially resulting in false negatives.

Consider a case of contraindication bias (panel C). When investigating the role of nivolumab in causing premature delivery, we find a negative statistical association, indicating no evidence of positive association according to pharmacovigilance standards (see footnote 2). Pregnancy is a necessary condition (and therefore a cause) for premature delivery, but it is also a partial contraindication to using nivolumab (due to lack of safety evidence[4]). This negative confounder means that nivolumab is less used in pregnancy and therefore less co-reported with premature delivery. Restricting the background to reports concerning pregnant patients reveals a positive association between nivolumab and premature delivery (gold). Not conditioning on the confounder would have led to missing a potential signal.

Similar considerations apply to the case of confounding by sex shown in panel D. When assessing the potential role of trastuzumab in causing prostate cancer, we find a negative association. But trastuzumab is primarily prescribed for breast cancer, and therefore mainly to females, who have a null risk of prostate cancer. Restricting to *males*, we address this negative confounding and observe an increase in the IC (gold), even if not of magnitude sufficient to identify a signal of disproportionate reporting.

3.3.2 Formalising Collider Bias

Collider biases can arise from various scenarios, including restricting to an outcome (Fig. 4, panel A) or a prophylactic measure that also serves as a treatment (panel B). Collider biases occur when both drug D and event E independently affect a third variable F (a common effect). When F is not conditioned upon, no information flows through it, and no bias is present. Conditioning on F opens a path and introduces bias.

A positive collider bias can occur, for example, when D reduces the chance of F, which is instead made more likely by E. Investigating the causal dependence between Chimeric Antigen Receptor T (CAR-T) therapy and infections shows a positive association in FAERS (panel A). If we restrict the analysis to fatal reports, we introduce a positive collider (red), as CAR-Ts improve survival, but infections pose a risk of death.

Conversely, a negative collider bias can occur when both D and E increase the chance of F. In this case the occurrence of F makes it more plausible for E to have occurred if D has not. Consider the known role of aspirin in inducing gastric ulcers (panel B). In the entire FAERS, we see a positive association. Knowing that proton pump inhibitors (PPIs) are used to prevent aspirin-related gastric ulcers, we might restrict the analysis to reports mentioning them. However, since they also treat existing gastric ulcers, patients taking PPIs, but not aspirin, are more likely already under treatment. This negative collider bias deflates the IC when restricting to PPIs, risking losing potential signals.

Conditioning on a collider is a well-known source of bias in causal inference, famously demonstrated in examples like the 'low birth weight paradox' [37]. This highlights the critical awareness that by incorrectly specifying the causal model before conditioning on variables might result in incorrect inferences. And yet specifying one's model even when incorrect allows for more transparent communication of the assumptions and an easier discussion and improvement of the assumed models within the scientific community.

3.3.3 Formalising Measurement Errors

So far, we have pedagogically assumed direct access to constructs (e.g., drug exposure, event occurrence), but reports only provide measurements (here denoted with an asterisk *;

[3] The observed negative association (i.e., an association with the upper limit of the credibility interval lower than the significance threshold) might suggest a protective effect of haloperidol against impulsivity. This seems plausible: by treating schizophrenia, which also causes impulsivity, haloperidol reduces impulsivity. However, in pharmacovigilance, a cautionary principle suggests not to interpret negative associations because: a) DPA aims to identify early warnings of ADRs, not unexpected therapeutic effects; b) DPA compares reports not with control but with non-cases (other reports of suspected ADR). Whether a causal inference informed pharmacovigilance could make better use of negative associations is outside the scope of this article.

[4] And, plausibly, there is also a lower frequency of cancer in reports recording pregnancy than in other individual case safety reports (ICSRs).

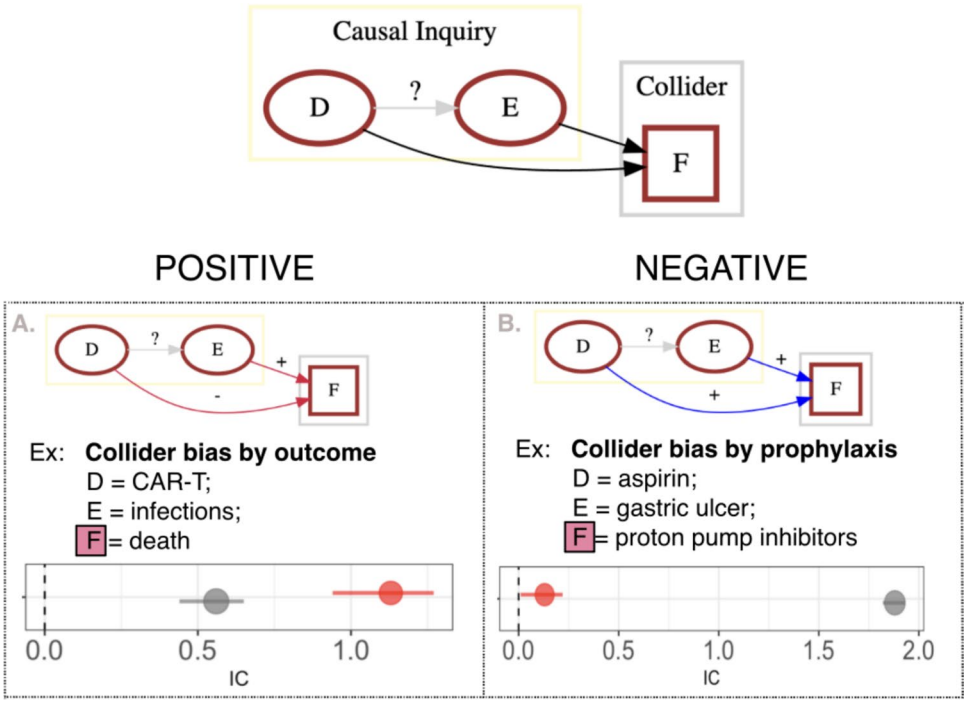


Fig. 4 Collider bias. Colliders can introduce either a positive (red) or negative (blue) distortion. Prototypical structures are shown here together with examples showing the effect of incautiously restricting on reports recording the collider (red) compared to the analysis on the entire Food and Drug Administration Adverse Event Reporting System (FAERS). Positive colliders increase the Information Component (IC), negative colliders decrease it. Avoiding conditioning on colliders is crucial to not introduce biases. The red square marks the collider used for conditioning

Fig. 5) [38–41]. Measurements act as surrogate variables for their underlying constructs, and disproportionality (association between D*–measurement of D– and E*–measurement of E) suggests a reaction when our conditioning makes it so that D causing E remains the only possible explanation. While all previous considerations still apply, measurements introduce additional complexity.

Conditioning on surrogate variables (e.g., C* or F*) has the same effect as conditioning on the underlying construct (e.g., C or F) to the extent that the construct and measurement share relevant variance. This is often not the case in reports due to measurement errors, which introduce noise because of inaccurate recording or, more often, incompleteness. As a rule of thumb, if something is reported, it likely happened; if not reported, we cannot draw firm conclusions. Incomplete reports can result from factors such as complexity and time demands of submitting a report [42–44], mandatory fields limited to the suspected drug and reaction, and heterogenous encoding (e.g., misclassification or ambiguity [45]). Conditioning on an unreliable proxy of a confounder fails to close the back-door (panel A, and C for similar considerations on colliders). The inadequacy of the proxy is a serious concern, particularly in pharmacovigilance, where the lack of a confounder cannot be inferred from the lack of its reporting (e.g., a drug not reported does not imply it was not taken). In such cases, restricting to reports in which the confounder was reported is more effective than restricting to reports in which it was not reported (note that the latter did not diverge from the crude analysis on the entire FAERS in panel A). Subsequently, even pooling together results from different strata, as in the case of pooled stratification or adjustment, should be avoided in pharmacovigilance whenever any of the strata are derived from lack of information. Still, restricting can result in a small number of cases, leading to wide credibility intervals and loss of significance. Sometimes it is possible to combine multiple measurements that are surrogates of the same confounder to increase the sample size when there are missing data. For instance, investigating whether cisplatin causes thrombosis shows an association, but cancer is a likely positive confounder (panel B). Conditioning just on the indication of cancer or on concomitant antineoplastic agents may reduce sample size and accuracy due to missing information. To address this challenge, we might restrict to reports that satisfy either criterion (gold).

