Available online www.ijpras.com

International Journal of Pharmaceutical Research & Allied Sciences, 2025, 14(4):24-34
https://doi.org/10.51847/5fa9e0Jz4R

Original Article

ISSN : 2277-3657
CODEN(USA) : IJPRPM

From Statistical Signal Detection to Causal Understanding of Drug Harm through Traceable, Multisource Pharmacovigilance Evidence for Decision-Making

Lina Hassan[1]*, Omar Khalaf[2], Reem Jaber[1], Rania Mostafa[3]

[1]Department of Causal Pharmacovigilance and Signal Detection, Faculty of Pharmacy, University of Jordan, Amman, Jordan.
[2]Department of Multisource Evidence and Decision-Making, Faculty of Pharmacy, Jordan University of Science and Technology, Irbid, Jordan.
[3]Department of Drug Harm Causal Inference, Faculty of Pharmacy, Al-Balqa Applied University, Salt, Jordan.

***Email:** lina.hassan@ju.edu.jo*

ABSTRACT

Pharmacovigilance systems can detect unusual patterns of reported adverse events, yet statistical prominence does not establish that a medicine caused the observed harm. Disproportionality measures, spontaneous reports, observational associations, case narratives, mechanistic findings, and computational predictions each address different questions and remain vulnerable to distinct biases. This Original Causal Safety Architecture Article proposes a traceable signal-to-causality evidence architecture for organizing those heterogeneous contributions without collapsing them into a single score or treating detection as causal confirmation. The architecture separates statistical alert generation, safety-signal formulation, case-level clinical and temporal assessment, source-specific causal analysis, mechanistic evaluation, multisource triangulation, causal-strength grading, and decision use. Triangulation is defined as the structured comparison of evidence generated through meaningfully different data sources and bias pathways, with contradictions and dependencies retained rather than averaged away. Proposed causal-strength grades characterize evidentiary support and residual uncertainty without representing validated probabilities. Decision thresholds are treated as context-dependent judgements influenced by harm seriousness, reversibility, preventability, exposure, evidence quality, and consequences of delayed or premature action. A traceability layer records data provenance, coding, analytical choices, evidence dependence, contradictions, version history, and updating triggers. The original contribution is an end-to-end conceptual architecture connecting pharmacovigilance detection with conditional causal reasoning and accountable decision-making. Its components require computational, epidemiological, clinical, temporal, and implementation validation. The architecture cannot eliminate missing data, reporting bias, confounding, measurement error, model drift, or expert disagreement, and it is not a validated clinical tool, regulatory pathway, or deployment-ready decision system.

Key words: *Pharmacovigilance, Drug safety, Safety signal, Causal inference, Adverse event, Real-world data*
Received: 19 June 2025; **Revised:** 18 October 2025; **Accepted:** 21 October 2025

INTRODUCTION

Pharmacovigilance begins with the recognition that information about a suspected drug-related harm is often incomplete, distributed, and generated for purposes other than causal inference. Characteristics of emerging safety

signals may be associated with subsequent changes to product information, but such downstream action reflects an assessment and decision process rather than direct proof that exposure caused the reported event [1].

National spontaneous-reporting systems similarly demonstrate that detected signals do not move uniformly toward regulatory intervention. Signals may be closed, monitored, communicated, investigated further, or translated into risk-management action according to the strength, seriousness, novelty, and practical implications of the available evidence [2]. The existence of a signal therefore indicates a question requiring structured evaluation, not a resolved causal relationship.

Contemporary signal management connects detection with validation, prioritization, assessment, recommendation, communication, and continuing surveillance [3]. However, the evidence used across those stages is frequently represented in separate analytical products: disproportionality outputs, case narratives, observational estimates, pharmacological arguments, experimental findings, or computational predictions. Without an explicit architecture linking these products, statistical results may acquire stronger causal meaning as they move through assessment and citation.

This article develops an original, non-empirical Signal-to-Causality Evidence Architecture for traceable, multisource assessment of suspected drug harm. The contribution defines distinct evidence layers, relationships, transition conditions, uncertainty registers, causal-strength grades, and decision boundaries. It does not offer an exhaustive review, estimate causal effects, validate a predictive model, prescribe clinical action, or reproduce an existing regulatory procedure. Instead, it proposes a testable structure for preserving the difference between hypothesis generation, causal evaluation, and decision use.

Why disproportionality signals are not causal conclusions

Disproportionality analysis compares the relative reporting of a drug–event combination with an internal reporting background. Its estimand concerns the distribution of reports under defined coding, database, comparator, and analytical conditions; it is not a counterfactual estimate of what would have occurred in the same population without exposure [4]. A high reporting ratio can therefore identify an unusual pattern while remaining compatible with causal and non-causal explanations.

