Available online **www.ijpras.com**

International Journal of Pharmaceutical Research & Allied Sciences, 2025, 14(4):57-65
https://doi.org/10.51847/RkUnYFkWDF

Original Article

ISSN : 2277-3657
CODEN(USA) : IJPRPM

An Adverse-Event Provenance Graph for Linking Case Narratives, Exposure Evidence, Clinical Context, Analytical Decisions, and Safety Conclusions

Saif Al-Hinai[1]*, Amal Al-Balushi[1], Mohammed Al-Maskari[2], Sultan Al-Mahruqi[3]

[1]Department of Adverse-Event Provenance and Safety Conclusions, College of Pharmacy, Sultan Qaboos University, Muscat, Oman.
[2]Department of Case Narratives and Exposure Evidence, College of Pharmacy, University of Nizwa, Nizwa, Oman.
[3]Department of Clinical Context and Analytical Decisions, College of Pharmacy, Dhofar University, Salalah, Oman.

***Email:** saif.alhinai@squ.edu.om*

ABSTRACT

Pharmacovigilance evidence is progressively transformed as source reports are received, linked, coded, interpreted, analyzed, corrected, and translated into safety conclusions. Existing systems frequently preserve the resulting records without preserving a complete, inspectable account of how each assertion, exposure estimate, contextual judgment, analytical result, and conclusion was derived. This article proposes an adverse-event provenance graph as an original conceptual architecture for linking source narratives, reporter and channel information, provisional case identities, product and exposure evidence, clinical context, competing explanations, coding operations, analytical decisions, and safety conclusions. The architecture separates original artifacts from extracted assertions, standardized codes, inferred relationships, statistical outputs, evidence interpretations, and decision statements. It further represents corrections, superseded records, human overrides, uncertainty, and historical graph states as first-class objects rather than undocumented workflow events. Evidence interpretation and safety conclusions are therefore modeled as traceable but contestable stages, not automatic consequences of signal detection. Validation would require source-to-assertion fidelity testing, duplicate-linkage assessment, temporal exposure reconstruction, coding reproducibility, clinical-context review, interpretation auditing, subgroup evaluation, change-control testing, and reconstruction of prior conclusions. Interoperability would require preservation of identifiers, uncertainty, semantic meaning, transformation history, and source linkage across systems. The principal contribution is a proposed assertion-centered architecture that treats provenance as part of the scientific evidence rather than administrative metadata. A complete provenance path would improve inspectability and error investigation, but it would not establish causal truth, clinical utility, regulatory acceptance, universal applicability, or deployment readiness.

Key words: *Pharmacovigilance, Drug safety, Safety signal, Causal inference, Adverse event, Real-world data*
Received: 04 August 2025; **Revised:** 09 November 2025; **Accepted:** 11 November 2025

INTRODUCTION

Pharmacovigilance encompasses the systematic collection, evaluation, interpretation, communication, and use of safety information generated from diverse clinical, observational, regulatory, and patient-derived evidence sources. Contemporary safety assessment is therefore not a single analytical event but a distributed process in which observations are transformed into increasingly structured forms of evidence. Although statistical signal

detection remains an important component of pharmacovigilance, safety conclusions cannot be reduced to the identification of disproportional reporting patterns or the management of a single reporting database. They depend on the interaction between reporting practices, clinical assessment, analytical methods, contextual interpretation, and human decision processes [1]. Consequently, the scientific reliability of a safety conclusion depends not only on what evidence is available but also on whether the pathway from original observation to final interpretation can be reconstructed.

The expansion of real-world data has increased both the opportunities and challenges associated with pharmacovigilance. Safety information may originate from spontaneous reports, electronic health records, administrative claims, registries, disease surveillance systems, patient-generated data, or other routinely collected sources. However, the meaning of these data depends on their origin, collection mechanism, population characteristics, observation period, missingness patterns, and intended analytical purpose. Therefore, apparently similar variables or outcomes may represent different underlying evidence structures when generated through different processes [2]. Without explicit representation of data origin and transformation history, downstream analyses may obscure important differences between observed events, derived variables, analytical constructs, and interpreted conclusions.