Another source of complexity emerging from measurements is ascertainment bias (panel D), which occurs when drug exposure increases the chance of detecting and reporting a specific event (or when manifesting an event increases the chance of reporting a specific drug). Clozapine's risk of agranulocytosis leads to frequent blood tests and a higher probability of detecting and reporting hyperkalaemia. Among reports not recording hyperkalaemia or clozapine, it is plausible that several cases were undetected due to the absence of blood tests, which inflated the IC. In fact, the crude analysis shows a higher (although non-significant) IC compared to when restricting to reports recording any laboratory investigations (gold). Beyond this example—which illustrates ascertainment bias due to increased monitoring in exposed patients—numerous other factors, often

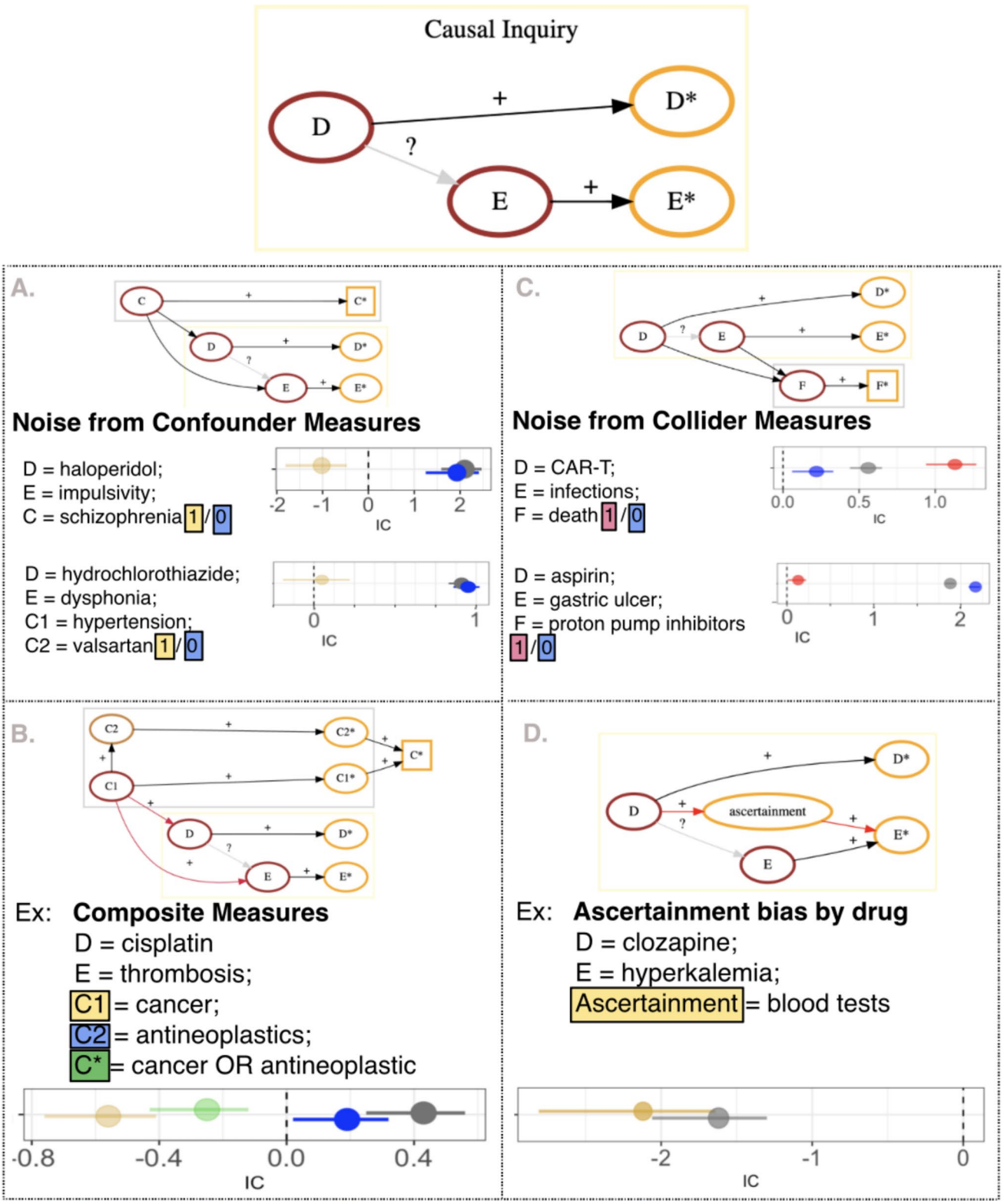


Fig. 5 Measurement errors. When analysing adverse event reports, we rely on measurements (surrogate variables) rather than actual constructs. Reporting provides more information than the lack of it, and restricting on the lack of reporting of a variable does not effectively address (or introduce) biases (**A** and **C**). When conditioning on the reporting of a variable results in too few cases and wide credibility intervals, composite measurements can be used to map constructs (**B**). In ascertainment bias (**C**), drug exposure increases the detection of an event (or vice versa). Coloured squares mark subgrouping variables

unmeasurable or unrecorded in adverse event report data, may influence the association between drug exposure and the reporting or coding of the outcome (or conversely, between the outcome and the reporting or coding of the drug), including differential access to healthcare.

3.3.4 Formalising Reporting Biases

A specific challenge in adverse event reports is the reporting process. For an instance to be reported, a drug must be taken, measured, and reported (Rd) and an event must be manifested, measured, and reported (Re). All drugs and events in the database contribute to Rd and Re, potentially creating colliders, already open due to reporting (i.e., drug exposure without adverse events and events without drug administration remain unobserved). Additionally, reporting depends on factors such as awareness, motivation, reporter occupation, pharmacovigilance system, and clinical setting [42–44]. The resulting reporting biases can introduce positive or negative distortions based on the affected cell of the contingency table (Fig. 6). Traditionally, the solution is to restrict the analysis to unaffected reports.

Reporting biases that inflate reports with D* and E* (panel A) positively distort the IC. An example is notoriety bias resulting from regulatory warnings or media attention [46, 47], which increases the reporting rate of both D* and E*. Restricting to reports collected before the warning mitigates this bias (gold) [26, 46]. As a proof of concept, restricting to reports following the regulatory warning further inflates the IC (blue). A similar example of notoriety bias resulting in the emergence of a spurious disproportionality occurred after the fraudulent publication on the association between autism and measles-mumps-rubella vaccine by Wakefield in 1998 [6]. Another example is precautionary bias, which involves an increased reporting rate from patient support programs [48]. Conversely, reports may occur without an actual event—as in cases of superficial diagnostic verification (e.g., reporting drug-induced liver injury without verifying all diagnostic criteria) and malicious reporting (as speculated for court litigations [49])—or may proliferate from a single report into many duplicates [50, 51]. Despite potential biases, all reports should be encouraged, and bias considerations should be limited to analysis, not data collection.

