INSIGHTIA JOURNAL OF RESEARCH AND INNOVATION
OPEN-ACCESS JOURNAL (ISSN: XXX)

Review
Volume 1, Issue 03, 2026, 01
https://doi.org/10.33/IJRI13.01

AI and ML Driven Pharmacovigilance: Advancing ADR Detection and Risk Management Using Real-World Data

Yash Saxena [1], Akash Sharma [2,*], Richa Kharkwal [3], Hritik Goswami [4], Tarun K. Yadav [5]

[1,2,3,5] Moradabad Educational Trust Group of Institutions Faculty of Pharmacy, Moradabad-244001, Uttar Pradesh, India; e-mail@e-mail.com (F.L.);
[4] Department of Pharmaceutical Sciences, Sir J.C Bose Technical Campus, Bhimtal, Kumaun University, Nainital- 263136, Uttarakhand, India; e-mail@e-mail.com (F.L.);
* Correspondence: akashsharma3669@gmail.com (A.S.)

Received: date; Accepted: date; Published: date

Abstract: Pharmacovigilance (PV) is increasingly defined by the ability to integrate high-volume, heterogeneous evidence generated during routine medicine use. Spontaneous individual case safety reports remain central to global monitoring, but under-reporting, duplicate and incomplete cases, stimulated reporting, variable terminology, and the absence of a stable denominator can constrain signal interpretation. Real-world data (RWD) from electronic health records (EHRs), claims, registries, literature, patient forums, and social media provide complementary temporal and clinical context. Artificial intelligence (AI), machine learning (ML), deep learning, natural language processing (NLP), and graph-based methods can automate parts of case triage, extract drug-event relations, identify non-linear patterns, prioritize candidate signals, and support predictive surveillance. In particular, NLP and transformer models improve the ability to interpret clinical narratives, while graph learning can represent interactions among medicines, diagnoses, patient factors, and adverse outcomes. The value of these methods depends not only on discrimination but also on data provenance, phenotype definition, calibration, external validation, causal reasoning, fairness, privacy, and regulatory credibility. This review synthesizes developments relevant to the 2016-2026 period while retaining earlier methodological foundations where needed. The emphasis is on a human-in-the-loop architecture in which AI expands surveillance capacity but does not substitute for clinical assessment, causal evaluation, or regulatory decision-making.

Keywords: Pharmacovigilance; Adverse Drug Reactions; Artificial Intelligence; Machine Learning; Natural Language Processing; Risk Management.

© 2026 by the authors. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Pharmacovigilance (PV) is the systematic science and set of activities used to detect, assess, understand, and prevent adverse effects and other medicine-related problems. Post-marketing surveillance is especially important because routine clinical practice exposes medicines to broader populations, longer treatment durations, more combinations, and more complex disease states than are usually represented in pivotal trials [1-3]. The World Health Organization (WHO) Programme for International Drug Monitoring established a global framework in which national centres contribute individual case safety reports to support signal generation and international safety assessment [2].

The principal analytical challenge is that modern safety information is both large and heterogeneous. Individual case safety reports contain structured fields and narrative text; EHRs add longitudinal diagnoses, laboratory results, prescriptions, procedures, and clinical notes;

claims provide broad healthcare-utilization patterns; registries can contribute disease- or product-specific follow-up; and patient-generated data capture experiences that may not reach formal reporting channels [4-20]. A 2023 systematic review of signal-detection methods using routinely collected healthcare data identified a wide variety of statistical, epidemiological, and machine learning (ML) approaches and emphasized heterogeneity in performance reporting and limited availability of common reference standards [5,18].

Within this environment, Artificial intelligence (AI) should be interpreted as a family of methods rather than a single technology. ML learns predictive or classificatory patterns from examples; deep learning learns high-level representations from complex data; natural language processing (NLP) converts free text into computable concepts; graph learning models relationships among entities; and generative or large language models (LLMs) can organize and summarize evidence. These capabilities complement, rather than replace, conventional disproportionality methods and expert review [21-30]. The utilization of ML in the healthcare sector is illustrated in Figure 1.

This review focuses on the transition from 2016 through 2026, with earlier references retained only where they provide foundational definitions or methods. The identification and selection of the studies included in this review are summarized in Figure 2. The central questions are how AI and ML contribute to adverse drug reaction (ADR) detection, how RWD can be combined across sources, how signals can be moved from screening to verification and risk management, and which governance requirements determine whether an algorithm is suitable for real-world deployment. Recent regulatory activity is increasingly framed around risk-based, context-of-use assessment for AI-generated evidence rather than blanket assumptions about model acceptability [3,31-33].

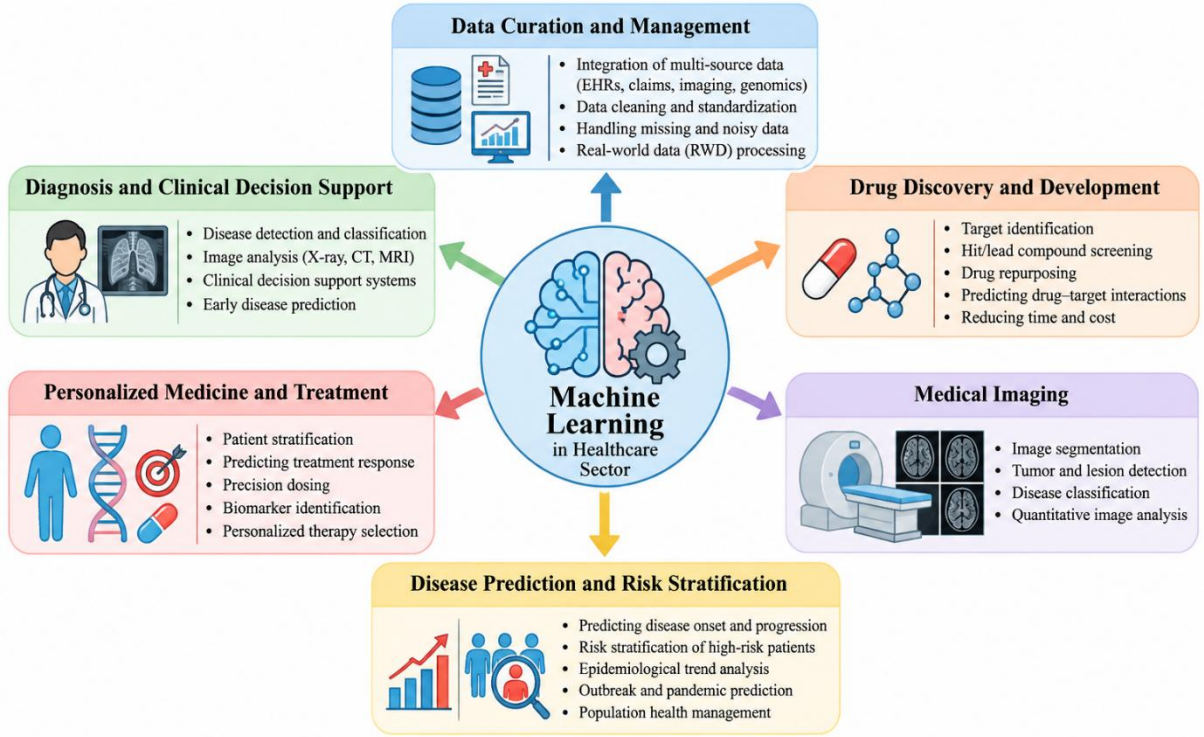

Figure 1. Utilization of machine learning (ML) in Healthcare sector.

2. Adverse Drug Reactions: Definitions, Classification, and Burden

An ADR is a harmful and unintended response to a medicinal product at normally used doses. In operational PV, ADRs must be distinguished from the broader concept of adverse

drug events, which can include medication errors, overdoses, and other preventable incidents. This distinction matters for AI because model labels define the outcome being learned: a system trained to identify any adverse event (AE) may not produce the same construct as a model trained on adjudicated drug-related causality.

The Rawlins-Thompson framework groups ADRs into six practical categories. Type A reactions are dose-related and pharmacologically predictable; Type B reactions are bizarre or idiosyncratic and less predictable; Type C reactions are chronic and associated with cumulative or time-dependent exposure; Type D reactions are delayed; Type E reactions occur during withdrawal; and Type F reactions represent unexpected failure of therapy. The classification retained in the source manuscript is reproduced in Table 1.

Table 1. Classification of ADRs

Types of Reaction	Type of effect	Characteristics	Frequency	Examples	Management
A	Augmented	Dose related; predictable pharmacology	Common	Respiratory depression caused by antihypertensive medications and orthostatic hypotension; digoxin toxicity; serotonin syndrome with selective serotonin reuptake inhibitors (SSRIs); warfarin hemorrhage; anticholinergic effects of tricyclic antidepressants.	Reduce or withdraw the dose when appropriate; assess concurrent treatment and drug-drug interactions.
B	Bizarre	Non-dose related; unpredictable	Uncommon	Anaphylaxis to penicillin and idiosyncratic reactions such as malignant hyperthermia with anesthetics.	Withdraw the offending medicine and avoid re-exposure when indicated.
C	Chronic	Cumulative; dose related; time related	Uncommon	Corticosteroid-induced suppression of the hypothalamic-pituitary-adrenal axis.	Dose reduction or withdrawal when required; management may need to be prolonged.
D	Delayed	Time-related and often dose-related	Uncommon	Tardive dyskinesia; teratogenesis; carcinogenesis.	Often difficult to reverse; prevention and early recognition are important.
E	End-of-treatment	Related to withdrawal time	Uncommon	Myocardial ischemia after beta-blocker cessation; withdrawal symptoms from benzodiazepines or opiates.	Gradual withdrawal and, where clinically appropriate, reintroduction of therapy.
F	Unexpected failure of therapy	Dose-related; interaction-related	Common	Antimicrobial resistance; inadequate oral-contraceptive exposure when combined with an enzyme inducer.	Adjust dose when appropriate and assess concurrent medicines and interactions.

The burden of medicine-related harm is clinically important. Prospective evidence synthesized by Lazarou et al. established that serious ADRs contribute to hospitalization and mortality, while subsequent studies have documented ADR burden in primary care and among older adults [34-38]. The epidemiology is not uniform: age, treatment setting, route of

administration, comorbidity, polypharmacy, and therapeutic class can modify observed patterns [34-37].

The supplied dataset also demonstrates the importance of multidimensional stratification. It reports significant associations between therapeutic category and age, administration route, and ADR system manifestation, with chi-square statistics of 110.022, 137.588, and 366.906, respectively, all reported with $p<0.0001$. These values are reproduced exactly in Table 2 from the supplied source dataset and should be interpreted only within that dataset; they are not a basis for generalizing national or global ADR epidemiology.

Table 2 illustrates how a safety signal can be hidden when exposure context is collapsed into a simple drug-event count. Within the supplied data, anti-tumor medicines dominate several severe ADR strata, intravenous administration is frequent among reported routes, and hematologic manifestations are prominent in the anti-tumor category. These observations motivate the integration of patient-level and contextual variables into ML pipelines, while preserving the distinction between descriptive association and causal interpretation [5,8,22,39].