Interpretability is further limited when reports omit essential information about data extraction, duplicate management, coding, comparator construction, thresholds, sensitivity analyses, or analytical decisions. Meta-epidemiological evidence indicates that insufficient transparency is common in published disproportionality analyses, restricting reproducibility and evaluation of whether a reported association is robust to reasonable methodological alternatives [5]. Better reporting improves inspectability but does not itself remove bias.

The evidentiary meaning of a disproportionality result also depends on how the analysis was conceived, conducted, reported, and interpreted [6]. Choice of database, event hierarchy, drug role, time period, comparator, subgroup, and statistical method can change the visibility of a signal. A statistical threshold should therefore function as an investigation trigger within a wider assessment process rather than as a universal boundary between causal and non-causal relationships.

The proposed architecture distinguishes an adverse event, meaning an observed medical occurrence without an implied causal relationship; a statistical alert, meaning an unusual quantitative pattern; a safety signal, meaning information supporting a new or incompletely documented drug–harm hypothesis; and a causal safety conclusion, meaning a conditional judgement integrating multiple evidence domains. These constructs are related, but none automatically transforms into the next. No disproportionality statistic, regardless of magnitude or precision, independently establishes individual-level or population-level causation.

Proposed signal-to-causality evidence architecture

The architecture begins with a traceable detection record containing the data source, observation period, coding hierarchy, exposure and event definitions, comparator, analytic method, missing-data treatment, sensitivity choices, and interpretation boundary. Reporting standards for disproportionality analyses support documentation of these elements, although compliance with a reporting standard does not establish causal validity [7]. The article newly proposes explicit transition gates separating statistical alert generation, safety-signal formulation, causal-hypothesis specification, evidence integration, and decision use. **Figure 1** presents the original conceptual synthesis for transition from statistical signal to causal safety conclusion, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

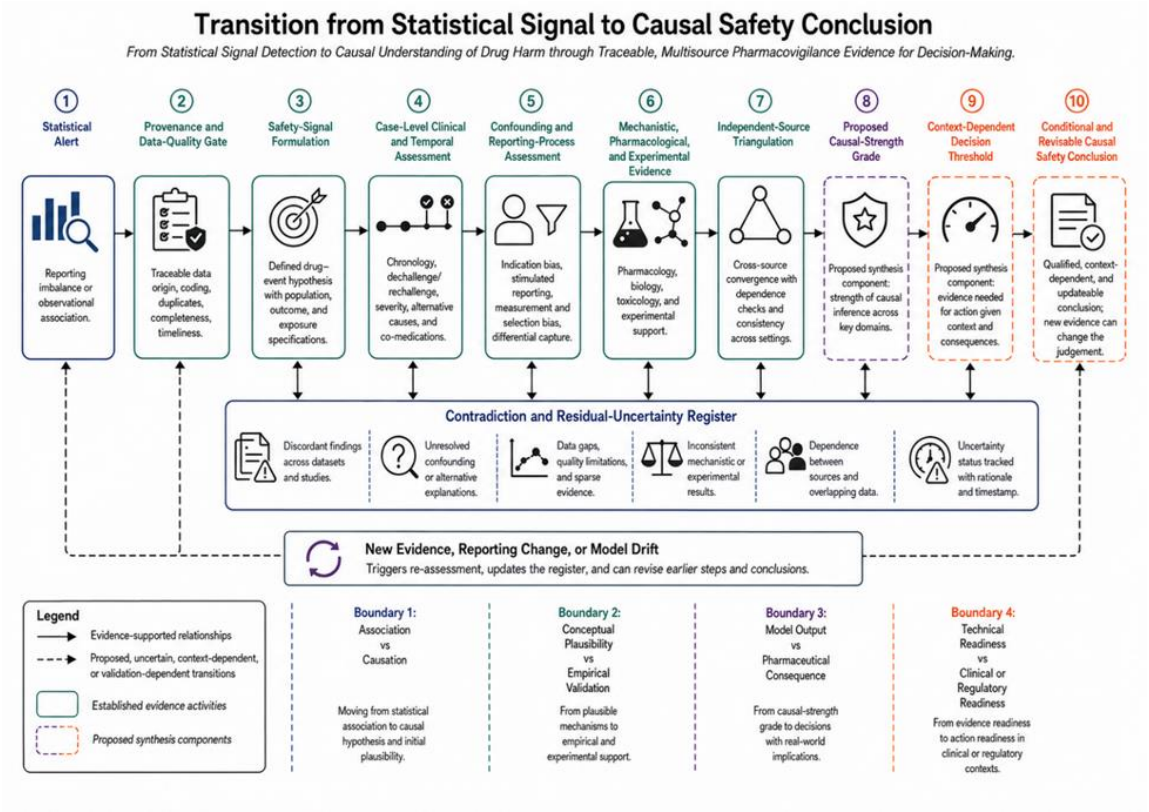

Figure 1. Transition from Statistical Signal to Causal Safety Conclusion.

