JAAFR
INTERNATIONAL RESEARCH JOURNAL
JOURNAL OF ADVANCE AND FUTURE RESEARCH
JAAFR.ORG | ISSN : 2984-889X
An International Open Access, Peer-reviewed, Refereed Journal

Real-World Evidence and Signal Detection: Emerging Methodologies in Indian Pharmacovigilance Practice

Aditi Tripathi[1], Abhash Kumar[1], Deep Das[1]
1. Department of Pharmacovigilance, Ichelon Consulting Group, Gurugram, 122018

Abstract: Real-World Evidence (RWE) derived from Real-World Data (RWD) has gained global prominence as a critical complement to traditional clinical trial data for post-marketing safety surveillance. Regulatory agencies including the U.S. Food and Drug Administration (FDA), European Medicines Agency (EMA), and CIOMS emphasize the value of RWE in enhancing signal detection, improving risk assessment, and informing regulatory decisions. In pharmacovigilance (PV), RWE enables longitudinal follow-up, supports detection of rare and delayed adverse drug reactions (ADRs), and enhances safety monitoring in diverse, real-world populations.

India's rapidly evolving healthcare ecosystem with a vast heterogeneous population, mixed public-private systems, and increasing digitalization presents a strong opportunity for RWE-based PV. Initiatives such as the Pharmacovigilance Programme of India (PvPI) have strengthened spontaneous reporting, but integration of RWD sources such as electronic health records (EHRs), claims data, registries, and digital health platforms remains nascent. The absence of standardized data architectures, limited penetration of EHRs, fragmented registries, and variable data quality present significant barriers to large-scale active surveillance.

Simultaneously, emerging analytical methodologies, including machine learning (ML) and artificial intelligence (AI), are being increasingly applied to large RWD sources for early detection of safety signals and prediction of risk patterns. These innovations hold promise for application within the Indian regulatory context.

This review synthesizes global and Indian perspectives on RWE for signal detection, outlines current frameworks, evaluates available RWD sources, and examines emerging methods including AI/ML. It also discusses operational, regulatory, and ethical challenges in India and proposes a roadmap for strengthening RWE-based pharmacovigilance systems. By consolidating insights from open-access scientific literature and regulatory guidance, this review aims to support PV scientists, regulators, and healthcare policymakers in harnessing RWE to enhance drug safety and public health in India.

Keywords: Real-World Evidence; Real-World Data; Pharmacovigilance; Signal Detection; India; Electronic Health Records; Machine Learning; Post-Marketing Surveillance

List of Abbreviations

Abbreviation	Full form
RWD	Real-World Data
RWE	Real-World Evidence
PV	Pharmacovigilance
ADR	Adverse Drug Reaction
PvPI	Pharmacovigilance Programme of India
CDSCO	Central Drugs Standard Control Organization
EHR	Electronic Health Record(s)
PSUR	Periodic Safety Update Report
ICSR	Individual Case Safety Report
ML	Machine Learning
AI	Artificial Intelligence
EMA	European Medicines Agency
FDA	US Food and Drug Administration
CIOMS	Council for International Organizations of Medical Sciences

1. Introduction

Pharmacovigilance (PV) is defined as the science and activities relating to the detection, assessment, understanding, and prevention of adverse effects or other drug-related problems. While spontaneous reporting systems have historically formed the backbone of PV, they suffer from under-reporting, incomplete case information, and limited ability to quantify risk. Global healthcare digitization has led to the rapid accumulation of Real-World Data (RWD), which includes data from electronic health records (EHRs), medical claims, registries, and digital health sources. Analyses of RWD generate Real-World Evidence (RWE), which contributes essential insights for understanding a medicine's safety profile in routine clinical practice [1, 2]. Regulatory bodies such as FDA and EMA now explicitly recognize the role of RWE in post-marketing surveillance and safety decision-making. CIOMS also recommends structured integration of RWD and RWE into pharmacovigilance frameworks to complement spontaneous reporting [3, 4]. In India, RWE has significant potential to support PV due to the size and clinical diversity of the population. The Pharmacovigilance Programme of India (PvPI) has strengthened spontaneous reporting, but substantial gaps remain in leveraging RWD sources for proactive signal detection [5, 6]. Thus, there is a growing need to examine global methodologies, assess the readiness of Indian PV systems, and explore how emerging tools (including machine learning) can enhance detection of safety signals in India.

2. PRISMA-Style Methodolog

A structured literature search and evidence selection process was conducted using PRISMA principles.

2.1 Data Sources and Search Strategy: Open-access databases including PubMed Central (PMC), WHO, FDA, EMA, CIOMS, PvPI, and CDSCO websites were searched using combinations of the terms such as "real-world evidence," "real-world data," "pharmacovigilance," "signal detection," "India," "post-marketing surveillance," "RWE methodology." Searches included documents published between January 2000 and Nov 2025.

