INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES
[ISSN: 0975-4725; CODEN(USA): IJPS00]
Journal Homepage: https://www.ijpsjournal.com

OPEN ACCESS

Review Article

Artificial Intelligence and Machine Learning in Pharmacovigilance: A Review of Applications, Challenges, and Future Perspectives

Sachin Kumbhar*, Dr. Sanjay Bais, Ashlesha Vasekar

Fabtech College of Pharmacy, Sangola.

ARTICLE INFO

Published: 27 Nov 2026

Keywords:
included pharmacovigilance, drug safety, machine learning, NLP, and signal detection, focusing on human medicines, automation methods, and validation metrics

DOI:
10.5281/zenodo.24582833

ABSTRACT

Background: Pharmacovigilance plays a pivotal role in post-marketing surveillance to detect, assess, understand, and prevent adverse drug reactions . As the volume of healthcare data scales exponentially, traditional manual case processing methods face critical bottlenecks, leading to delayed safety signals. Objective: The primary objective of this narrative review is to comprehensively summarize the emerging applications of Artificial Intelligence , Machine Learning , and Natural Language Processing in modern pharmacovigilance, evaluate their absolute benefits alongside systemic limitations, and discuss global regulatory and future technological directions. Methods: A comprehensive literature search was conducted using PubMed/MEDLINE, Google Scholar, and Scopus, including research studies, reviews, and guidelines from WHO, CIOMS, FDA, and EMA up to 2026.

INTRODUCTION:

1.1 Overview of Pharmacovigilance

1.1.1 Definition of Pharmacovigilance

According to the World Health Organization (WHO), pharmacovigilance (PV) encompasses the scientific research and practical measures involved in identifying, evaluating, comprehending, and mitigating adverse events or any issues associated with pharmaceuticals. PV is vital for maintaining medication safety and efficacy across a product's entire lifespan from early clinical evaluation through post-approval monitoring. Its primary objectives include detecting adverse drug reactions (ADRs), assessing safety risks, weighing benefits against potential harms, and disseminating crucial safety insights to clinicians and patients. By facilitating the reporting and analysis of potential side effects, pharmacovigilance fosters the safe, logical administration of therapies. Furthermore, its purview extends beyond standard pharmaceuticals to cover vaccines, biologics, herbal and traditional remedies, blood derivatives, and medical devices[1].

***Corresponding Author:** Sachin Kumbhar
Address: Fabtech College of Pharmacy, Sangola
Email : ashleshavasekar@gmail.com

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

1.1.2 Objective of pharmacovigilance:

The main goal of pharmacovigilance (PV) is to enhance patient safety while encouraging the safe and therapeutic application of pharmaceuticals. It focuses on identifying, evaluating, comprehending, and mitigating adverse drug reactions (ADRs) and related safety concerns. Crucially, PV aids in uncovering previously undetected side effects, tracking established risks, and analysing shifts in a drug's safety profile post-approval. It continuously weighs a product's overall benefits against its risks, supplying dependable safety data to clinicians, patients, and regulatory agencies. Furthermore, PV drives the reporting of suspected adverse reactions, informs necessary regulatory interventions, and helps minimize preventable harm. Through ongoing monitoring across both clinical trials and real-world post-marketing environments, PV ensures that medications maintain an acceptable safety profile [1].

1.1.3 Prevention and Detection Medicine related harm:

Table 1: Prevention and Detection of Medicine Related Harm

Stage	Key activities
Data collection & detection	Spontaneous reports (patients & healthcare professionals); electronic health records (EHRs) and claims data; clinical trials and epidemiological studies; literature reviews; product defect reports
Signal detection & data mining	Statistical analysis; pattern recognition
Signal evaluation & assessment	Causality assessment; medical review; plot (case series) analysis; scientific review
Regulatory authorities (e.g., FDA)	Review of confirmed risks
Prevention strategies	Informed prescribing → patient education; patient education → enhanced product design; medication error reduction
Harm-reduction outcomes	Fewer adverse events; reduced morbidity and mortality; improved patient safety
Goals of pharmacovigilance	Detect new safety issues; assess benefit–risk balance; understand causality; communicate risks; prevent harm

1.1.3.1 Adverse event:

An adverse event is any unwanted or unexpected health problem that happens to a patient while they are taking a medication. It does not automatically mean the drug caused the issue it simply refers to any negative medical event that occurred during treatment[2].

1.1.3.2 Adverse drug reaction (ADR):

An adverse drug reaction is a harmful or unwanted side effect caused by a medication. Unlike a general event, an ADR specifically happens because the medicine itself is suspected to have triggered the negative response.

1.1.3.3 Serious adverse event/reaction:

A serious adverse event is a critical health crisis that occurs while someone is receiving medical treatment. It is considered "serious" because it results in a devastating outcome, such as needing emergency hospital care, facing a life-threatening situation, suffering a lasting

disability, or even death. This also includes any birth defects in a child whose parent was treated. Basically, if an event puts your life at risk or changes it significantly, doctors will classify it as serious.

1.1.3.4 Safety signal:

A safety signal is an early warning sign that a medicine might be causing a new, unexpected side effect or that a known side effect is behaving differently than expected. It's not proof of harm yet, but rather a red flag that prompts experts to investigate further to protect patients.

1.1.3.5 Individual case safety report (ICSR):

An Individual Case Safety Report (ICSR) is a detailed record of a single patient's experience with a suspected side effect. It brings together all the crucial details about the patient, the medicine they took, and the health problem that occurred so safety experts can review what happened[1,2].

