



Original Article

ISSN : 2277-3657
CODEN(USA) : IJPRPM

Drug Safety Surveillance Needs Explicit Models for Indication Bias, Channeling, and Treatment Switching

Andreas Müller^{1*}, Stefan Weber¹, Julia Hoffmann², Lukas Schneider³, Tobias Klein⁴

¹Department of Drug Safety Surveillance and Bias Modeling, Faculty of Pharmacy, Heidelberg University, Heidelberg, Germany.

²Department of Indication Bias and Channeling, Faculty of Pharmaceutical Sciences, Technical University of Munich, Munich, Germany.

³Department of Treatment Switching and Safety, Faculty of Pharmacy, University of Freiburg, Freiburg, Germany.

⁴Department of Pharmacoepidemiology, Faculty of Pharmacy, University of Stuttgart, Stuttgart, Germany.

***Email:** andreas.mueller@uni-heidelberg.de

ABSTRACT

Drug safety surveillance increasingly combines spontaneous reports, longitudinal healthcare data, distributed data networks, and computational methods, yet the inferential structure connecting these resources remains under-specified. A recurring problem is that treatment is not allocated independently of prognosis. Indication, disease severity, prior treatment history, channeling toward newer or alternative therapies, switching, treatment-confounder feedback, and informative discontinuation can each reshape the exposed population and the observed risk set. Treating these mechanisms as a single generic category of “confounding” obscures when they arise and which design or analytical response they require. This Current Opinion argues that safety surveillance should explicitly represent treatment allocation and treatment-state evolution before interpreting an exposure–event association. Baseline indication should be distinguished from severity within an indication; channeling should be distinguished from ordinary incident treatment initiation; switching should be treated as both an estimand and exposure-definition problem; and post-baseline confounders affected by earlier treatment should not be handled as ordinary baseline covariates. We further argue that discontinuation and censoring require separate consideration because the reasons patients leave treatment or observation can themselves carry prognostic information. The article develops a proposed bias-diagnostic approach in which causal structure precedes statistical adjustment, source triangulation does not substitute for identification, and computational assistance remains bounded by clinical and epidemiological interpretation. The objective is not to produce a universal correction algorithm, but to make the assumptions connecting treatment selection, treatment history, exposure state, event occurrence, observation, and regulatory interpretation explicit.

Keywords: Pharmacovigilance, Pharmacoepidemiology, Confounding by indication, Channeling bias, Treatment switching, Causal inference

Received: 28 July 2026; **Revised:** 30 September 2026; **Accepted:** 03 October 2026

INTRODUCTION

Post-marketing drug safety surveillance now operates across evidence systems that differ fundamentally in what they observe. Spontaneous reports can identify unusual or clinically salient drug–event combinations without requiring a prespecified cohort, whereas longitudinal healthcare databases can support explicit denominators, treatment histories, temporal risk windows, and comparative epidemiological analyses. These sources are

therefore complementary rather than interchangeable, and sequential analysis of routinely collected data can extend surveillance beyond one-time retrospective studies [1, 2]. Their combination, however, does not eliminate the need to specify how treatment exposure arose in the first place.

This distinction matters because a statistical association between a medicine and an adverse event can be generated, attenuated, or reversed by structures operating before and after treatment initiation. Causal graphs have become particularly useful for making such structures explicit in pharmacovigilance because they force investigators to distinguish common causes, selection mechanisms, conditioning variables, mediators, and reporting processes rather than treating every covariate as a generic adjustment candidate [3]. The graph is not proof of the causal structure; it is a transparent statement of the structure being assumed.

Interpretive discipline is equally important. Meta-epidemiological evidence indicates that causal wording originating from disproportionality analyses can become stronger as findings move through the citation literature [4]. A numerical signal may therefore acquire an apparent mechanistic or causal status that its generating design never established. This problem is not semantic only: language can influence which safety concerns are escalated, how regulators and clinicians understand uncertainty, and whether alternative explanations remain visible.

Reporting initiatives such as READUS-PV improve transparency by requiring clearer specification of data, design, analysis, and interpretation in disproportionality studies [5, 6]. Nevertheless, complete reporting and causal identification answer different questions. A perfectly reported analysis can remain confounded or selected. The central argument of this Current Opinion is therefore that drug safety surveillance should model treatment allocation and subsequent treatment-state evolution explicitly. Indication bias, disease severity, channeling, switching, time-varying confounding, and informative discontinuation should be treated as related but analytically separable mechanisms whose consequences depend on the data source, exposure definition, clinical setting, event phenotype, and time scale.

Why treatment allocation distorts safety evidence

Many drug-safety analyses begin conceptually at the moment an exposure is recorded. Clinically, however, the causal history begins earlier. Treatment selection reflects diagnosis, prognosis, prior response, contraindications, prescriber preference, healthcare access, patient behavior, and competing therapeutic options. Pharmacoepidemiological bias taxonomies consequently distinguish mechanisms arising from the source population, study design, exposure definition, follow-up, and measurement process [7]. If this upstream structure is ignored, later statistical adjustment may be aimed at the wrong problem.

Time zero is particularly important. New-user, active-comparator new-user, and prevalent-new-user designs create different populations because they encode different assumptions about prior treatment and the clinical decision being compared [8]. An incident user who has never received therapy, a patient initiating a drug after failure of another agent, and a long-term user switching within a therapeutic class do not enter the same causal state merely because each has a new prescription on the index date.

Comparator choice is consequently part of causal design. Large-scale empirical methods can help identify candidate comparators with similar observed treatment-use contexts [9], but statistical similarity alone does not establish clinical exchangeability. A comparator must represent a plausible alternative at the same decision point. Otherwise, differences in prognosis, healthcare contact, monitoring, contraindication burden, or treatment stage may be built into the contrast before outcome follow-up starts.