3. Real-World Data Streams for Pharmacovigilance

3.1. Spontaneous Reporting Systems

Spontaneous reporting systems (SRSs), including the U.S. Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS), EudraVigilance, and VigiBase, remain indispensable because they can capture rare, unexpected, and geographically diverse events. Their limitation is that reporting behavior is not a controlled sampling mechanism. Under-reporting, duplicate submissions, stimulated reporting, selective publicity, missing denominators, and variable clinical detail can alter the apparent frequency of drug-event pairs [4,9,10,40-43].

Modern signal management therefore treats SRS data as hypothesis-generating evidence. Deduplication, case-quality scoring, provenance tracking, missing-data characterization, and explicit review of reporting artifacts should precede escalation. The Uppsala Monitoring Centre (UMC) workflow described in Section 7.2 reflects this logic: statistical screening is followed by manual assessment and then deeper clinical review [1-3,9,10,31,39,41,43].

3.2. Electronic Health Records, Claims, and Registries

EHRs can supplement spontaneous reports by providing longitudinal clinical context, including diagnoses, medication exposure, laboratory measurements, procedures, outcomes, and narrative notes. Claims databases broaden healthcare utilization capture, while registries can supply disease- or product-specific follow-up. Evidence from systematic reviews indicates that EHR-based signal detection can support complementary perspectives on safety, especially when combined with other sources [5,7,8,18,19,22].

A major opportunity lies in combining structured and unstructured EHR information. A medication order, laboratory trend, diagnosis code, and clinical narrative may jointly reconstruct exposure, temporal sequence, dechallenge, and clinical reasoning. However, the same EHR can also contain duplicated encounters, institutional coding differences, missing documentation, and informative missingness, requiring careful phenotype definition and transportability assessment [7,8,18,28,44]. The EHR modules and free-text ADR notes referred to above are shown in Figure 3.

Table 2. Significance analysis of the influence of drug type on other factors of severe ADRs

Domain	Subgroup	Anti-tumor drugs	Anti-infectious drugs	Circulating Systems Drugs	Respiratory Drugs	Blood System drugs	Others	Total	X^2	P
Age	<18	118	7	0	0	0	14	139	110.022	<0.0001
	19-45	18	7	1	3	0	11	40	110.022	<0.0001
	46-65	71	13	10	9	5	42	150	110.022	<0.0001
	66-79	31	11	4	6	1	10	63	110.022	<0.0001
	>80	0	4	2	0	3	7	16	110.022	<0.0001
Administration route	Intravenous drip	145	30	7	13	0	25	220	137.588	<0.0001
	Oral	6	0	4	2	4	12	28	137.588	<0.0001
	Nebulization inhalation	0	0	0	2	0	0	2	137.588	<0.0001
	Subcutaneous injection	6	0	0	0	2	1	9	137.588	<0.0001
	Others	81	12	6	1	3	46	149	137.588	<0.0001
ADR system	Hematologic	201	4	2	1	1	8	217	366.906	<0.0001
	Mucocutaneous System	4	31	3	12	1	30	81	366.906	<0.0001
	Digestive System	18	2	4	0	3	4	31	366.906	<0.0001
	Others	6	2	2	1	3	14	28	366.906	<0.0001
	Systemic damage	5	2	0	0	1	13	21	366.906	<0.0001
	Respiratory System	4	1	2	1	0	9	17	366.906	<0.0001
	Cardiovascular System	0	0	4	3	0	6	13	366.906	<0.0001

 © 2026. InsightiaPubs. All rights reserved.

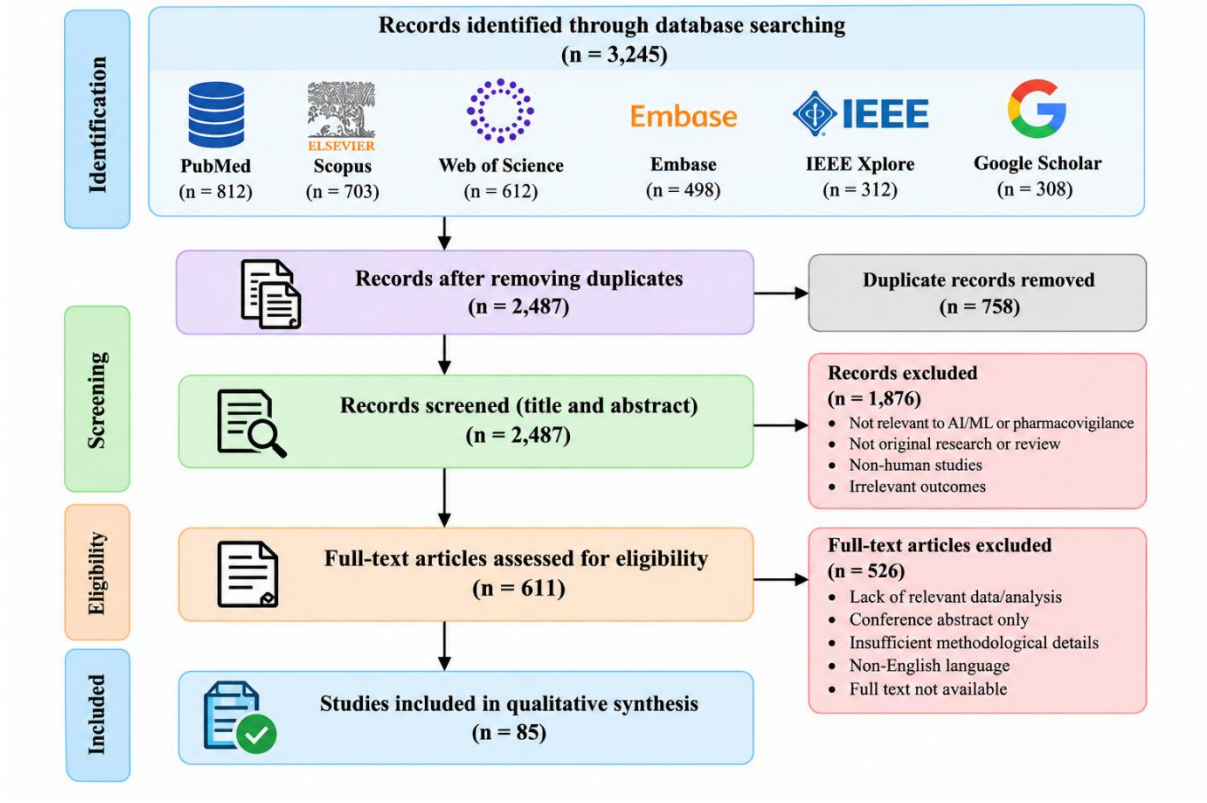

Figure 2. Identification of studies via databases.

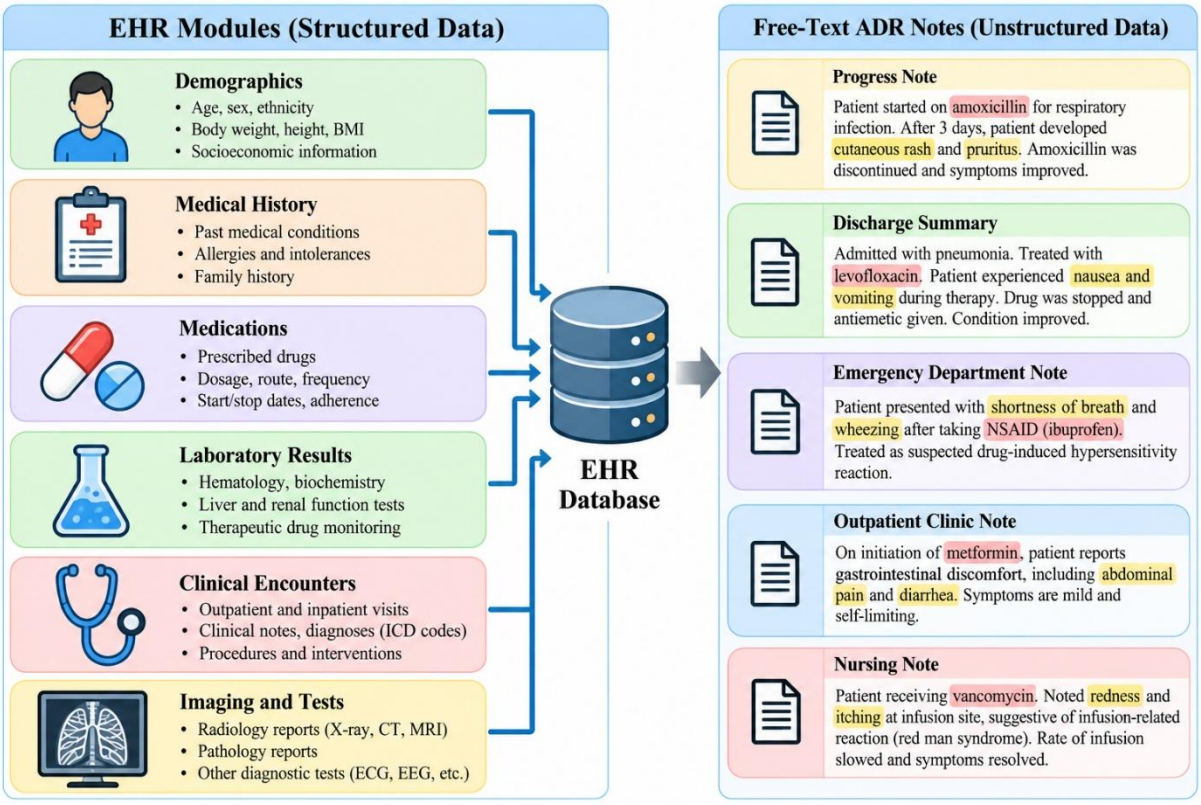

Figure 3. EHR modules and free-text ADR notes.

3.3. Clinical Trials and Pre-Approval Evidence

Randomized clinical trials remain essential for estimating common adverse effects under controlled exposure, but they often have limited power for rare events and may exclude older adults, children, multimorbidity, and complex polypharmacy. Linking trial evidence to RWD can provide continuity across the product life cycle. Trial annotations can seed

supervised models, while post-marketing data can reveal low-frequency or context-dependent phenotypes that were not visible during development [3,16,45].

3.4. Social Media and Patient-Generated Data

Patient forums and social platforms offer a patient-centered surveillance layer. They can contain symptom narratives, quality-of-life effects, suspected interactions, and experiences that never enter an SRS. NLP-based studies have shown that sequence labeling and related methods can recover ADR mentions from social content, while online-forum research demonstrates potential use for patient engagement and hypothesis generation [4,14,15,19,20,46].

The principal limitation is evidentiary context. Posts may be duplicated, multilingual, colloquial, incomplete, or temporally ambiguous, and demographic information may be absent. Social-media signals should therefore be triangulated against EHRs, claims, registries, clinical literature, or case reports rather than interpreted as stand-alone causal evidence [14,19,20,46].