A safety signal represents an important stage in this evidence pathway but is not equivalent to an adverse event, an adverse drug reaction, a causal relationship, or a regulatory decision. Signal identification generally indicates that further evaluation may be warranted rather than establishing the presence, magnitude, or mechanism of harm. The interpretation of safety information involves multiple stakeholders, including clinicians, epidemiologists, regulators, data scientists, and pharmaceutical safety specialists, whose available evidence, analytical perspectives, and decision responsibilities may differ [3]. A scientifically defensible safety conclusion therefore requires traceability across multiple epistemic transitions: from source report to extracted information, from extracted information to structured evidence, from structured evidence to analytical output, and from analytical output to safety interpretation.

Despite advances in signal detection, natural language processing, real-world evidence generation, and knowledge representation, current safety workflows frequently separate these stages into disconnected systems. Narrative information may be transformed into coded concepts without preserving the original context. Exposure information may be reconstructed from incomplete or heterogeneous sources without retaining assumptions regarding timing, dose, or product identity. Analytical outputs may be archived without a complete record of the transformations, model versions, inclusion criteria, or interpretive decisions that produced them. These discontinuities create a provenance problem: the inability to determine how an individual assertion, evidence element, or conclusion was generated, modified, challenged, or accepted over time.

This Original Provenance Architecture Article proposes an adverse-event provenance graph as a conceptual framework for linking source reports, narrative assertions, exposure episodes, clinical context, competing explanations, coding transformations, analytical operations, evidence interpretation, and safety conclusions. The proposed architecture treats provenance as an integral component of scientific evidence rather than as administrative metadata. It introduces a structured representation of relationships among evidence nodes, transformations, uncertainties, corrections, and decision boundaries. The contribution is an original conceptual synthesis intended to define architectural requirements, evaluation criteria, and future validation pathways. It is not an empirically validated model, clinical recommendation, regulatory standard, or deployment specification.

The central premise of this article is that improving safety assessment requires preserving the lineage of evidence throughout its transformation rather than focusing exclusively on the final analytical output. A provenance-aware architecture may help distinguish reported information from inferred information, analytical association from causal interpretation, and computational processing from human judgement. However, the ability to reconstruct evidence pathways does not by itself establish correctness, causality, clinical utility, regulatory acceptance, or universal applicability. The proposed framework therefore defines a research direction whose components require empirical investigation, interoperability testing, and context-specific evaluation before any operational use can be considered.

Proposed adverse-event provenance graph

The proposed architecture is a typed, versioned assertion-and-provenance graph. Existing pharmacovigilance scholarship supports knowledge graphs as structures for integrating heterogeneous safety entities and relationships, while also showing that implementations and evidentiary purposes remain heterogeneous [4]. **Figure 1** presents the original conceptual synthesis for adverse-event provenance graph and principal evidence