1.1.4 Traditional PV Workflow:

The end-to-end pharmacovigilance process starts when an individual whether a patient, doctor, or clinical trial researcher reports a suspected drug side effect. These initial Individual Case Safety Reports (ICSRs) undergo an intake and triage phase to ensure all necessary details are present and valid; any non-compliant or duplicate reports are archived. Once verified, the case details are logged into a central safety database and standardized using MedDRA medical terminology. Next, qualified physicians perform a medical review to evaluate the event's severity and determine if there is a direct causal link between the medication and the reaction. Ultimately, the centralized data powers downstream safety workflows, including submitting reports to regulatory agencies (such as the FDA, EMA, or DCGI), implementing risk mitigation strategies like label revisions and Dear Healthcare Professional Letters (DHPCs), and

conducting ongoing signal detection to uncover emerging safety trends[3,4].

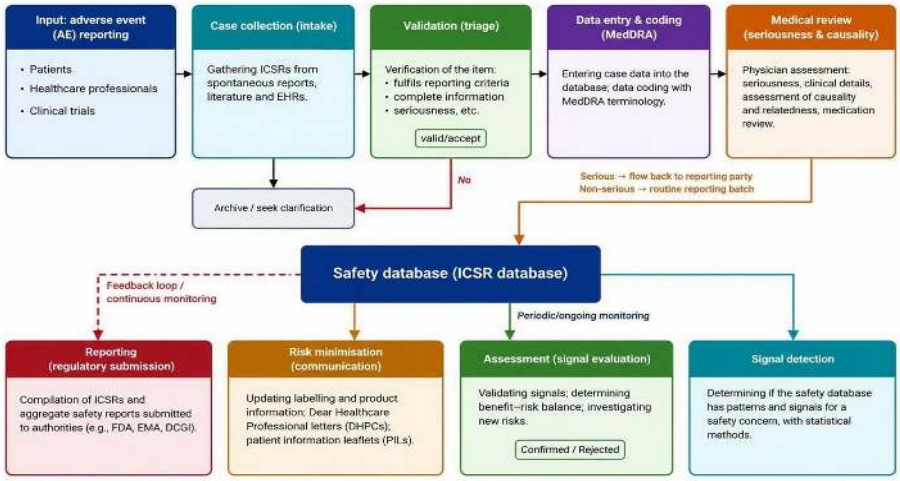


Figure 1: Traditinal PV Workflow

1.1.5 Limitations of conventional workfow

Traditional pharmacovigilance workflows face major operational hurdles. Safety teams are routinely overwhelmed by the high volume of incoming safety reports, while simultaneously dealing with widespread under-reporting of actual adverse events. Furthermore, incoming cases often lack essential clinical details, which limits the precision of safety assessments. Because these traditional workflows are heavily manual, they are inherently labor-intensive, expensive, and prone to processing delays that can slow down regulatory submissions. Finally, managing duplicate reports from multiple sources adds another layer of

redundancy, ultimately reducing overall efficiency and elevating potential risks to patient safety[7,10].

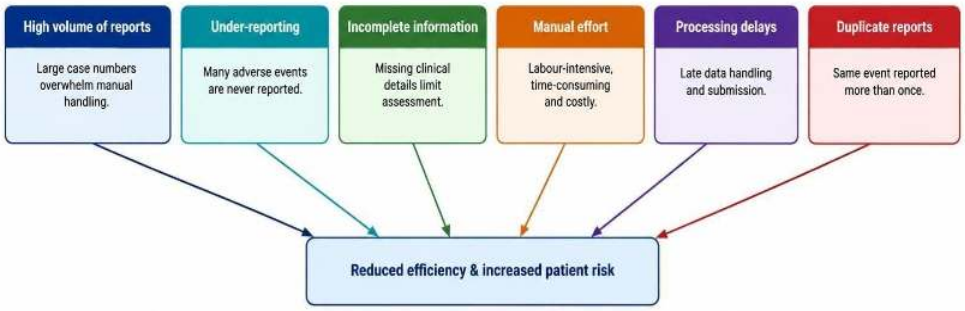


Figure 2: Limitations of Conventional Pharmacovigilance Workflows That Reduce Efficiency and Increase Patient Risk.

1.1.2 Introduction to AI, ML and NLP :

2. Artificial intelligence:

2.1 Definition

Artificial Intelligence (AI) is a core branch of computer science focused on building intelligent systems capable of performing tasks that traditionally require human cognitive abilities such as logical reasoning, learning, problem-solving, pattern recognition, and natural language understanding[3].

2.1.1 Role in Pharmacovigilance:

AI significantly enhances drug safety operations by efficiently sifting through vast volumes of data. It helps identify potential adverse drug reactions (ADRs), surface emerging safety signals, and automate routine tasks in individual case safety report (ICSR) processing. Furthermore, AI tools extract valuable safety insights from divers [3].

2..1.2Machine Learning (ML):

sources such as electronic health records and published literature. Crucially, however, AI outputs must always be validated and paired with human oversight to ensure reliability[3].

2.1.3 Definition:

Machine Learning (ML) is a branch of artificial intelligence in which algorithms discover patterns and correlations directly from data. Instead of relying on rigid, pre-programmed rules, ML models use historical information to make predictions, categorize data, and spot emerging trends autonomously[4].

2.1.4 Role of Machine learning :

Machine learning (ML) has become essential to modern pharmacovigilance by efficiently managing the influx of drug safety data. Rather than relying on rigid, pre-written rules, ML systems uncover hidden trends and correlations directly from past health data. These models help triage high-risk cases for expert review, categorize adverse event reports, highlight potential drug–event links, and pick up faint safety signals across clinical notes and medical literature. Ultimately, ML speeds up routine safety monitoring while relying on human clinical judgment to ensure accuracy[4].

2.2.1 Types of Machine learning:

Supervised Learning: It uses labelled data to predict outcomes. Common techniques include classification and regression.

Unsupervised Learning: It identifies hidden patterns and relationships in unlabelled data using clustering and association techniques.

Semi-supervised Learning: It combines a small amount of labelled data with a large amount of unlabelled data to improve learning.