4. Conventional Statistical Signal Detection

Disproportionality analysis remains the first analytical layer in many PV environments. In a 2x2 contingency table, the frequency of a target drug-event combination is compared with the reporting background. The proportional reporting ratio (PRR) and reporting odds ratio (ROR) are frequentist measures, whereas the Information Component (IC) and related Bayesian approaches use shrinkage to stabilize estimates when counts are small [10-12,40,42,43]. The main features of these three measures are compared in Table 3.

Table 3. Comparison of PRR, ROR, and IC.

Feature	PRR	ROR	IC (Bayesian Confidence Propagation Neural Network, BCPNN)
Type of Method	Frequentist	Frequentist	Bayesian
Basic concept	Compares a drug's percentage of a certain AE with that of all other medications.	Odds of reporting an AE with a drug versus odds with other drugs.	Measures deviation from the expected drug-AE combination using Bayesian shrinkage.
Formula (simplified)	(a/(a+b))/(c/(c+d))	(a/b)/(c/d) = ad/bc	log2(observed/expected)
Data requirement	2x2 contingency table	2x2 contingency table	2x2 contingency table
Sensitivity	Moderate	High	Moderate
Specificity	Moderate	Moderate	High
False Positives	High risk	Higher risk	Lower risk
Complexity	Very simple	Simple	More complex
Regulatory Use	Used by the Medicines and Healthcare products Regulatory Agency (MHRA, UK)	Widely used in research and regulatory practice	Used by WHO-UMC (VigiBase)
Database Compatibility	FAERS, EudraVigilance	FAERS, EudraVigilance	VigiBase (WHO), large datasets
Strengths	Easy to calculate and interpret	Statistically robust and widely used	Handles small counts and variability well
Limitations	Sensitive to reporting bias and sparse cells	Similar limitations to PRR	Harder to interpret and more computationally intensive

No disproportionality metric is universally optimal. PRR is transparent and easy to compute but can be unstable with sparse cells. ROR is familiar and useful across many databases but remains sensitive to reporting behavior. Bayesian approaches reduce volatility in

low-count settings by shrinking estimates toward expected background levels [10-12,42,43]. Contemporary reporting guidance emphasizes explicit definitions of the reporting denominator, case selection, comparator, and data-processing rules because analytic choices materially affect reproducibility [10,40].

The deeper limitation is that drug-event pairs do not exist in isolation. Patients receive multiple medicines, may have several diseases, and change treatment over time. Confounding by indication, channeling, differential follow-up, changes in healthcare utilization, and stimulated reporting can all create associations that are not causal [5,9-12,40]. Thus conventional statistics remain valuable as screening tools, but their outputs are best incorporated into a broader evidence-generation workflow.

5. Machine Learning Paradigms for Pharmacovigilance

ML introduces a flexible pattern-recognition layer to PV. Supervised learning uses labelled examples to predict an outcome; unsupervised learning searches for latent structure in unlabelled observations; and reinforcement learning learns policies from feedback. Their PV roles differ, and the source framework is summarized in Table 4 [26,27,47-52].

Table 4. Comparison of Supervised Learning, Unsupervised Learning and Reinforcement Learning

Aspects	Supervised Learning	Unsupervised Learning	Reinforcement Learning
Definition	Learns from labelled data (input-output pairs)	Finds hidden patterns in unlabeled data	Learns by interacting with an environment using rewards/penalties.
Data Requirement	Requires labelled datasets (e.g., ADR confirmed vs not confirmed)	Uses unlabelled datasets (e.g., spontaneous reports without classification)	Does not require labelled datasets; needs feedback or reward signals.
Typical Tasks	ADR detection; signal classification	Signal detection; clustering	Adaptive drug-safety monitoring
Examples in PV	Classifying reports as serious/non-serious	Grouping similar AEs	Optimizing alert thresholds in safety systems
Handling New/Unknown ADRs	Limited; depends on training data	Strong; may uncover hidden or novel patterns	Moderate; depends on whether the reward captures new risks
Human Involvement	High (for labelling data)	Moderate (interpretation of clusters)	Moderate (design of reward functions)
Strengths	High accuracy when labelled data are available; often familiar to reviewers	Useful for discovery of unlabelled or novel patterns	Can adapt policies dynamically when the environment and reward are well specified
Limitations	Requires reliable labels and may miss unknown ADRs	Harder to interpret; clusters can be ambiguous	Complex to design; requires well-defined reward mechanisms
Real-world Use in PV systems	Automated case triage; regulatory reporting support	Signal detection in large safety databases such as FAERS	Potentially adaptive monitoring and alert optimization

5.1. Supervised Learning for Case Triage and Signal Classification

Supervised learning is particularly useful when historical reports, adjudicated records, or curated benchmarks provide a stable target label. Logistic regression remains an interpretable baseline, and its coefficients can be evaluated alongside regularization, calibration, and subgroup performance [48,49]. The number of research papers using logistic regression over time is shown in Figure 4. Tree ensembles can model nonlinear interactions and heterogeneous feature effects. Random forests use bootstrap aggregation and randomized

feature selection, while gradient-boosting methods iteratively reduce residual error and can integrate structured PV features with NLP-derived representations [50-52].

Support vector machines remain useful in high-dimensional settings, particularly when the sample size is moderate and the feature representation is carefully controlled [51]. However, a high area under the receiver-operating characteristic curve is not enough for operational deployment. The target outcome, decision threshold, action associated with an alert, and consequences of false positives and false negatives must be defined before the model is evaluated. Reporting frameworks such as Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) + AI support more transparent communication of these decisions [48].

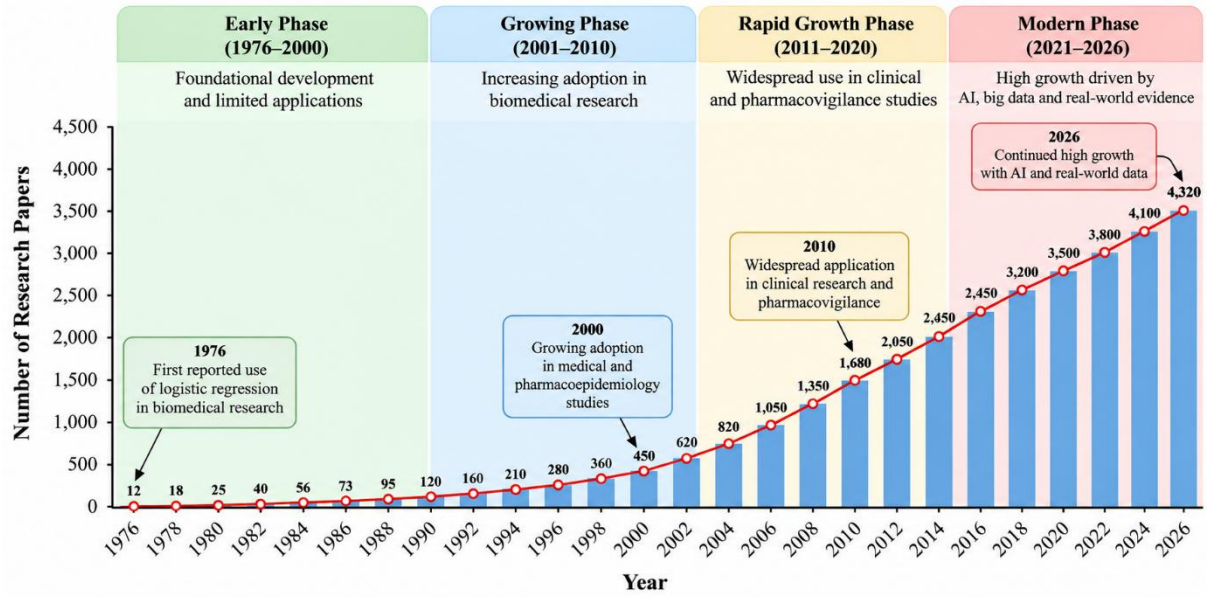

Figure 4. Research papers using logistic regression from 1971 to 2026.

5.2. Neural Networks and Deep Learning

Artificial neural networks approximate complex nonlinear mappings through interconnected layers of weighted units (Figure 5). Deep learning extends this principle by learning hierarchical representations from increasingly complex data structures [27]. In PV, deep models can integrate large feature spaces spanning structured variables, text embeddings, sequences, and relational information. Their flexibility comes with increased requirements for training data, computational resources, calibration, model monitoring, and interpretability. [53] The relationships among ML concepts and classes are summarized in Figure 6.

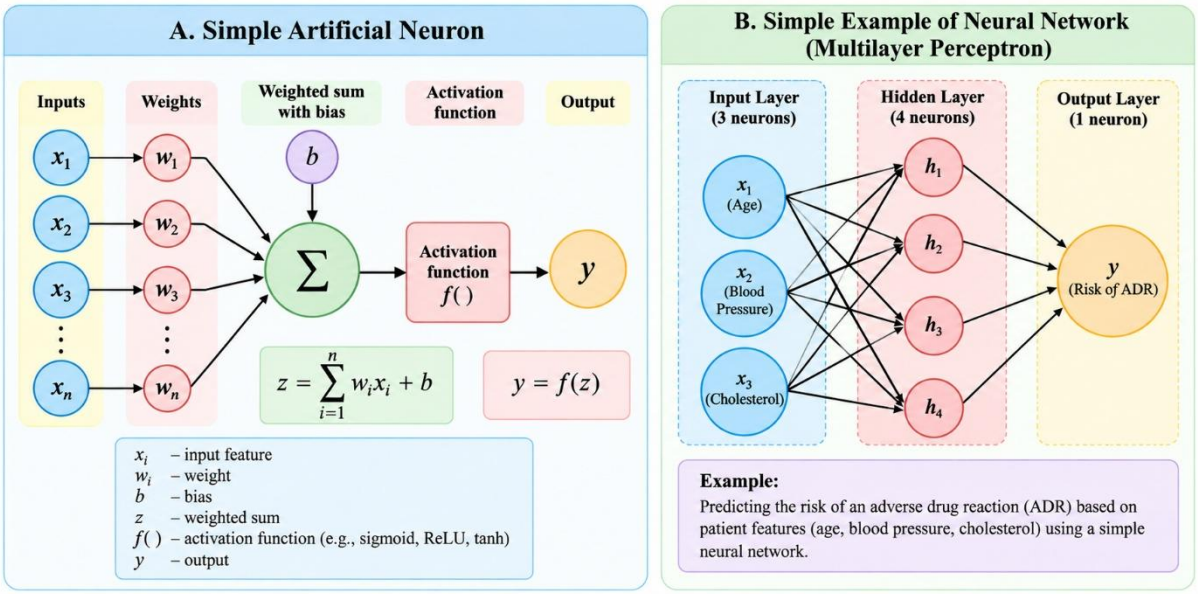

Figure 5. A) Simple artificial neuron and B) Simple example of neural network.