nodes, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

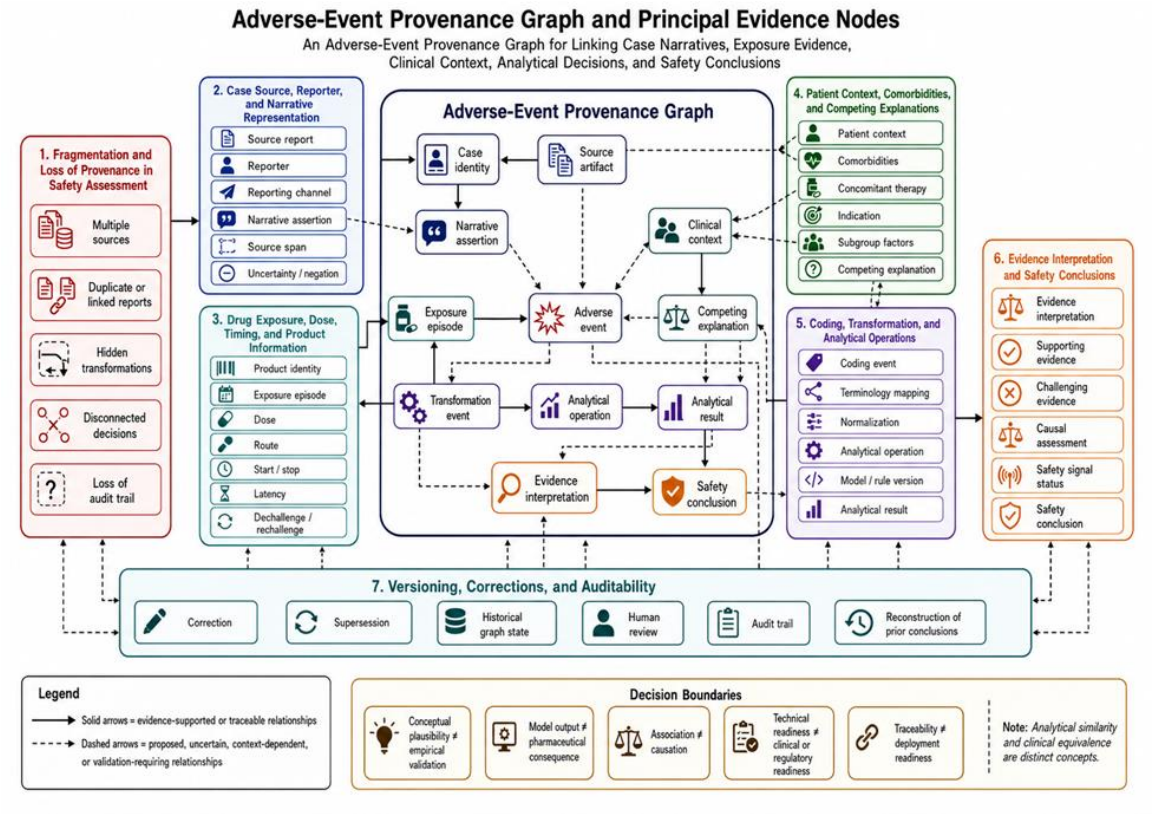

Figure 1. Adverse-Event Provenance Graph and Principal Evidence Nodes.

The figure is an original conceptual synthesis developed for "An Adverse-Event Provenance Graph for Linking Case Narratives, Exposure Evidence, Clinical Context, Analytical Decisions, and Safety Conclusions." 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 graph places source reports, reporter identities, reporting channels, and provisional case identities upstream of atomic assertions. Domain-specific guidance supports explicit definition of graph entities, relationships, identifiers, and construction procedures [5]. Applied adverse-reaction knowledge graphs further demonstrate that drug–event relationships can be represented computationally and assessed against external clinical evidence [6].

Together, existing pharmacovigilance graph scholarship and an implemented adverse-reaction knowledge graph show that heterogeneous entities and relationships can be integrated, but they do not validate the decision-provenance semantics proposed here [4-6]. The present architecture therefore separates established representational feasibility from its proposed end-to-end lineage model.

Principal node classes include source artifact, case identity, narrative assertion, exposure episode, product identity, adverse event, clinical context, competing explanation, coding event, analytical operation, analytical result, evidence interpretation, safety conclusion, correction, and supersession. Each node carries source, time, author or system, version, and uncertainty properties where applicable.

Typed edges distinguish what was reported, extracted, inferred, coded, calculated, supported, challenged, corrected, or superseded. This distinction prevents an inferred medication–event relation from appearing equivalent to original testimony and prevents a statistical association from appearing equivalent to causal assessment. The integrated taxonomy and edge semantics remain proposed and require empirical validation.

Case source, reporter, and narrative representation

A source report is the originating artifact received through a defined channel; a reporter is the person or organization supplying information; and a narrative assertion is an atomic statement anchored to a source location.

These elements should remain separate because patients and healthcare professionals may provide differently distributed but clinically useful information [7].

Reporting channel is part of provenance because capture interfaces, prompts, follow-up opportunities, and participation patterns can influence report characteristics. Evaluation of mobile adverse-reaction reporting demonstrated channel-specific differences in report quality and contribution [8]. Channel metadata should therefore contextualize evidence without automatically determining credibility.

Narratives should not disappear after coding. Natural-language processing can identify medications, events, temporality, and clinical context, but semantic ambiguity, negation, incomplete documentation, and integration limitations remain material [9]. Every extracted assertion should retain its source span, extraction method, model or rule version, uncertainty status, and human-review history.