Reinforcement Learning: It learns through interaction with an environment using rewards and feedback to improve decisions[4,7].

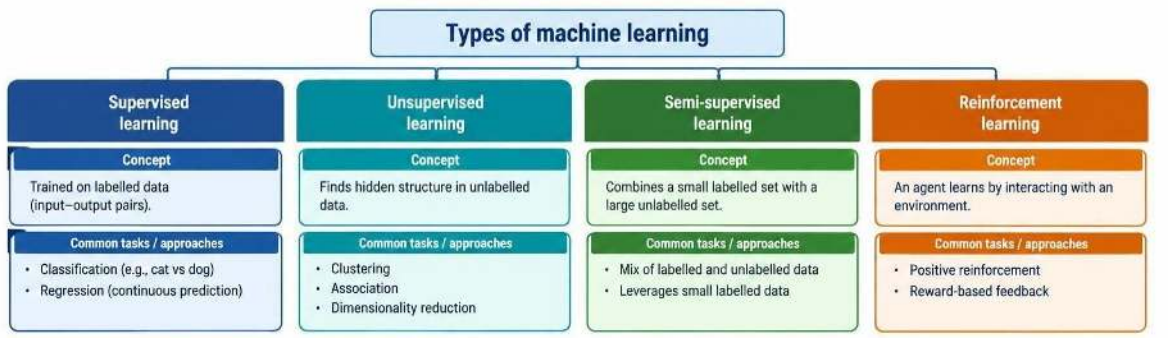


Figure 3: Types of Machine Learning: Supervised, Unsupervised, Semi-Supervised and Reinforcement Learning.

2.2.2 Deep Learning (DL):

2.2.3 Definition

Deep Learning (DL) is an advanced branch of machine learning that utilizes multi-layered artificial neural networks. These complex architectures allow models to automatically discover high-level patterns and intricate relationships directly from massive dataset[5].

2.2.4 Role in Pharmacovigilance

In drug safety, deep learning architectures such as Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), and Transformer models—are used to process complex clinical narratives, map intricate drug–adverse event relationships, and handle large volumes of unstructured safety data. However, the accuracy and reliability of these models remain strictly dependent on the quality and representativeness of their training data[5].

2.3.1 Natural Language Processing (NLP):

2.3.2 **Definition** -

Natural Language Processing (NLP) is a specialized domain of artificial intelligence that empowers computers to analyse, comprehend, and extract structured meaning from human language in both text and speech formats[6].

2. .3.3Why NLP is Essential for Drug Safety:

- **Workflow Efficiency:** Automatically classifies incoming reports, extracts essential patient and medication details, and checks for incomplete data.
- **Standardization & Speed:** Assists with MedDRA coding, prompts follow-up questions, and detects duplicate records—saving time and ensuring consistent data entry.
- **Human-in-the-Loop Safeguards:** While automation handles repetitive tasks, expert human verification is critical to prevent missed cases, inaccurate data extraction, and false duplicate matches.

Much of the world's most critical medical data doesn't live in neatly organized spreadsheets it is locked away in free-text clinical notes, published literature, physician narratives, and online patient forums. NLP solves this challenge by automatically identifying key safety elements like dosages, symptoms, drug names, and suspected adverse events. By structuring this raw narrative data, NLP accelerates adverse event tracking, drug–event connection mapping, and case intake workflows[6].

2.3.4Different Key technologys

Artificial Intelligence (AI)

What it is: The umbrella field focused on creating smart machines that mimic human intelligence.

Core Function: Enables computers to reason, solve problems, and make decisions.

Machine Learning (ML)

What it is: A subset of AI where systems learn from data automatically.

Core Function: Finds patterns in information to make predictions without direct programming.

Deep Learning (DL)

What it is: An advanced subset of ML modeled on the human brain.

Core Function: Uses multi-layered neural networks to analyze massive, complex data (like images or speech) [3,7].

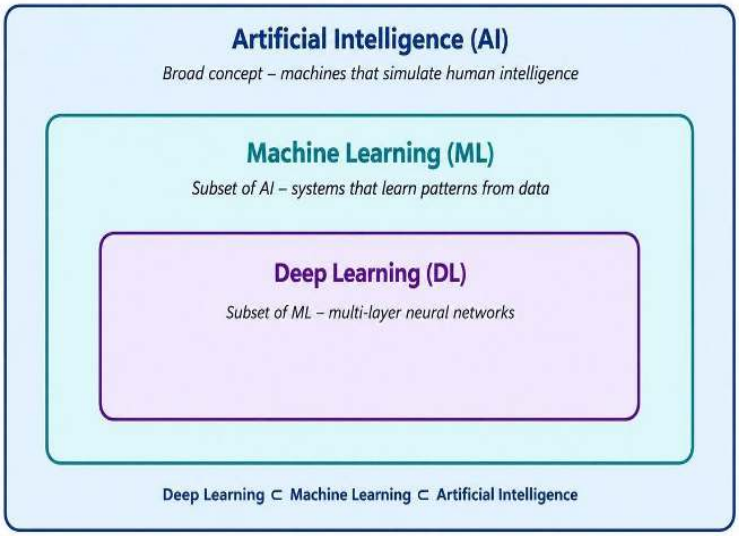


Figure 4: Relationship Between Artificial Intelligence (AI), Machine Learning (ml) and Deep Learning (dl): dl Is a Subset of ml, Which Is in Turn a Subset of AI.

2.4.1 Rationale for AI/ML in Pharmacovigilance:

The absolute rationale for deploying AI/ML architectures in modern pharmacovigilance stems from the rapid expansion and structural shift of the global data ecosystem. Drug safety professionals no longer monitor simple, structured forms alone; they face a massive influx of diverse information spanning unstructured electronic health records (EHRs), healthcare claims databases, spontaneous registry databases, thousands of weekly biomedical literature publications, and extensive patient-generated text on digital health forum[7].