Nor is every allocation difference confounding by indication. Healthy-user mechanisms illustrate why. Patients who initiate preventive or chronic therapy can differ from nonusers in health behavior and prognosis, while later adherence may introduce a second, longitudinal selection process [10]. These baseline and post-initiation mechanisms require different responses. Drug safety surveillance therefore needs a more specific diagnostic vocabulary than “adjusted for confounders.”

Confounding by indication

Confounding by indication arises when the clinical reason for treatment is itself associated with the outcome under study. The indication may be a diagnosis, symptom pattern, prognostic profile, prior treatment failure, or other clinical feature influencing prescribing. Active comparators can improve exchangeability when they represent realistic alternatives for the same clinical indication, particularly when combined with new-user design [11]. They do not, however, guarantee balance in the determinants of why one treatment was chosen over another.

Residual confounding therefore remains relevant even when the headline design appears strong. A systematic review of observational active-comparator studies found heterogeneous use of sensitivity analyses for

uncontrolled confounding [12]. Such analyses are valuable because they turn an implicit assumption—“remaining confounding is probably small”—into an explicit robustness problem. Their output is still conditional on assumptions about the prevalence and strength of unmeasured factors.

Quantitative bias analysis, high-dimensional disease-risk approaches, and machine-learning implementations of prognostic or propensity models provide complementary ways to interrogate residual bias [13-15]. Their performance is not invariant to the data-generating structure. Treatment prevalence, outcome frequency, proxy quality, model specification, and the strength of unmeasured common causes all matter. More variables or a more complex algorithm should therefore not be interpreted as automatic evidence of better causal control.

This point is illustrated by high-dimensional proxy analyses in multiple sclerosis, where estimates changed only modestly across alternative adjustment strategies [16]. Such stability has at least two interpretations: confounding not captured by conventional covariates may have been limited, or the available proxy variables may have been insufficient to represent the clinically relevant latent differences. Safety surveillance should preserve that ambiguity rather than treating estimate stability as proof that residual confounding has been removed.

Disease severity

Disease severity deserves explicit separation from the mere presence of an indication. Two patients can share the same diagnosis yet occupy very different prognostic and therapeutic states. Severity may determine escalation to a newer or more intensive therapy while independently increasing risks such as hospitalization, organ dysfunction, infection, cardiovascular events, or mortality. Thus, an active comparator that successfully aligns diagnostic indication may still compare patients at unequal points along a disease trajectory [11].

Administrative and claims-based data frequently approximate severity through prior diagnoses, utilization, procedures, medication history, laboratory availability, or other prognostic proxies. High-dimensional disease-risk methods can use these variables to improve prognostic balancing where the relevant information is captured [14]. Yet proxy density should not be confused with severity validity. A large covariate set can still omit clinical features that drove the prescribing decision.

The multiple-sclerosis proxy analysis reinforces this boundary: limited movement after high-dimensional adjustment cannot reveal by itself whether severity was already well controlled or remained inadequately represented [16]. We therefore propose severity residuality as a practical diagnostic state: after standard indication alignment, investigators should ask whether clinically meaningful variation in prognosis remains capable of influencing both treatment assignment and the safety event. This is a proposed interpretive distinction, not a validated score or measurable threshold.

The need for such separation becomes clearer when competing explanations for an observed safety association are compared side by side. The purpose of **Table 1** is to distinguish what is observed from what must still be inferred when disease severity is a plausible common cause.

Table 1. Distinguishing treatment-related safety signals from residual disease-severity structure

Scientific configuration	Observable support	Mechanistic test and consequence
surveillance case / Event context that can mimic a treatment effect / Background-risk structure examination / Permitted safety interpretation	What is directly observed	Discriminating article-level test
Manuscript-derived surveillance case: Same indication, different treatment intensity Event context that can mimic a treatment effect: The more intensive therapy is used in patients with poorer prognosis Background-risk structure requiring examination: Severity influences both escalation and event susceptibility Permitted safety interpretation: Association remains compatible with residual severity confounding until the treatment-stage explanation is addressed	Patients share a diagnosis but receive therapies representing different treatment stages	Determine whether severity indicators and treatment-stage history are balanced or credibly proxied
Manuscript-derived surveillance case: Apparently balanced conventional covariates	Standard measured covariates	Examine whether available proxies represent the

Event context that can mimic a treatment effect: Clinically decisive severity information may be absent from routine data Background-risk structure requiring examination: Latent severity can survive measured balance Permitted safety interpretation: Covariate balance alone does not establish severity exchangeability	are similar after adjustment	variables actually driving the prescribing decision
Manuscript-derived surveillance case: High-dimensional adjustment produces little estimate change Event context that can mimic a treatment effect: Either residual confounding was small or available proxies failed to capture it Background-risk structure requiring examination: Competing explanations produce the same apparent stability Permitted safety interpretation: Estimate stability is diagnostically ambiguous	Expanded proxy adjustment does not materially alter the association	Test plausibility of proxy capture using clinical knowledge and sensitivity analyses
Manuscript-derived surveillance case: Event is also a complication of worsening disease Event context that can mimic a treatment effect: Worsening disease may independently produce the same event Background-risk structure requiring examination: Background event hazard changes with disease progression Permitted safety interpretation: Temporal association is insufficient without accounting for evolving disease state	Adverse outcome occurs more often among treated patients	Align risk windows and severity trajectory with treatment initiation or escalation
Manuscript-derived surveillance case: Treatment choice follows prior failure or intolerance Event context that can mimic a treatment effect: Prior failure may identify a more difficult-to-treat population Background-risk structure requiring examination: Treatment history becomes a marker of prognosis Permitted safety interpretation: Interpretation should transition from simple severity confounding toward channeling and switching analysis	Current treatment begins after an earlier therapeutic course	Reconstruct prior treatment state rather than treating the index prescription as a de novo decision

The interaction among clinical severity, recorded proxies, algorithmic assistance, and safety interpretation is not adequately represented by a simple adjustment pipeline. **Figure 1** therefore places disease severity at a decision boundary between observable treatment-allocation data and the causal interpretation that surveillance systems are asked to produce.