2.2 Inclusion and Exclusion Criteria: This review included freely accessible, full-text publications written in English that were directly relevant to real-world data (RWD), real-world evidence (RWE), pharmacovigilance, signal detection, and the Indian regulatory context**.** Eligible sources comprised peer-reviewed scientific literature as well as official documents issued by regulatory authorities or recognized organizations**.** Publications were excluded if they were paywalled and not available in full text**,** constituted non-methodological opinion or commentary articles**,** or focused exclusively on randomized controlled trial (RCT) efficacy outcomes without relevance to pharmacovigilance or post-marketing safety evaluation

2.3 Screening Process and Synthesis Approach: The study selection followed a two-stage screening process, in which titles and abstracts were initially reviewed, followed by a comprehensive full-text assessment of potentially eligible publications. To ensure broader coverage of relevant literature, snowballing techniques were applied to reference lists to identify additional open-access sources. Data synthesis was conducted using a narrative thematic approach, focusing on key domains including global real-world evidence (RWE) regulatory frameworks, the Indian pharmacovigilance ecosystem, available real-world data sources, RWE-based methodologies for signal detection, the application of artificial intelligence and machine learning, and identified challenges with future research and regulatory needs. Due to the heterogeneity in study designs and evidence types, a quantitative meta-analysis was not undertaken.

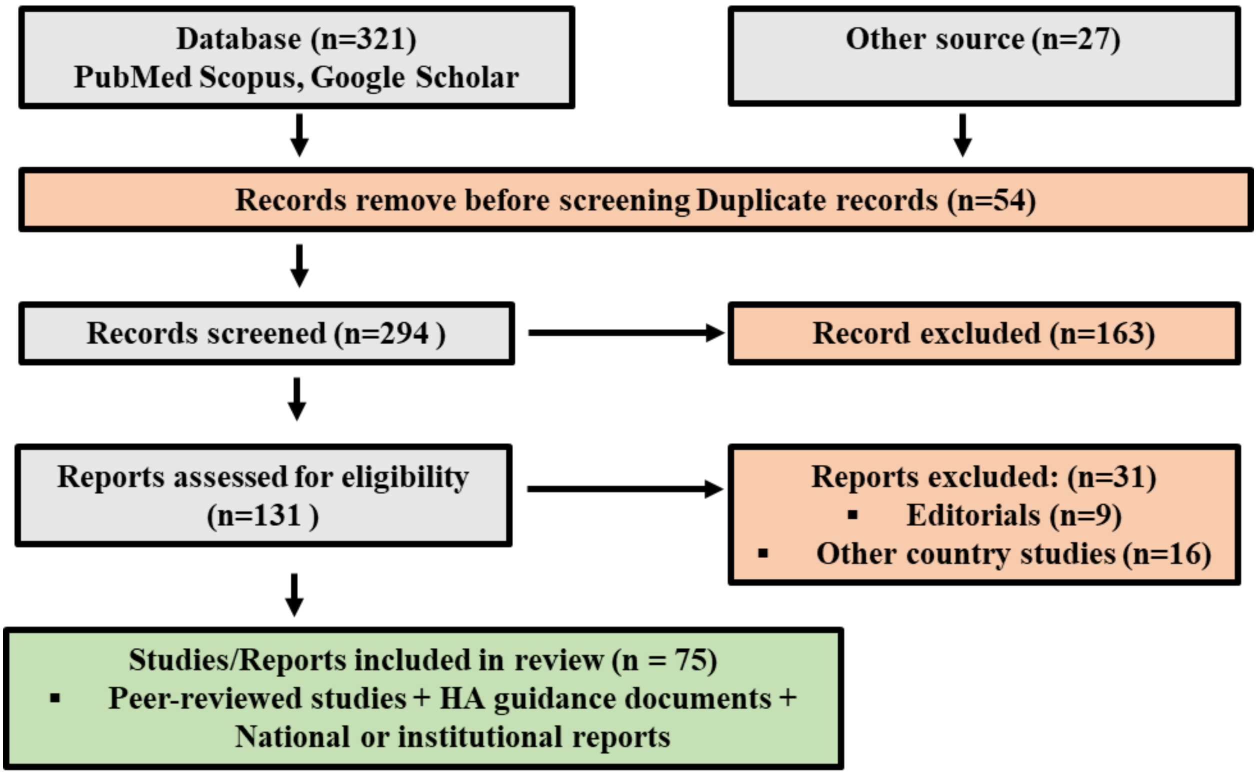

Figure 1: PRSIMA Methodology

3. Background: RWE in Global and Indian Context

3.1 Global Evolution of RWE:

RWE is defined as evidence regarding the usage and potential benefits or risks of a medical product derived from analysis of RWD. FDA's RWE framework emphasizes the role of RWE in safety surveillance, regulatory decisions, and monitoring of post-marketing commitments. EMA also outlines methodological and governance considerations for the use of RWE in regulatory assessments. CIOMS guidance similarly stresses data quality, transparency, and appropriate analytical methodology when using RWD for PV [1, 2]. RWD sources such as EHRs, administrative claims, registries, and patient-generated data can support active surveillance, estimate incidence rates, characterize high-risk populations, and detect rare or delayed adverse reactions [4, 5]. RWE complements traditional PV systems by providing denominator data and longitudinal follow-up not available in spontaneous reporting.