2.4.2 Potential benefits of AI in pharmacovigilance

- Rapid Case Triage: Automatically screens incoming reports and flags urgent issues instantly.
- Effortless Data Extraction: Uses NLP to quickly gather key details on drugs, side effects, and patient outcomes from unstructured text.
- Smart Prioritization: Ranks cases by risk level so experts focus on critical issues first.
- Deep Pattern Recognition: Detects hidden safety signals across massive datasets that human review might miss.
- Smarter Decisions: Gives pharmacovigilance teams data-driven insights to make faster, informed safety evaluations[7].
- Balancing efficiency with clinical validity, explain ability, data protection, and patient safety:
- Clinical Validity: AI must be validated against real-world clinical data to ensure accurate, trustworthy results.
- Explain ability: Pharmacovigilance experts must understand the logic behind AI predictions before relying on them for safety assessments.
- Data Protection: Robust privacy safeguards and strict access controls are mandatory when handling sensitive patient health records.
- Patient Safety: Human experts must review AI outputs to catch potential biases, false positives, or missed adverse events.
- Efficiency vs. Accuracy: Smart automation should lighten administrative workloads while maintaining strict quality and compliance standard[7].

3. Aim, Objectives and Scope :

3.1 **Aim : Artificial Intelligence and Machine Learning in Pharmacovigilance: A Review of Applications, Challenges, and Future Perspectives.**

3.1.1 Objective :

i. **Uncover Safety Issues Early:** Swiftly spot unexpected adverse drug reactions (ADRs) and hidden safety concerns before they escalate.

ii. **Streamline Safety Reporting:** Automate routine tasks like data entry, info extraction, and individual case safety report (ICSR) processing.

iii. **Smart Signal Detection:** Pick up subtle patterns, trends, and unusual links between medications and adverse events across vast datasets.

iv. **Spot Drug–Drug Interactions:** Analyse complex safety data to catch dangerous interactions between co-administered medications.

v. **Elevate Data Quality:** Automatically flag duplicate records, fill in missing details, and ensure accurate coding for reliable safety data.

vi. **Make Sense of Large Datasets:** Seamlessly pull insights from EHRs, medical literature, clinical notes, and real-world drug safety sources.

vii. **Cut Time and Expenses:** Slash manual workloads to make overall pharmacovigilance operations faster, leaner, and more efficient.

viii. **Empower Clinical Decisions:** Help drug safety teams evaluate complex data and prioritize high-risk cases for rapid review.

ix. **Protect Patient Health:** Catch medicine-related risks early to prevent avoidable harm and keep patients safe[8].

3.1.2.Application:

i. **Spotting Adverse Reactions:** Scans medical records and safety reports to catch potential side effects early.

ii. **Uncovering Safety Signals:** Identifies hidden patterns and links between specific medications and unexpected health events.

iii. **Streamlining Case Processing:** Automatically reads and extracts key details from Individual Case Safety Reports (ICSRs), saving valuable time.

iv. **Live Safety Tracking:** Continuously monitors live health data to catch emerging safety concerns as they happen.

v. **Scanning Literature & Social Media:** Digs through scientific journals and online platforms to pull out relevant drug safety insights.

vi. **Catching Duplicates:** Uses smart AI tools like vigiMatch to find and remove repeated report entries before they distort data.

vii. **Evaluating Patient Risk:** Helps doctors and researchers pinpoint high-risk individuals and make better decisions about drug safety[9].

4. AI/ML Workflow in Pharmacovigilance 4.1Introduction to AI/ML Integration

Artificial intelligence (AI) and machine learning (ML) are rapidly transforming pharmacovigilance (PV), turning the tide on the ever-growing mountains of safety data associated with adverse drug reactions (ADRs). Traditionally, sorting through these safety reports has been a painstakingly manual, time-consuming challenge. Today, advanced techniques like natural language processing, machine learning models, and deep learning architectures empower PV professionals to systematically organize safety signals, parse complex data points, and spot hidden trends across massive datasets. Yet, despite these technological leaps, automation is not a silver bullet; any insights generated by AI must still undergo rigorous, hands-on review by qualified healthcare experts before shaping clinical actions or regulatory policies[10,11].

4.1.1 Data Sources in Pharmacovigilance

Ensuring medicine safety relies on drawing from a rich mosaic of data streams, each carrying distinct strengths and inherent flaws:

4.1.2 Spontaneous Reporting Systems & EHRs:

Spontaneous reporting systems capture suspected ADR accounts straight from patients, clinicians, and pharmaceutical companies, serving as vital early-warning tools for unexpected hazards—though they frequently

suffer from underreporting and duplicate entries. Complementing this, electronic health records (EHRs) offer a real-world clinical lens into patient histories and treatment trajectories, despite occasional gaps or formatting inconsistencies.

4.1.3 Claims Data & Registries:

Meanwhile, medical claims databases provide a broad population-level view of medicine utilization, while patient registries track long-term outcomes for specific conditions, though sometimes struggling with generalizability.

4.1.4.Literature, Poison Centres, & Social Media:

Furthermore, medical literature, poison-centre databases, and even unverified social media chatter or patient forums offer auxiliary insights into adverse symptoms, provided they are handled with robust privacy safeguards and critical scrutiny[10].

4.2 Data Preparation and Pre-processing

Before any of this multi-source data can feed into a machine learning model, it must undergo meticulous preparation and pre-processing, because the golden rule of "garbage in, garbage out" rings especially true in medicine safety. Raw PV datasets are notoriously messy, riddled with typos, inconsistent drug naming conventions, missing timelines, and unstructured clinical narratives:

4.2.1 De-identification & Normalization:

To tame this chaos, data engineers must strip away patient identifiers to safeguard privacy while complying with strict regulations. They then normalize disparate formats, standardize measurement units, and handle missing data thoughtfully—never treating absent information as a confirmed clinical fact.