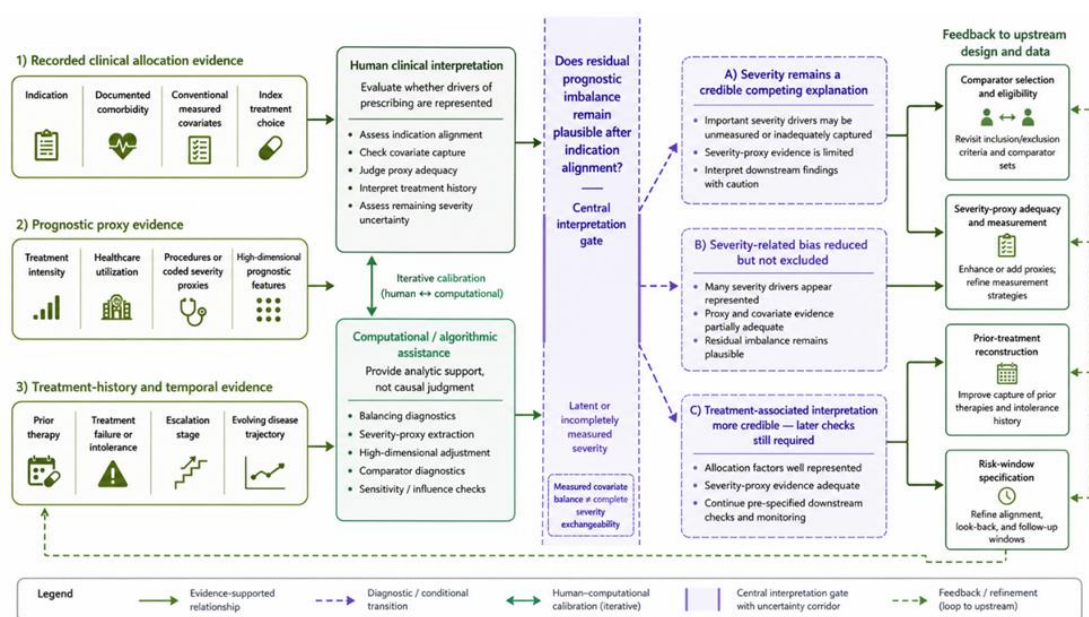


Figure 1. Disease severity as the hidden decision boundary between treatment allocation and safety-signal interpretation

Channeling bias

Channeling bias adds treatment history and clinical selection to the indication problem. New therapies are often adopted unevenly: they may be directed toward patients who failed existing treatment, experienced intolerance, have particular contraindications, are managed by specialists, or differ in prognosis from patients retained on established therapy. The prevalent-new-user literature demonstrates that implementation choices vary substantially and that simply describing a study as PNU does not establish that prior-treatment structure was handled adequately [17].

The core advantage of the PNU approach is that it can retain clinically relevant switchers while matching them to users of the established comparator at a comparable point in treatment history. The original time-conditional propensity-score formulation illustrated this problem in diabetes, where many initiators of a newer therapy had already used the established comparator [18]. Excluding all such patients can narrow the study to atypical initiators; including them without reconstructing prior treatment can introduce a different distortion. PNU design is therefore best understood as one channeling-aware solution, not as a universal correction.

Comparator structure also matters in spontaneous-report data, although the mechanism is not identical to clinical channeling. Analyses of FAERS have shown that restricting the comparator set can materially alter disproportionality estimates and that no single comparator strategy performs optimally across all drug–event structures [19]. In cohort studies, channeling concerns who receives which treatment; in spontaneous reporting, comparator composition concerns which reports define the expected background. Both demonstrate why a safety estimate is inseparable from the population against which it is constructed.

Treatment switching

Treatment switching is often handled as a nuisance that occurs after cohort entry, but this framing is inadequate when switching reflects treatment failure, adverse effects, disease progression, access changes, or clinician reassessment. Empirical work examining switchers in pharmacoepidemiological analyses has shown that alternative handling strategies can change the estimated treatment contrast [20]. The prior treatment state is therefore part of the exposure history, not merely a historical covariate.

A broader scoping review similarly demonstrates that statistical methods used to address switching correspond to different estimands and assumptions [21]. Intention-to-treat-like strategies, censoring approaches, time-varying exposure models, weighting strategies, and other adjustments need not answer the same causal question. Choosing an adjustment because it is familiar, without first specifying whether the target concerns treatment initiation, sustained treatment, actual treatment history, or another strategy, risks producing a precise answer to an unintended question.

Switching is simultaneously an exposure-measurement problem. Routinely collected prescribing and dispensing data provide records from which exposure episodes are inferred; they do not directly measure ingestion or biological exposure [22]. Overlapping supplies, stockpiling, delayed initiation, gaps, concomitant therapies, and uncertain discontinuation can all affect when a patient should be considered exposed. A recorded switch can therefore represent a clinical decision accurately while still leaving uncertainty about the pharmacologically relevant transition.

Even apparently modest algorithmic choices deserve scrutiny. Simulation work in self-controlled drug–drug interaction analyses showed that grace periods intended to accommodate uncertainty can under some circumstances increase exposure misclassification rather than reduce it [23]. The precise direction depends on which exposure is misclassified and on the temporal structure of the analysis. The broader lesson is that risk windows should be pharmacologically and clinically motivated, then stress-tested rather than inherited mechanically.