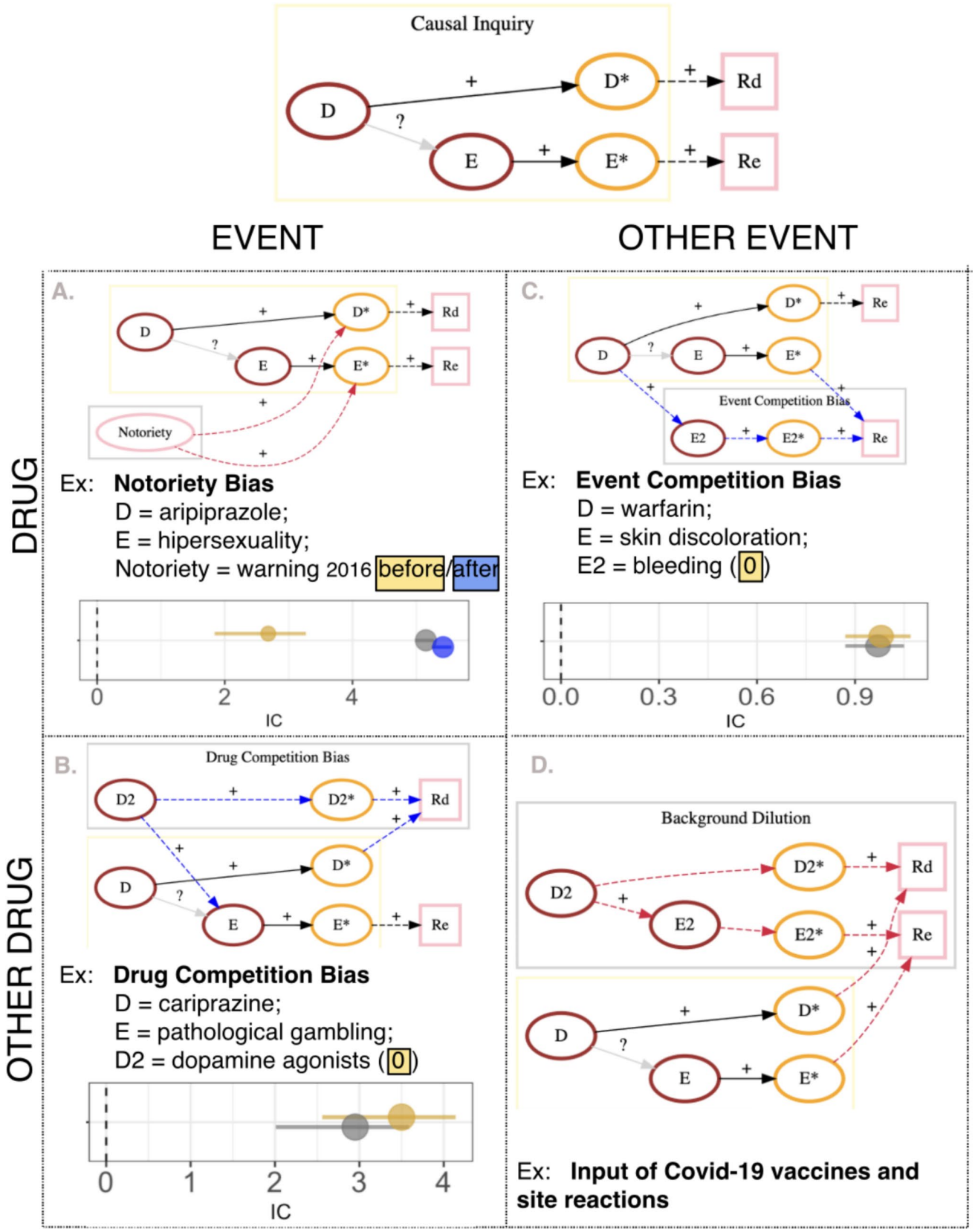
Masking/cloaking (panels B and C) includes drug competition bias (inflating reports with E* related to another drug) and event competition bias (inflating reports with D* and another event) [52], which can be formalised as colliders at the level of Rd or Re. Masking introduces a negative distortion and can be tackled by excluding competing drug-event pairs from the database [53]. For instance, the statistical dependence between cariprazine and gambling increases when removing competing dopamine agonists reports (panel B). Similarly, the association between warfarin and skin discoloration increases when removing competing bleeding reports (panel C). These masking effects are less pronounced than other discussed biases, consistent with studies questioning their significance [54], but unmasking can still allow an earlier detection of an ADR [6].

Background dilution (panel E) is a less characterised bias where the IC is inflated by a sudden influx of reports concerning another drug and another event. It can be formalised as a double collider, where the drug and the event are associated because another drug in the database causes another event, and these two are linked to the drug and event of interest through the colliders Rd and Re. The presence of the double collider inverts the sign twice, resulting in a positive distortion. While more common in small databases, background dilution was also observed in the WHO global database of adverse event reports for medicines and vaccines, VigiBase: the solicited reporting of events following immunisation with COVID-19 vaccines at the same time tapered signals for other drugs and related events (drug-competition bias) but also boosted associations for other drugs and unrelated events [55].

3.4 A Workflow for Systematically Using DAGs in Complex Scenarios

Having introduced DAGs for both simple and complex causal scenarios, this section presents a workflow for their systematic use, applied to a realistic example involving an established ADR: aripiprazole-induced impulse control disorders [56, 57] (Figs 7 and 8).

The first steps (a1–2) involve defining the drug (e.g., aripiprazole) and event under investigation (e.g., impulse control disorders). Next, consider the possibility that the drug might be prescribed to treat the event, indicating reverse causality (a3). Although aripiprazole may be prescribed off-label to prevent impulsivity in head trauma [58], this is uncommon and unlikely to explain an association. Aripiprazole is also prescribed for conditions that commonly manifest with impulsivity (e.g., psychotic and mood disorders). Even if impulsivity is not generally the reason for administering aripiprazole, disproportionality cannot exclude an association due to peculiar coding practices (e.g., reporting impulse control issues as adverse events even if they might have pre-existed the drug administration). A case-by-case analysis may be essential to understand the relationship between aripiprazole and impulsivity when reported together. Still, the indication is a common cause (i.e., a confounder), and to reduce its variability we focus on bipolar disorder. Confounders are placed on the leftmost layer of the graph (a4). Common effects (a5, colliders) are added to the right to preserve the left-to-right temporal flow. In this case, no colliders were identified. Unobserved


Causal Inquiry
D
?
+
D*
Rd
E
E*
Re
EVENT
OTHER EVENT
DRUG
OTHER DRUG
A.
Notoriety
Ex: Notoriety Bias
D = aripiprazole;
E = hipersexuality;
Notoriety = warning 2016 before/after
IC
C.
Event Competition Bias
E2
E2*
Ex: Event Competition Bias
D = warfarin;
E = skin discoloration;
E2 = bleeding (0)
IC
B.
Drug Competition Bias
D2
D2*
Ex: Drug Competition Bias
D = cariprazine;
E = pathological gambling;
D2 = dopamine agonists (0)
IC
D.
Background Dilution
Ex: Input of Covid-19 vaccines and site reactions

◂**Fig. 6** Reporting bias. Acknowledging in causal graphs that reporting at least one drug (Rd) and at least one event (Re) are minimal criteria for reports helps address reporting bias. Notoriety bias (**A**) increases reports with D* and E*, inflating the information component (IC). Masking results from drug competition bias (increased reports with E* and another drug, **B**) or event competition bias (increased reports with D* and another event, **C**). Background distortion (**D**) occurs particularly in small databases: the influx of COVID-19 vaccine reports with site reactions (i.e., another drug and another event) increases the IC of other drugs with events unrelated to site reactions

variables are explored and linked to surrogate variables capturing part of their variability (a6). Since the indication is often unreported, lithium is used as a surrogate of bipolar disorder. Other drugs used for bipolar disorder are not included because they are less specific. The measurements layer is added on the right. Then, we should assess how well measurements capture the underlying constructs (a7). Drug names in FAERS are reported as free text and mapped to active ingredients during preprocessing, introducing subjectivity. For aripiprazole, 443 terms including "Abilify", "aripiprazole", and "aripipazole" were used. Similarly, ICDs are coded using MedDRA, and we referred to a multiterm query from a scoping review tailored to FAERS [59] to address redundancy and coding variability. For the confounder, a composite measure looking for either the high-level group term (HLGT) manic and bipolar mood disorders in the indication or lithium among the concomitants was used. A reporting bias layer was introduced (a8), completing the DAG. We included notoriety bias due to a 2016 FDA regulatory warning [56] and masking by dopamine agonists (ATC N04C), commonly used in Parkinson's disease and known to cause impulsivity [60]. Finally, arrows are annotated with signs to indicate assumed monotonic relationships where applicable (a9). Bipolar disorder increases the chance of both being prescribed aripiprazole and manifesting impulsivity, marked with a +.