Deduplication, Terminology Mapping, & Quality Checks: Sophisticated duplicate-detection scripts help eliminate redundant reports, while standard medical terminologies like MedDRA are integrated to ensure uniform coding of adverse events and pharmaceutical agents. Finally, data quality checks ensure logical sequencing and clinical accuracy before training begins[11].

4.2.2 Feature Extraction & Variable Creation

Once the data is clean, the workflow moves into feature extraction, bridging the gap between raw records and actionable variables that an algorithm can actually parse:

4.2.3.Drug & Patient Attributes:

This stage involves pulling out critical details such as the suspected drug's name, dosage, administration route, treatment duration, and frequency of use. Concurrently, systems extract core event characteristics including reported symptoms, severity, and clinical outcomes alongside vital patient demographics like age, gender, and pre-existing medical conditions.

4.2.4NLP & Entity Recognition:

By leveraging natural language processing and named-entity recognition, AI can seamlessly scan dense clinical notes and scientific papers to isolate relevant medical entities, establishing a solid foundation for subsequent case classification and signal prioritization[11].

4.3 Model Development & Validation

Building a robust predictive model requires a clear definition of the specific pharmacovigilance hurdle the system is designed to tackle, whether that means flagging high-risk adverse event reports, classifying complex clinical narratives, or prioritizing safety signals:

4.3.1. Algorithm Selection & Data Partitioning:

Once the goal is locked in, data scientists select appropriate algorithms—ranging from classic logistic regression and random forests to advanced deep learning architectures—and partition the dataset into training, validation, and test subsets.

4.3.2. Data Leakage & Benchmarking:

Guarding against data leakage is paramount during this phase to prevent the model from cheating by recognizing duplicated cases across splits or peeking at future information. Ultimately, developers benchmark their creation against simpler baseline rules or current manual workflows to prove it delivers genuine value over status-quo practices[12].

4.4 Model Evaluation and Deployment

Evaluating an AI model goes far beyond a simple accuracy score; it demands a nuanced look at metrics tailored to the high-stakes nature of drug safety:

4.4.1 Performance Metrics & Calibration:

Sensitivity and recall ensure that few genuine adverse reactions slip through the cracks, while specificity and precision help minimize false alarms that could overwhelm review boards. Additionally, metrics like the F1-score balance skewed datasets, and calibration checks ensure that predicted probabilities align with real-world frequencies.

4.4.2 Operational Impact & Drift Monitoring:

Once deployed, the system cannot be left on autopilot. Because healthcare landscapes evolve patient demographics, new prescribing trends, or changing reporting behaviours can trigger data or model drift—ongoing real-world monitoring, periodic revalidation, and occasional retraining are absolute necessities[13].

4.5 The Human-in-the-Loop Framework:

Ultimately, the most effective pharmacovigilance frameworks rely on a symbiotic "human-in-the-loop" model that pairs the blazing processing speed of artificial intelligence with the nuanced clinical judgment of human experts:

4.5.1 Decision-Support & Expert Review:

Rather than letting algorithms make autonomous regulatory decisions, the AI system acts as a sophisticated cognitive assistant organizing safety files, surfacing hidden patterns, and flagging cases that demand a closer look.

4.5.2.Clinical Oversight & Accountability:

Pharmacovigilance specialists then step in to cross-verify extracted details against original source records, weigh alternative medical explanations, assess concomitant medications, and evaluate the plausibility of a causal link. This human oversight ensures accountability, catches algorithmic blind spots, and safeguards patient well-being[14].

5. Major AI/ML Methods Used in pharmacovigilance:

INTRODUCTION:

Pharmacovigilance (PV) canters on the ongoing collection, detection, assessment, monitoring, and prevention of adverse drug reactions (ADRs) and broader medication safety issues. Traditional pharmacovigilance workflows such as manual case processing, periodic spontaneous reporting system (SRS) reviews, and classical disproportionality metrics—struggle to keep pace with today's massive volume of health data. They face inherent hurdles due to global underreporting and the unorganized nature of clinical notes, biomedical literature, and electronic health records (EHRs).To address these operational bottlenecks, artificial intelligence (AI) and machine learning (ML) have become indispensable tools. These technologies streamline data ingestion, literature screening, key term extraction, and early signal identification. Below is an overview of the primary AI/ML methodologies applied in current pharmacovigilance practices, including how they work, their key challenges, and how they compare across core tasks[15].

5.1.1 Major AI/ML methods used in pharmacovigilance:

1. Rule-Based Systems

- **How They Work:** Rule-based tools rely on defined dictionaries, curated domain logic, and direct Boolean rules (for example, flagging a case when a specific medication and symptom appear in the same entry).
- **Application in PV:** These systems excel at straightforward, standardized activities—such as mapping reported verbatim terms directly to Medical Dictionary for Regulatory Activities (MedDRA) terminology or verifying required fields in Individual Case Safety Reports (ICSRs).
- **Key Challenges:** They lack flexibility. As medical jargon evolves, or when syntax, colloquial expressions, and typos crop up in real-world clinical records, rule-based systems tend

to miss critical data and require frequent manual upkeep[16].

2. Supervised Machine Learning

- **How They Work:** Models such as logistic regression, decision trees, random forests, support vector machines (SVMs), and gradient boosting algorithms learn from historical, human-labelled data to classify new entries or project outcomes.
- **Training & Challenges:** Effective performance relies on high-quality historical training labels (such as verified causality determinations). A core challenge in safety surveillance is severe **class imbalance**—true adverse events make up only a small fraction of overall medical records compared to normal exposures. Overcoming this

hurdle calls for specialized sampling techniques (like SMOTE), adjusted class weightings, and precision-focused evaluation metrics like the Area Under the Precision-Recall Curve (AUPRC), rather than standard accuracy scores.