Finally, censoring at switching can create posttreatment selection. Real-world and simulation analyses comparing alternative handling strategies show that estimates can diverge when switching or censoring is related to prognosis [24]. A patient who switches because of worsening disease, emerging toxicity, or inadequate response carries information about both prior treatment experience and future outcome risk. Removing that patient from observation does not erase the information; it changes the composition of the risk set.

Time-varying confounding

Longitudinal drug safety surveillance becomes more difficult when a variable both predicts subsequent treatment and has itself been affected by earlier treatment. In that setting, conventional adjustment can block part of the treatment pathway while simultaneously attempting to control confounding. Marginal structural modelling has

been applied to this treatment-confounder feedback structure in longitudinal multiple-sclerosis data, illustrating why baseline-only causal reasoning is insufficient when treatment and prognosis co-evolve [25].

Related work comparing marginal structural approaches with G-estimation demonstrates the same principle from another angle: when a time-varying factor influences subsequent exposure yet is altered by prior exposure, standard time-dependent regression does not automatically estimate the desired causal contrast [26]. G-methods can address such structures under explicit assumptions concerning exchangeability, positivity, measurement, and model specification. They should therefore be selected because the causal structure requires them, not because longitudinal data are intrinsically more sophisticated when analysed with weighting.

Time-varying causal models also inherit the quality of the exposure information supplied to them. Claims- or EHR-derived treatment states remain inferred constructs [22]. Sophisticated weighting cannot recover a clinically meaningful exposure history if treatment episodes, overlap, switching, or discontinuation are systematically misclassified. Identification and measurement are related but noninterchangeable problems.

Nor should every postbaseline distortion be renamed time-varying confounding. Switching and censoring may change who remains under observation even when they are not acting as conventional confounders [24]. The analytical task is therefore to determine which time-varying variables are common causes of later treatment and outcome, which lie on prior treatment pathways, which define the treatment strategy itself, and which govern selection into continued observation.

These distinctions are difficult to compare in prose because the same observed event—such as dose escalation, switching, treatment interruption, or worsening disease—can occupy different causal roles depending on its temporal position. **Table 2** therefore separates the classification basis from clinical seriousness, reversibility, and the unresolved inferential question.

Table 2. Classifying evolving treatment states before selecting a longitudinal safety analysis

State and entry conditions		Evidence through time
Classification basis in the causal history / Clinical seriousness relevant to surveillance	Preventability or reversibility dimension / What standard adjustment risks doing / Unresolved drug-safety question that must be answered first	exposure–event scenario
Classification basis in the causal history: Time-varying prognostic factor influencing subsequent treatment Clinical seriousness relevant to surveillance: Potentially high because event risk may reflect disease progression rather than toxicity	Preventability or reversibility dimension: May be partly reversible clinically without resolving inferential bias What standard adjustment risks doing: Treating postbaseline severity as an ordinary covariate can block pathways affected by earlier treatment Unresolved drug-safety question that must be answered first: Is the factor a treatment-affected confounder requiring a longitudinal causal method?	Disease worsening precedes treatment escalation and also predicts the adverse event
Classification basis in the causal history: Treatment-confounder feedback Clinical seriousness relevant to surveillance: Depends on whether the evolving state is itself clinically consequential	Preventability or reversibility dimension: Biological state may change after treatment modification What standard adjustment risks doing: Conventional regression may adjust away part of prior treatment action while incompletely controlling later selection Unresolved drug-safety question that must be answered first: Which treatment strategy and longitudinal estimand are being targeted?	Prior treatment changes a biomarker or clinical state that later influences switching
Classification basis in the causal history: Exposure-state measurement problem Clinical seriousness relevant to surveillance: Seriousness depends on the suspected event and overlap of therapies	Preventability or reversibility dimension: Exposure misclassification may be reducible through alternative episode definitions What standard adjustment risks doing: A sophisticated causal model can still be fitted to the wrong exposure history Unresolved drug-safety question that must be answered first: What risk window and treatment-state definition correspond to plausible biological exposure?	Prescription records show a switch but actual pharmacological transition is uncertain
Classification basis in the causal history: Posttreatment selection process Clinical seriousness relevant to surveillance: Potentially serious if	Preventability or reversibility dimension: Analytically reversible only if selection mechanism can be represented credibly What standard adjustment risks doing: Censoring can preferentially remove patients with different future event risk	Patient is censored immediately after switching

switching follows toxicity or deterioration	Unresolved drug-safety question that must be answered first: Does censoring create informative loss from the risk set rather than control switching?	
Classification basis in the causal history: Time-dependent feedback between event evolution and treatment behavior Clinical seriousness relevant to surveillance: High when prodromal symptoms trigger treatment change	Preventability or reversibility dimension: Often not reversible through simple reclassification What standard adjustment risks doing: Adjustment can condition on an early consequence of the event process Unresolved drug-safety question that must be answered first: Is treatment change a cause, consequence, or marker of the developing event?	Treatment interruption follows an emerging symptom that may be an early manifestation of the outcome
Classification basis in the causal history: Joint identification-and-measurement limitation Clinical seriousness relevant to surveillance: Context dependent	Preventability or reversibility dimension: Measurement improvement may be possible; causal assumptions remain What standard adjustment risks doing: Model complexity can create false reassurance about bias control Unresolved drug-safety question that must be answered first: Are the data sufficiently informative for the assumed causal structure to be credible?	Longitudinal model is correctly specified but severity or treatment state is poorly measured

Informative discontinuation

Treatment discontinuation is informative when the probability of remaining on therapy or under observation is related to prognosis. The distinction is important because discontinuation may represent toxicity, inadequate effectiveness, disease progression, treatment burden, access barriers, or a clinical decision unrelated to the safety event of interest. In anticoagulation research, predictors of censoring have been shown to differ across treatment groups, with corresponding consequences for inverse-probability-of-censoring models [27]. A common censoring model is therefore not automatically appropriate merely because treatment groups share the same outcome definition.