The proposed representation also permits conflicting assertions to coexist. A reporter may describe an uncertain start date, while a later record supplies a different date or product. The graph should preserve both statements, their sources, and any reconciliation decision. Reporter identity and channel provide context; neither establishes factual correctness or causal attribution.

Drug exposure, dose, timing, and product information

Medication exposure should be represented as a time-bounded episode rather than a binary drug-present variable. Exposure definitions require explicit assumptions about initiation, discontinuation, persistence, switching, and permissible gaps [10].

Prescribed, dispensed, administered, and patient-reported exposure are not interchangeable. Each reflects a different observation process and source of error, so the graph should retain the provenance, temporal precision, and uncertainty of every exposure statement [11].

Medication and adverse-event relations may also be extracted from clinical narratives. Such extraction can support structured analysis, but its outputs remain derived assertions linked to the text span, extraction method, and model version that produced them [12].

The proposed exposure node therefore records ingredient, product, formulation, dose, route, frequency, start and stop intervals, latency, interruption, dechallenge, rechallenge, and concomitant treatment when available. These properties improve traceability but do not independently establish causal attribution.

Patient context, comorbidities, and competing explanations

Patient context includes age, sex, indication, disease severity, comorbidities, organ function, pregnancy status, concomitant therapy, prior events, and other factors that may alter exposure or interpretation. Reported adverse-reaction patterns can differ across patient subgroups [13].

Observed subgroup differences may reflect biology, prescribing, healthcare access, reporting behavior, exposure opportunity, or analytical choices. They should therefore be represented as contextual evidence rather than automatically interpreted as susceptibility or causal effect.

Competing explanations require similarly explicit representation. Causality-assessment methods vary in structure and evidentiary requirements, particularly regarding temporality, alternative causes, dechallenge, rechallenge, and prior knowledge [14].

The proposed graph links each competing explanation to the assertions that support or challenge it and records whether it was considered, unresolved, or rejected. **Table 1** organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for definitions, components, relationships, and intended uses in an adverse-event provenance graph for linking case narratives, exposure evidence.

Table 1. Definitions, components, relationships, and intended uses in An Adverse-Event Provenance Graph for Linking Case Narratives, Exposure Evidence.

Component or scientific dimension	Problem addressed	Inputs or determinants	Proposed mechanism or relationship	Expected contribution	Evidence required	Failure or uncertainty risk	Boundary statement
Source report	Loss of originating evidence	Narrative, form, channel, timestamp	Proposed immutable artifact linked to	Source-level reconstruction	Source-to-record concordance	Missing or altered source	Preservation does not prove accuracy

			later assertions				
Case identity	Duplicate or fragmented reports	Identifiers, dates, products, narratives	Proposed same-case-as or possible-duplicate-of relation	Transparent identity resolution	Curated duplicate pairs [15]	False merge or missed duplicate	Linkage is probabilistic
Reporter and channel	Context stripped from content	Reporter type, role, interface	Proposed contextual link, not credibility score	Interpretive transparency	Reporter and channel comparisons [7, 8]	Systematic reporting differences	Metadata does not establish truth
Narrative assertion	Coding replaces original meaning	Source span, subject, negation, uncertainty	Proposed atomic assertion anchored to text	Inspectable extraction	Clinician-annotated narratives [9]	Ambiguity or extraction error	Extracted text is not ground truth
Exposure episode	Binary or imprecise exposure	Product, dose, route, dates, adherence	Proposed uncertain temporal episode	Better temporal reasoning	Source-adjudicated exposure data [10, 11]	Misclassification and missing timing	Exposure does not prove causation
Clinical context	Competing factors omitted	Comorbidities, indication, subgroup variables	Proposed context nodes linked to assertions	More explicit alternative explanations	Clinician review and subgroup analysis [13, 14]	Confounding or incomplete context	Context cannot eliminate residual bias
Transformation event	Hidden coding or normalization	Input, terminology, rule, version	Proposed first-class derivation node	Reproducibility and auditability	Re-executable transformations	Version loss or semantic distortion	Reproducibility is not validity
Safety conclusion	Result conflated with decision	Evidence, uncertainty, analyst rationale	Proposed conclusion node linked to support and challenge	Inspectable reasoning	Independent decision audit	Unsupported or irreversible judgement	Traceability does not confer authority

Coding, transformation, and analytical operations

Coding converts free-text descriptions into standardized concepts. Automated narrative-to-MedDRA mapping demonstrates the feasibility of this transformation but also shows that coding is an interpretive operation rather than a verbatim copy of the source [16].