3. Unsupervised and Semi-Supervised Learning

- **How They Work:** Unsupervised tools including clustering methods (e.g., k-means, hierarchical clustering) and outlier detection models —group dataset points or flag irregularities without relying on pre-annotated labels.
- **Application in PV:** These techniques are ideal for exploratory research, spotting hidden patient phenotypes, and surfacing unusual safety signals across large, unstructured repositories.
- **Key Challenges:** Because these algorithms identify mathematical clusters rather than confirmed medical cause-and-effect patterns, they carry a noticeable risk of false positives and always require thorough follow-up review by medical safety experts.

4. Deep Learning

- **How They Work:** Deep neural networks and multi-layered transformer models automatically discover intricate, non-linear patterns within high-dimensional datasets without requiring manual feature design.
- **Application in PV:** Deep learning excels at analysing multi-layered patient data, helping researchers pick up on subtle drug-drug interactions (DDIs), polypharmacy risks, and broad clinical trends hidden across large-scale EHR networks.

5.1.2. Natural Language Processing (NLP) and Named-Entity Recognition (NER)

- **How They Work:** NLP and NER tools read and structure free-form clinical narrative text—such as doctor visit notes, hospital discharge summaries, patient support forums, and research papers.
- **Application in PV:** These tools automatically pinpoint and extract crucial safety variables, including drug names, adverse events, dosage details, administration routes, timestamps, patient demographics, and clinical outcomes.
- Example-
- Drug name: telmisartan
- Adverse reaction: dizziness
- Dose: 40 mg
- Route: oral
- Date: date of treatment or reaction, when documented
- Patient details: age and sex, when available
- Outcome: recovered, recovering, or other documented status

5.1.3Large Language Models (LLMs) and Generative AI

- **How They Work:** Pre-trained transformer architectures use broad text training to perform zero-shot or few-shot tasks, enabling rapid text summarization, structured data capture, literature triage, and initial safety narrative drafting.
- **Key Challenges:** Despite their adaptability, LLMs introduce notable risks into safety workflows. These include **hallucinations** (generating confident but incorrect clinical details), high sensitivity

to prompt structure, variable outputs, and data privacy concerns regarding sensitive personal health information[16].

6Application 1: Automated ICSR Intake and Case Processing 6.1Introduction

Individual Case Safety Reports are at the heart of pharmacovigilance, serving as key records of potential adverse drug reactions. Received from diverse channels such as doctors, patients, clinical trials, medical journals, emails, and online forms processing these reports traditionally demands extensive manual work. Safety teams must handle intake, triage, data entry, completeness checks, medical coding, duplicate detection, and quality control. When case volumes surge, this manual workflow can quickly become time-consuming and resource-heavy. Modern AI, Machine Learning, and Natural Language Processing (NLP) tools offer a way to streamline this process. By classifying incoming reports, extracting key data from unstructured text, flagging missing details, assisting with medical coding, and catching duplicates, AI helps speed up workflows and improve consistency. Crucially, these technologies don't replace human experts; instead, they take on the repetitive heavy lifting so pharmacovigilance professionals can focus on critical medical assessments and regulatory decisions [17].

6.1.1 Application :

- **Workflow Efficiency:** Automatically classifies incoming reports, extracts essential patient and medication details, and checks for incomplete data.
- **Standardization & Speed:** Assists with MedDRA coding, prompts follow-up questions, and detects duplicate records saving time and ensuring consistent data entry.
- **Human-in-the-Loop Safeguards**: While automation handles repetitive tasks, expert human verification is critical to prevent missed cases, inaccurate data extraction, and false duplicate matches[17].

6.1.2 Benefits and limitations of automated ICSR processing:

The primary advantage of AI-assisted ICSR processing lies in taking over repetitive, manual tasks. Automating document classification, data extraction, completeness checks, code suggestions, and duplicate detection allows safety teams to work with greater efficiency and consistency. As report volumes continue to rise, automation helps manage the workload, freeing up safety experts to focus their energy on complex medical evaluations. However, incorporating automation brings notable risks. A bad data extraction can alter a report's meaning, a missed case can delay critical safety reviews, and a false duplicate match might cause distinct reports to be incorrectly merged. Furthermore, an AI's performance often hinges on input quality, language nuances, spelling errors, terminology variations, and data formats. To mitigate these risks, AI systems must be thoroughly validated using representative datasets, continuously monitored for errors, and designed with human oversight at the core. Current research consistently underscores a key principle: AI should enhance human expertise, not replace it[17,18].

6.1.3 Evaluation of AI-based ICSR systems

Table 2: Evaluation of AI Based ICSR System

Evaluation measure	Purpose
Sensitivity (recall)	Measures how many genuinely reportable cases are correctly identified.
Precision	Measures how many cases flagged by the system are genuinely relevant.
Agreement with expert reviewers	Assesses whether extracted information, classifications, and coding suggestions match expert decisions.
Processing time	Compares the time required for automated and manual workflows.

Evaluation measure	Purpose
Error rate	Measures incorrect extraction, classification, coding, or duplicate identification.
Duplicate detection performance	Assesses the number of correctly identified duplicates and false matches.
Data completeness and quality	Determines whether required fields are captured accurately and consistently.

7. Application 2: NLP for ADR Detection :

7.1 Natural Language Processing (NLP) for Adverse Drug Reaction (ADR) Detection:

Natural Language Processing (NLP) is a branch of artificial intelligence that helps computers read, understand, and make sense of human language. In pharmacovigilance, NLP plays a key role in uncovering potential adverse drug reactions (ADRs) buried within unstructured text such as electronic health records , doctor's notes, discharge summaries, scientific journals, and patient reports. By sifting through vast amounts of medical text automatically, NLP takes the heavy lifting off manual reviews and helps safety teams catch potential drug hazards much faster[19].