Analytical censoring should also be distinguished from clinical discontinuation. The former is an investigator-defined rule governing follow-up; the latter is an observed or inferred treatment-state transition. If a patient stops therapy because of an emerging symptom, censoring at discontinuation may selectively remove precisely the person-time most informative about the developing event. Conversely, retaining exposure indefinitely after true cessation can misclassify the relevant pharmacological risk period. The correct response depends on why treatment stopped, whether exposure persists biologically, and which treatment strategy the analysis is intended to represent.

Self-controlled designs provide a useful complementary perspective because stable characteristics that differ between patients are controlled by construction. Their strength does not extend to every longitudinal bias. Time-varying disease activity, treatment changes, event-dependent exposure, and altered observation can still threaten interpretation [28]. They therefore provide an alternative design under suitable exposure–outcome structures rather than a universal solution to confounding by indication or discontinuation.

We propose that safety surveillance explicitly distinguish three post-initiation states: treatment cessation as a clinical event, change in biologically relevant exposure, and loss or restriction of analytical observation. These states may coincide, but they need not. Treating them as synonymous can obscure whether a changing estimate reflects pharmacology, clinical selection, exposure misclassification, or modification of the observed risk set.

The practical value of this distinction is adjudicative rather than taxonomic. **Table 3** organizes recurrent discontinuation claims around what would need to be observed before a stronger safety interpretation is justified.

Table 3. Adjudicating safety interpretations when discontinuation changes exposure or observation

Evidence configuration			Adjudication requirement
Discontinuation case / Preventability or reversibility	Cross-source concordance / Signal source	Exposure denominator problem / Escalation rule	Minimum validation need / Drug-safety synthesis consequence
Discontinuation case: Therapy stops after a symptom that may represent an early adverse event	Cross-source concordance required: Confirm whether longitudinal records and spontaneous reports show compatible timing	Exposure denominator problem: Post-discontinuation person-time may be incorrectly excluded or attributed	Minimum validation need: Event-onset and treatment-stop chronology Drug-safety synthesis consequence: Proposed interpretation:

Preventability or reversibility: Clinical action may prevent further exposure, but the original event process may already be underway	Signal source: Longitudinal data with possible spontaneous-report support	Escalation rule: Escalate only if temporal reconstruction supports symptom onset before or around cessation	discontinuation may be part of the event pathway rather than independent censoring
Discontinuation case: Therapy stops because disease control deteriorates Preventability or reversibility: Potentially reversible through alternative treatment, but underlying prognosis remains changed	Cross-source concordance required: Compare disease-activity trajectory across treatment alternatives Signal source: Longitudinal healthcare data	Exposure denominator problem: Remaining treated population becomes progressively selected for better response Escalation rule: Do not interpret lower post-discontinuation event rates as treatment protection without selection assessment	Minimum validation need: Time-varying disease-severity reconstruction Drug-safety synthesis consequence: Worsening disease can generate informative discontinuation without implying treatment toxicity
Discontinuation case: Switch occurs after suspected toxicity Preventability or reversibility: Exposure may be reversible; biological risk can persist after the recorded switch	Cross-source concordance required: Seek agreement among prescribing history, event timing, and clinically relevant follow-up Signal source: Clinical RWD plus pharmacovigilance where available	Exposure denominator problem: New and previous therapies may overlap biologically Escalation rule: Escalate when event timing remains plausible under realistic washout and overlap assumptions	Minimum validation need: Pharmacologically credible risk-window analysis Drug-safety synthesis consequence: Switching should not automatically terminate attribution to the previous treatment
Discontinuation case: Administrative follow-up ends near treatment cessation Preventability or reversibility: Not a clinical state and not inherently reversible analytically once data are absent	Cross-source concordance required: Cross-source evidence may be needed to distinguish care migration from true cessation Signal source: Claims/EHR	Exposure denominator problem: Denominator may shrink for reasons unrelated to treatment biology Escalation rule: Treat as a selection problem until the loss mechanism is characterized	Minimum validation need: Sensitivity analysis for censoring assumptions Drug-safety synthesis consequence: Absence of recorded events after loss to follow-up is not evidence of safety
Discontinuation case: Treatment persists in records despite probable nonuse Preventability or reversibility: Measurement may be improved through episode reconstruction but true adherence can remain unknown	Cross-source concordance required: Pharmacy, prescribing, and clinical-contact patterns should be checked where available Signal source: Routinely collected medication data	Exposure denominator problem: Nominal exposed person-time may exceed biologically exposed time Escalation rule: Avoid stronger causal interpretation when exposure-state uncertainty dominates	Minimum validation need: Alternative exposure definitions and risk windows Drug-safety synthesis consequence: Exposure measurement and discontinuation uncertainty remain separate from causal adjustment
Discontinuation case: Different treatment groups discontinue for different prognostic reasons Preventability or reversibility: Some mechanisms may be clinically modifiable, but analytic consequences remain	Cross-source concordance required: Treatment-specific longitudinal patterns are required Signal source: Comparative RWD	Exposure denominator problem: Risk sets cease to represent comparable prognostic populations Escalation rule: Consider treatment-specific censoring models when clinically and statistically justified	Minimum validation need: Positivity and censoring-model diagnostics Drug-safety synthesis consequence: Proposed interpretation: discontinuation mechanism should be part of the bias diagnosis before effect estimation

These distinctions also require a visual representation in which discontinuation can redirect, rather than simply terminate, the surveillance process. **Figure 2** therefore treats discontinuation as a set of feedback pathways connecting treatment state, observation, event evolution, and signal interpretation.