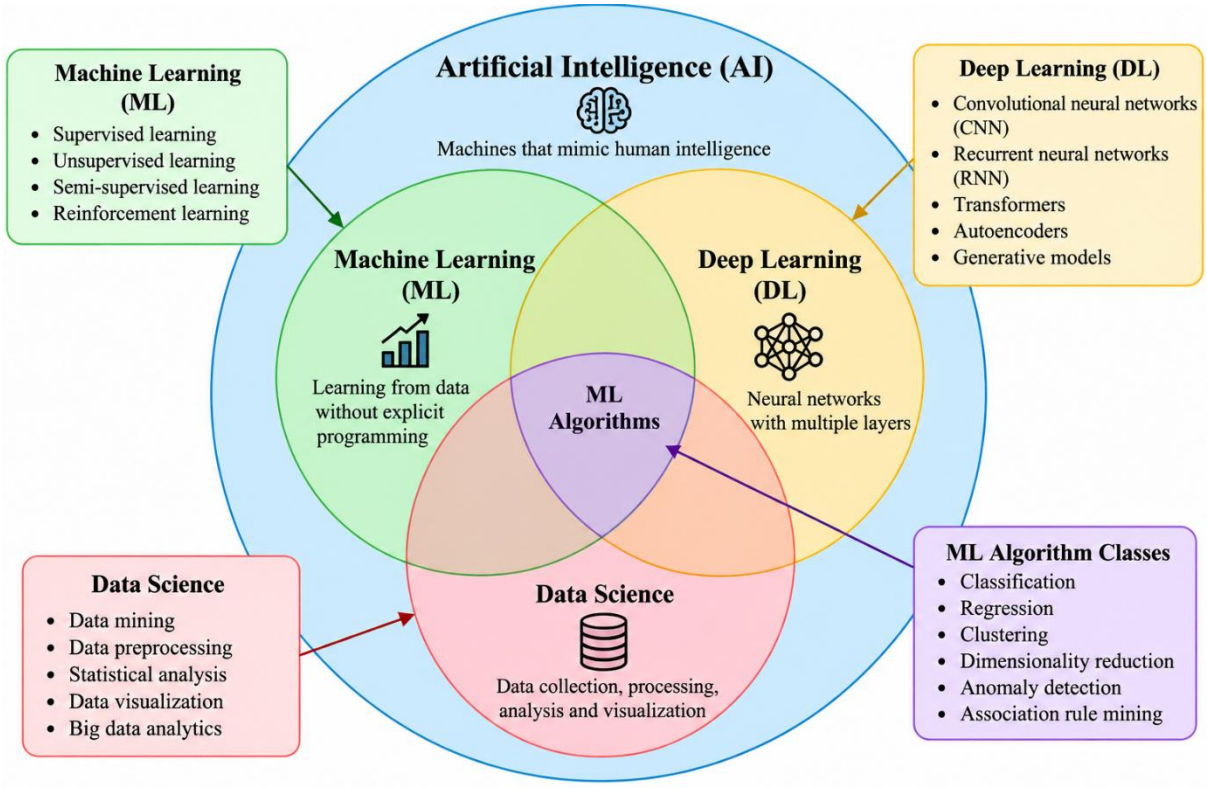

Figure 6. Venn diagram of ML concepts and classes.

5.3. Graph Neural Networks and Relational Pharmacovigilance

Many drug-safety questions are naturally relational. Drugs interact with biological targets, patients receive multiple medicines, diseases influence prescribing, and adverse outcomes arise within clinical and molecular contexts. Graph-based ML represents such relationships explicitly, with nodes for drugs, events, diagnoses, targets, or patient states and edges for observed or biologically plausible relationships. Message-passing and attention-based graph methods can then generate embeddings that capture neighborhood context [22,52].

Graph representations are also relevant to drug-drug interaction and molecular safety hypotheses because the same medicine can occupy multiple clinical and biological networks. They can therefore prioritize candidate relationships for investigation, but the resulting graph

score remains predictive evidence and requires external validation, clinical review, and causal assessment before regulatory action [22,52].

6. Natural Language Processing and Advanced Feature Engineering

6.1. Clinical Text, Entity Extraction, and MedDRA Mapping

A major fraction of actionable safety information is stored in free text. NLP can identify drug names, doses, symptoms, diagnoses, temporal expressions, treatment status, outcomes, and relationships among these entities [28,54-57]. Mapping extracted concepts to the Medical Dictionary for Regulatory Activities (MedDRA) provides a shared safety vocabulary across local documentation and international reporting, while negation and assertion detection are required to distinguish a present ADR from a denied symptom or hypothetical instruction [54,55].

The transition from rules to contextual language models is important because the same lexical token can mean different things depending on context. Clinical phrases such as "history of rash," "no rash," or "rash after starting the medicine" cannot be reliably interpreted by keyword frequency alone. NLP pipelines therefore benefit from a combination of terminology resources, contextual models, relation extraction, temporal reasoning, and post-processing safeguards [28,45,56-60].

6.2. Transformer Models and Biomedical Language Representations

Transformer architectures use self-attention to represent relationships among tokens across a sequence. Bidirectional Encoder Representations from Transformers (BERT) established a broadly transferable bidirectional language representation framework; BioBERT adapted the architecture to biomedical literature, and ClinicalBERT-oriented models extend contextual representations toward clinical notes [58-60]. These models can improve ADR extraction when training data and evaluation tasks are appropriately defined, but performance remains sensitive to annotation quality, domain shift, class imbalance, and clinical context [28,45,57,59,60].

6.3. Patient-Generated Text and Probabilistic ADR Estimation

Patient-generated text can be converted into safety features by first extracting ADR-related expressions and then integrating them with structured evidence. Similarity-based or embedding-based approaches can prioritize semantically related events, after which supervised or probabilistic models may estimate downstream risk [4,14,19,28]. Such systems should be treated as hypothesis-generation tools because patient-generated content usually lacks validated exposure denominators and may not provide laboratory confirmation or adjudicated causality. A schematic workflow for predicting future ADR probabilities is shown in Figure 7.

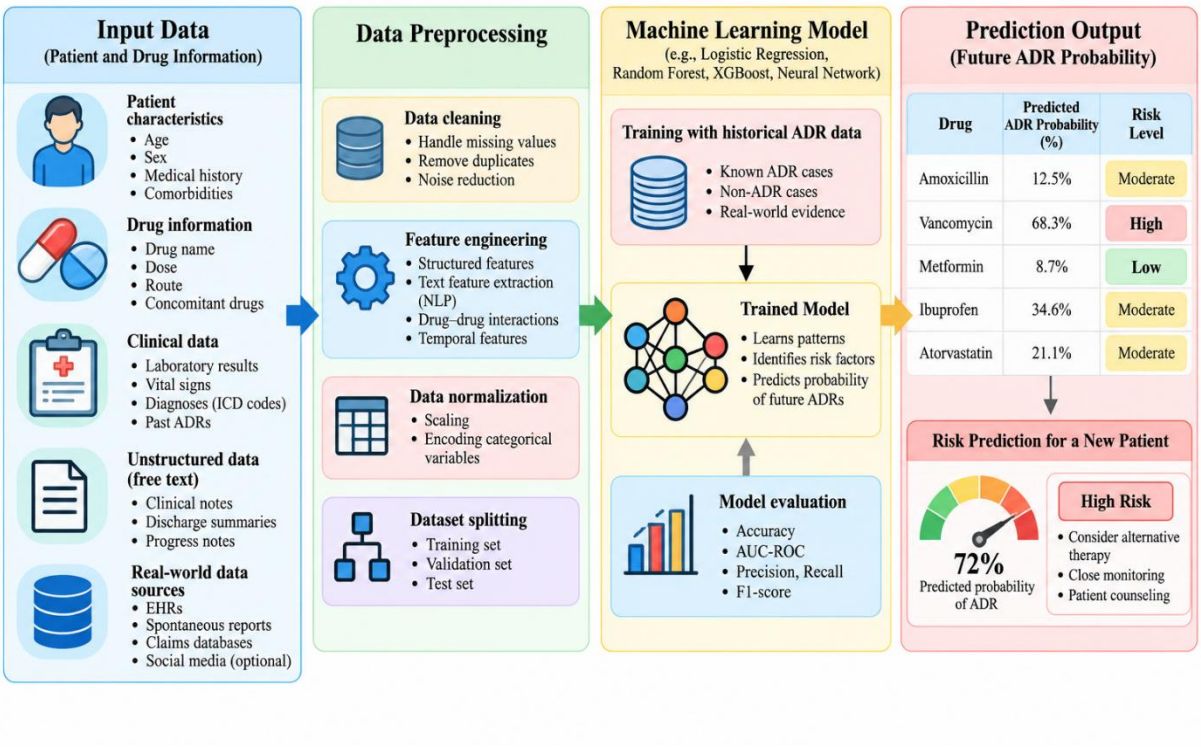

Figure 7. Predicting Future ADRs Probabilities.

6.4. Anomaly Detection and Rare-Event Triage

Rare but serious ADRs may remain below conventional disproportionality thresholds for some time. Anomaly-detection algorithms, including isolation forests, can identify observations that deviate strongly from learned background patterns and therefore deserve early review [61]. The operational rule is simple: an anomaly is a reason to investigate, not proof that a drug caused harm. Outliers may represent true rare reactions, coding artifacts, unusual disease states, or unmeasured co-exposures [42,43,61].

7. Operational Pharmacovigilance, Signal Management, and Regulation

7.1. Closed-Loop Signal Management

Signal management is an iterative process rather than a one-time prediction task. A candidate signal is detected, validated, clinically assessed, investigated in relation to alternative explanations, and then followed after any risk-minimization or regulatory action. AI can support several steps by triaging cases, extracting structured evidence from narratives, ranking drug-event pairs, searching literature, and monitoring changes in signal strength over time [9-11,16,30,62]. The overall signal management process is shown in Figure 8.

A robust architecture should keep signal generation separate from signal confirmation. Each algorithmic output should be accompanied by an evidence packet describing the source population, data window, phenotype definition, model version, salient features, comparator, uncertainty, and human-review status. This design aligns AI with established Good Pharmacovigilance Practices (GVP) principles and makes post-deployment monitoring auditable [10,31,42,43,63].