The figure is an original conceptual synthesis developed for "From Statistical Signal Detection to Causal Understanding of Drug Harm through Traceable, Multisource Pharmacovigilance Evidence for Decision-Making." Arrows and grouping indicate proposed or evidence-supported relationships rather than measured effect sizes. The figure does not represent a validated predictive model, clinical recommendation, regulatory determination, or deployment-ready system.

The second layer assigns evidence according to the question it can answer. Real-world evidence may contribute to signal identification, refinement, evaluation, or monitoring when its data, design, timing, and population are fit for the intended decision [8]. Spontaneous reports may reveal rare, serious, unexpected, or clinically distinctive events, whereas longitudinal data may provide denominators, temporal comparisons, covariates, and comparator groups. Their outputs should remain analytically distinct until source-specific biases and dependencies have been characterized.

The third layer structures case-level clinical evidence. Machine-assisted systems may extract chronology, comedications, alternative causes, dechallenge, rechallenge, and other features relevant to individual-case assessment [9]. Such systems can support organization and prioritization, but their outputs remain conditional on feature quality, reference labels, missingness, and the clinical meaning assigned to each variable. Human review is therefore treated as an accountable interpretive function rather than a procedural endorsement of algorithmic conclusions.

The fourth layer evaluates bias and plausibility before synthesis. For each evidence source, the architecture records exposure and outcome validity, selection processes, reporting incentives, confounding pathways, measurement error, missing data, analytical flexibility, and possible dependence on other sources. Mechanistic evidence is represented as a distinct domain that can support coherence or identify incompatibility but cannot erase weak chronology, invalid measurement, or severe residual confounding.

The final layer performs bias-aware triangulation and produces a proposed causal-strength profile. Evidence is not combined by counting supporting sources or averaging unlike measures. Instead, the synthesis asks whether results converge across methods with meaningfully different biases, whether contradictions have credible explanations, and whether conclusions are sensitive to unresolved assumptions. The architecture's gates, grades, contradiction register, and decision rules are original proposed constructs requiring empirical validation before operational use.

Signal detection across spontaneous and observational data

Spontaneous-reporting and observational-healthcare systems detect different manifestations of possible drug harm. Reviews of routinely collected healthcare data identify diverse approaches based on temporal surveillance, self-controlled designs, cohort comparisons, data mining, and other analytical structures, each with question-specific assumptions and limitations [10]. Active surveillance can therefore complement spontaneous reporting, but it does not provide a universally superior or automatically causal form of evidence.

Common-data-model infrastructures can standardize vocabularies, transformations, analytical routines, and distributed execution across heterogeneous healthcare databases [11]. This may improve repeatability and make differences between sites more visible. Nevertheless, harmonized data can preserve harmonized error: coding variation, incomplete capture, channeling, outcome misclassification, and unmeasured confounding remain possible even when identical software is applied.

Spontaneous systems can also identify abrupt or unexpected increases in reporting frequency. Algorithms designed to detect such changes may help distinguish emerging temporal patterns from stable background reporting [12]. However, an increase can arise from changes in exposure, reporting requirements, publicity, stimulated reporting, coding practice, duplicate submission, or clinical awareness. Report counts and reporting proportions must not be interpreted as incidence without an appropriate population denominator.

The architecture consequently treats spontaneous and observational detection as parallel hypothesis-generating routes. Each route produces a source-specific evidence object containing its estimand, population, time window, comparator, assumptions, diagnostics, and unresolved biases. Advancement to causal assessment requires a clearly specified drug–harm hypothesis rather than an unqualified database association. Detection prioritizes what should be investigated; it does not determine whether the medicine caused the harm.

Case-level clinical assessment and temporal plausibility

Case-level assessment asks whether the clinical sequence is compatible with the suspected drug–harm relationship. Patient reports may contribute symptom detail, onset timing, severity, functional consequences, and treatment experience that are absent from professionally submitted reports [13]. Such information complements rather than replaces clinical documentation.

Chronology should represent exposure initiation, dose changes, latency, event onset, withdrawal, recovery, re-exposure, concomitant treatment, and disease progression. Machine-assisted methods may organize these variables and identify missing evidence, but their conclusions remain dependent on data quality, feature construction, and reference labels [14].

Structured causality tools promote consistent questioning, yet different instruments may classify the same case differently because they weight chronology, alternatives, prior knowledge, and rechallenge evidence differently [15]. A numerical or categorical case score should therefore be interpreted as an evidence summary rather than an objective causal verdict.

Within the proposed architecture, case-level assessment functions as a plausibility and contradiction gate. Compatible timing may support further evaluation, while implausible latency or a strong alternative cause may weaken the hypothesis. Neither pattern alone establishes or excludes population-level causation.