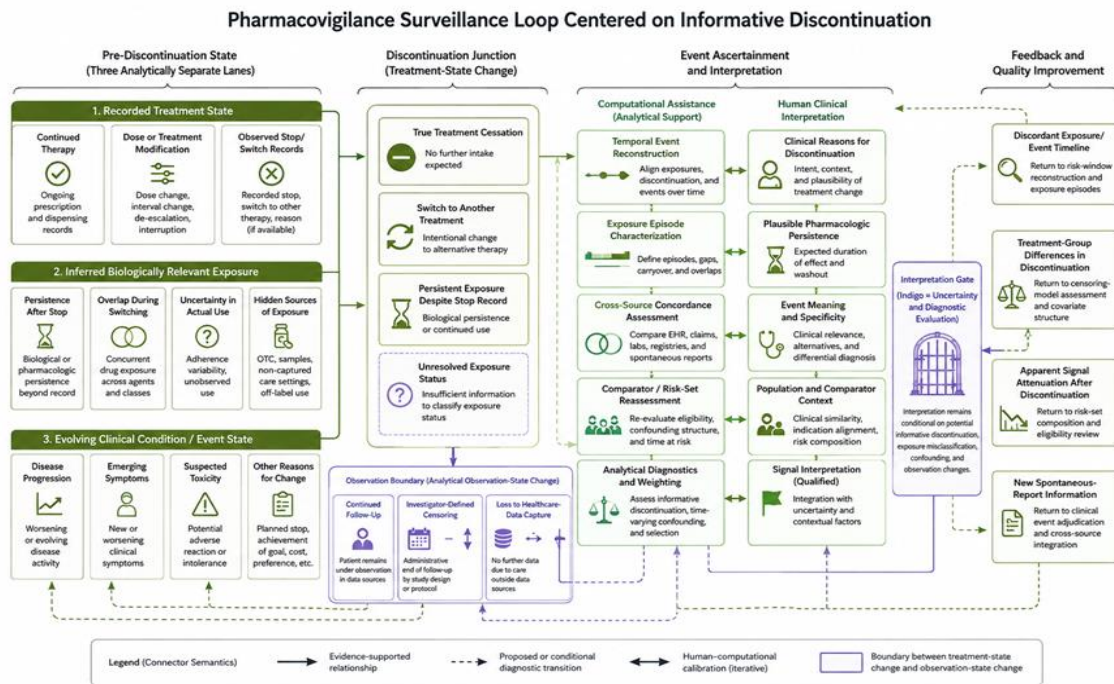


Figure 2. When leaving treatment changes the risk set: feedback paths from discontinuation to safety interpretation

Causal graphs for safety surveillance

Causal graphs are useful because they require the analyst to specify the temporal and causal roles of variables before adjustment. In safety surveillance, treatment indication, severity, prior therapy, switching, evolving disease state, adverse-event development, healthcare contact, discontinuation, and reporting can occupy different positions in the graph. Pharmacovigilance-specific work demonstrates how conditioning choices can create or amplify bias in disproportionality analyses [3]. The same discipline should extend to longitudinal safety analyses. The proposed use is not a single universal graph. Each drug–event question requires a context-specific structure. A variable representing disease activity might be a baseline confounder in one study, a treatment-affected time-varying confounder in another, or an intermediate on the pathway to an outcome in a third. Similarly, discontinuation can represent a treatment decision, an event consequence, a selection mechanism, or several of these simultaneously.

Causal language should follow this identification logic. Systematic evaluation of observational health research has shown substantial inconsistency in how causal and associational language is used [29]. The appropriate response is not to prohibit causal questions in observational research, but to make the required assumptions visible. We therefore propose that every major safety association be accompanied conceptually by an explicit statement of treatment allocation, treatment history, time-varying confounding, selection, exposure measurement, and reporting pathways that could generate the observed result.

Design and analytical countermeasures

No single analytical method resolves every bias described here. Distributed data networks can expand sample size, population diversity, and the reproducibility of multi-database surveillance, yet common data models and shared protocols do not make the underlying patient populations or prescribing processes causally identical [30]. Cross-database consistency should increase confidence only after differences in coding, treatment selection, exposure ascertainment, event capture, and healthcare context have been examined.

Artificial intelligence and machine-learning methods can support phenotyping, covariate construction, comparator discovery, prioritization, and workflow automation, but they inherit the causal limitations of the data and design in which they are used [31]. Predictive discrimination does not establish exchangeability, resolve treatment-confounder feedback, or distinguish a prodromal event from a cause of discontinuation. Human clinical and epidemiological judgment therefore remains part of the analytical system rather than an optional validation layer. Spontaneous-report disproportionality likewise requires structural interpretation. Recent methodological work shows how confounding by indication or comedication, masking, notoriety, effect modification, misclassification,

and collider structures can create or obscure disproportionality [32]. Stratification and comparator restriction can sometimes help but can also introduce new distortions. Signal magnitude should therefore trigger investigation, not substitute for it.

Interpretation itself is a countermeasure. Meta-epidemiological evidence documents substantial spin in published disproportionality research [33]. A mature safety workflow should consequently audit not only the model but also the inferential language applied to its output. “Detected,” “associated,” “compatible with,” “supports investigation,” and “causal” are not interchangeable states.