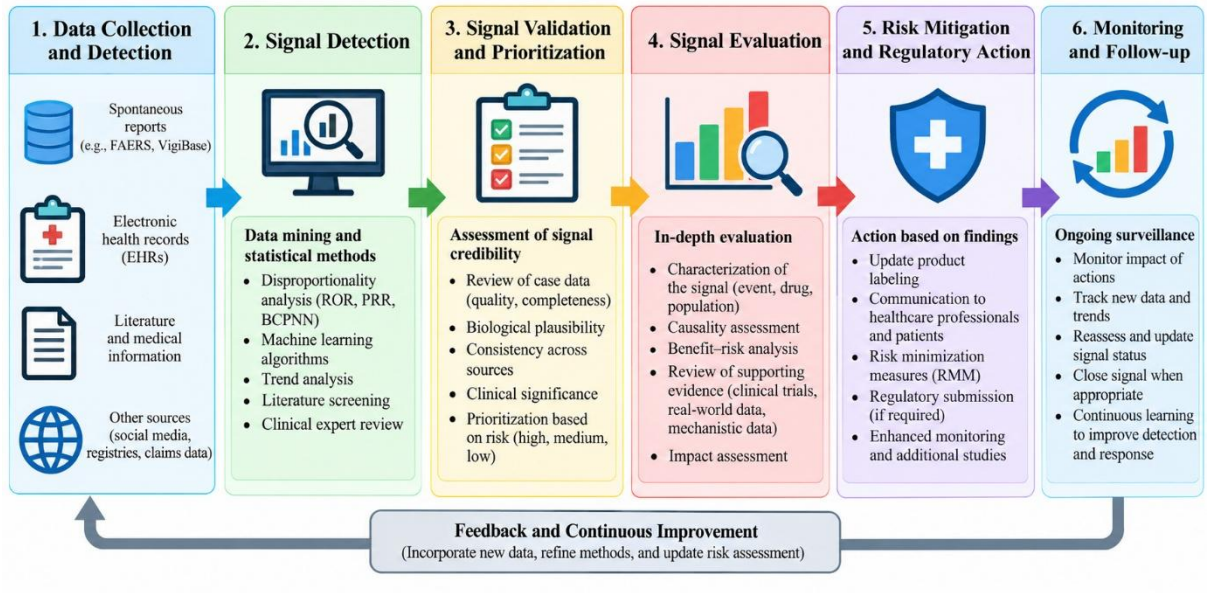

Figure 8. The signal management process's workflow.

7.2. Uppsala Monitoring Centre and VigiBase

VigiBase, maintained by the UMC, provides a global collection of individual case safety reports. Its international breadth supports cross-country signal generation, but heterogeneity in reporting systems and case quality necessitates layered screening and expert review [2,41,43]. Tools for case completeness, such as vigiGrade-related approaches, can help prioritize reports with stronger documentation, but completeness scoring should never obscure access to the original case information [43]. The stepwise UMC screening workflow is summarized in Table 5.

Table 5. ML in UMC.

Screening Type	Description
First-pass statistical screening	A list of drug-event pairings (with or without a subgroup) or drug-drug events is processed by data-driven algorithms and filters while study-specific inclusion and exclusion criteria are applied.
Initial manual assessment	UMC evaluators review the prioritized drug-event list for indicators requiring further research. The review checks whether product characteristics appropriately describe the event and examine individual reports to remove cases that are insufficiently documented or have a more likely alternative explanation.
In-depth manual assessment	UMC staff or external skilled scientists and physicians evaluate the causation of individual case reports and review relevant literature. The strength of the case series is then considered when deciding whether a signal warrants communication.

7.3. FDA, Sentinel, and Software Classification

The FDA combines spontaneous reporting with active surveillance systems such as Sentinel and uses RWD and real-world evidence (RWE) within regulatory science. The distinction between passive reporting and distributed active surveillance is important because active systems can define exposed populations and comparators for targeted analyses [3,64]. The FDA also describes growing use of AI across the drug life cycle and has published a risk-based draft framework for the credibility of AI-generated information used in regulatory decision-making [32,33,65]. The FDA definitions of Software in a Medical Device (SiMD), Software as a Medical Device (SaMD), and Clinical Decision Support (CDS) are summarized in Table 6.

Table 6. FDA Definitions of SiMD, SaMD, and CDS.

FDA Term	Definition	Example

SiMD	Software that works primarily with a device and/or is unable to function independently.	Bluetooth or Wi-Fi connectivity software in an ultrasound; image-enhancer software in an X-ray machine.
SaMD	Software designed for one or more medical uses that fulfills those functions independently of a physical medical device.	Electrocardiogram analysis software; X-ray image-processing software.
CDS	A program that, in the stated use case, presents or analyses medical information without independently diagnosing or selecting therapy.	Automated reminders, notifications, or EHR-based clinical information tools.

The practical regulatory classification of software depends on intended use, autonomy, context, risk, and jurisdiction. An AI model used to prioritize a safety-review queue may present a different regulatory profile from software that directly recommends a clinical intervention. Validation should therefore include analytical performance, intended-use validity where applicable, human factors, cybersecurity, change control, and post-deployment performance monitoring [32,33,65,66].

7.4. A Multilayered AI Pharmacovigilance Ecosystem

A contemporary AI-enabled PV ecosystem can be represented as four interacting layers. The data layer contains SRS data, EHRs, claims, registries, literature, and patient-generated information. The analytical layer contains disproportionality methods, supervised ML, unsupervised learning, anomaly detection, NLP, transformers, and graph learning. The governance layer addresses privacy, provenance, model monitoring, standards, auditability, and change control. The action layer links evidence to clinical review, risk minimization, labeling, and regulatory communication [3,16,31,63,67,68]. A summary of PV difficulties, data sources, and pertinent AI solutions is provided in Figure 9.

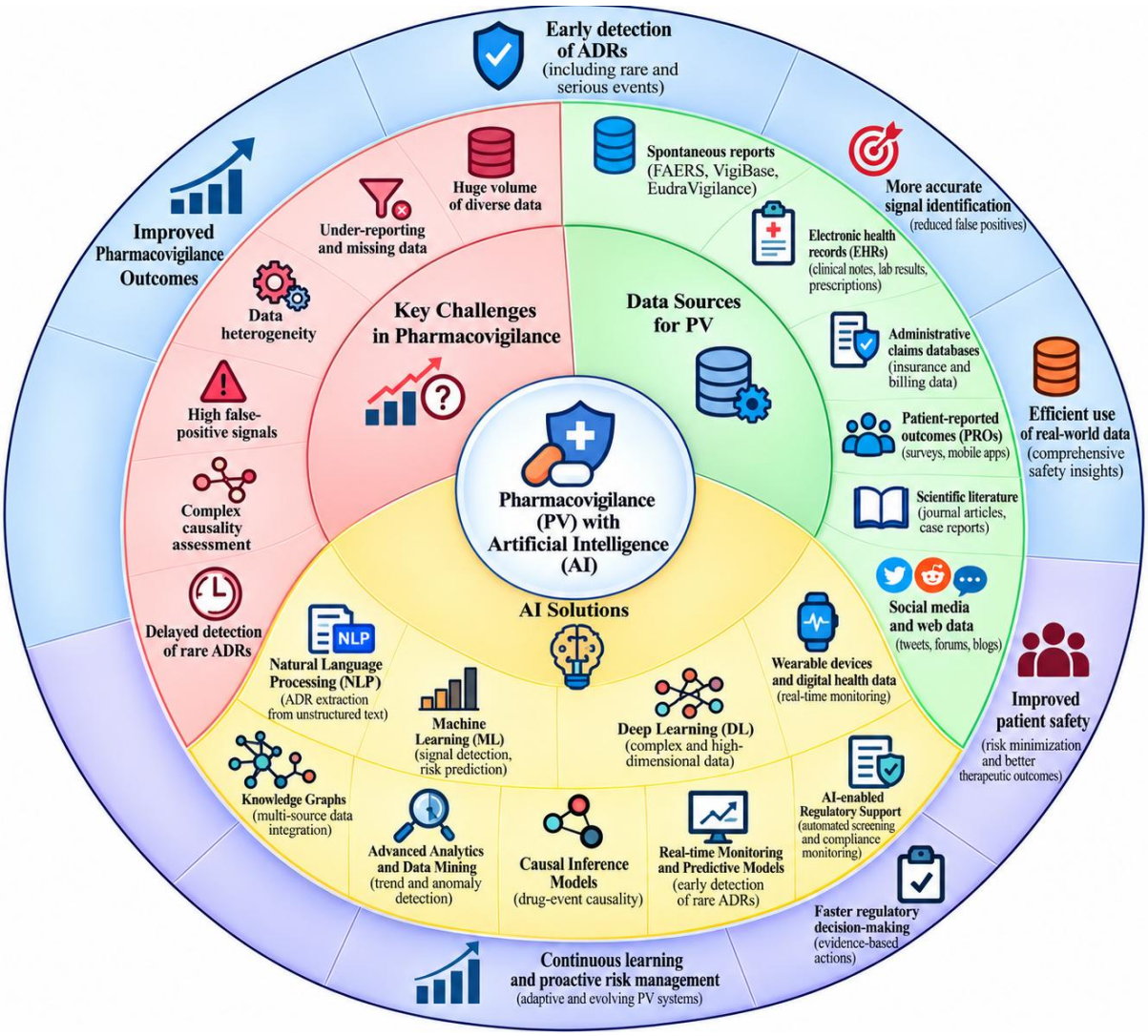

Figure 9. A thorough summary of PV difficulties, data sources, and pertinent AI solutions to solve the issues. PV with AI.

7.5. Active Surveillance During Global Health Crises

Near-real-time surveillance became particularly visible during the coronavirus disease 2019 (COVID-19) pandemic. Rapid Cycle Analysis (RCA) and related sequential approaches can update safety analyses at regular intervals using active healthcare data, allowing observed event rates to be compared with historical or concurrent comparators [64,68]. A 2025 review of published RCA studies found that the approach is increasingly used for post-approval vaccine safety and can complement passive reporting, although most available studies have been conducted in the United States [64].

8. Critical Challenges: Data Quality, Bias, Causality, Privacy, and Trust

8.1. Data Quality and Heterogeneity

AI systems inherit the data-generating processes from which they learn. Missing values, duplicated records, inconsistent coding, changing documentation practices, non-random observation, and variable data latency can all influence predictions [9,18,40,42,43]. The challenge is particularly important for RWD because the same clinical concept can be represented differently across institutions. Common data models, standardized terminology, documented preprocessing, and provenance are therefore core scientific requirements rather than purely technical conveniences [18,67-69].

Evaluation should begin with the safety question and target definition. Investigators should specify inclusion criteria, index date, observation window, exposure definition, outcome phenotype, reference standard, and downstream clinical action. These choices determine what the model is actually learning and whether performance is meaningful outside the development dataset [18,48,49].

8.2. Causality and the Black-Box Problem

Predictive association is not equivalent to causal attribution. A model may learn that a drug is associated with an outcome because the drug is preferentially prescribed to patients already at high baseline risk. Confounding by indication, treatment channeling, differential follow-up, and surveillance intensity can all produce associations that are predictive yet not causal [5,8,10,40]. Causal inference methods such as propensity-based adjustment, negative controls, target-trial thinking, and explicit counterfactual definitions should therefore operate alongside prediction rather than beneath it.

Explainability has a related but distinct role. Interpretable models can expose their decision structure directly, while post-hoc approaches such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) quantify feature contributions or construct local approximations [70-73]. These methods can assist reviewers, but a plausible explanation does not prove that the prediction is correct. The central governance objective is an auditable chain linking raw evidence, model output, uncertainty, reviewer assessment, and final action [70,73,74].

8.3. Demographic Fairness and Representation

AI-enabled PV can reproduce or amplify the demographic composition of the source data. Differences in care access, documentation, treatment patterns, genetic variation, and participation in digital platforms can affect observed ADR signals [66,75-77]. Evaluation should therefore include subgroup-specific discrimination, calibration, missingness, error rates, and false-negative burden rather than relying solely on a single global metric [48,49,66].