Seriousness, expectedness, relatedness, and case priority may also be assigned by rules or computational systems. Automated seriousness determination illustrates why model output, confidence, version, reviewer intervention, and final disposition should remain distinct [17].

The proposed architecture represents every coding, mapping, normalization, linkage, query, and model execution as a transformation event. Required properties include input identifiers, software or terminology version, parameters, execution time, responsible actor, output, and any human modification.

A reproducible transformation may still be clinically inappropriate or semantically incomplete. The graph therefore distinguishes whether an operation can be repeated from whether its assumptions, outputs, and intended decision use have been validated.

Evidence interpretation and safety conclusions

Analytical results depend on data construction, comparator selection, statistical method, thresholds, and handling of duplication and missingness. Comparative evaluation of spontaneous-reporting methods shows that discovery results can differ across analytical approaches [18].

A statistical association is therefore represented as an analytical result, not as a causal conclusion. Safety decision-making additionally requires reasoning about evidence strength, uncertainty, counterevidence, clinical plausibility, and differing stakeholder perspectives [19].

The proposed graph separates analytical result, evidence interpretation, causal assessment, signal status, and safety conclusion. A conclusion can be linked simultaneously to supporting evidence, challenging evidence, unresolved questions, decision context, and the person or body responsible for the judgement.

Table 2 organizes the evidence, constructs, relationships, uncertainties, and boundary conditions required for testable propositions, evidence requirements, and validation criteria for an adverse-event provenance graph for linking case narratives, exposure evidence.

Table 2. Testable propositions, evidence requirements, and validation criteria for An Adverse-Event Provenance Graph for Linking Case Narratives, Exposure Evidence.

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
Source capture	Preserve originating evidence	Report to immutable artifact	Feeds assertion extraction	Source content remains inspectable	Incomplete or altered capture	Data steward verifies integrity	Technical capture only
Identity resolution	Relate possible duplicate reports	Reports to provisional case identity	Preserves source-specific differences	Linked and unlinked cases match adjudication	False merge or missed linkage	Trained reviewers adjudicate samples	No assumption of perfect identity
Assertion extraction	Structure narrative content	Text span to atomic assertion	Connects source, exposure, event, and context	Assertions retain source, negation, and uncertainty	Hallucinated or omitted assertion	Clinical annotators assess fidelity	Extraction is not clinical truth
Exposure reconstruction	Organize dose and timing	Multiple exposure statements to episode	Interacts with event timing and context	Temporal precision and conflicts are preserved	Incorrect latency or product assignment	Pharmacoepidemiology review	No causal inference from timing alone
Transformation layer	Record coding and derivation	Inputs through rules or models to outputs	Supplies analytical datasets	Operation is reproducible with recorded version	Semantic loss or undocumented override	Terminology and model owners	Reproducibility is not validity
Analytical layer	Produce method-specific evidence	Curated graph state to analytical result	Feeds interpretation, not conclusion directly	Inputs, assumptions, and outputs are recoverable	Method-dependent or biased result	Analysts document rationale	Association remains noncausal
Interpretation layer	Weigh evidence and uncertainty	Results and context to reasoned assessment	Supports or challenges conclusions	Independent reviewer can reconstruct reasoning	Hidden assumptions or selective evidence	Qualified safety assessors	No automated decision authority
Version and audit layer	Preserve change history	Corrections and supersession across graph states	Applies to every node and edge	Historical conclusions can be reconstructed	Overwriting or missing lineage	Governance body controls changes	Auditability does not prove correctness

Versioning, corrections, and auditability

Transparent pharmacoepidemiologic reporting requires explicit description of data sources, definitions, processing, design, and analysis [20]. Comparable transparency is needed within the provenance graph so that every conclusion can be reconstructed from the applicable evidence state.