3.2 RWE and PV: Complementary Roles

Traditional spontaneous reporting systems are essential but limited by under-reporting and variable data completeness. RWE enhances signal detection when combined with spontaneous reporting by providing context for exposure, comorbidities, and time-anchored clinical events. Studies have shown that RWD combined with electronic monitoring can improve detection of drug-event associations [4, 5].

3.3 Indian Context for RWE Integration

India's healthcare structure comprising a highly heterogeneous population, diverse diseases, and mixed public-private health systems creates strong potential for RWE applications. Several Indian authors highlight the need for India-specific RWD infrastructures, including registries and digital health systems [6, 7]. The Pharmacovigilance Programme of India (PvPI), operational since 2010, coordinates spontaneous reporting through adverse drug reaction monitoring centres (AMCs) and contributes to global safety databases [8]. However, reliance on spontaneous reporting alone limits detection of rare ADRs or those occurring in under-represented regions. Digital health initiatives under Ayushman Bharat, such as the ABHA health ID and proposed national digital health ecosystem, may facilitate future RWD capture but are not yet integrated into PV workflows [9, 10]. Substantial gaps remain in EHR penetration, standard coding systems, and linkage across hospitals or states.

4. Current methodologies for signal detection (global frameworks)

Signal detection is a foundational component of pharmacovigilance and refers to the process of identifying new or incompletely documented adverse drug reactions from multiple data streams. The World Health Organization (WHO), Uppsala Monitoring Centre (UMC), FDA, EMA, and CIOMS have developed standardized methodologies for signal detection using spontaneous reports, formal statistical methods, active surveillance tools, and real-world data [11, 12, 13].

4.1 Traditional Spontaneous Reporting and Qualitative Assessment

Historically, signal detection relied heavily on manual clinical review of Individual Case Safety Reports (ICSRs). The WHO's VigiBase its global ICSRs database receives reports from more than 130 countries and supports qualitative clinical assessment [14]. Clinical judgement remains essential for causality assessment, pattern recognition, and identifying unusual event clusters [15]. However, dependence on spontaneous reporting alone is limited by under-reporting, reporting biases, and delayed identification of rare or long-latency events [13]. These limitations have led to the development of quantitative, automated, and data-driven approaches.

4.2 Quantitative Signal Detection Methods in Large Databases

UMC and the EMA introduced formal statistical methods to evaluate disproportionality in spontaneous reporting systems. These methods quantify whether a drug–event combination occurs disproportionately more often than expected relative to the overall database.

4.2.1 Proportional Reporting Ratio (PRR)

PRR compares the proportion of a specific ADR for a drug with the proportion of the same ADR for all other drugs in the database [16]. PRR thresholds (e.g., PRR ≥ 2 with ≥ 3 cases) are widely adopted by regulatory agencies for initial automated detection [16, 17].

4.2.2 Reporting Odds Ratio (ROR)

The ROR, used extensively in European systems, evaluates the odds of a drug–ADR pair versus all other pairs [17]. When the lower bound of the 95% confidence interval exceeds 1, a signal may be indicated.

4.2.3 Bayesian Confidence Propagation Neural Network (BCPNN)

UMC's BCPNN method applies Bayesian statistics to detect unexpected drug–event combinations. It identifies signals by estimating *Information Components (IC)*; an IC lower bound above zero suggests disproportionate reporting [14, 16]. BCPNN is a cornerstone of VigiBase signal detection.

4.2.4 Multi-Item Gamma Poisson Shrinker (MGPS)

FDA's MGPS algorithm, used in the FDA Adverse Event Reporting System (FAERS), calculates Empirical Bayes Geometric Mean (EBGM) scores to identify elevated reporting rates [15]. EBGM05 (lower 5% confidence bound) > 2 indicates a potential signal. Quantitative disproportionality methods have significantly enhanced early detection and consistency across global regulatory systems.

4.3 Active Surveillance and Quantitative Methods Using Healthcare Databases

Growing availability of large healthcare databases, especially in the U.S. and Europe, has expanded the scope of active surveillance.

4.3.1 FDA Sentinel System

The FDA Sentinel Initiative is a distributed data network enabling real-time, population-level safety assessments using EHRs, claims data, and registries [18]. Sentinel uses:

- Propensity-score based cohort studies
- Self-controlled case series
- Sequential monitoring (e.g., maximized sequential probability ratio tests)

These methods quantify incidence, risk changes over time, and causal associations more effectively than spontaneous reports.

4.3.2 EMA's EudraVigilance and DARWIN EU

EudraVigilance applies disproportionality analysis and automated screening for drug–event combinations [17]. DARWIN EU, launched by EMA, is a federated real-world data network to support regulatory decisions using large-scale RWD analytics [19].

4.3.3 WHO Global Surveillance and VigiLyze Tools and Structured Benefit–Risk Assessment Frameworks

WHO-UMC provides VigiLyze and VigiFlow, enabling global regulators to monitor ADR patterns, evaluate signal strength, and benchmark country-level data [14].