7.1.1 Extraction of ADR information

NLP tools automatically scan medical documents to extract crucial details like drug names, dosages, symptoms, and side effects. For instance, if a clinical note mentions that a patient came down with a skin rash shortly after starting a new medication, NLP can connect the dots between the drug and the reaction. Pharmacovigilance specialists can then use these findings to monitor safety trends and dig deeper into prospective risks[19].

7.1.2.Identification of drug ADR relationships:

NLP is smart enough to read context it knows whether a patient actually took a medication or if the doctor was just discussing it. It also filters out unrelated health issues so safety teams can focus on real side effects, all while building a clear timeline of events. That said, correlation isn't causation. Just because a symptom shows up after starting a pill doesn't guarantee the pill caused it, which is why human experts still need to evaluate the findings and make the final call.

8.1 Handling negation, uncertainty and clinical context:

To accurately spot true side effects, NLP must understand the full context of a medical note.

It looks out for nuances like:

8.8.1Negation:

Identifying statements like "the patient has no rash" so an absent symptom isn't misclassified as an active reaction.

8.1.2 Uncertainty:

Distinguishing potential risks from proven facts for example, "the drug may have caused dizziness" signals a suspected issue, not a confirmed one.

8.1.3 Timing & History:

Recognizing past events so a note like "the patient had nausea last year" isn't mistaken for a current side effect.

8.1.4 Actual Use vs. Prescription:

Knowing that "the patient was prescribed Drug A" doesn't guarantee they actually took the medication.

8.2 Major NLP techniques used in ADR detection:

8.2.1 Named-Entity Recognition (NER):

Finds and tags critical medical terms within the text such as drug names, symptoms, conditions, and suspected side effects.

8.2.2.Relation Extraction:

Connects the dots between those terms to determine if a specific drug is actually linked to a specific adverse event.

8.2.3.Text Classification:

Filters and categorizes entire reports or records, sorting relevant drug safety information from unrelated medical noise based on how the model was trained.

Working together, these techniques transform disorganized clinical notes into clean, structured safety data giving pharmacovigilance teams the accurate insights they need to protect patient health[20].

8.3 Benefits and limitations:

NLP for ADR detection streamlines safety data screening, uncovers adverse events often buried in unformatted medical notes, and cuts down on repetitive manual reviews. However, AI accuracy can still be hampered by sloppy text, missing documentation, vague terminology, or complex drug-event linkages. Because of these limitations, expert human review remains essential before making any final pharmacovigilance decisions[19,20].

9.Application 3: Signal Detection and Prioritization
9.1Definition of safety signal

A safety signal alerts experts to potential new side effects or unexpected patterns with a drug. Today, AI and machine learning process huge volumes of safety data to quickly flag and prioritize these potential risks, enabling pharmacovigilance teams to review critical safety issues faster and more effectively[21].

9.1.1. Signal detection, validation, assessment and regulatory action

Monitoring drug safety is a structured, four-part journey:

9.1.2. Signal Detection:

Identifying early flags or patterns in patient data, safety reports, and clinical records that hint at a potential drug-related side effect.

9.1.3. Signal Validation:

Screening those flags to confirm the data is reliable, relevant, and worth spending resources to investigate further.

9.1.4 Signal Assessment:

Conducting a deep-dive evaluation weighing clinical findings, statistical trends, and epidemiological data to measure the true strength of the association.

9.1.5 Regulatory Action:

Implementing official safeguards based on the findings, such as issuing safety alerts, updating product warnings, or limiting patient use[21].

10AI/ML methods used for signal detection

Modern pharmacovigilance utilizes several AI and machine learning techniques to flag potential safety concerns:

10.1 Supervised Classification:

Models leverage pre-labelled datasets to evaluate drug –event combinations and assign them probability scores, helping teams prioritize meaningful signals over random noise.

10.1.1 Anomaly Detection:

By tracking baseline reporting trends, anomaly detection flags statistical outliers and sudden surges in data, offering an early warning system for novel adverse events.

Graph None: Based Networks:

By modelling biological and clinical relationships as interconnected networks, graph analysis uncovers complex, multi-layered pathways linking drugs to unexpected adverse events.

10.1.3 Ensemble Modelling:

By synthesizing outputs from multiple algorithms, ensemble approaches deliver more accurate and stable risk-prioritization results than any individual model could achieve alone[22]. **10.1.4Advantages and limitation:**

AI and machine learning are transforming drug safety. By scanning massive databases in seconds, these tools can spot complex patterns, catch potential side effects much earlier, and dramatically lighten the heavy manual workload of sorting through safety reports. But AI isn't fool proof its accuracy depends heavily on the quality of real-world data. Things like underreporting,

biased reporting, or sudden spikes in reports after a media headline can trick an algorithm into seeing a problem where there isn't one (or missing one altogether). For instance, if a public safety warning goes out, people start reporting that side effect much more often simply because they're aware of it, not because it's actually happening more frequently.

Because of these nuances, AI predictions must always be thoroughly tested, transparently audited, and paired with expert human judgment[22].

11.Challenges, Ethics, and Regulatory Considerations

Deploying AI tools in highly regulated medical safety environments requires strict guardrails. One major hurdle is that severe adverse events and true drug toxicities are extremely rare. This heavy imbalance in the data creates a real risk of false negatives meaning an AI might overlook a critical safety issue. Additionally, complex "black-box" deep learning models must be paired with Explainable AI (XAI) frameworks like SHAP or LIME. These tools open up the AI's inner decision-making process, making it completely transparent, understandable, and auditable for international health regulators [23].