Corrections should create new versioned events rather than silently overwrite prior records. Postmarketing use of artificial intelligence similarly requires verification, reproducibility, uncertainty management, and meaningful human oversight [21].

The proposed architecture assigns persistent identifiers to nodes and edges and records creation, modification, correction, reviewer action, and supersession. Historical graph states should remain recoverable, including conclusions later revised after follow-up evidence or corrected case linkage.

Auditability enables investigation of who changed what, when, why, and with which evidence. It does not prevent poor judgement, guarantee data completeness, or convert a documented decision into a scientifically or regulatorily valid one.

Validation and interoperability requirements

Validation must occur at the level of individual graph functions and intended uses. Clinical concepts and outcomes derived from real-world data require purpose-specific validation against appropriate reference standards rather than acceptance based solely on coding or availability [22].

Interoperability must preserve more than field names. Canonical FHIR repositories can support reproducible data transformation, but exchange must retain identifiers, source linkage, uncertainty, terminology versions, and derivation history [23]. Validation must combine transparent specification, independently checked clinical representations, and reproducible transformation across interoperable information models [20, 22, 23].

Evaluation should proceed through schema-conformance tests, clinician-annotated narratives, duplicate-linkage studies, exposure reconstruction, coding reproducibility, interpretation audits, historical-state recovery, cross-system transfer tests, and prospective shadow-mode assessment.

Successful technical exchange would not establish clinical utility or decision benefit. Separate studies would be required to determine whether the architecture improves error detection, reasoning quality, timeliness, consistency, or outcomes in a defined operational setting.

Limitations and implementation priorities

The graph cannot recover facts that were never observed, reported, or retained. It remains vulnerable to missing dates, ambiguous products, conflicting narratives, selective reporting, uncertain linkage, inaccessible records, ontology mismatch, and privacy constraints.

Its analytical components may also degrade as data sources, populations, coding practices, and workflows change. Empirical studies show that data drift must be detected rather than assumed absent [24]. Aggregate performance can additionally conceal unequal error burdens across patient subgroups [25].

Figure 2 presents the original conceptual synthesis for traceable pathway from source report to safety conclusion, showing how the article's principal components, evidence relationships, uncertainties, and decision boundaries are connected.