Table 1 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for definitions, components, relationships, and intended uses in from statistical signal detection to causal understanding of drug harm through.

Table 1. Definitions, components, relationships, and intended uses in From Statistical Signal Detection to Causal Understanding of Drug Harm through.

Component or scientific dimension	Problem addressed	Inputs or determinants	Proposed mechanism or relationship	Expected contribution	Evidence required	Failure or uncertainty risk	Boundary statement
Adverse event	Observation may be mistaken for causation	Clinical occurrence, exposure history, reporter narrative	Event enters assessment without causal presumption	Preserves hypothesis-generating status	Verifiable event and exposure documentation	Missing chronology or diagnosis	Temporal occurrence alone does not establish causation
Statistical alert	Quantitative prominence may	Reporting distributions or	Flags an unusual pattern for evaluation	Prioritizes investigation	Reproducible definitions and diagnostics	Comparator, coding, or	Statistical strength is not causal strength

	be overinterpreted	observational estimates				reporting distortion	
Safety signal	Alert lacks a clinically formulated question	Novelty, seriousness, plausibility, prior knowledge	Converts an alert into a defined drug–harm hypothesis	Focuses subsequent assessment	Traceable signal formulation	Vague exposure, outcome, or population	A signal remains an investigational proposition
Case-level plausibility	Population patterns may lack clinical coherence	Chronology, alternatives, dechallenge, rechallenge	Organizes compatibility and contradiction	Refines plausibility and information needs	Complete clinical narrative and expert review [13]	Missingness and subjective weighting	A case score is not a causal verdict
Structured case assessment	Complex narratives are difficult to compare	Extracted clinical features and decision rules	Supports reproducible evidence organization	Improves consistency of review [14]	External and inter-rater validation	Automation bias and label error	Machine assistance does not replace accountable judgement
Causal-strength profile	Heterogeneous evidence cannot be reduced safely to one statistic	Case, observational, mechanistic, and contextual evidence	**Proposed synthesis:** grades support and residual uncertainty	Supports transparent decision discussion	Prospective validation of grade reliability	False precision or concealed contradiction	Grades are not validated probabilities
Decision threshold	Identical evidence may imply different actions	Seriousness, preventability, reversibility, exposure, uncertainty	**Proposed relationship:** decision context modifies action threshold	Separates causal judgement from response urgency	Context-specific governance evaluation	Premature or delayed action	Thresholds are not universal regulatory rules

Confounding, indication bias, and reporting processes

Observational drug-safety associations can arise because treatment is selected in response to disease, prognosis, severity, previous failure, or contraindications. Active comparators may reduce some differences, but comparator validity depends on the clinical decision being represented and cannot remove all confounding [16].

Spontaneous-reporting data are shaped by reporting probability as well as event occurrence. Changes in regulation, communication, public attention, or professional awareness can alter report volume independently of changes in underlying incidence [17]. These processes must be represented as competing explanations rather than treated as background noise.

Statistical discovery methods also differ in their susceptibility to co-prescribing, indication-linked events, competition among reports, and innocent-bystander associations [18]. Agreement between methods may reflect shared data and assumptions rather than independent corroboration.

The proposed architecture therefore requires an explicit bias model for every source. The model should identify plausible common causes, selection pathways, measurement processes, reporting incentives, and analytic decisions. Adjustment complexity alone cannot demonstrate that these pathways have been controlled adequately.

Mechanistic, pharmacological, and experimental evidence

Mechanistic evidence can connect a statistical pattern to pharmacology, biological pathways, tissue susceptibility, exposure timing, or interaction processes. Knowledge graphs may generate previously unrecognized adverse-reaction hypotheses and support evaluation against clinical data, but their predictions remain dependent on graph construction and available knowledge [19].

Graph-based models can also represent polypharmacy relationships in which combinations produce effects not attributable readily to either medicine alone [20]. These models are useful for prioritization and relational hypothesis generation, but a predicted edge is neither direct experimental evidence nor causal confirmation.

Combining statistical detection with biological plausibility may improve prioritization of suspected drug–drug interactions [21]. Plausibility is strongest when predicted relationships are compatible with exposure, timing, target biology, metabolism, and observed clinical patterns.

Mechanistic evidence should therefore enter the architecture as an independent evidentiary domain. It may strengthen coherence, expose incompatibility, or define an experiment, but it cannot override implausible chronology, invalid measurement, severe confounding, or contradictory human evidence.

Table 2 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for testable propositions, evidence requirements, and validation criteria for from statistical signal detection to causal understanding of drug harm through.

Table 2. Testable propositions, evidence requirements, and validation criteria for From Statistical Signal Detection to Causal Understanding of Drug Harm through.