Precision PV is increasingly linked to pharmacogenomics and individualized risk estimation. Genomic data can provide mechanistic information about susceptible subgroups, while RWD supplies the longitudinal context in which those risks are expressed [78,79]. Importantly, genetic risk should not be reduced to broad demographic labels; the value of personalization depends on biologically relevant markers, calibrated prediction, clinical actionability, and appropriate consent and governance.

8.4. Privacy, Federated Learning, and Data Governance

Longitudinal patient-level data raise privacy and cybersecurity concerns. De-identification, access controls, secure computing environments, federated learning, and synthetic data can reduce exposure of raw records but do not eliminate all disclosure risks [66,77,80-83]. Federated learning keeps local training data at participating institutions and exchanges model updates, enabling some forms of multi-site analysis without centralizing the raw dataset [82]. Its performance can nevertheless be affected by non-identical local data distributions, and update information itself can leak information if protections are inadequate.

Blockchain or distributed-ledger approaches can provide tamper-evident records and provenance for multi-party data workflows, but they should be viewed as infrastructure rather than a substitute for clinical validation, statistical assessment, or governance [80,81,84]. Privacy-enhancing technologies should be selected according to the actual threat model, the

required analytic utility, and the regulatory context. Synthetic data similarly require joint assessment of utility and privacy rather than acceptance based on realism alone [83].

9. Future Perspectives and Technological Horizons

9.1. Explainable AI and Mechanistic Transparency

Future PV systems will need explanations that are useful for a concrete review task. Analysts may need to know why a case was prioritized; clinical reviewers may need the temporal evidence supporting a suspected ADR; and regulators may require traceability of how an algorithmic result entered the evidentiary chain. SHAP, LIME, counterfactual explanations, feature attribution, and interpretable model classes offer complementary strategies [70-74].

9.2. Blockchain and Auditable Safety Records

Data provenance becomes increasingly important as safety evidence moves across organizations. Distributed ledgers can support tamper-evident event histories, while conventional audit trails, secure logs, and controlled versioning can provide similar functions when a blockchain is unnecessary [80,81,84]. The design principle is purpose fit: immutability is useful when multi-party traceability is genuinely required, but it does not resolve semantic interoperability, missingness, or causal uncertainty.

9.3. Real-Time Global Surveillance and Common Data Models

Scalable surveillance requires standards that permit analytic methods to move across institutions without changing the meaning of variables. Observational Medical Outcomes Partnership (OMOP)-based infrastructures and related common-data-model strategies illustrate how data can be normalized while preserving local governance [67]. The Data Analysis and Real World Interrogation Network (DARWIN EU) of the European Medicines Agency (EMA) demonstrates a complementary RWE infrastructure in which data from multiple healthcare databases can support safety, effectiveness, utilization, and regulatory studies across the medicine life cycle [68].

9.4. Personalized Pharmacovigilance and Pharmacogenomics

Population-level signal detection may gradually be complemented by individualized risk estimation. Models can combine age, organ function, prior reactions, co-medications, laboratory trajectories, treatment history, and genomic information to identify patients who may warrant closer monitoring [24,78,79]. Personalized PV should remain probabilistic and calibrated; the clinical question is not whether a model can predict an event, but whether its output is sufficiently reliable and actionable to change monitoring or treatment decisions safely.

9.5. Large Language Models and Human-in-the-Loop Pharmacovigilance

LLMs introduce a new capability for summarizing narratives, extracting timelines, mapping free text to concepts, organizing literature, and generating preliminary evidence tables. Their principal risk is that fluent text can obscure unsupported assertions, missing evidence, or incorrect causal interpretations [29,30]. For safety-critical use, LLMs should be constrained by retrieval from authoritative records, explicit source traceability, terminology controls, uncertainty handling, and mandatory human review. The current regulatory trajectory similarly emphasizes intended use, model credibility, and context-specific validation [32,33].

10. Implementation Framework for AI/ML-Enabled Pharmacovigilance

A practical implementation pathway can be organized into seven linked stages. First, define the safety question, intended use, target population, and clinical action. Second, characterize the RWD sources, coding systems, missingness, data latency, and exposure definitions. Third, build reproducible phenotypes and a reference standard using expert review. Fourth, develop models with train/validation/test separation and transparent preprocessing. Fifth, perform external validation across sites, populations, time periods, and clinically important subgroups. Sixth, connect alerts to a documented signal-management workflow with predefined escalation criteria. Seventh, continue post-deployment monitoring for calibration drift, data-shift, fairness, cybersecurity, and operational performance [3,10,31-33,45,48,49,63,64,66,82].

The deployment target should be an evidence pipeline rather than a stand-alone prediction service. A model should generate a traceable evidence packet containing data-source identifiers, cohort definitions, model version, features or extracted concepts, uncertainty, comparator analyses, and the human assessment status. This approach supports auditability and makes it possible to distinguish an algorithmic alert from a validated PV signal [9-11,30,42,43,74].

Operational success should be measured against meaningful PV outcomes: time from event appearance to review, reviewer workload, false-negative burden, calibration, external reproducibility, incremental signal yield, and the extent to which model-supported evidence changes a justified risk-management decision. Model accuracy alone is insufficient when the cost of missed rare safety events differs dramatically from the cost of additional review.

11. Discussion

The 2016-2026 period shows a clear movement from isolated analytics toward integrated computational PV. Early applications emphasized automated extraction and classical ML; more recent work increasingly combines RWD, deep learning, language models, graph representations, and explainable or causal AI. The literature also shows that the main barrier to operationalization is often not model novelty but evidence quality, external validation, workflow integration, privacy, and governance.

A key implication is that AI should be layered onto existing PV practice rather than positioned as a replacement for disproportionality analysis or expert review. Conventional methods provide transparent screening; ML can prioritize complex patterns; NLP can recover information hidden in narrative records; graph approaches can add relational context; causal methods can address attribution; and explainable AI (XAI) and governance mechanisms can support traceability. Another implication is that source diversity is useful only when the evidence streams are analytically compatible. A patient-generated symptom, an EHR diagnosis, a laboratory value, and a spontaneous report may describe the same clinical episode at different levels of detail. Data linkage therefore requires patient-episode identity, temporal alignment, standardized terminology, and explicit uncertainty. The future value of RWD will depend as much on interoperability and governance as on increasing data volume.

Finally, the regulatory environment is evolving toward context-specific credibility rather than one universal definition of an acceptable AI model. FDA's 2025 draft guidance proposes a risk-based framework for establishing model credibility for a particular context of use, while its 2026 AI-for-drug-development program and EMA initiatives signal continued integration of AI and RWD into life-cycle evidence generation. This trajectory reinforces the

need for transparent validation, model documentation, human oversight, and continuous monitoring.

12. Conclusions

AI and ML can expand PV by increasing the scale and speed at which structured and unstructured RWD are screened for possible ADRs. The strongest use cases are not limited to automated classification. They include case triage, narrative extraction, signal prioritization, anomaly detection, longitudinal risk prediction, network-based reasoning, and multimodal evidence synthesis. The evidence also supports a clear boundary: an algorithmic association is not itself a confirmed ADR or proof of causality. Reliable implementation requires standardized phenotypes, high-quality data, calibration, external validation, subgroup assessment, privacy protections, auditable workflows, and appropriately trained human reviewers. Conventional disproportionality analysis and expert review remain foundational, while AI can serve as a scalable analytical layer around them. Across the 2016-2026 period, the field has therefore moved toward a hybrid model of PV in which statistical screening, ML, NLP, graph learning, causal inference, RWD infrastructures, and regulatory governance operate together. The practical objective is not automation for its own sake, but a shorter and more reliable path from an emerging safety pattern to evidence that can be clinically interpreted, externally verified, and translated into proportionate risk-management action.

Author Contributions

The authors confirm their contribution to the paper as follows: study conception and design: YS; AK draft manuscript: RK, HG, TKY. All authors reviewed the results and approved the final version of the manuscript.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable

Data Availability Statement

No new data were created or analyzed in this study. Data sharing is not applicable.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Abbreviation	Full term

ADR	Adverse Drug Reaction
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
PV	Pharmacovigilance
NLP	Natural Language Processing
EHR	Electronic Health Record
RWD	Real-World Data
RWE	Real-World Evidence
FAERS	FDA Adverse Event Reporting System
FDA	U.S. Food and Drug Administration
EMA	European Medicines Agency
WHO	World Health Organization
MedDRA	Medical Dictionary for Regulatory Activities
UMC	Uppsala Monitoring Centre
VigiBase	WHO Global Individual Case Safety Report Database
ICSR	Individual Case Safety Report
ADE	Adverse Drug Event
AE	Adverse Event
NER	Named-Entity Recognition
PROs	Patient-Reported Outcomes
ROR	Reporting Odds Ratio
PRR	Proportional Reporting Ratio
BCPNN	Bayesian Confidence Propagation Neural Network
XAI	Explainable Artificial Intelligence
LLM	Large Language Model
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
SVM	Support Vector Machine
RF	Random Forest
XGBoost	Extreme Gradient Boosting
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
RMM	Risk-Minimisation Measure
CDS	Clinical Decision Support
SaMD	Software as a Medical Device
SiMD	Software in a Medical Device
EDA	Exploratory Data Analysis
RCT	Randomized Controlled Trial
EBM	Evidence-Based Medicine
GDPR	General Data Protection Regulation
PHI	Protected Health Information
PII	Personally Identifiable Information
UMLS	Unified Medical Language System
SHAP	SHapley Additive exPlanations

LIME	Local Interpretable Model-Agnostic Explanations
IC	Information Component
SSRIs	Selective Serotonin Reuptake Inhibitors
SRS	Spontaneous Reporting System
TRIPOD	Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis
BERT	Bidirectional Encoder Representations from Transformers
GVP	Good Pharmacovigilance Practices
MHRA	Medicines and Healthcare products Regulatory Agency
RCA	Rapid Cycle Analysis
COVID-19	Coronavirus Disease 2019
OMOP	Observational Medical Outcomes Partnership
DARWIN EU	Data Analysis and Real World Interrogation Network

References

1. World Health Organization. Pharmacovigilance: Ensuring the Safe Use of Medicines; World Health Organization: Geneva, Switzerland, 2004.
2. World Health Organization. WHO Programme for International Drug Monitoring; World Health Organization: Geneva, Switzerland, 2026.
3. U.S. Food and Drug Administration. Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products: Guidance for Industry; U.S. Food and Drug Administration: Silver Spring, MD, USA, 2023.
4. Nikfarjam, A.; Sarker, A.; O'Connor, K.; Ginn, R.; Gonzalez, G. Pharmacovigilance from social media: Mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. *J. Am. Med. Inform. Assoc.* **2015**, *22*, 671–681. https://doi.org/10.1093/jamia/ocu041.
5. Coste, A.; Wong, A.; Bokern, M.; Bate, A.; Douglas, I.J. Methods for drug safety signal detection using routinely collected observational electronic health care data: A systematic review. *Pharmacoepidemiol. Drug Saf.* **2023**, *32*, 28–43. https://doi.org/10.1002/pds.5548.
6. Pane, J.; Verhamme, K.M.C.; Villegas, D.; Gamez, L.; Rebollo, I.; Sturkenboom, M.C.J.M. *Challenges associated with the safety signal detection process for medical devices. Med. Devices Evid. Res.* **2021**, *14*, 43–57. https://doi.org/10.2147/MDER.S278868.
7. Choudhury, R. *Integrating EHRs for proactive pharmacovigilance. J. Pharmacovigil.* **2025**, *13*, 1000502. https://doi.org/10.35248/2329-6887.25.13.502.
8. Kim, J.V.; Davis, S.E.; Matheny, M.E.; Smith, J.C. Integrating electronic health records with other data sources for postmarket drug safety signal identification: A review. *Front. Drug Saf. Regul.* **2024**, *4*, 1428831. https://doi.org/10.3389/fdsfr.2024.1428831.
9. Mugada, V.; Suryadevara, V.; Cheekurumilli, M.; Yarguntla, S.R. Signal detection in pharmacovigilance: Methods, tools, and workflows from case identification to adverse drug reaction database entry. *Przegl. Epidemiol.* **2025**, *79*, 404–414. https://doi.org/10.32394/pe/211665.
10. Fusaroli, M.; Salvo, F.; Begaud, B.; AlShammari, T.M.; Bate, A.; Battini, V.; et al. The REporting of A Disproportionality Analysis for DrUg Safety Signal Detection Using Individual Case Safety Reports in PharmacoVigilance (READUS-PV): Explanation and elaboration. *Drug Saf.* **2024**, *47*, 585–599. https://doi.org/10.1007/s40264-024-01423-7.