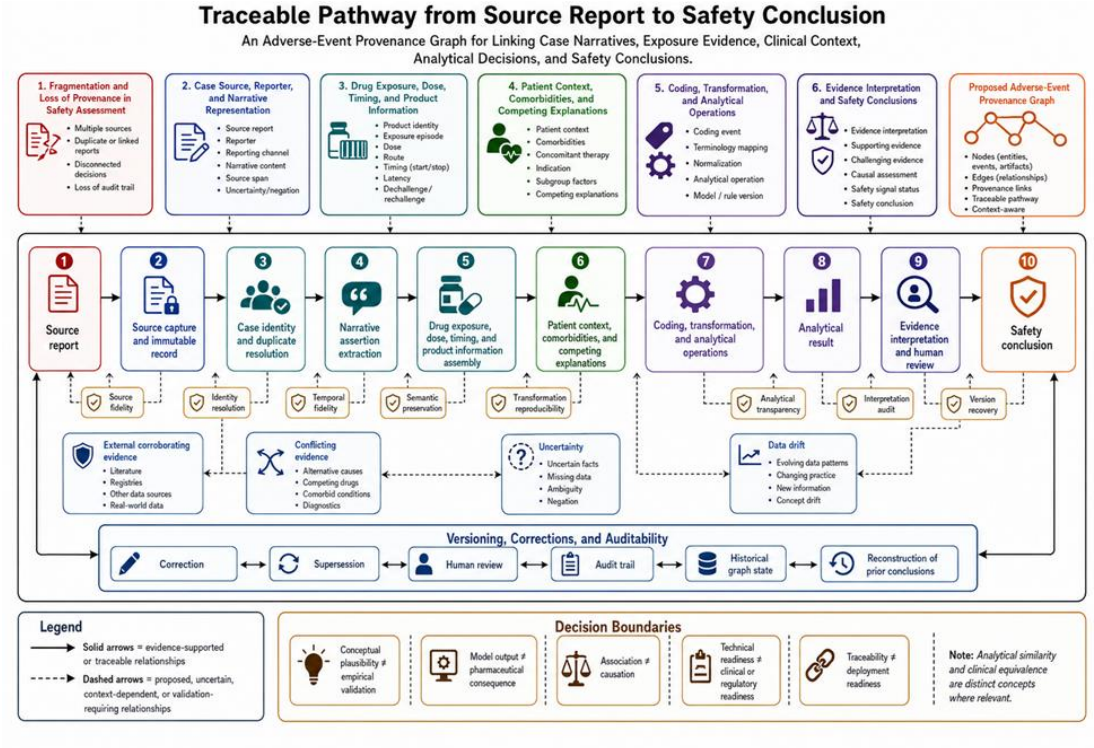

Figure 2. Traceable Pathway from Source Report to Safety Conclusion.

The figure is an original conceptual synthesis developed for "An Adverse-Event Provenance Graph for Linking Case Narratives, Exposure Evidence, Clinical Context, Analytical Decisions, and Safety Conclusions." 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 priorities are to define a minimal provenance schema, establish reference corpora, test source and temporal fidelity, evaluate interoperability, quantify reviewer burden, and determine whether provenance improves decisions relative to existing practice.

Any operational evaluation should include human verification, drift assessment, subgroup analysis, change control, and rollback procedures [21, 24, 25]. These requirements define research priorities; they do not demonstrate that the architecture is technically mature, clinically beneficial, regulator acceptable, or ready for deployment.

CONCLUSION

The proposed adverse-event provenance graph treats lineage as part of safety evidence rather than administrative metadata. By separating source artifacts, assertions, exposure episodes, contextual factors, transformations, analytical results, interpretations, conclusions, and corrections, it offers a structured basis for reconstructing how a safety judgement was formed and revised. Its value remains conditional on empirical validation, interoperable implementation, human governance, and evidence that traceability improves defined safety-assessment tasks. The architecture does not itself establish causality, correctness, clinical utility, regulatory acceptance, universal applicability, or deployment readiness.

ACKNOWLEDGMENTS: None

CONFLICT OF INTEREST: None

FINANCIAL SUPPORT: None

ETHICS STATEMENT: None

REFERENCES

1. Santoro A, Genov G, Spooner A, Raine J, Arlett P. Promoting and protecting public health: How the European Union pharmacovigilance system works. Drug Saf. 2017;40(10):855-69. doi:10.1007/s40264-017-0572-8
2. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W. What is real-world data? A review of definitions based on literature and stakeholder interviews. Value Health. 2017;20(7):858-65. doi:10.1016/j.jval.2017.03.008
3. Sartori D, Aronson JK, Norén GN, Onakpoya IJ. Signals of adverse drug reactions communicated by pharmacovigilance stakeholders: A scoping review of the global literature. Drug Saf. 2023;46(2):109-20. doi:10.1007/s40264-022-01258-0
4. Hauben M, Rafi M, Abdelaziz I, Hassanzadeh O. Knowledge graphs in pharmacovigilance: A scoping review. Clin Ther. 2024;46(7):544-54. doi:10.1016/j.clinthera.2024.06.003
5. Hauben M, Rafi M. Knowledge graphs in pharmacovigilance: A step-by-step guide. Clin Ther. 2024;46(7):538-43. doi:10.1016/j.clinthera.2024.03.006
6. 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
7. Rolfes L, van Hunsel F, van der Linden L, Taxis K, van Puijenbroek E. The quality of clinical information in adverse drug reaction reports by patients and healthcare professionals: A retrospective comparative analysis. Drug Saf. 2017;40(7):607-14. doi:10.1007/s40264-017-0530-5
8. Oosterhuis I, Taavola H, Tregunno PM, Mas P, Gama S, Newbould V, et al. Characteristics, quality and contribution to signal detection of spontaneous reports of adverse drug reactions via the WEB-RADR mobile application: A descriptive cross-sectional study. Drug Saf. 2018;41(10):969-78. doi:10.1007/s40264-018-0679-6