4.4 Structured Benefit–Risk Assessment Frameworks

Frameworks such as FDA's Benefit-Risk Framework and EMA's PRAC adopt structured models for integrating quantitative signal data with clinical judgement [20]. These frameworks ensure:

- Systematic evaluation of evidence
- Transparency of regulatory reasoning
- Integration of RWE with spontaneous report signals

4.5 Real-World Data and Hybrid Signal Detection Approaches

As RWE becomes integral to regulatory science, hybrid designs combining spontaneous reporting with active surveillance have evolved.

Examples include:

- Sequential signal monitoring in claims/EHR databases [18]
- Linkage of spontaneous reports with real-world exposure data to calculate incidence rates [5]
- Target trial emulation for evaluating drug safety outcomes [21]

Global regulators increasingly recognize that combining spontaneous reports, RWE, and advanced analytics offers the most comprehensive signal detection.

5. Emerging Indian Approaches (PvPI, CDSCO, Regional Centres)

5.1 Overview of India's National PV Architecture

India's national pharmacovigilance architecture has matured substantially since the formal launch of the Pharmacovigilance Programme of India (PvPI) in 2010. PvPI is coordinated at the National Coordination Centre (NCC) housed within the Indian Pharmacopoeia Commission (IPC) and works in collaboration with the Central Drugs Standard Control Organization (CDSCO) and a network of Adverse Drug Reaction Monitoring Centres (AMCs) distributed across tertiary hospitals, medical colleges, and other institutions [22, 23, 24]. The PvPI aims to strengthen spontaneous reporting, improve ADR signal detection, support regulatory actions, and promote awareness and training in PV across healthcare professions [22, 23]. PvPI's structure is designed around a hub-and-spoke model: the NCC at IPC provides national coordination, technical support, and data collation; regional and state-level centres conduct training, reporting, and local signal assessment; and AMCs collect individual case safety reports (ICSRs) which are submitted to the NCC and shared with international databases when appropriate [24, 25, 26]. PvPI also collaborates with disease-specific national programmes (e.g., tuberculosis control programmes) to capture therapy-specific safety information [24].

5.2 Reporting Mechanisms and Data Flow

The PvPI encourages ADR reporting from healthcare professionals, patients, and industry. Multiple reporting modalities exist: paper ADR reporting forms, electronic reporting forms via the PvPI/IPC portal, a dedicated toll-free helpline, and mobile/web reporting options [24, 25, 26, 27]. Standardized ICSR forms and coding practices are promoted to improve completeness and usability of reports [25] Notably, PvPI has worked to improve reporting completeness and timeliness key determinants of signal detection capacity [22, 23]. Collected ICSRs are quality-checked at AMCs and the NCC, entered into national databases, and periodically communicated to global pharmacovigilance repositories (e.g., WHO-UMC's VigiBase) as appropriate. PvPI employs case-level clinical review processes and contributes to national and regional signal assessment activities; however, methods for automated quantitative screening (e.g., disproportionality analyses within national databases) have historically been less mature than in established global systems [22, 23].

5.3 Regulatory Interfaces: CDSCO and Guidance for MAHs

CDSCO serves as the national regulatory authority for medicines and works closely with PvPI to translate safety intelligence into regulatory action when necessary [28]. CDSCO has issued guidance documents for Marketing Authorization Holders (MAHs) and stakeholders detailing pharmacovigilance obligations, Periodic Safety Update Report (PSUR) requirements, and processes for signal notification and risk-minimization activities [25, 28]. These regulatory guidance documents increasingly reference the need for robust PV systems within MAHs and clarify timelines and responsibilities for reporting serious and unexpected ADRs [25]. Recent CDSCO/IP C guidance explicitly outlines MAH responsibilities for establishing PV systems, conducting signal detection, and cooperating with national Pharmacovigilance authorities. The guidance also underscores the role of registries and real-world data in fulfilling post-marketing commitments, although operational specifics for integrating RWD into formal signal detection workflows are still evolving [25].

5.4 Expansion of ADR Monitoring Centres and Capacity Building

Since its inception, PvPI has steadily expanded the AMC network. The IPC publishes and updates lists of active AMCs, and enrollment procedures for new AMCs are open and documented [26]. The AMC network supports case collection, preliminary causality assessment, and local capacity building through training modules and workshops. Capacity-building initiatives aim to improve ADR recognition, standardized documentation, and timely reporting from across diverse healthcare settings [22, 23, 27]. Training programs both online and in-person have been implemented to educate clinicians, pharmacists, and allied health professionals about ADR reporting procedures, causality assessment, and signal interpretation. These initiatives are important for improving data quality and bridging regional disparities in reporting rates [27].

5.5 PvPI's Signal Detection and Collaboration with Global Systems

PvPI participates in international pharmacovigilance through data sharing and collaboration with WHO-UMC and adoption of global case-management tools [22, 24]. India's contribution to global ICSR databases has increased over time; published reports indicate progressive growth in the number and completeness of reports contributed to global repositories [22, 29]. Clinical review of aggregates and country-level signal reports is performed, and significant safety signals identified nationally can prompt CDSCO review and regulatory actions. However, compared to systems like FDA-FAERS or EMA-EudraVigilance, India's resources for automated, large-scale quantitative signal detection using nationwide linked RWD remain relatively nascent [18, 19, 22]. PvPI has prioritized improving case completeness and expanding the AMC network as precursors to more advanced, database-driven surveillance capabilities.