Once knowledge-based DAGs are established, we should be careful to still consider them a "work in progress", a tentative step towards a better understanding of the true underlying mechanisms. While keeping this caution in mind, knowledge-informed DAGs can drive the design of the DPA by identifying potential confounders and colliders, thereby streamlining conditioning techniques. The resulting DAG indicates the best possible analysis design based on given data and causal assumptions, identifying variables for subgrouping to close back-doors and open colliders (b1) without opening new colliders (b2). The DAG of the example reveals positive confounding by bipolar disorder—conditioned on using the composite measure of bipolar disorder and lithium—, positive reporting bias from the FDA warning—addressed restricting to reports before the warning—, and negative collider bias from masking by other dopamine agonists—addressed by removing competing reports. Finally, sensitivity analyses are run and compared using a forest plot (b3).

Results are interpreted based on the specified assumptions (c1). Assuming the appropriateness of the underlying DAG, the observed associations (consistent even when addressing the expected biases) align more closely with a scenario in which aripiprazole causes impulse control disorders than with one where it does not. The forest plot confirms that subgroup analyses effectively correct the association in the expected direction, with almost no effect of removing the masking. Despite these analyses, uncharted biases and reliance on surrogate variables introduce residual bias (c2), indicating the need for further evaluation using additional data sources (d1), not affected by reporting biases and with more reliable measurements of exposure, outcome, and confounder (potentially measuring impulsivity before and after the exposure).

4 Discussion

Pharmacovigilance relies on adverse event reports to infer potential ADRs, but biases can distort association and impact the generation of valid signals for further assessment and evaluation using complementary data sources. This creates a dilemma: how can findings utility be maximised while avoiding overstatements? The proposed causal inference framework for DPA using DAGs grounds analysis on existing knowledge, identifies necessary assumptions for causal interpretation, and guides future study designs.

Causal inference techniques can be integrated in DPA in two ways. First, DPA can more traditionally be used for agnostic hypothesis generation in routine drug monitoring, and causal inference can be used to structure the evaluation of emerging associations in the light of clinical knowledge and existing evidence. While many scientific publications discuss biases [61], we still lack a standardised framework to describe and tackle them systematically [6]. Knowledge-based DAGs drive DPA design by identifying potential confounders and colliders, ensuring alignment with underlying assumptions. Formalising biases within a unified DAG framework reduces their complexity to a few causal mechanisms that can be combined to build highly specific complex diagrams with formalised rules for addressing them. This approach minimises errors when accounting for multiple concurrent biases, strengthens the theoretical foundation of the analysis, and combats p-hacking (arbitrarily altering parameters to produce significant results [62, 63]). Thus, DAGs can enhance hypothesis-driven analyses, promoting DPA's utility in evidence synthesis while preserving its traditional exploratory role.

Once the study is run, DAGs help interpret results by contextualising them within expected causal scenarios. Despite

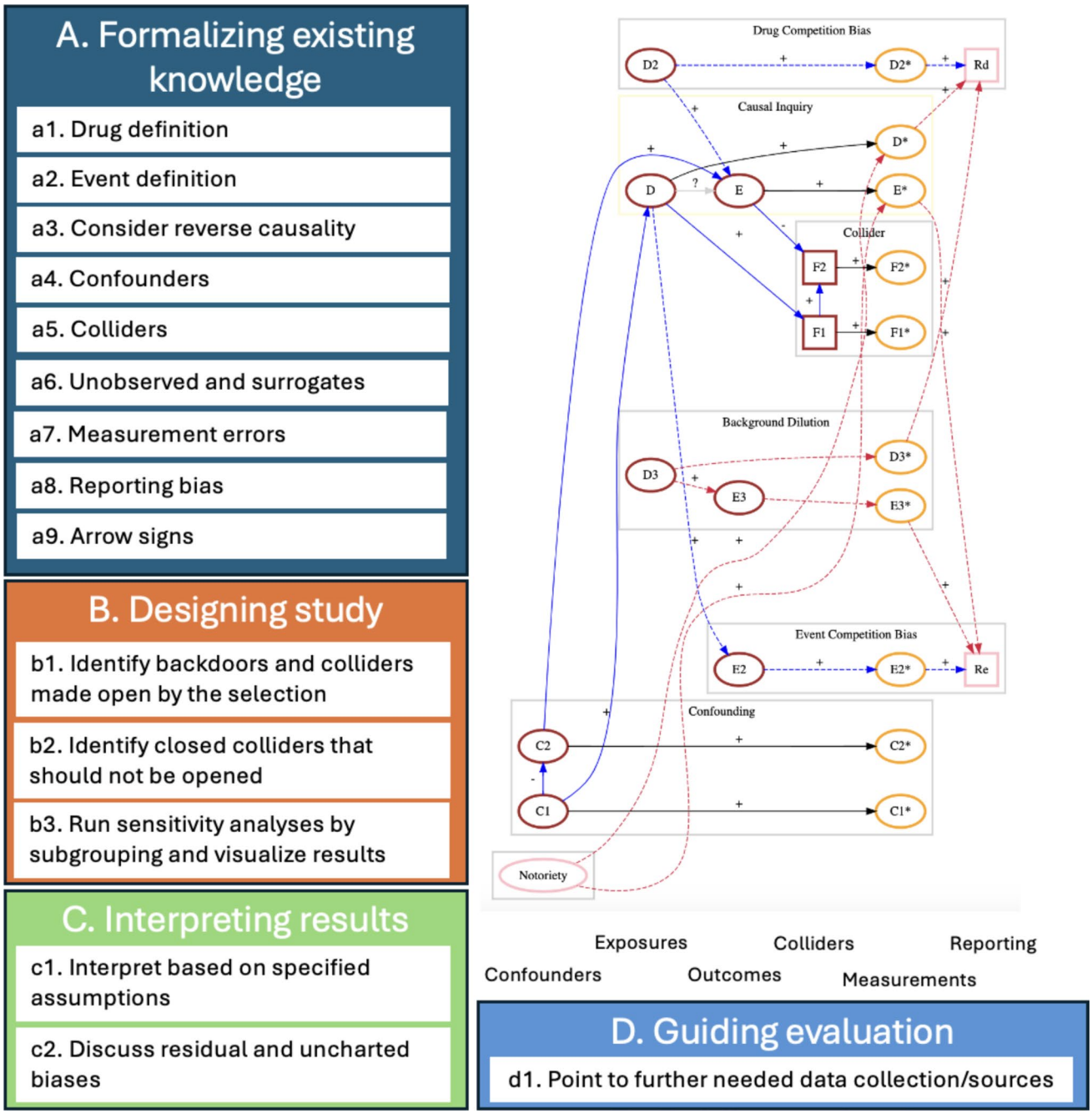


Fig. 7 Workflow for drawing directed acyclic graphs (DAGs) and systematic stratification of DAGs visualisation. The PVdagger R package simplifies the formalisation of causal inquiries and assumptions in pharmacovigilance, showing the stage at which the distortion is introduced by graphically distinguishing between constructs (confounders, exposures, outcomes, colliders), measurements (coding), and reporting (selection at the generation of the reports)

the inherent limitations of DPA, by mapping the conditions at which the inference would be causal, we can better assess the presence of confounders and colliders, and the potential impact of unobserved biases can be visualised and better assessed. Thus, the gap between correlation and causation can be better qualified and, in some cases, reduced, although not eliminated due to measurement error and reporting. The DAGs clarify the limits of causal interpretation by showing how residual and unobserved biases could undermine ADR inference and highlighting biases from not knowing the true non-exposed population, indicating that DPA results should not be treated as precise risk estimates [64–66].