9. Luo Y, Thompson WK, Herr TM, Zeng Z, Berendsen MA, Jonnalagadda SR, et al. Natural language processing for EHR-based pharmacovigilance: A structured review. Drug Saf. 2017;40(11):1075-89. doi:10.1007/s40264-017-0558-6
10. Clary A, Lin ND, Lasky T, Reynolds MR, Chokkalingam A, Rodriguez-Watson C, et al. Considerations for defining medication exposure when analyzing real-world data. Pharmacoepidemiol Drug Saf. 2023;32(8):933-7. doi:10.1002/pds.5613
11. 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
12. Wei Q, Ji Z, Li Z, Du J, Wang J, Xu J, et al. A study of deep learning approaches for medication and adverse drug event extraction from clinical text. J Am Med Inform Assoc. 2020;27(1):13-21. doi:10.1093/jamia/ocz063
13. Watson S, Caster O, Rochon PA, den Ruijter H. Reported adverse drug reactions in women and men: Aggregated evidence from globally collected individual case reports during half a century. EClinicalMedicine. 2019;17:100188. doi:10.1016/j.eclinm.2019.10.001
14. Mascolo A, Scavone C, Sessa M, di Mauro G, Cimmaruta D, Orlando V, et al. Can causality assessment fulfill the new European definition of adverse drug reaction? A review of methods used in spontaneous reporting. Pharmacol Res. 2017;123:122-9. doi:10.1016/j.phrs.2017.07.005
15. Kreimeyer K, Menschik D, Winiecki S, Paul W, Barash F, Woo EJ, et al. Using probabilistic record linkage of structured and unstructured data to identify duplicate cases in spontaneous adverse event reporting systems. Drug Saf. 2017;40(7):571-82. doi:10.1007/s40264-017-0523-4
16. Combi C, Zorzi M, Pozzani G, Moretti U, Arzenton E. From narrative descriptions to MedDRA: Automagically encoding adverse drug reactions. J Biomed Inform. 2018;84:184-99. doi:10.1016/j.jbi.2018.07.001
17. Routray R, Tetarenko N, Abu-Assal C, Mockute R, Assuncao B, Chen H, et al. Application of augmented intelligence for pharmacovigilance case seriousness determination. Drug Saf. 2020;43(1):57-66. doi:10.1007/s40264-019-00869-4
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. Hammad TA, Davies S. Navigating diverging perspectives: Reasoning, evidence, and decision-making in drug safety. Drug Saf. 2025;48(6):587-93. doi:10.1007/s40264-025-01537-6
20. 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
21. Ball R, Talal AH, Dang O, Muñoz M, Markatou M. Trust but verify: Lessons learned for the application of AI to case-based clinical decision-making from postmarketing drug safety assessment at the US Food and Drug Administration. J Med Internet Res. 2024;26. doi:10.2196/50274
22. Weinstein EJ, Ritchey ME, Lo Re V 3rd. Core concepts in pharmacoepidemiology: Validation of health outcomes of interest within real-world healthcare databases. Pharmacoepidemiol Drug Saf. 2023;32(1):1-8. doi:10.1002/pds.5537
23. Lenert LA, Ilatovskiy AV, Agnew J, Rudisill P, Jacobs J, Weatherston D, et al. Automated production of research data marts from a canonical Fast Healthcare Interoperability Resource data repository: Applications to COVID-19 research. J Am Med Inform Assoc. 2021;28(8):1605-11. doi:10.1093/jamia/ocab108
24. Kore A, Abbasi Bavil E, Subasri V, Abdalla M, Fine B, Dolatabadi E, et al. Empirical data drift detection experiments on real-world medical imaging data. Nat Commun. 2024;15(1):1887. doi:10.1038/s41467-024-46142-w
25. Seyyed-Kalantari L, Zhang H, McDermott MBA, Chen IY, Ghassemi M. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. Nat Med. 2021;27(12):2176-82. doi:10.1038/s41591-021-01595-0