5.6 Innovations and Recent Policy Moves

There are nascent policy and operational innovations aimed at broadening RWD capture and public engagement in ADR reporting. For example, digital facilitation of reporting (mobile apps, web portals) and public awareness campaigns have been introduced to increase voluntary reporting [24, 27]. More recently, CDSCO-led initiatives to display QR codes and toll-free reporting information at pharmacies have been publicized as measures to make ADR reporting more accessible to the public, indicating a policy-level push toward wider engagement [30, 31, 32]. Such measures, if implemented at scale, could increase reporting rates and broaden the types of events captured (e.g., consumer-reported side effects).

5.7 Limitations and Areas for Development

Key gaps remain for fully operationalizing RWE-driven signal detection in India:

1. **Data fragmentation and limited EHR penetration:** Many hospitals and clinics still lack interoperable EHR systems with standardized coding, limiting the availability of structured RWD for cohort or database studies [9, 22].
2. **Heterogeneous data quality across AMCs and regions:** Variability in report completeness and diagnostic detail constrains the use of ICSRs for robust signal quantification [22, 23].
3. **Limited nationwide linkage:** Lack of nationally linked claims or EHR data impedes calculation of denominators and incidence rates necessary for epidemiologic signal evaluation [18, 19, 22].
4. **Analytic capacity and automation:** While PvPI performs clinical review and manual signal assessment, large-scale automated disproportionality screening and advanced analytics (including ML) are still emerging areas [22, 27].

Addressing these gaps through digitization, standardization of data capture, expansion of registries, and investment in analytic infrastructure would enable PvPI and CDSCO to evolve from predominantly spontaneous-reporting systems to hybrid surveillance models that integrate RWE for proactive signal detection and regulatory decision-making.

Table 1: Overview of Emerging Indian Pharmacovigilance Framework and RWE Integration

Domain	Key Components	Description
National PV Architecture	PvPI, NCC (IPC), CDSCO, AMCs	India’s pharmacovigilance system operates through a hub-and-spoke model coordinated by the National Coordination Centre at IPC, supported by CDSCO and a nationwide network of ADR Monitoring Centres.
ADR Reporting Sources	HCPs, patients, industry	Adverse drug reactions are reported by healthcare professionals, patients, and marketing authorization holders, enabling broad post-marketing safety surveillance.
Reporting Modalities	Paper forms, online portal, mobile apps, toll-free helpline	Multiple reporting channels facilitate accessibility and improve ADR reporting volume and timeliness across healthcare settings.
Data Collection & Processing	ICSRs, standardized coding	Individual case safety reports are collected at AMCs, quality-checked, coded using standardized practices, and transmitted to the NCC for national collation.
Signal Detection Approach	Clinical review, aggregate assessment	Signal detection primarily relies on expert clinical review of individual and aggregated reports, with limited routine use of automated quantitative methods.
Regulatory Interface	CDSCO–PvPI collaboration	Safety signals identified through PvPI are reviewed by CDSCO to inform regulatory decisions, risk minimization measures, and post-marketing actions.
Guidance for MAHs	PV systems, PSURs, signal reporting	CDSCO guidance outlines pharmacovigilance obligations for MAHs, including PSUR submission, signal detection, and cooperation with national PV authorities.

Domain	Key Components	Description
Capacity Building	AMC expansion, training programs	Ongoing expansion of AMCs and structured training programs aim to improve ADR recognition, report quality, and regional reporting equity.
Global Collaboration	WHO-UMC, VigiBase	PvPI contributes Indian safety data to global pharmacovigilance systems, supporting international signal detection and benefit–risk evaluation.
Policy Innovations	Digital tools, QR-based reporting	Recent initiatives promote digital reporting and public engagement to enhance voluntary ADR reporting and data capture.
Current Limitations	Data fragmentation, limited linkage	Challenges include fragmented health data systems, variable report quality, lack of national EHR/claims linkage, and limited advanced analytics.
Future Direction	Hybrid RWE-driven surveillance	Strengthening digitization, standardization, registries, and analytic capacity is essential for integrating RWE into proactive signal detection and regulatory decision-making.

6. Use of RWD Sources: EHRs, Claims Data, Registries, Digital Health, Patient- Reported Outcomes

Real-World Data (RWD) for pharmacovigilance comprises multiple complementary sources electronic health records (EHRs), administrative/claims databases, disease registries, digital health platforms, and patient-reported outcomes (PROs). Each source has unique strengths and limitations for signal detection and safety assessment; leveraging them in combination (hybrid approaches) often yields the most robust surveillance [33, 34, 6, 35].