Even in the common scenario where not all confounding can be removed, DAGs help predict the direction of residual

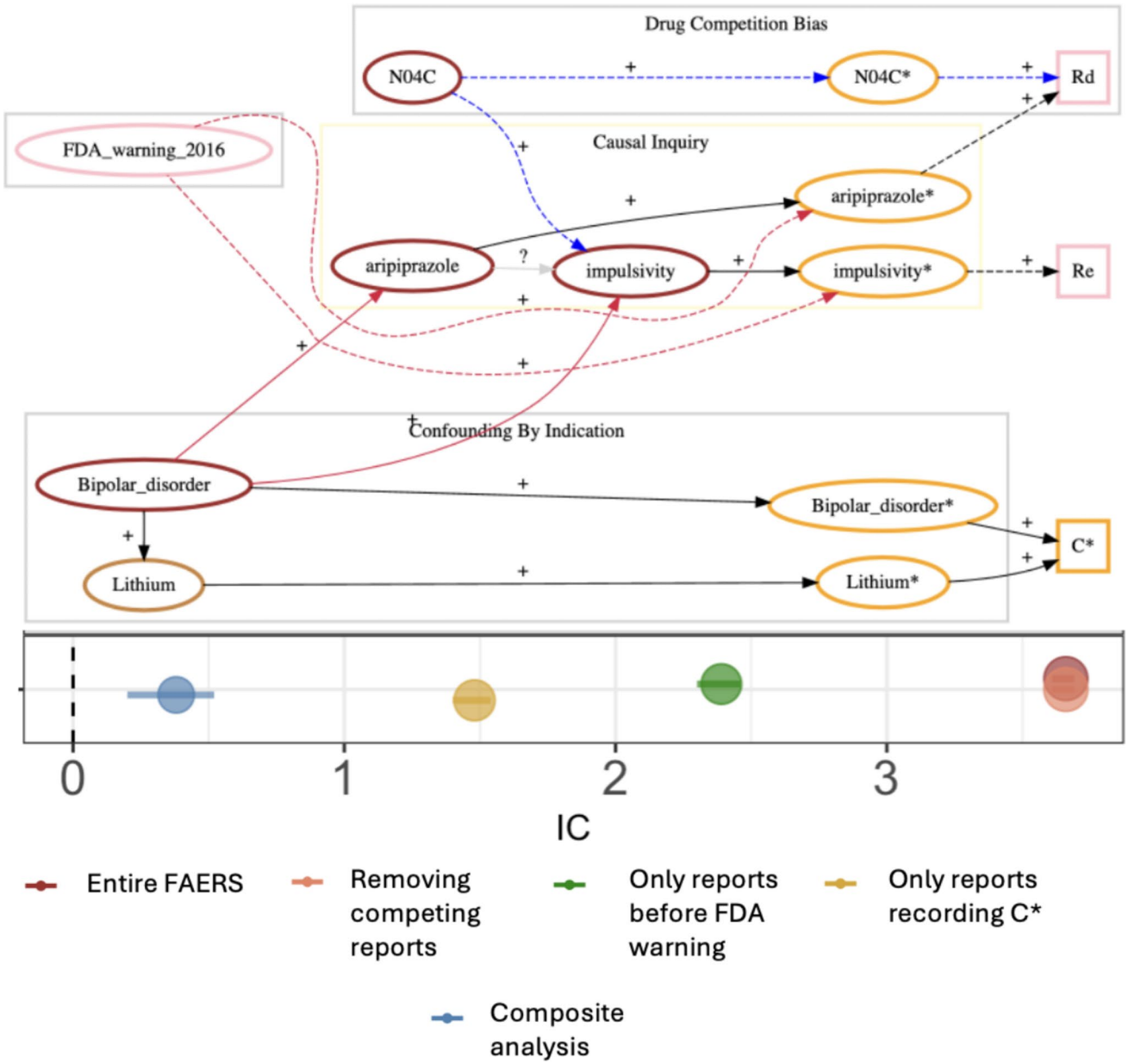


Fig. 8 Complex example, showing the causal model underlying our investigation into the causal inquiry of whether aripiprazole causes impulse control disorders, and comparing the results of the individual and composite sensitivity analyses.

confounding, providing insights into which signals and confounds require further assessment. A statistical dependence with unaccounted negative bias is not expected to disappear when removing biases. A statistical independence with unaccounted positive bias is not expected to appear when removing bias. Residual bias can be further addressed in evaluation studies, ensuring key variables are measured accurately for proper adjustment (i.e., pooling different strata) rather than relying solely on subgroup analyses. Quantifying the magnitude of a causal effect requires evaluation studies unaffected by reporting biases, a defined background population, accurate mapping of exposure, non-exposure, and temporality. When additional evaluation studies are necessary, DAGs identify confounders and colliders that should be considered, guiding researchers toward existing data sources (e.g., electronic health records [67]) or suggesting additional tailored data collection.

However, DAGs on their own, do not ensure valid causal inferences, since they can be misunderstood, misapplied and even misused. Causal inference techniques are quickly evolving and challenging to master (see paragraph 4.1.8 with recommendations for further readings). Further, for the use

of DAGs to be effective, certain assumptions should not be critically violated, such as the accurate reporting of adverse events and the correct specification of the causal model. For this reason, inference on adverse event reports is fraught, due to the highly biased data-generating mechanisms and the complex nature of the relations between drugs and potential adverse events—see Fusaroli et al. [6] for an overview on common pitfalls in these contexts. As a result, incomplete or even mis-specified DAGs could be quite common. Further research will help refine the proposed (canonical) DAG structures, so they more accurately capture our evolving understanding of pharmacovigilance data-generating processes.

Nevertheless, DAGs empower researchers in formalising and visualising causal questions, existing knowledge, and assumptions [68], and help readers to understand and challenge them. By doing so, honest attempts at DAGs, even in presence of incompleteness and incorrectness, can provide a framework for reviewers and readers to more easily challenge the assumptions in an analysis, propose alternative underlying causal scenarios, and suggest different analyses or interpretation. Thus, DAGs facilitate critical collective scientific developments [62, 63, 67, 69, 70].

4.1 Limitations and Further Methodological Advancements

4.1.1 Ontological Versus Pragmatic Tools

Directed acyclic graphs might seem to overstate hypotheses and current states of knowledge as ontological causal structures. Writing down assumptions in a diagram might feel like condoning subjective choices. However, transparency allows us to more easily criticise and build upon the work, improving it. The choice in any analysis is not between a perfectly specified DAG and no model, but between an explicit model whose assumptions are open to debate, and an implicit one whose assumptions are hidden. While a mis-specified DAG can be harmful, making it explicit at least allows for its identification and correction through scientific discourse. Directed acyclic graphs are representations of current knowledge and assumptions, serving as pragmatic tools for causal discovery through iterative design and collective scientific efforts [69, 71–73]. To ensure that the DAG is an appropriate, even if not completely accurate, model of reality, it is crucial to overcome the traditional separation between data scientists (performing quantitative signal detection) and healthcare experts (performing clinical causal assessments), formalising both adverse event reports' generative mechanisms (e.g., reporting bias and measurement) and biases related to specific drug-event combinations (e.g., confounders, colliders, biological plausibility) [74]. Incorrect models can increase residual bias, particularly when key variables are omitted or incorrectly specified. This is a significant concern in both epidemiology and pharmacovigilance. However, explicitly modelling the knowledge justifying a specific analysis, even if the model is initially incorrect, can still be beneficial in the collaborative process of science. It allows other researchers to critique, question, and propose alternative analyses, serving as a starting point for further refinement and improvement of models and a better understanding of the data.

Ultimately, DAGs are not a panacea. Their application requires substantial domain knowledge, and they do not eliminate uncertainty or the problem of unmeasured confounders. Their primary utility is in enforcing intellectual honesty by compelling researchers to translate their qualitative beliefs into a formal, testable structure. This transparency is, in our view, the first and most crucial step toward more reliable and knowledge-based safety signals. Drawing DAGs requires a high level of expertise in causal modelling. Our tutorial aims to lower the barriers to accessing DAGs and to introduce important aspects that need to be further mastered, thus providing a starting point for further reading and practice (see Sect. 4.1.8).

4.1.2 Upfront Investment

Directed acyclic graphs may require substantial upfront investment and epistemic labour. While large-scale semi-automated analyses might seem more suitable for prompt ADR warnings, this assumes that the current cost of false negatives and positives is acceptable. Better integration of DPA results into the scientific process could reduce miscommunication risks [12], improve the quality and clarity of safety signals, and lessen the regulatory burden of unqualified signals needing resource-intensive follow-up. Creating repositories of documented assumptions could streamline future analyses, reducing redundancy and epistemic labour. Efforts in this direction are already underway in fields such as economics [75] and language evolution [68].