12Conclusion and future prospects

CONCLUSION:

This narrative review demonstrates that Artificial Intelligence and Machine Learning represent a transformative paradigm shift in modern pharmacovigilance. The digital transformation of global healthcare has rendered traditional manual case processing obsolete. AI-driven models particularly transformer-based NLP architectures and advanced supervised classification networks demonstrate mature operational efficiency in automating data intake, extracting clinical entities from unstructured text, removing duplicate reports, and prioritizing safety signals. However, technology cannot function in a clinical vacuum. The integration of AI/ML is bounded by challenges regarding data quality, algorithmic bias, the critical need for explain ability, and strict international regulatory standards. The clear consensus across modern literature emphasizes that AI/ML architectures should serve as powerful decision-support systems to augment human capability. They cannot replace the ultimate clinical causality assessment, ethical judgment, and regulatory accountability of qualified human medical experts.

12.2 Future Perspectives

The next decade of automated pharmacovigilance will be defined by the transition from retrospective, database-driven tracking to real-time active safety surveillance ecosystems. The integration of Explainable AI will bridge the current trust gap, providing completely transparent, auditable pathways for automated safety decisions. Future breakthroughs will rely on multimodal AI frameworks capable of simultaneously analysing genomic datasets, chemical properties, laboratory values, and unstructured clinical notes to predict personalized patient toxicity risks. To achieve this, global drug safety operations must prioritize multicentre validation studies, establish open-source benchmark safety datasets, and build collaborative,

multidisciplinary networks connecting data scientists, pharmacists, clinical practitioners, and international regulatory policymakers.

REFERENCES

1. Pansare K, Sonawane G, Tapadiya P, Tapadiya R, Patil C, Vaghela JS, et al. History, current status and future aspects of pharmacovigilance in India. Nat Volatiles Essent Oils. 2021;8 (6):1558–1565.
2. Potter E. FDA Adverse Event Reporting System (FAERS) Essentials. Pharmacotherapy / PMC.
3. Aronson JK. Artificial intelligence in pharmacovigilance: An introduction to terms, concepts, applications, and limitations.
4. Pilipiec P, et al. Using Machine Learning for Pharmacovigilance: A Systematic Review.
5. Tatonetti NP, et al. Artificial Intelligence for Drug Toxicity and Safety.
6. Murphy MF, et al. Adverse drug event detection using natural language processing: A scoping review of supervised learning methods .
7. Kompa B, Hakim JB, Palepu A, et al. Artificial Intelligence Based on Machine Learning in Pharmacovigilance: A Scoping Review. Drug Saf.

2022;45:477–491.Useful for: data sources, heterogeneous safety reports, AI applications, and signal detection. Read full paper — Open access

8. Desai MK. Artificial intelligence in pharmacovigilance – Opportunities and challenges. Perspect Clin Res. 2024;15 (3):116-121. Source: Pages 1–3, especially the introduction, definition of AI, and opportunities and benefits sections.
9. Nagar A, Gobburu J, Chakravarty A. Artificial intelligence in pharmacovigilance: advancing drug safety monitoring and regulatory integration. Ther Adv Drug Saf. 2025;16. Doi: 10.1177/20420986251361435
10. Kompa B, Hakim JB, Palepu A, et al. Artificial intelligence based on machine learning in pharmacovigilance: a scoping review. Drug Saf. 2022;45:477–491.
11. Dimitsaki S, Natives P, Jaulent MC. Applying AI to Structured Real-World Data for Pharmacovigilance Purposes: Scoping Review. J Med Internet Res. 2024;26:e57824.
12. Desai MK. Artificial intelligence in pharmacovigilance – Opportunities and challenges. Perspect Clin Res. 2024;15 (3):116–121.
13. Nagar A, Gobburu J, Chakravarty A. Artificial intelligence in pharmacovigilance: advancing drug safety monitoring and regulatory integration. Ther Adv Drug Saf. 2025;16.
14. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med. 2018;169 (7):467–473. Article.
15. Painter, J. L., et al. (2022). Artificial Intelligence Based on Machine Learning in Pharmacovigilance: A Scoping Review. Drug Safety, 45 (5), 477 –491.
16. Teleb, K., et al. (2026). Artificial Intelligence and Machine Learning in Pharmacovigilance: A Systematic Review of Models for Drug Safety in Polypharmacy. Preprints.org.
17. Shah P, et al. Artificial intelligence in pharmacovigilance – Opportunities and challenges. Perspect Clin Res. 2024. Supports: ICSR workflow, automated data entry, NLP extraction, MedDRA coding, duplicate checking, opportunities and limitations.Free full text: Read or download the full article.
18. Ball R, Dal Pan G. "Artificial Intelligence" for pharmacovigilance: Ready for prime time? Drug Saf. 2022;45 (5):429–438.
19. Murphy, A. Et al. (2023). Adverse drug event detection using natural language processing: A scoping review of supervised learning methods.
20. 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):1075–1089.
21. Warner, J., Prada Jardim, A., & Albera, C. (2025). Artificial Intelligence: Applications in Pharmacovigilance Signal Management. Pharmaceutical Medicine, 39, 183–198. Relevant for AI-assisted signal management, signal prioritization, challenges and human oversight.
22. Ibrahim H, Abdo A, El Kerdawy AM, Sharaf Eldin A. Signal detection in pharmacovigilance: A review of informatics-driven approaches for the discovery of drug-drug interaction signals in different data sources. Artif Intell Life Sci. 2021;1:100005.
23. Bate, A., & Hobbiger, S. F. (2021). Artificial intelligence, machine learning and deep learning in pharmacovigilance: Current status and future perspectives. Drug Safety, 44 (2), 145–162.

HOW TO CITE: Sachin Kumbhar, Dr. Sanjay Bais, Ashlesha Vasekar, Artificial Intelligence and Machine Learning in Pharmacovigilance: A Review of Applications, Challenges, and Future Perspectives, Int. J. of Pharm. Sci., 2026, Vol 4, Issue 11, 5691-5705. https://doi.org/10.5281/zenodo.24582833