Architecture layer or process stage	Core function	Information or material flow	Interaction with other components	Validation criterion	Potential failure mode	Human or experimental responsibility	Readiness boundary
Knowledge-graph hypothesis generation	Identify plausible drug–event relations	Curated entities and relations to candidate hypotheses	Supplies hypotheses to clinical and observational assessment	External replication in independent clinical data [19]	Missing, biased, or circular knowledge	Curate sources and inspect relation provenance	Prediction does not establish mechanism
Polypharmacy representation	Model combination-specific harm patterns	Drug pairs, biological relations, and adverse effects	Informs interaction assessment	Temporal and external validation [20]	Sparse combinations and unstable embeddings	Review pharmacological plausibility	Model performance is not causal proof
Biological plausibility assessment	Test compatibility with known pharmacology	Targets, metabolism, exposure, pathways, timing	Challenges or supports detected associations	Reproducible convergence with human evidence [21]	Post hoc mechanistic storytelling	Define falsifiable experiments	Plausibility cannot replace causal analysis
Contradiction register	Preserve incompatible evidence	Positive, negative, and uncertain findings	Feeds triangulation and grading	**Proposed criterion:** contradictions remain traceable across updates	Selective omission	Adjudicate explanations transparently	Unresolved contradiction limits readiness
Causal-strength profile	Integrate distinct evidence domains	Source-specific assessments to qualitative profile	Informs context-dependent decisions	**Proposed criterion:** reproducibility across independent assessors	False precision	Maintain accountable multidisciplinary review	Profile is not a validated probability
Evidence-update loop	Reassess conclusions when evidence changes	New reports, studies, mechanisms, or drift signals	Returns evidence to earlier architecture layers	**Proposed criterion:** timely, versioned reassessment	Stale conclusions	Define ownership and review triggers	Updating does not guarantee correctness

Triangulation across independent data sources

Multisource assessment is valuable because different sources can illuminate different parts of a drug–harm hypothesis. In one signal assessment, spontaneous reports and longitudinal healthcare data supplied complementary clinical and epidemiological information rather than interchangeable confirmation [22].

Triangulation requires more than repetition. Its inferential value increases when methods have meaningfully different biases and still produce compatible findings [23]. Reanalyses of the same patients, coding systems, reporting incentives, or underlying databases should not be counted as independent corroboration.

Spontaneous reports may contribute clinical novelty, rare-event recognition, and detailed narratives, whereas longitudinal databases may contribute denominators, comparators, covariates, and temporal estimation [24]. Either source may also contradict or qualify the other.

The architecture therefore preserves source-specific results before synthesis. It documents shared data, shared assumptions, common measurement pathways, and possible dependence. Convergence strengthens a causal hypothesis only when plausible common errors are insufficient to explain the agreement.

Causal-strength grading and decision thresholds

Causal language can become stronger than the underlying evidence as disproportionality findings are cited and reused [25]. The architecture addresses this problem by separating the wording of a conclusion from the urgency or importance of a decision.

The proposed grades are hypothesis-generating, clinically plausible, analytically supported, triangulated but uncertain, and decision-relevant under stated conditions. These labels describe evidence organization and residual uncertainty; they are not causal probabilities or validated regulatory categories.

Large-scale evidence generation requires standardized designs, diagnostics, transparency, and systematic evaluation across databases [26]. Accordingly, progression between proposed grades should require reproducibility, measurement validity, sensitivity analysis, source-independence assessment, and transparent accounting of contradictory results.

Sensitivity analysis can indicate how strong unmeasured confounding would need to be to explain an observational association [27]. It cannot demonstrate that such confounding is absent. Decision thresholds must therefore reflect both evidence strength and the consequences of acting too early or too late.

Table 3 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for failure modes, boundary conditions, governance needs, and implementation priorities for from statistical signal detection to causal understanding of drug harm through..

Table 3. Failure modes, boundary conditions, governance needs, and implementation priorities for From Statistical Signal Detection to Causal Understanding of Drug Harm through.

Decision scenario or design trade-off	Trigger or context	Competing objective	Evidence threshold	Uncertainty treatment	Corrective or escalation action	Expected implication	What cannot be concluded
Strong statistical alert with weak clinical evidence	Large disproportionality estimate	Early warning versus false alarm	Traceable signal plus case review	Preserve non-causal wording [25]	Seek chronology, comparators, and alternative explanations	Prioritized investigation	Magnitude proves causation
Replication across databases	Similar estimates in several datasets	Reproducibility versus hidden dependence	Comparable definitions and diagnostics [26]	Map overlapping populations and assumptions	Conduct independent or differently biased analyses	Stronger evidentiary coherence	Database count equals independence
Residual confounding	Observational association persists	Timely interpretation versus uncertainty	Design diagnostics and sensitivity analysis	Quantify vulnerability where possible [27]	Refine design or restrict claim strength	More transparent causal boundary	Sensitivity analysis removes confounding
Serious but incompletely characterized harm	Potentially severe or irreversible outcome	Precaution versus premature attribution	**Proposed:** lower action threshold with explicit uncertainty	Separate action urgency from causal grade	Escalate surveillance and expert review	Earlier protective consideration	Urgent action confirms causality
Suspected subgroup harm	Effect appears concentrated in a subgroup	Detection sensitivity versus sparse-data instability	**Proposed:** valid measurement, positivity, comparator adequacy, replication	Report instability and multiplicity	Seek independent confirmation	More credible heterogeneity assessment	Subgroup difference is established
Conflicting evidence streams	Mechanistic and clinical evidence disagree	Coherence versus selective interpretation	**Proposed:** documented contradiction analysis	Retain unresolved conflict visibly	Reassess assumptions and collect discriminating evidence	Prevents artificial consensus	One evidence domain automatically dominates