4.1.3 Complexity

The causal structures presented may not fully capture the complexity of drug-adverse event relations. This manuscript introduces key DAG techniques for pharmacovigilance but does not exhaust available techniques. G-methods can conceptually remove variables from DAGs, blocking surrogate confounders without opening colliders [76]. Mediation-analysis decomposes the total exposure-outcome effect into indirect effects, elucidating composite mechanisms [77]. Conditioning on competing exposures can enhance precision by addressing heterogeneity sources [78]. While outside the scope of this paper, the generative models embodied by DAGs could be further specified in their mathematical

form (e.g., using structural equation modelling) beyond monotonic relationships. This is a beneficial development, which will rely on considerable efforts of evidence synthesis (e.g. including pharmaco-dynamic knowledge) and targeted studies [79]. Measurement errors and reporting biases could be better mapped (e.g., tracking which ADRs are reported and why) to include observed mechanisms and estimates of measurement error in the model as mitigation factors. Using DAGs, the impact of various biases on the sign and magnitude of the association can be simulated, illustrating different levels of reporting bias even without knowledge of actual levels [80]. Simulation studies could also be used to better investigate the effect of the violation of assumptions and of unobserved biases, and semi-automated tools could be developed to identify contexts in which the assumption violation could become problematic [6]. In the meantime, simplified DAGs, while not suitable for structural equations, can still be useful for addressing expected biases, but should not be blindly applied as they are incomplete.

Longitudinal information from individual case safety reports (ICSRs) or longitudinal medical records could be included in temporal DAGs to address cyclical feedback loops and bidirectionality [67, 81]. Data on the actual number of exposed and the background incidence rate can provide better insight into the frequency and magnitude of the effect (e.g., observed over expected analyses [82]). Finally, DPAs do not remove the need for case series assessment but aim to reduce the number of spurious associations needing in-depth assessment.

4.1.4 Sample Size and Bias Correction

There is a risk of losing important signals due to limited sample sizes when conditioning on under-reported confounders. Using the IC instead of the reporting odds ratio partially addresses this concern, as the credibility interval depends only on the number of cases, not the background size [30]. Future research should explore balancing sample size and bias correction effectively.

4.1.5 Better Measurements

Improving report completeness, developing standardised queries, and implementing algorithms to reliably identify reports with specific variables (e.g., pregnancy [83, 84]) could further mitigate residual bias. This requires acting from above, promoting standardisation of report-based data through global harmonisation (ICH E2B). Region-specific differences in coding and follow-up persist [85], affecting the reliability of data processing and subsequent analyses, but increasing digitalisation of reports will ultimately improve harmonisation [86]. It also requires acting from below, promoting completeness, accuracy, and relevance of reporting among market authorisation holders, health professionals and the public.

4.1.6 Generalisability

Restricting the database may sacrifice generalisability of the findings (e.g., would the causal dependencies inferred in a subset of reports apply outside of that subset?). Adjustment would allow the association between D and E to vary according to the confounder's value, preserving generalisability. However, under-reporting in the confounding variable invalidates adjustment, often making restriction the only possible conditioning.

4.1.7 Importance for Reporting

The biggest potential of DAG techniques and causal inference is in improving reporting in the quickly evolving pharmacovigilance literature [87]. Directed acyclic graphs support authors in their responsibility to make explicit assumptions and study design, reviewers in their responsibility to assess and challenge these assumptions, and the scientific community in understanding, criticising, and using DPA results.

4.1.8 Further Readings

Causal inference is a fast-evolving field of inquiry, with many complex techniques to master. This manuscript provides a pedagogical entry point, but we highly recommend the reader to further improve their understanding of causal inference. More extensive generic introductions to causal inference can be found in Hernan and Robbin's "Causal Inference: What If" book [88], in Hernan's "Causal Diagrams: Draw Your Assumptions Before Your Conclusions" online course (see https://www.harvardonline.harvard.edu/course/causal-diagrams-draw-your-assumptions-your-conclusions#), and in Huntington-Klein's book "The Effect: An Introduction to Research Design and Causality" [89]. The mathematical treatment of DAG-related issues can be further explored in Pearl's book "Causality" [90] and in Hũnermund and Bareinboim's paper "Causal inference and data fusion in econometrics" [91]. For a mathematical formulation of the assumptions underlying causal inference in DPA see Scosyrev et al paper "Disproportionality Analysis and Causal Inference in Drug Safety" [5]. For an example-driven pedagogical paper on the results of violating such assumptions see Fusaroli et al "Charting and Sidestepping the Pitfalls of Disproportionality Analysis" [6].

5 Conclusion

Pharmacovigilance faces the challenge of promptly and reliably identifying potential causal dependencies between drugs and adverse events using information-rich albeit sparse and indirect evidence. While there is still much work to do to provide solid causal inference in pharmacovigilance, such as better mapping reporting biases and improving triangulation across evidence sources, DAGs offer a structured approach to begin this process. Building DAGs is complex, requiring deep domain and methodological knowledge but it can lead to more principled DPA and improved signal reliability. Not least, DAGs can foster tighter integration between pharmacovigilance and epidemiological studies, contributing to a more comprehensive evidence synthesis in drug safety.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40264-025-01628-4.

Acknowledgements We thank Judith Brand, Paul Hünermund, Daniele Sartori, Carlos Melgarejo, Charles Khouri, and Peter Hjelmström for their critical review and constructive comments on the preprint (https://osf.io/preprints/osf/h5w9u). We would like to thank the anonymous reviewers for their substantial suggestions for improving the paper. The views expressed are those of the authors, and not necessarily those of their respective organisations. 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).

Funding The authors did not receive support from any organisation for the submitted work.

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval Not applicable because spontaneous reports of FAERS are anonymous and publicly available.

Consent to participate and publication Not applicable because spontaneous reports of FAERS are anonymous and publicly available.

Data availability statement The data that support the findings of this study are publicly available in their raw form at https://fis.fda.gov/extensions/FPD-QDE-FAERS/FPD-QDE-FAERS.html, and in their clean version through DiAna. An open-access toolkit for DIsproportionality ANAlysis and other pharmacovigilance investigations in FAERS (https://github.com/fusarolimichele/DiAna_package).

Script availability statement The script, using functions from the DiAna and PVdagger R packages, is made available as supplementary material, and as a tutorial at the website of the DiAna package (https://fusarolimichele.github.io/DiAna_package/articles/Causal_inference.html).

Author contributions MF, RF conceptualised and designed the study. The formal analyses and visualisation were performed by MF. MF, RF wrote the original draft. MF, JM, AR, ER, RF strongly contributed to the systematisation of the framework, and to the review and editing of the draft. MF, JM, AR, ER, RF read and approved the final version. MF was the creator of the related Pvdagger R package.

Adherence to reporting guidelines We followed the checklist from the READUS-PV guidelines [87, 92] as much as applicable, taking into account that this is a methodological study, using several drug-event combinations as examples.

References

1. Faillie J-L. Case-non-case studies: principle, methods, bias and interpretation. Therapies. 2019;74:225–32.
2. 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 [Internet]. 2018 [cited 2019 Feb 3]. https://www.intechopen.com/online-first/evolving-roles-of-spontaneous-reporting-systems-to-assess-and-monitor-drug-safety.
3. Härmark L, Raine J, Leufkens H, Edwards IR, Moretti U, Sarinic VM, et al. Patient-reported safety information: a renaissance of pharmacovigilance? Drug Saf. 2016;39:883–90.
4. Abraham J. International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use. In: Tietje C, Brouder A, editors. Handbook of Transnational Economic Governance Regimes [Internet]. Brill | Nijhoff; 2010 [cited 2025 Jan 12]. p. 1041–53. https://brill.com/view/book/edcoll/9789004181564/Bej.9789004163300.i-1081_085.xml.
5. Scosyrev E, Behr S, Jain D, Ponnuru A, Michel C. Disproportionality analysis and causal inference in drug safety. Pharm Med. 2025;39:97–107.
6. Fusaroli M, Sartori D, van Puijenbroek EP, et al. Chartingand Sidestepping the Pitfalls of Disproportionality Analysis. Drug Saf (2025). https://doi.org/10.1007/s40264-025-01604-y.
7. Anjum RL, Chandler RE, Rocca E. Dispositions and causality assessment in pharmacovigilance: proposing the Dx3 approach for assessing causality with small data sets. Pharm Med. 2022;36:153–61.
8. Bégaud B, Judith KJ. Assessing causality from case reports. Textbook of Pharmacoepidemiology. 3rd ed. 2021. p. 246–56.
9. Meyboom RHB, Hekster YA, Egberts ACG, Gribnau FWJ, Edwards IR. Causal or casual? Drug Saf. 1997;17:374–89.
10. Fusaroli M, Raschi E, Poluzzi E, Hauben M. The evolving role of disproportionality analysis in pharmacovigilance. Expert Opin Drug Saf. 2024;23:981–94.
11. Fusaroli M, Simonsen A, Borrie SA, Low DM, Parola A, Raschi E, et al. Identifying medications underlying communication atypicalities in psychotic and affective disorders: a pharmacovigilance