6.1 Electronic Health Records (EHRs)

Strengths: EHRs capture longitudinal clinical information (diagnoses, laboratory results, prescriptions, procedures) at the point of care, enabling temporal linkage between drug exposures and clinical events and providing rich covariate data needed for confounding adjustment [33, 10]. EHR data can support cohort and self-controlled study designs, incident-rate estimation, and detection of complex or delayed adverse events that may not be reported through spontaneous systems [33, 36].

Applications in signal detection: Methodological work has shown how EHRs enable automated screening for signals using structured data and natural language processing (NLP) of free-text clinical notes to identify potential ADRs not captured by diagnostic codes alone [34]. Examples include EHR-based identification of drug-related acute kidney injury, liver enzyme elevations, and bleeding events using combinations of lab trends and clinical notes [34].

Indian experience and readiness: India's EHR adoption has historically been heterogeneous. Early roadmap analyses identified infrastructural, interoperability, and policy barriers limiting EHR scale-up [35]. More recently, national digital health initiatives such as the Ayushman Bharat Digital Mission (ABDM) and the Ayushman Bharat Health Account (ABHA) have created foundational components (unique health identifiers, health data consent frameworks, and digital infrastructure) that could facilitate standardized EHR aggregation and PV applications if widely implemented and interoperable [37, 38]. Recent studies indicate growing public uptake of digital health IDs and mixed but improving provider adoption of EHR systems [39]. Nevertheless, current EHR coverage and standardization vary by region and facility type, limiting immediate nationwide EHR-based surveillance [35, 39].

Challenges: Typical challenges include incomplete data capture (missing outpatient prescriptions), inconsistent use of standardized terminologies (e.g., ICD, SNOMED), fragmented provider networks, and variable data quality all of which complicate case ascertainment and risk estimation [33, 35].

6.2 Administrative Claims and Insurance Databases

Strengths: Claims databases offer large, population-level coverage of billed diagnoses, procedures, and pharmacy dispensing records, making them useful for incidence estimation and longitudinal follow-up in settings with high insurance penetration [36]. They commonly include dates of service and drug dispensing information that support exposure definition and outcome ascertainment.

Indian context: India's claims databases are currently limited relative to high-income countries because a large fraction of healthcare in India is paid out-of-pocket and fragmented across public and private sectors. Corporate and insurer databases (e.g., empanelled private hospital insurance claims) exist but are not nationally representative [8, 35]. Therefore, while claims sources can be valuable for studies within insured subpopulations (e.g., corporate employees, government insurance schemes), they cannot yet serve as a comprehensive national PV denominator [8].

6.3 Disease and Therapeutic Registries

Strengths: Disease-specific registries (e.g., cancer registries, cardiovascular registries) collect high-quality, longitudinal clinical data and are particularly useful for monitoring safety in specific patient populations or therapies [40]. Registries often capture detailed staging, outcomes, and treatment patterns that support outcome validation and nested pharmacoepidemiologic studies.

Indian registries: India has longstanding registries in selected domains. The ICMR-NCDIR National Cancer Registry Programme (NCRP) operates population-based and hospital-based cancer registries that provide standardized incidence and outcome data across multiple centers [41, 42]. These registries are valuable for oncology drug safety monitoring but are limited to cancer populations and to the geographic coverage of participating registries. Other registries (e.g., tuberculosis programme registries, HIV registries) exist under national disease control programmes and can be leveraged for therapy-specific safety surveillance where linkage and access are permitted [43].

Limitations: Many registries are condition-specific, lack uniform data dictionaries across regions, and may not systematically capture concomitant medications or detailed safety outcomes unless designed for pharmacovigilance.

6.4 Digital Health Platforms, Mobile Apps and Patient-Reported Outcomes (PROs)

Strengths: Mobile health apps, patient portals, and dedicated ADR reporting apps capture patient-reported outcomes and consumer reports in near real-time. PROs can detect symptomatic ADRs (e.g., nausea, fatigue) that are under-represented in clinical records and can improve completeness of safety profiles [44].

Indian initiatives: PvPI's mobile app ("ADR PvPI") and web reporting portals facilitate direct reporting by healthcare professionals and patients; analyses have documented reporting features, uptake patterns, and areas for UX improvement [45]. The national digital health ecosystem (ABDM/ABHA) also provides citizen-facing tools (e.g., ABHA accounts and health record viewers) that could be integrated to enable patient-reported safety data capture and consent-based sharing with PV authorities [37, 38]. Recent local innovations (QR codes at pharmacies, hospital QR systems) aim to simplify patient access to ADR reporting and broaden consumer participation [31, 32].

Challenges: PROs pose validation challenges (subjectivity, potential for duplicates), and integrating patient reports with clinical EHRs or ICSRs requires robust de-duplication, standardization, and data governance mechanisms.

6.5 Hybrid and Linkage Approaches

Creating robust RWE for PV often requires linkage across data sources (EHR ↔ registries ↔ claims ↔ PROs) to combine numerator event data with population denominators and contextual clinical information [18, 21]. India's evolving digital identity frameworks (ABHA) and increasing adoption of hospital information systems provide a plausible technical route for secure, consented linkages. However, legal, privacy, and interoperability frameworks must be operationalized to realize such linkages at scale [37, 39].