A bias-diagnostic framework

We propose a bias-diagnostic framework that organizes the foregoing mechanisms by where in the treatment-and-observation process they enter. The framework begins with baseline allocation: indication, disease severity, comparator choice, and treatment history. It then asks whether channeling or prior therapy changes the clinical meaning of treatment initiation. Next, it reconstructs switching and the biologically relevant exposure state. Longitudinal assessment then distinguishes treatment-confounder feedback from other postbaseline processes. Finally, discontinuation, censoring, event ascertainment, and reporting are examined before signal interpretation. The framework is deliberately diagnostic rather than algorithmic. New-user design may address one allocation problem but not another [8]. Sensitivity analysis may reveal fragility to unmeasured confounding without correcting it [12]. Treatment-switching methods estimate different targets [21]. Self-controlled methods remove stable between-person differences while retaining important temporal assumptions [28]. Distributed networks increase evidential reach but do not create automatic transportability [30]. AI and machine learning can assist these tasks but do not select the correct causal estimand independently [31].

For each observed drug–event association, the proposed framework asks five questions. First, what clinical process generated the treated and comparator groups? Second, what prior treatment and disease-severity information is needed to make their risk states comparable? Third, how did treatment and exposure evolve after index time? Fourth, did switching, worsening disease, emerging events, or discontinuation alter subsequent treatment or observation? Fifth, what evidence remains after these plausible alternative structures are considered? These questions should be answered before assigning interpretive weight to a numerical signal.

The output is therefore not “causal” versus “noncausal.” It is a qualified inference state describing which competing explanations have been addressed, which remain plausible, and which data would discriminate among them. This restraint is particularly important in pharmacovigilance, where disproportionality signals can be structurally biased [32] and where causal language itself can outrun identification assumptions [29].

Regulatory and safety-prioritization decisions also require more than association magnitude. Reviews of drug-safety prioritization frameworks show that strength of evidence, public-health importance, seriousness, novelty, and related considerations have been used to prioritize emerging issues, while validation of formal frameworks remains limited [34]. The proposed bias-diagnostic framework should therefore inform expert and regulatory judgment rather than replace it. It has not been prospectively validated, does not assign a numerical certainty score, and should not be interpreted as a regulatory decision rule.

CONCLUSION

Drug safety surveillance cannot be made causally reliable simply by adding more data, more covariates, more databases, or more sophisticated algorithms. The critical problem is whether the analysis represents the clinical and temporal processes that generated treatment exposure and continued observation. Indication, disease severity, channeling, prior treatment, switching, treatment-confounder feedback, discontinuation, censoring, and reporting are related but distinct mechanisms.

The proposed bias-diagnostic framework makes those mechanisms explicit before stronger interpretation of a safety association. Its purpose is to align causal structure with comparator choice, exposure definition, longitudinal analysis, source triangulation, and interpretive language. It is a conceptual aid rather than a validated prediction or regulatory system. Its value will depend on future empirical evaluation across therapies, outcomes, databases, and surveillance settings in which the competing explanations for an observed drug–event association are sufficiently measurable to be tested.

ACKNOWLEDGMENTS: None

CONFLICT OF INTEREST: None

FINANCIAL SUPPORT: None

ETHICS STATEMENT: None

REFERENCES

1. Crisafulli S, Bate A, Brown JS, Candore G, Chandler RE, Hammad TA, et al. Interplay of spontaneous reporting and longitudinal healthcare databases for signal management: Position statement from the Real-World Evidence and Big Data Special Interest Group of the International Society of Pharmacovigilance. *Drug Saf.* 2025;48(9):959-76. doi:10.1007/s40264-025-01548-3
2. Abbasi SH, Lund LC, Hallas J, Pottegård A. Sequential epidemiological analyses of real-world data: A tool for prospective drug safety surveillance from the rofecoxib example. *Drug Saf.* 2025;48(5):489-502. doi:10.1007/s40264-024-01512-7
3. 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. *Drug Saf.* 2026;49(4):449-69. doi:10.1007/s40264-025-01628-4
4. Bernardeau C, Revol B, Salvo F, Fusaroli M, Raschi E, Cracowski JL, 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.* 2025;48(6):679-88. doi:10.1007/s40264-025-01524-x
5. 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(6):575-84. doi:10.1007/s40264-024-01421-9
6. 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(6):585-99. doi:10.1007/s40264-024-01423-7
7. Acton EK, Willis AW, Hennessy S. Core concepts in pharmacoepidemiology: Key biases arising in pharmacoepidemiologic studies. *Pharmacoepidemiol Drug Saf.* 2023;32(1):9-18. doi:10.1002/pds.5547
8. Her QL, Rouette J, Young JC, Webster-Clark M, Tazare J. Core concepts in pharmacoepidemiology: New-user designs. *Pharmacoepidemiol Drug Saf.* 2024;33(12). doi:10.1002/pds.70048
9. Bohn J, Gilbert JP, Knoll C, Kern DM, Ryan PB. Large-scale empirical identification of candidate comparators for pharmacoepidemiological studies. *Drug Saf.* 2025;48(11):1229-41. doi:10.1007/s40264-025-01569-y
10. Olawore O, Stürmer T, Glynn RJ, Lund JL. The healthy user effect in pharmacoepidemiology. *Am J Epidemiol.* 2025;194(7):2023-9. doi:10.1093/aje/kwae358
11. Sendor R, Stürmer T. Core concepts in pharmacoepidemiology: Confounding by indication and the role of active comparators. *Pharmacoepidemiol Drug Saf.* 2022;31(3):261-9. doi:10.1002/pds.5407
12. Latour CD, Delgado M, Su IH, Wiener C, Acheampong CO, Poole C, et al. Use of sensitivity analyses to assess uncontrolled confounding from unmeasured variables in observational, active comparator pharmacoepidemiologic studies: A systematic review. *Am J Epidemiol.* 2025;194(2):524-35. doi:10.1093/aje/kwae234
13. Brown JP, Hunnicutt JN, Ali MS, Bhaskaran K, Cole A, Langan SM, et al. Core concepts in pharmacoepidemiology: Quantitative bias analysis. *Pharmacoepidemiol Drug Saf.* 2024;33(10). doi:10.1002/pds.70026
14. Hossain MB, Wong H, Sadatsafavi M, Cook VJ, Johnston JC, Karim ME. High-dimensional disease risk score for dealing with residual confounding bias in estimating treatment effects with a survival outcome. *Pharmacoepidemiol Drug Saf.* 2025;34(7). doi:10.1002/pds.70172
15. Guo Y, Strauss VY, Khalid S, Prieto-Alhambra D. Use of machine learning to compare disease risk scores and propensity scores across complex confounding scenarios: A simulation study. *Pharmacoepidemiol Drug Saf.* 2025;34(6). doi:10.1002/pds.70165
16. Karim ME, Hossain MB, Ng HS, Zhu F, Frank HA, Tremlett H. Evaluating the role of high-dimensional proxy data in confounding adjustment in multiple sclerosis research: A case study. *Pharmacoepidemiol Drug Saf.* 2025;34(2). doi:10.1002/pds.70112