Traceability, governance, and evidence updating

Traceability requires documentation of data origin, linkage, coding, eligibility, design, analysis, and interpretation. Reporting standards for routinely collected pharmacoepidemiological data support this level of transparency [28]. The proposed architecture extends it into a versioned evidence-to-decision record.

Each conclusion should be linked to its source-specific analyses, assumptions, contradictions, reviewers, decision context, and date. Changes to a grade or recommendation should preserve the previous version and explain what new evidence altered the judgement.

Predictive and classificatory models may lose calibration when populations, measurements, workflows, or clinical environments change [29]. A model that performed acceptably at one time or site cannot be assumed to remain reliable without temporal and external evaluation.

Governance should therefore specify ownership, access control, review frequency, update triggers, conflict management, and error escalation. Traceability improves accountability and correction, but an auditable process can still produce an incorrect causal conclusion.

Limitations and implementation priorities

Automation may improve the speed of intake, deduplication, coding, prioritization, and evidence retrieval, while also creating requirements for validation, explainability, oversight, and accountability [30]. Automation should support defined tasks rather than convert an uncertain evidence process into autonomous causal adjudication.

Incomplete adverse-event reports constrain detection and case assessment because absent chronology, indication, dose, comedication, outcome, or reporter information may prevent competing explanations from being evaluated [31]. Improving completeness does not guarantee accuracy, representativeness, or causal validity.

Figure 2 presents the original conceptual synthesis for multisource triangulation architecture for drug-harm assessment, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

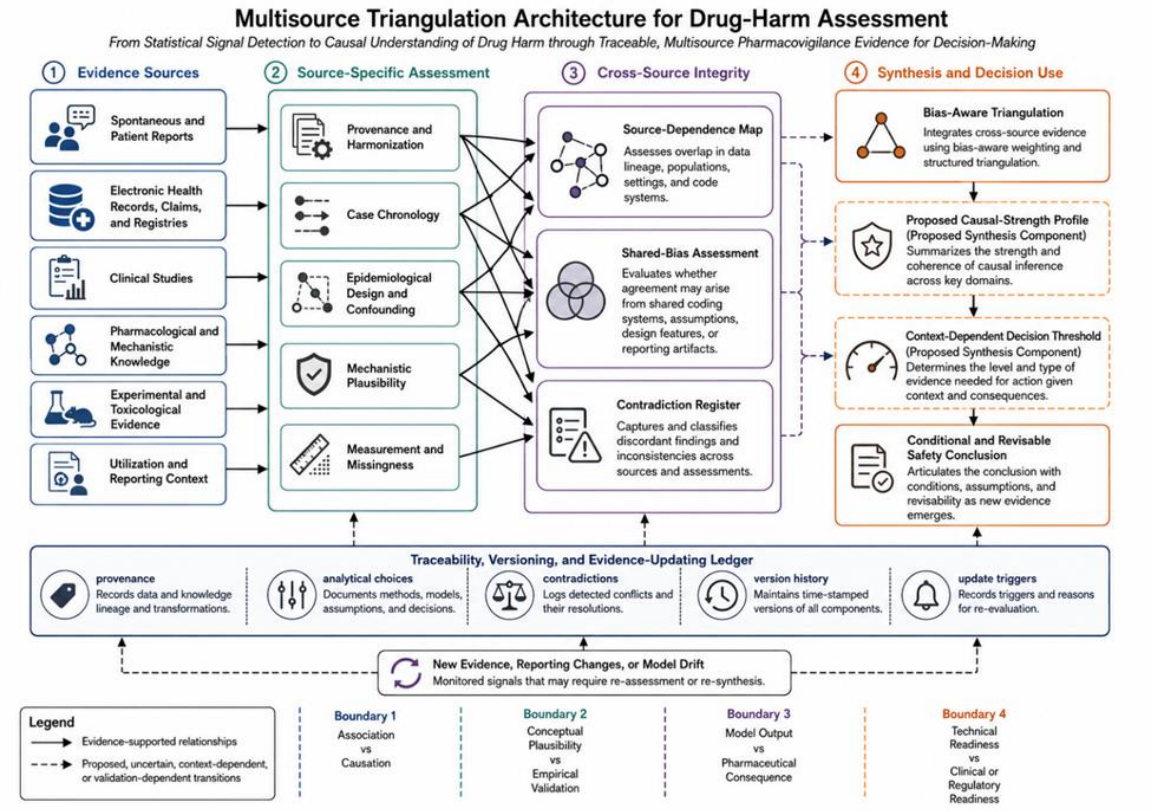

Figure 2. Multisource Triangulation Architecture for Drug-Harm Assessment.