11. Algarvio, R.C.; Conceição, J.; Pereira, P.; Inês, R. Artificial intelligence in pharmacovigilance: A narrative review and practical experience with an expert-defined Bayesian network tool. *Int. J. Clin. Pharm.* **2025**, *47*, 932–944. https://doi.org/10.1007/s11096-025-01975-3.
12. DuMouchel, W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *Am. Stat.* **1999**, *53*, 177–190. https://doi.org/10.2307/2686103.
13. Rudnisky, E.; Paudel, K. Pharmacovigilance in the era of artificial intelligence: Advancements, challenges, and considerations. *Cureus* **2025**, *17*, e86972. https://doi.org/10.7759/cureus.86972.
14. Sarker, A.; Ginn, R.; Nikfarjam, A.; O'Connor, K.; Smith, K.; Jayaraman, S.; et al. Utilizing social media data for pharmacovigilance: A review. *J. Biomed. Inform.* **2015**, *54*, 202–212. https://doi.org/10.1016/j.jbi.2015.02.004.
15. Powell, G.; Kara, V.; Painter, J.L.; Schifano, L.; Merico, E.; Bate, A. Engaging patients via online healthcare fora: Three pharmacovigilance use cases. *Front. Pharmacol.* **2022**, *13*, 901355. https://doi.org/10.3389/fphar.2022.901355.
16. Lavertu, A.; Vora, B.; Giacomini, K.M.; Altman, R.; Rensi, S. A new era in pharmacovigilance: Toward world data and digital monitoring. *Clin. Pharmacol. Ther.* **2021**, *109*, 1197–1202. https://doi.org/10.1002/cpt.2172.
17. Dimitsaki, S.; Natsiavas, P.; Jaulent, M.C. Applying AI to structured real-world data for pharmacovigilance purposes: Scoping review. *J. Med. Internet Res.* **2024**, *26*, e57824. https://doi.org/10.2196/57824.
18. Davis, S.E.; Zabotka, L.; Desai, R.J.; Wang, S.V.; Maro, J.C.; Coughlin, K.; et al. Use of electronic health record data for drug safety signal identification: A scoping review. *Drug Saf.* **2023**, *46*, 725–742. https://doi.org/10.1007/s40264-023-01325-0.
19. Golder, S.; O'Connor, K.; Wang, Y.; Klein, A.; Gonzalez Hernandez, G. The value of social media analysis for adverse events detection and pharmacovigilance: Scoping review. *JMIR Public Health Surveill.* **2024**, *10*, e59167. https://doi.org/10.2196/59167.
20. Caster, O.; Dietrich, J.; Kürzinger, M.L.; Lerch, M.; Maskell, S.; Norén, G.N.; Tcherny-Lessenot, S.; Vroman, B.; Wisniewski, A.; van Stekelenborg, J. Assessment of the utility of social media for broad-ranging statistical signal detection in pharmacovigilance: Results from the WEB-RADR project. *Drug Saf.* **2018**, *41*, 1355–1369. https://doi.org/10.1007/s40264-018-0699-2.
21. Verma, A.; Awasthi, A. Revolutionizing drug discovery: The role of artificial intelligence and machine learning. *Curr. Pharm. Des.* **2024**, *30*, 807–810. https://doi.org/10.2174/0113816128298691240222054120.
22. Barbieri, M.A.; Abate, A.; Balogh, O.M.; Pétervári, M.; Ferdinandy, P.; Ágg, B.; et al. Network analysis and machine learning for signal detection and prioritization using electronic healthcare records and administrative databases: A proof of concept in drug-induced acute myocardial infarction. *Drug Saf.* **2025**, *48*, 513–526. https://doi.org/10.1007/s40264-025-01515-y.
23. Ahmed, Z.; Mohamed, K.; Zeeshan, S.; Dong, X.Q. *Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. Database* **2020**, *2020*, baaa010. https://doi.org/10.1093/database/baaa010.
24. Li, L.; Jing, S.; Zhao, X.; Zhu, X.; Liang, C.; Shi, L.; et al. *Computational pharmacovigilance of lifitegrast in dry eye disease using machine learning and network toxicology. NPJ Digit. Med.* **2026**, *9*, 30. https://doi.org/10.1038/s41746-025-02173-1.

25. Lee, J.E.; Kim, J.H.; Bae, J.H.; Song, I.; Shin, J.Y. *Detecting early safety signals of infliximab using machine learning algorithms in the Korea adverse event reporting system. Sci. Rep.* **2022**, *12*, 14869. https://doi.org/10.1038/s41598-022-18522-z.
26. Pilipiec, P.; Liwicki, M.; Bota, A. Using machine learning for pharmacovigilance: A systematic review. *Pharmaceutics* **2022**, *14*, 266. https://doi.org/10.3390/pharmaceutics14020266.
27. LeCun, Y.; Bengio, Y.; Hinton, G. *Deep learning. Nature* **2015**, *521*, 436–444. https://doi.org/10.1038/nature14539.
28. Golder, S.; Xu, D.; O'Connor, K.; Wang, Y.; Batra, M.; Gonzalez Hernandez, G. Leveraging natural language processing and machine learning methods for adverse drug event detection in electronic health/medical records: A scoping review. *Drug Saf.* **2025**, *48*, 321–337. https://doi.org/10.1007/s40264-024-01505-6.
29. Salvo, F.; Crupi, L.; Cholle, C. Language models in pharmacovigilance: Applications, promises and limits. *Therapies* **2026**, *81*, 140–146. https://doi.org/10.1016/j.therap.2025.12.003.
30. Barbieri, M.A.; Battini, V.; Carnovale, C.; Cocco, M.; Papoutsi, D.G.; Heckmann, N.S.; et al. Artificial intelligence in pharmacovigilance signal management: A review of tools, implementations, research, and regulatory landscape. *Expert Opin. Drug Saf.* **2026**, *25*, 207–222. https://doi.org/10.1080/14740338.2025.2545926.
31. European Medicines Agency. Good Pharmacovigilance Practices (GVP): Module IX – Signal Management; European Medicines Agency: Amsterdam, The Netherlands, 2017.
32. U.S. Food and Drug Administration. Considerations for the Use of Artificial Intelligence To Support Regulatory Decision-Making for Drug and Biological Products: Draft Guidance for Industry and Other Interested Parties; U.S. Food and Drug Administration: Silver Spring, MD, USA, 2025.
33. U.S. Food and Drug Administration. Artificial Intelligence for Drug Development; U.S. Food and Drug Administration: Silver Spring, MD, USA, 2026.
34. Zazzara, M.B.; Palmer, K.; Vetrano, D.L.; Carfì, A.; Graziano, O. Adverse drug reactions in older adults: A narrative review of the literature. *Eur. Geriatr. Med.* **2021**, *12*, 463–473. https://doi.org/10.1007/s41999-021-00481-9.
35. Srisuriyachanchai, W.; Cox, A.R.; Jarernsiripornkul, N. Patients' identification, management and prevention of adverse drug reactions: A cross-sectional survey of patients with severe adverse drug reactions. *J. Clin. Med.* **2024**, *13*, 4165. https://doi.org/10.3390/jcm13144165.
36. Insani, W.N.; Whittlesea, C.; Stockley, I.H.; et al. Prevalence of adverse drug reactions in the primary care setting: A systematic review and meta-analysis. *PLoS ONE* **2021**, *16*, e0252161. https://doi.org/10.1371/journal.pone.0252161.
37. Sultana, J.; Cutroneo, P.; Trifirò, G. Clinical and economic burden of adverse drug reactions. J. Pharmacol. Pharmacother. 2013, 4 (Suppl. 1), S73–S77. https://doi.org/10.4103/0976-500X.120957.
38. Lazarou, J.; Pomeranz, B.H.; Corey, P.N. Incidence of adverse drug reactions in hospitalized patients: A meta-analysis of prospective studies. *JAMA* **1998**, *279*, 1200–1205. https://doi.org/10.1001/jama.279.15.1200.
39. Sartori, D.; Aronson, J.K.; Norén, G.N.; Onakpoya, I.J. Signals of adverse drug reactions communicated by pharmacovigilance stakeholders: A scoping review. *Drug Saf.* **2023**, *46*, 109–120. https://doi.org/10.1007/s40264-022-01258-0.
40. Hammad, T.A.; Naylor, M.; Ely, D.M.; Davies, S. Exploring the complexities of disproportionality analysis in pharmacovigilance: Reflections on the READUS-PV guideline and a call to action. *Front. Pharmacol.* **2025**, *16*, 1573353. https://doi.org/10.3389/fphar.2025.1573353.