study within the FDA adverse event reporting system. J Speech Lang Hear Res. 2023;66:3242–59.
12. Mouffak A, Lepelley M, Revol B, Bernardeau C, Salvo F, Pariente A, et al. High prevalence of spin was found in pharmacovigilance studies using disproportionality analyses to detect safety signals: a meta-epidemiological study. J Clin Epidemiol. 2021;138:73–9.
13. Khouri C, Fusaroli M, Salvo F, Raschi E. Interpretation of pharmacovigilance disproportionality analyses. Clin Pharmacol Ther. 2023. https://doi.org/10.1002/cpt.2951.
14. Haber NA, Wieten SE, Rohrer JM, Arah OA, Tennant PWG, Stuart EA, et al. Causal and associational language in observational health research: a systematic evaluation. Am J Epidemiol. 2022;191:2084–97.
15. Bernardeau C, Revol B, Salvo F, Fusaroli M, Raschi E, Cracowski J-L, et al. Are causal statements reported in pharmacovigilance disproportionality analyses using individual case safety reports exaggerated in related citations? A meta-epidemiological study. Drug Saf [Internet]. 2025 [cited 2025 Apr 23]. https://doi.org/10.1007/s40264-025-01524-x.
16. Khouri C, Fusaroli M, Salvo F, Raschi E. Transparency and robustness of safety signals. BMJ. 2022;379:o2588.
17. Pacurariu AC, Coloma PM, Gross-Martirosyan L, Sturkenboom MC, Straus SM. Decision making in drug safety—a literature review of criteria used to prioritize newly detected safety issues. Pharmacoepidemiol Drug Saf. 2017;26:327–34.
18. Hauben M, Aronson JK. Defining 'signal' and its subtypes in pharmacovigilance based on a systematic review of previous definitions. Drug-Safety. 2009;32:99–110.
19. Hill AB. The environment and disease: association or causation? Proc R Soc Med. 1965;58:295–300.
20. Andersen F, Anjum RL, Rocca E. When decisions must be based on partial causal knowledge: analyzing causality and evidence for health policy. The Routledge handbook of causality and causal methods. Milton Park: Routledge; 2024.
21. Hernán MA. The c-word: scientific euphemisms do not improve causal inference from observational data. Am J Public Health. 2018;108:616–9.
22. Dahabreh IJ, Bibbins-Domingo K. Causal inference about the effects of interventions from observational studies in medical journals. JAMA. 2024;331:1845–53.
23. Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. Int J Epidemiol. 2021;50:620–32.
24. Rohrer JM. Causal inference for psychologists who think that causal inference is not for them. Soc Pers Psychol Compass. 2024;18:e12948.
25. Chang H. Inventing temperature: measurement and scientific progress. Oxford: Oxford University Press; 2007.
26. Fusaroli M, Raschi E, Giunchi V, Menchetti M, Rimondini Giorgini R, De Ponti F, et al. Impulse control disorders by dopamine partial agonists: a pharmacovigilance-pharmacodynamic assessment through the FDA adverse event reporting system. Int J Neuropsychopharmacol. 2022;25:727–36.
27. 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.
28. Gallow JD. The metaphysics of Causation. In: Zalta EN, Nodelman U, editors. The Stanford Encyclopedia of Philosophy [Internet]. Fall 2022. Metaphysics Research Lab, Stanford University; 2022 [cited 2025 Jan 25]. https://plato.stanford.edu/archives/fall2022/entries/causation-metaphysics/.
29. Halpern JY. Actual causality. Cambridge: The MIT Press; 2016.
30. 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.
31. Fusaroli M, Mitchell J, Rudolph A, Rocca E, Fusaroli R. PVdagger: visualization of causal structures in pharmacovigilance data using DAGs [Internet]. 2025. https://pvverse.github.io/pv_dagger/.
32. Hernan MA, Robins JM. Causal inference: what if. Boca Raton: CRC Press; 2024.
33. Smith SW, Hauben M, Aronson JK. Paradoxical and bidirectional drug effects. Drug Saf. 2012;35:173–89.
34. Deffner D, Rohrer JM, McElreath R. A causal framework for cross-cultural generalizability. Adv Methods Pract Psychol Sci. 2022;5:25152459221106370.
35. Cinelli C, Forney A, Pearl J. A crash course in good and bad controls. Sociol Methods Res. 2022. https://doi.org/10.1177/00491241221099552.
36. VanderWeele TJ, Robins JM. Signed directed acyclic graphs for causal inference. J R Stat Soc Ser B Stat Methodol. 2010;72:111.
37. Hernández-Díaz S, Schisterman EF, Hernán MA. The birth weight "paradox" uncovered? Am J Epidemiol. 2006;164:1115–20.
38. Hernán MA, Cole SR. Invited commentary: causal diagrams and measurement bias. Am J Epidemiol. 2009;170:959–62.
39. Berrie L, Arnold KF, Tomova GD, Gilthorpe MS, Tennant PWG. Depicting deterministic variables within directed acyclic graphs (DAGs): an aid for identifying and interpreting causal effects involving tautological associations, compositional data, and composite variables [Internet]. 2023 [cited 2024 Feb 29]. arXiv:http://arxiv.org/abs/2211.13201.
40. Kuroki M, Pearl J. Measurement bias and effect restoration in causal inference. Biometrika. 2014;101:423–37.
41. Finkelstein N, Adams R, Saria S, Shpitser I. Partial identifiability in discrete data with measurement error [Internet]. 2020 [cited 2024 Jun 11]. arXiv:http://arxiv.org/abs/2012.12449
42. Gahr M, Connemann B, Zeiss R, Schönfeldt-Lecuona C, Dreyhaupt J, Lazik C. Reporting, handling, and subjective importance of adverse drug reactions among general practitioners: an exploratory cross-sectional survey. Expert Opin Drug Saf. 2021;20:979–85.
43. van Hunsel F, van der Welle C, Passier A, van Puijenbroek E, van Grootheest K. Motives for reporting adverse drug reactions by patient-reporters in the Netherlands. Eur J Clin Pharmacol. 2010;66:1143–50.
44. van Hunsel FPAM, ten Berge EAAM, Borgsteede SD, van Grootheest K. What motivates patients to report an adverse drug reaction? Ann Pharmacother. 2010;44:936–7.
45. Kralova K, Wilson C-A, Richebourg N, D'souza J. Quality of MedDRA® coding in a sample of COVID-19 vaccine medication error data. Drug Saf. 2023;46:501–7.
46. 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-Safety. 2007;30:891–8.
47. Chouchana L, Pariente A, Pannier E, Tsatsaris V, Treluyer J-M. Dolutegravir and neural tube defects: a new insight. Lancet Infect Dis. 2020;20:405–6.
48. 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.
49. Rogers JR, Sarpatwari A, Desai RJ, Bohn JM, Khan NF, Kesselheim AS, et al. Effect of lawyer-submitted reports on signals of disproportional reporting in the Food and Drug Administration's adverse event reporting system. Drug Saf. 2019;42:85–93.