The figure is an original conceptual synthesis developed for "From Statistical Signal Detection to Causal Understanding of Drug Harm through Traceable, Multisource Pharmacovigilance Evidence for Decision-Making." Arrows and grouping indicate proposed or evidence-supported relationships rather than measured effect sizes. The figure does not represent a validated predictive model, clinical recommendation, regulatory determination, or deployment-ready system.

Implementation should begin with retrospective reconstruction of completed signal assessments, followed by inter-rater testing, external and temporal validation, prospective workflow evaluation, subgroup analysis, drift monitoring, and governance testing. Each stage should evaluate whether the architecture improves transparency without creating false confidence or excessive procedural burden.

The architecture remains limited by incomplete observation, residual confounding, source dependence, changing reporting behavior, mechanistic uncertainty, and disagreement among experts. It is not universally applicable, clinically prescriptive, regulator-approved, or deployment-ready. Context-specific validation is required before any operational use.

CONCLUSION

The proposed Signal-to-Causality Evidence Architecture separates statistical detection from causal assessment while connecting spontaneous reports, observational data, case chronology, bias analysis, mechanistic evidence, triangulation, causal-strength grading, and decision use through a traceable structure. Its principal contribution is not a new causal score, but an explicit account of how heterogeneous evidence should remain source-specific, contradiction-aware, and updateable before supporting a conditional drug-safety judgement. The architecture requires empirical evaluation of its reliability, transportability, usability, governance, and decision consequences. It cannot transform incomplete evidence into certainty, replace expert judgement, or establish clinical or regulatory readiness.

ACKNOWLEDGMENTS: None

CONFLICT OF INTEREST: None

FINANCIAL SUPPORT: None

ETHICS STATEMENT: None

REFERENCES

1. Insani WN, Pacurariu AC, Mantel-Teeuwisse AK, Gross-Martirosyan L. Characteristics of drugs safety signals that predict safety related product information update. Pharmacoepidemiol Drug Saf. 2018;27(7):789-96. doi:10.1002/pds.4446
2. van Hunsel F, de Jong E, Gross-Martirosyan L, Hoekman J. Signals from the Dutch national spontaneous reporting system: Characteristics and regulatory actions. Pharmacoepidemiol Drug Saf. 2021;30(8):1115-22. doi:10.1002/pds.5246
3. Potts J, Genov G, Segec A, Raine J, Straus S, Arlett P. Improving the safety of medicines in the European Union: From signals to action. Clin Pharmacol Ther. 2020;107(3):521-9. doi:10.1002/cpt.1678
4. Scosyrev E, Behr S, Jain D, Ponnuru A, Michel C. Disproportionality analysis and causal inference in drug safety. Pharm Med. 2025;39(2):97-107. doi:10.1007/s40290-024-00549-4
5. Khouri C, Revol B, Lepelley M, Mouffak A, Bernardeau C, Salvo F, et al. A meta-epidemiological study found lack of transparency and poor reporting of disproportionality analyses for signal detection in pharmacovigilance databases. J Clin Epidemiol. 2021;139:191-8. doi:10.1016/j.jclinepi.2021.07.014
6. Raschi E, Salvo F, Khouri C. Conceiving, conducting, reporting, interpreting, and publishing disproportionality analyses: A call to action. Br J Clin Pharmacol. 2022;88(7):3535-6. doi:10.1111/bcp.15269
7. 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
8. Norén GN, Donegan K, Muñoz MA, Alshammari TM, Pratt N, Candore G, et al. Recommendations to enable broader use of real-world evidence to inform decision-making throughout pharmacovigilance signal management. Pharmacoepidemiol Drug Saf. 2025;34(10). doi:10.1002/pds.70231
9. Kreimeyer K, Dang O, Spiker J, Muñoz MA, Rosner G, Ball R, et al. Feature engineering and machine learning for causality assessment in pharmacovigilance: Lessons learned from application to the FDA