17. Tazare J, Gibbons DC, Bokern M, Williamson EJ, Gillespie IA, Cunningham M, et al. Prevalent new user designs: A literature review of current implementation practice. *Pharmacoepidemiol Drug Saf.* 2023;32(11):1252-60. doi:10.1002/pds.5656
18. Suissa S, Moodie EEM, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. *Pharmacoepidemiol Drug Saf.* 2017;26(4):459-68. doi:10.1002/pds.4107
19. Gravel CA, Bai W, Douros A. Comparators in pharmacovigilance: A quasi-quantification bias analysis. *Drug Saf.* 2024;47(8):809-19. doi:10.1007/s40264-024-01433-5
20. Wang T, Wang J, Song Z, Miller E, Pate V, Her Q, et al. Caution in handling switchers in pharmacoepidemiologic studies estimating treatment effects: The example of dipeptidyl peptidase-4 inhibitors and inflammatory bowel disease. *Am J Epidemiol.* 2026;195(3):768-82. doi:10.1093/aje/kwaf044
21. Collet RJ, Ben AJ, Varga AN, van Leth F, El Alili M, Esser J, et al. Statistical methods to adjust for treatment switching in real-world clinical studies: A scoping review and descriptive comparison. *Clin Pharmacol Ther.* 2025;118(5):1022-36. doi:10.1002/cpt.70013
22. Thai TN, Winterstein AG. Core concepts in pharmacoepidemiology: Measurement of medication exposure in routinely collected healthcare data for causal inference studies in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf.* 2024;33(3). doi:10.1002/pds.5683
23. Zhang H, Bilker WB, Leonard CE, Hennessy S, Miano TA. Grace periods and exposure misclassification in self-controlled case-series studies of drug-drug interactions. *Am J Epidemiol.* 2025;194(3):802-10. doi:10.1093/aje/kwae231
24. Belviso N, Zhang Y, Aronow HD, Wyss R, Barbour M, Kogut S, et al. Addressing posttreatment selection bias in comparative effectiveness research, using real-world data and simulation. *Am J Epidemiol.* 2022;191(2):331-40. doi:10.1093/aje/kwab242
25. Karim ME, Tremlett H, Zhu F, Petkau J, Kingwell E. Dealing with treatment-confounder feedback and sparse follow-up in longitudinal studies: Application of a marginal structural model in a multiple sclerosis cohort. *Am J Epidemiol.* 2021;190(5):908-17. doi:10.1093/aje/kwaa243
26. Madden JM, Leacy FP, Zgaga L, Bennett K. Fitting marginal structural and G-estimation models under complex treatment patterns: Investigating the association between de novo vitamin D supplement use after breast cancer diagnosis and all-cause mortality using linked pharmacy claim and registry data. *Am J Epidemiol.* 2020;189(3):224-34. doi:10.1093/aje/kwz243
27. Sinyavskaya L, Schnitzer M, Renoux C, Guertin JR, Talbot D, Durand M. Evidence of the different associations of prognostic factors with censoring across treatment groups and impact on censoring weight model specification: The example of anticoagulation in atrial fibrillation. *Am J Epidemiol.* 2021;190(12):2671-9. doi:10.1093/aje/kwab186
28. Bots SH, Brown J, Wong AYS, Martin I, Douglas I, Klungel OH, et al. Core concepts: Self-controlled designs in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf.* 2025;34(1). doi:10.1002/pds.70071
29. 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(12):2084-97. doi:10.1093/aje/kwac137
30. Haber R, Webster-Clark M, Pratt N, Barclay N, Li X, Maro JC, et al. Core concepts in pharmacoepidemiology: Multi-database distributed data networks. *Pharmacoepidemiol Drug Saf.* 2025;34(7). doi:10.1002/pds.70177
31. Rough K, Rashidi ES, Tai CG, Lucia RM, Mack CD, Largent JA. Core concepts in pharmacoepidemiology: Principled use of artificial intelligence and machine learning in pharmacoepidemiology and healthcare research. *Pharmacoepidemiol Drug Saf.* 2024;33(11). doi:10.1002/pds.70041
32. Fusaroli M, Sartori D, van Puijenbroek EP, Norén GN. Charting and sidestepping the pitfalls of disproportionality analysis. *Drug Saf.* 2026;49(2):147-62. doi:10.1007/s40264-025-01604-y
33. 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. doi:10.1016/j.jclinepi.2021.06.022
34. 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(3):327-34. doi:10.1002/pds.4128