7. Artificial Intelligence and Machine Learning in Signal Detection

Artificial intelligence (AI) and machine learning (ML) are increasingly applied across the pharmacovigilance (PV) lifecycle to improve efficiency, scale, and sensitivity of signal detection from large, heterogeneous real-world data (RWD) sources. AI/ML methods span rule-based automation, natural language processing (NLP) for unstructured text, supervised learning for classification/prediction, and unsupervised or anomaly-detection methods for hypothesis-free screening [46, 47, 48, 49, 50]. This section summarizes applications, representative methods, performance considerations, regulatory perspectives, and key limitations relevant to implementation in India.

7.1 Primary AI/ML Use-Cases in Pharmacovigilance

7.1.1 Case intake automation, triage and coding: AI systems (often combining NLP and classification models) can automatically extract key fields from spontaneous reports and free-text clinical narratives e.g., suspected drug(s), adverse events (mapped to MedDRA terms), seriousness, and temporal relationships reducing manual effort and turnaround time for case processing [46, 50]. Industry surveys and reviews report broad adoption of automation tools for case intake and narrative coding among large MAHs and contract organizations [10].

7.1.2 Duplicate detection and case linkage: Duplicate ICSRs and overlapping reports create noise for signal detection. ML-based record linkage and fuzzy-matching algorithms improve de-duplication by identifying similar narratives and matching demographic/temporal attributes beyond exact string matches [50, 51].

7.1.3 NLP for ADE/ADR extraction from unstructured text: NLP techniques from rule-based lexicons and pattern matching to modern deep-learning models extract adverse event mentions from clinical notes, discharge summaries, social media, and patient forums. Early systems (e.g., MedLEE) demonstrated feasibility for extracting adverse events from discharge summaries [52]. Recent systematic reviews show substantial progress: transformer-based models and task-specific architectures now achieve high accuracy on benchmark ADE extraction tasks, enabling EHR-driven signal detection pipelines [47, 53].

7.1.4 Supervised prediction models for ADR risk: Supervised ML models trained on labeled datasets have been used to predict risk of specific ADRs (e.g., predicting drug-induced liver injury or bleeding) using structured covariates and derived features [46]. These models support risk stratification and targeted monitoring but require high-quality labeled outcome data and careful confounding control.

7.1.5 Unsupervised and anomaly detection for hypothesis-free signal discovery: Unsupervised ML methods (clustering, outlier detection, autoencoders) can flag unusual patterns in high-dimensional RWD that merit clinical review. Such approaches may complement disproportionality analyses by identifying novel, unexpected drug-event associations in multi-modal datasets [46, 48].

7.1.6 Social listening and patient-generated data mining: AI/NLP pipelines applied to social media, forums, and consumer reviews can surface patient-reported symptoms and experiences that may signal emerging safety concerns, although data quality and representativeness issues require careful interpretation [47].

7.2 Representative Methods and Technological Evolution

AI/ML in PV has progressed from early rule-based NLP to statistical ML and, more recently, deep learning and transformer architectures (e.g., BERT variants) that excel at extracting context and relations in clinical text [52, 53]. Techniques commonly applied include:

- **Named Entity Recognition (NER)** to locate drug and event mentions.
- **Relation Extraction** to link drugs with events and temporal markers.
- **Sequence tagging and token classification** for identifying event spans. [53]
- **Supervised classifiers (random forests, gradient boosting, neural networks)** for event prediction or seriousness classification. [46]
- **Graph-based and probabilistic models (Bayesian networks)** for causal modeling and integrating heterogeneous evidence. [46]

Recent reviews suggest transformer models fine-tuned on ADE tasks outperform traditional approaches when sufficient annotated data is available; however, resource and data constraints remain an implementation challenge outside large academic/industry centers [47, 53].

7.3 Performance, Evaluation and Validation

Robust evaluation is essential before deploying AI/ML systems in regulatory PV. Key practices include:

- **Use of gold-standard annotated datasets** for training and validation. Where possible, benchmark datasets (e.g., from shared tasks) permit cross-study comparison [53].
- **Reporting of performance metrics** (precision, recall/sensitivity, F1 score) tailored to PV priorities e.g., high sensitivity may be prioritized for initial screening, with human triage controlling specificity [47].
- **External validation and prospective monitoring** to evaluate model generalizability across sites, populations, and data capture systems [49].

Regulatory guidance underscores the need for lifecycle performance monitoring, documentation of training sets, and clear specifications for intended use topics addressed in FDA discussion papers and WHO guidance on AI for health [49, 54].

7.4 Regulatory and Ethical Considerations

Regulatory authorities have signaled active engagement with AI/ML in healthcare. The FDA's discussion paper and subsequent action plan propose a total-product-lifecycle approach for AI/ML-based software (SaMD), emphasizing change-control plans, real-world performance monitoring, and transparency [49, 55]. WHO guidance on ethics and governance highlights principles such as human-centred design, transparency, equity, and accountability for AI in health all relevant to PV implementations that may affect public safety [54].