Adverse Event Reporting System. Comput Biol Med. 2021;135:104517. doi:10.1016/j.compbiomed.2021.104517
10. Coste A, Wong A, Bokern M, Bate A, Douglas IJ. Methods for drug safety signal detection using routinely collected observational electronic health care data: A systematic review. Pharmacoepidemiol Drug Saf. 2023;32(1):28-43. doi:10.1002/pds.5548
11. Shin H, Lee S. An OMOP-CDM based pharmacovigilance data-processing pipeline (PDP) providing active surveillance for ADR signal detection from real-world data sources. BMC Med Inform Decis Mak. 2021;21(1):159. doi:10.1186/s12911-021-01520-y
12. Pinheiro LC, Candore G, Zaccaria C, Slattery J, Arlett P. An algorithm to detect unexpected increases in frequency of reports of adverse events in EudraVigilance. Pharmacoepidemiol Drug Saf. 2018;27(1):38-45. doi:10.1002/pds.4344
13. Inácio P, Cavaco A, Airaksinen M. The value of patient reporting to the pharmacovigilance system: A systematic review. Br J Clin Pharmacol. 2017;83(2):227-46. doi:10.1111/bcp.13098
14. Cherkas Y, Ide J, van Stekelenborg J. Leveraging machine learning to facilitate individual case causality assessment of adverse drug reactions. Drug Saf. 2022;45(5):571-82. doi:10.1007/s40264-022-01163-6
15. Um SH, Abuzgaia A, Rieder M. Comparison of the Liverpool Causality Assessment Tool vs. the Naranjo Scale for predicting the likelihood of an adverse drug reaction: A retrospective cohort study. Br J Clin Pharmacol. 2023;89(8):2407-12. doi:10.1111/bcp.15704
16. 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
17. Miyazaki M, Sakai T, Obara T, Mano N. The impact of regulation changes in the spontaneous reporting system for vaccines on reporting trends and signal detection in Japan. Pharmacoepidemiol Drug Saf. 2021;30(8):1091-100. doi:10.1002/pds.5231
18. Dijkstra L, Garling M, Foraita R, Pigeot I. Adverse drug reaction or innocent bystander? A systematic comparison of statistical discovery methods for spontaneous reporting systems. Pharmacoepidemiol Drug Saf. 2020;29(4):396-403. doi:10.1002/pds.4970
19. Bean DM, Wu H, Iqbal E, Dzahini O, Ibrahim ZM, Broadbent M, et al. Knowledge graph prediction of unknown adverse drug reactions and validation in electronic health records. Sci Rep. 2017;7(1):16416. doi:10.1038/s41598-017-16674-x
20. Zitnik M, Agrawal M, Leskovec J. Modeling polypharmacy side effects with graph convolutional networks. Bioinformatics. 2018;34(13). doi:10.1093/bioinformatics/bty294
21. Kontsioti E, Maskell S, Anderson I, Pirmohamed M. Identifying drug-drug interactions in spontaneous reports utilizing signal detection and biological plausibility aspects. Clin Pharmacol Ther. 2024;116(1):165-76. doi:10.1002/cpt.3258
22. Chandler RE. Nintedanib and ischemic colitis: Signal assessment with the integrated use of two types of real-world evidence, spontaneous reports of suspected adverse drug reactions, and observational data from large health-care databases. Pharmacoepidemiol Drug Saf. 2020;29(8):951-7. doi:10.1002/pds.5022
23. Hammerton G, Munafò MR. Causal inference with observational data: The need for triangulation of evidence. Psychol Med. 2021;51(4):563-78. doi:10.1017/S0033291720005127
24. 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
25. 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
26. Schuemie MJ, Ryan PB, Pratt N, Chen R, You SC, Krumholz HM, et al. Principles of large-scale evidence generation and evaluation across a network of databases (LEGEND). J Am Med Inform Assoc. 2020;27(8):1331-7. doi:10.1093/jamia/ocaa103
27. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: Introducing the E-value. Ann Intern Med. 2017;167(4):268-74. doi:10.7326/M16-2607
28. Langan SM, Schmidt SAJ, Wing K, Ehrenstein V, Nicholls SG, Filion KB, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). BMJ. 2018;363. doi:10.1136/bmj.k3532

29. Davis SE, Lasko TA, Chen G, Siew ED, Matheny ME. Calibration drift in regression and machine learning models for acute kidney injury. J Am Med Inform Assoc. 2017;24(6):1052-61. doi:10.1093/jamia/ocx030
30. Wadhwa D, Kumar K, Batra S, Sharma S. Automation in signal management in pharmacovigilance-an insight. Brief Bioinform. 2021;22(4). doi:10.1093/bib/bbaa363
31. Lee I, Jokinen JD, Crawford SY, Calip GS, Kilpatrick RD, Lee TA. Exploring completeness of adverse event reports as a tool for signal detection in pharmacovigilance. Ther Innov Regul Sci. 2021;55(1):142-51. doi:10.1007/s43441-020-00199-z