50. Hauben M, Reich L, DeMicco J, Kim K. Extreme duplication" in the US FDA adverse events reporting system database. Drug Saf. 2007;30:551–4.
51. Hauben M, Zou C, Bright S, Hung E. More extreme duplication in the U.S. FDA FAERS database and a suggested check point for disproportionality analysis. Pharmacoepidemiol Drug Saf. 2021;30:1140–1.
52. de Boissieu P, Kanagaratnam L, Abou Taam M, Roux M-P, Dramé M, Trenque T. Notoriety bias in a database of spontaneous reports: the example of osteonecrosis of the jaw under bisphosphonate therapy in the French national pharmacovigilance database. Pharmacoepidemiol Drug Saf. 2014;23:989–92.
53. Juhlin K, Ye X, Star K, Norén GN. Outlier removal to uncover patterns in adverse drug reaction surveillance—a simple unmasking strategy. Pharmacoepidemiol Drug Saf. 2013;22:1119–29.
54. Hauben M, Maignen F. Does serious consequential masking exist? An update. Pharmacoepidemiol Drug Saf. 2017;26:727–9.
55. . 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)00039-8.
56. Center for Drug Evaluation and Research. FDA Drug Safety Communication: FDA warns about new impulse-control problems associated with mental health drug aripiprazole (Abilify, Abilify Maintena, Aristada). FDA. 2019.
57. Wolfschlag M, Håkansson A. Increased risk for developing gambling disorder under the treatment with pramipexole, ropinirole, and aripiprazole: a nationwide register study in Sweden. PLoS ONE. 2021;16:e0252516.
58. Plummer NR, Tam AW, Mulvaney CA, Preston NJ, Laha SK. Dopamine agonists for traumatic brain injury. Cochrane Database Syst Rev [Internet]. 2018 [cited 2023 Apr 14]. https://doi.org/10.1002/14651858.CD013062/full.
59. Fusaroli M, Raschi E, Contin M, Sambati L, Menchetti M, Fioritti A, et al. Impulsive conditions in Parkinson's disease: a pharmacosurveillance-supported list. Parkinsonism Relat Disord. 2021;90:79–83.
60. Grall-Bronnec M, Victorri-Vigneau C, Donnio Y, Leboucher J, Rousselet M, Thiabaud E, et al. Dopamine agonists and impulse control disorders: a complex association. Drug Saf. 2018;41:19–75.
61. Fusaroli M, Salvo F, Bernardeau C, Idris M, Dolladille C, Pariente A, et al. Mapping strategies to assess and increase the validity of published disproportionality signals: a meta-research study. Drug Saf. 2023;46:857–66.
62. Khouri C, Nguyen T, Revol B, Lepelley M, Pariente A, Roustit M, et al. Leveraging the variability of pharmacovigilance disproportionality analyses to improve signal detection performances. Front Pharmacol. 2021;12:668765.
63. Bruns SB, Ioannidis JPA. P-curve and p-hacking in observational research. PLoS ONE. 2016;11:e0149144.
64. Fusaroli M, Khouri C, Poluzzi E, De Ponti F, Salvo F, Raschi E. Meta-analyzing results from disproportionality analysis of individual case safety reports: a note of caution. J Clin Psychopharmacol. 2024;44:443.
65. Williams B, Lee K, Ewah SO, Neelam K. Response to comments by Dr Fusaroli and colleagues. J Clin Psychopharmacol. 2024;44:445.
66. Williams BD, Lee K, Ewah SO, Neelam K. Aripiprazole and other third-generation antipsychotics as a risk factor for impulse control disorders: a systematic review and meta-analysis. J Clin Psychopharmacol. 2024;44:39–48.
67. Gauffin O, Brand JS, Vidlin SH, Sartori D, Asikainen S, Català M, et al. Supporting pharmacovigilance signal validation and prioritization with analyses of routinely collected health data: lessons learned from an EHDEN network study. Drug Saf. 2023;46:1335–52.
68. Roberts SG, Killin A, Deb A, Sheard C, Greenhill SJ, Sinnemäki K, et al. CHIELD: the causal hypotheses in evolutionary linguistics database. J Lang Evol. 2020;5:101–20.
69. Tylén K, Fusaroli R, Bjørndahl JS, Raczaszek-Leonardi J, Østergaard S, Stjernfelt F. Diagrammatic reasoning: abstraction, interaction, and insight. Pragmat Cogn. 2015;22:264–83.
70. Woensdregt M, Fusaroli R, Rich P, Modrák M, Kolokolova A, Wright C, et al. Lessons for theory from scientific domains where evidence is sparse or indirect. Comput Brain Behav. 2024;7:588–607.
71. Woensdregt M, Fusaroli R, Rich P, Modrák M, Kolokolova A, Wright C, et al. Lessons for theory from scientific domains where evidence is sparse or indirect. Comput Brain Behav. 2024;7:588–607.
72. De Pretis F, Landes J, Osimani B. E-Synthesis: a Bayesian framework for causal assessment in pharmacosurveillance. Front Pharmacol [Internet]. 2019 [cited 2024 May 28];10. https://doi.org/10.3389/fphar.2019.01317/full.
73. Landes J, Osimani B, Poellinger R. Epistemology of causal inference in pharmacology. Eur J Philos Sci. 2018;8:3–49.
74. Rocca E, Copeland S, Ralph Edwards I. Pharmacovigilance as scientific discovery: an argument for trans-disciplinarity. Drug Saf. 2019;42:1115–24.
75. Garg P, Fetzer T. Causal claims in economics [Internet]. Rochester: Social Science Research Network; 2024 [cited 2025 Jan 10]. https://papers.ssrn.com/abstract=5009061.
76. Igelström E, Craig P, Lewsey J, Lynch J, Pearce A, Katikireddi SV. Causal inference and effect estimation using observational data. J Epidemiol Community Health. 2022;76:960–6.
77. Rijnhart JJM, Lamp SJ, Valente MJ, MacKinnon DP, Twisk JWR, Heymans MW. Mediation analysis methods used in observational research: a scoping review and recommendations. BMC Med Res Methodol. 2021;21:226.
78. Tennant PWG, Textor J, Gilthorpe MS, Ellison GTH. OP87 dagitty and directed acyclic graphs in observational research: a critical review. J Epidemiol Community Health. 2017;71:A43–A43.
79. Kunicki ZJ, Smith ML, Murray EJ. A primer on structural equation model diagrams and directed acyclic graphs: when and how to use each in psychological and epidemiological research. Adv Methods Pract Psychol Sci. 2023;6:25152459231156085.
80. Tay E, Dymock M, Lopez L, Glover C, Huang YA, Bakar KS, et al. Applying causal inference and Bayesian statistics to understanding vaccine safety signals using a simulation study. npj Vaccines. 2024;9:1–8.
81. Murray EJ, Kunicki Z. As the wheel turns: causal inference for feedback loops and bidirectional effects. 2024 [cited 2024 Feb 14]. https://osf.io/9em5q.
82. van der Boom MDX, van Eekeren R, van Hunsel FPAM. Observed-over-expected analysis as additional method for pharmacovigilance signal detection in large-scaled spontaneous adverse event reporting. Pharmacoepidemiol Drug Saf. 2023;32:783–94.
83. 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.
84. 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.
85. Lomeli-Silva A, Contreras-Salinas H, Barajas-Virgen MY, Romero-Lopez MS, Rodríguez-Herrera LY. Harmonization of individual case safety reports transmission requirements among

PAHO reference authorities: a review of their current regulation. Ther Adv Drug Saf. 2024;15:20420986241228119.
86. Pan American Health Organization. System for evaluation of the National Regulatory Authorities for medicines [Internet]. 2018. https://www3.paho.org/hq/index.php?option=com_content&view=article &id=1615:2009-sistema-evaluacion-autoridadesreguladoras-nacionales-medicamentos&Itemid=1179&lang=en#gsc.tab=0.
87. 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.
88. Hernán M, Robins J. Causal inference: what if [Internet]. Boca Raton: Chapman & Hall/CRC; 2020 [cited 2025 Apr 28]. https://miguelhernan.org/whatifbook.
89. Huntington-Klein N. The effect: an introduction to research design and causality [Internet]. CRC Press. Taylor and Francis Group; 2023 [cited 2025 Apr 28]. https://theeffectbook.net/.
90. Pearl J. Causality [Internet]. 2nd ed. Cambridge: Cambridge University Press; 2009 [cited 2025 May 3]. https://www.cambridge.org/core/books/causality/B0046844FAE10CBF274D4ACBDAEB5F5B.
91. Hünermund P, Bareinboim E. Causal inference and data fusion in econometrics. Econom J. 2025;28:41–82.
92. 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): development and statement. Drug Saf. 2024;47:575–84.