41. Juhlin, K.; Karimi, G.; Ande, M.; Camilli, S.; Dheda, M.; Har, T.S.; et al. Using VigiBase to identify substandard medicines: Detection capacity and key prerequisites. *Drug Saf.* **2015**, *38*, 373–382. https://doi.org/10.1007/s40264-015-0271-2.
42. Dal Pan, G.J. *Ongoing challenges in pharmacovigilance. Drug Saf.* **2014**, *37*, 1–8. https://doi.org/10.1007/s40264-013-0123-x.
43. Bergvall, T.; Norén, G.N.; Lindquist, M. vigiGrade: A tool to identify well-documented individual case reports and highlight systematic data quality issues. *Drug Saf.* **2014**, *37*, 65–77. https://doi.org/10.1007/s40264-013-0131-x.
44. Henry, S.; Buchan, K.; Filannino, M.; Stubbs, A.; Uzuner, O. 2018 n2c2 shared task on adverse drug events and medication extraction in electronic health records. *J. Am. Med. Inform. Assoc.* **2020**, *27*, 3–12.
45. Liu, X.; Rivera, S.C.; Moher, D.; Calvert, M.J.; Denniston, A.K.; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. *BMJ* **2020**, *370*, m3164. https://doi.org/10.1136/bmj.m3164.
46. Lardon, J.; Abdellaoui, R.; Bellet, F.; Asfari, H.; Souvignet, J.; Texier, N.; et al. Adverse drug reaction identification and extraction in social media: A scoping review. *J. Med. Internet Res.* **2015**, *17*, e171. https://doi.org/10.2196/jmir.4304.
47. De Abreu Ferreira, R.; Zhong, S.; Moureaud, C.; Le, M.T.; Rothstein, A.; Li, X.; et al. A pilot, predictive surveillance model in pharmacovigilance using machine learning approaches. *Adv. Ther.* **2024**, *41*, 2435–2445. https://doi.org/10.1007/s12325-024-02870-5.
48. Collins, G.S.; Moons, K.G.M.; Dhiman, P.; et al. TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* **2024**, *385*, e078378. https://doi.org/10.1136/bmj-2023-078378.
49. Dey, D.; Haque, M.S.; Islam, M.M.; Aishi, U.I.; Shammy, S.S.; Mayen, M.S.A.; et al. The proper application of logistic regression model in complex survey data: A systematic review. *BMC Med. Res. Methodol.* **2025**, *25*, 15. https://doi.org/10.1186/s12874-024-02454-5.
50. Breiman, L. *Random forests. Mach. Learn.* **2001**, *45*, 5–32. https://doi.org/10.1023/A:1010933404324.
51. Cortes, C.; Vapnik, V. *Support-vector networks. Mach. Learn.* **1995**, *20*, 273–297. https://doi.org/10.1007/BF00994018.
52. Xuan, P.; Sun, C.; Zhang, T.; Ye, Y.; Shen, T.; Dong, Y. *Gradient boosting decision tree-based method for predicting interactions between target genes and drugs. Front. Genet.* **2019**, *10*, 459. https://doi.org/10.3389/fgene.2019.00459.
53. Dsouza, V.S.; Leyens, L.; Kurian, J.R.; Brand, A.; Brand, H. Artificial intelligence (AI) in pharmacovigilance: A systematic review on predicting adverse drug reactions (ADR) in hospitalized patients. *Res. Soc. Adm. Pharm.* **2025**, *21*, 453–462. https://doi.org/10.1016/j.sapharm.2025.02.008.
54. Ly, T.; Pamer, C.; Dang, O.; Brajovic, S.; Haider, S.; Botsis, T.; et al. Evaluation of natural language processing (NLP) systems to annotate drug product labeling with MedDRA terminology. *J. Biomed. Inform.* **2018**, *83*, 73–86. https://doi.org/10.1016/j.jbi.2018.05.019.
55. Wasylewicz, A.; van de Burgt, B.; Weterings, A.; Jessurun, N.; Korsten, E.; Egberts, T.; et al. *Identifying adverse drug reactions from free-text electronic hospital health record notes. Br. J. Clin. Pharmacol.* **2022**, *88*, 1235–1245. https://doi.org/10.1111/bcp.15068.
56. Mahendran, D.; McInnes, B.T. Extracting adverse drug events from clinical notes. AMIA Jt. Summits Transl. Sci. Proc. 2021, 420–429.

57. Li, Z.; et al. Artificial intelligence-powered pharmacovigilance: A review of machine and deep learning in clinical text-based adverse drug event detection for benchmark datasets. *J. Biomed. Inform.* **2024**, *152*, 104621. https://doi.org/10.1016/j.jbi.2024.104621.
58. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; Association for Computational Linguistics: Minneapolis, MN, USA, 2019; pp. 4171–4186. https://doi.org/10.18653/v1/N19-1423.
59. Lee, J.; Yoon, W.; Kim, S.; et al. BioBERT: A pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics* **2020**, *36*, 1234–1240. https://doi.org/10.1093/bioinformatics/btz682.
60. Huang, K.; Altosaar, J.; Ranganath, R. ClinicalBERT: Modeling clinical notes and predicting hospital readmission. arXiv 2019, arXiv:1904.05342.
61. Liu, F.T.; Ting, K.M.; Zhou, Z.-H. Isolation forest. In Proceedings of the 2008 IEEE International Conference on Data Mining; IEEE: Pisa, Italy, 2008; pp. 413–422. https://doi.org/10.1109/ICDM.2008.17.
62. Trifirò, G.; Crisafulli, S. A new era of pharmacovigilance: Future challenges and opportunities. *Front. Drug Saf. Regul.* **2022**, *2*, 866898. https://doi.org/10.3389/fdsfr.2022.866898.
63. European Medicines Agency. Good Pharmacovigilance Practices (GVP): Module XVI – Risk Minimisation Measures: Selection of Tools and Effectiveness Indicators, Revision 3; European Medicines Agency: Amsterdam, The Netherlands, 2024.
64. Gandhi, S.; Iannacone, M.R.; Leapley, A.; Wang, L.; Mtenga, M.; Younus, M.; Wu, J. Real-time pharmacovigilance: Transforming population-based monitoring of post-approval vaccine safety through rapid cycle analysis (RCA)—A review of the published literature. *Pharmaceuticals* **2025**, *18*, 80. https://doi.org/10.3390/ph18010080.
65. Ebinger, J. *Artificial intelligence and digital health in the United States Food and Drug Administration context. Mayo Clin. Proc. Digit. Health* **2025**, *3*, 100231. https://doi.org/10.1016/j.mcpdig.2025.100231.
66. World Health Organization. Ethics and Governance of Artificial Intelligence for Health; World Health Organization: Geneva, Switzerland, 2021; ISBN 978-92-4-002920-0.
67. Hripcsak, G.; Duke, J.D.; Shah, N.H.; Reich, C.G.; Huser, V.; Schuemie, M.J.; et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for observational researchers. *Stud. Health Technol. Inform.* **2015**, *216*, 574–578. https://doi.org/10.3233/978-1-61499-564-7-574.
68. European Medicines Agency. Data Analysis and Real World Interrogation Network (DARWIN EU); European Medicines Agency: Amsterdam, The Netherlands, 2026.
69. Merotiwon, D.O.; Akomolafe, O.O.; Okoli, A.O.; Afrihyia, E. *Reviewing pharmacovigilance strategies using real-world data for drug safety monitoring and management. Int. J. Pharma Growth Res. Rev.* **2025**, *2*, 12–17. https://doi.org/10.54660/ijpgrr.2025.2.3.12-17.
70. Rudin, C. *Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat. Mach. Intell.* **2019**, *1*, 206–215. https://doi.org/10.1038/s42256-019-0048-x.
71. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems 30 (NIPS 2017); Curran Associates, Inc.: Long Beach, CA, USA, 2017.
72. Ribeiro, M.T.; Singh, S.; Guestrin, C. "Why should I trust you?": Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on

Knowledge Discovery and Data Mining; ACM: New York, NY, USA, 2016; pp. 1135–1144. https://doi.org/10.1145/2939672.2939778.

73. Jung, J.; Lee, H.; Jung, H.; Kim, H. Essential properties and explanation effectiveness of explainable artificial intelligence in healthcare: A systematic review. *Heliyon* **2023**, *9*, e16110. https://doi.org/10.1016/j.heliyon.2023.e16110.
74. Ferreira, R.; Cruz, R.; Inês, C. Beyond black boxes: Using explainable causal artificial intelligence to separate signal from noise in pharmacovigilance. *Int. J. Clin. Pharm.* **2026**, *48*, 677–681. https://doi.org/10.1007/s11096-025-02004-z.
75. Jedličková, A. Enhancing pharmacovigilance practices: Insights from audits and ethical connotations. Česká Slov. *Farm.* **2024**, *73*, E13–E22. https://doi.org/10.36290/csf.2024.033.
76. Jiang, F.; Jiang, Y.; Zhi, H.; Dong, Y.; Li, H.; Ma, S.; et al. Artificial intelligence in healthcare: Past, present and future. *Stroke Vasc. Neurol.* **2017**, *2*, 230–243. https://doi.org/10.1136/svn-2017-000101.
77. Rumbold, J.M.M.; Pierscionek, B. *The effect of the General Data Protection Regulation on medical research. J. Med. Internet Res.* **2017**, *19*, e47. https://doi.org/10.2196/jmir.7100.
78. Giacon, M.; Terrazzino, S. *Precision pharmacovigilance integrates genomics and real world evidence to overcome drug safety barriers. Discov. Pharm. Sci.* **2025**, *1*, 29. https://doi.org/10.1007/s44395-025-00032-2.
79. De Pretis, F.; Peden, W.; Landes, J.; Osimani, B. Pharmacovigilance as personalized evidence. In Personalized Medicine in the Making; Beneduce, C., Bertolaso, M., Eds.; Springer: Cham, Switzerland, 2022; pp. 147–171. https://doi.org/10.1007/978-3-030-74804-3_8.
80. Agbo, C.C.; Mahmoud, Q.H.; Eklund, J.M. Blockchain technology in healthcare: A systematic review. *Healthcare* **2019**, *7*, 56. https://doi.org/10.3390/healthcare7020056.
81. Bawany, N.Z.; Qamar, T.; Tariq, H.; Adnan, S. *Integrating healthcare services using blockchain-based telehealth framework. IEEE Access* **2022**, *10*, 36505–36517. https://doi.org/10.1109/ACCESS.2022.3161944.
82. Teo, Z.L.; Jin, L.; Liu, N.; Li, S.; Miao, D.; Zhang, X.; et al. Federated machine learning in healthcare: A systematic review on clinical applications and technical architecture. *Cell Rep. Med.* **2024**, *5*, 101419. https://doi.org/10.1016/j.xcrm.2024.101419.
83. Kaabachi, B.; Despraz, J.; Meurers, T.; Otte, K.; Halilovic, M.; Kulynych, B.; Prasser, F.; Raisaro, J.L. *A scoping review of privacy and utility metrics in medical synthetic data. NPJ Digit. Med.* **2025**, *8*, 60. https://doi.org/10.1038/s41746-024-01359-3.
84. Kuo, T.T.; Kim, H.E.; Ohno-Machado, L. *Blockchain distributed ledger technologies for biomedical and health care applications. J. Am. Med. Inform. Assoc.* **2017**, *24*, 1211–1220. https://doi.org/10.1093/jamia/ocx068.

Publisher's Note & Disclaimer

The statements, opinions, and data presented in this publication are solely those of the individual author(s) and contributor(s) and do not necessarily reflect the views of the publisher and/or the editor(s). The publisher and/or the editor(s) disclaim any responsibility for the accuracy, completeness, or reliability of the content. Neither the publisher nor the editor(s) assume any legal liability for any errors, omissions, or consequences arising from the use of the information presented in this publication. Furthermore, the publisher and/or the editor(s) disclaim any liability for any injury, damage, or loss to persons or property that may result from the use of any ideas, methods, instructions, or products mentioned in the content. Readers are encouraged to independently verify any information before relying on it, and the publisher assumes no responsibility for any consequences arising from the use of materials contained in this publication.