Regulatory expectations relevant to PV implementations include:

- **Traceability of model development** (data provenance, labeling procedures). [49]
- **Explainability and interpretability** to support regulatory review and clinician trust, especially when models influence decisions about safety signals or regulatory actions. [54]
- **Bias and fairness assessment** to ensure models do not amplify disparities (e.g., under-representation of specific demographic groups). [54]
- **Post-deployment monitoring** to detect model drift and ensure continued performance in changing data environments. [49, 55]

7.5 Practical Limitations and Challenges

Despite promise, AI/ML adoption for signal detection faces important constraints:

- **Data quality and representativeness:** RWD in many settings (including parts of India) can be incomplete, inconsistently coded, or biased by healthcare access patterns; ML models trained on such data may underperform or mislead [33, 35, 54].
- **Scarcity of labelled training data:** Supervised approaches require annotated corpora (drug-event labels), which are costly to create and may not exist for many Indian contexts [53].
- **Explainability and regulatory acceptance:** Deep learning models offer high performance but limited interpretability; regulators and clinical reviewers often require explainable outputs for safety decisions [54, 55].
- **Privacy and governance:** Using sensitive health data for model training/monitoring requires robust consent, de-identification, and legal frameworks areas still evolving in India [37, 54].
- **Infrastructure and skills gaps:** Implementing and maintaining ML systems demands computational resources and specialized personnel (data scientists, ML engineers), which may be limited outside major centers [10].

7.6 Opportunities for India

India can pragmatically adopt AI/ML in PV by focusing on achievable, high-value use cases:

1. Automation of routine case processing (intake, triage, MedDRA coding) to free clinician time and improve timeliness of ICSRs [10].
2. NLP pipelines for ADR extraction within large hospital EHRs or registries where available, starting with high-yield conditions (e.g., drug-induced liver injury in tertiary hepatology centres) [52, 53].
3. Hybrid models combining disproportionality outputs with ML-based features (e.g., temporal patterns, lab signals) to prioritize signals for clinical review [63].
4. Capacity building and shared annotated corpora: national initiatives could fund creation of Indian-language and Indian-context annotated datasets to enable locally robust models. [53, 54]
5. Regulatory sandboxing and pilot projects under PvPI/CDSCO to evaluate performance and governance models before national scale-up [49, 55].

Table 2: Role of AI/ML across Signal Detection in Pharmacovigilance

Signal Detection Stage	AI/ML Methods	Purpose in PV
Case Intake & Coding	NLP, classifiers, rule-based automation	Automated extraction, triage, MedDRA coding, and reduced manual case-processing burden.
Duplicate Management	ML-based record linkage, fuzzy matching	Identification and removal of duplicate or overlapping ICSRs to reduce analytical noise.
Signal Screening	Supervised ML, unsupervised clustering, anomaly detection	Detection of potential drug–event associations beyond traditional disproportionality analyses.
Signal Prioritization	Hybrid statistical–ML models, risk prediction	Ranking and stratification of signals by clinical relevance, severity, and temporal patterns.
Clinical Review	Explainable AI, decision-support models	Supports expert assessment while maintaining interpretability and regulatory trust.
Regulatory Oversight	Lifecycle monitoring, drift detection	Ensures sustained model performance, governance, and alignment with regulatory expectations.
Key Constraints	Data quality, limited labels, explainability, privacy	Challenges affecting reliability, scalability, and regulatory acceptance, particularly in LMIC settings.
Indian Opportunities	Automation, NLP in EHRs, pilot sandboxes	Pragmatic adoption through high-impact use cases under PvPI/CDSCO governance.

8. Case Examples and Landmark Studies from India

This section presents representative Indian examples where pharmacovigilance (PV) practice, RWE collection, or signal detection activities have produced actionable findings or illustrate important lessons for RWE-based PV in India. Examples span national PvPI reporting and safety alerts, hospital-based observational studies that function as RWD sources, evaluations of consumer-facing tools (ADR cards, mobile app), vaccine safety surveillance, and registry-linked opportunities (oncology registries).

8.1 PvPI national reporting and drug-safety alerts

The Pharmacovigilance Programme of India (PvPI) has systematically collated Individual Case Safety Reports (ICSRs) and issued drug-safety alerts and advisories based on national signal assessments. PvPI annual and performance reports summarise reporting trends, signal investigations, and safety communications to healthcare professionals and the public [56]. Analyses of PvPI-driven safety communications (e.g., national safety alerts issued over defined periods) demonstrate the programme's role in identifying product- or class-specific safety concerns and in raising awareness across AMCs and clinicians [57, 58]. These national summaries function as prime examples of how a strengthened spontaneous reporting network can generate national safety intelligence and initiate regulatory or operational follow-up.

Key lessons: national coordination (PvPI/NCC-IPC), standardized ICSR workflows, and active dissemination mechanisms enable country-level signal identification even before large-scale RWD systems are mature [56, 57, 58].

8.2 Hospital-based prospective and surveillance studies as RWD

Tertiary-care hospital studies in India have long provided valuable RWD for PV. Prospective observational studies documenting ADR incidence, risk factors, seriousness, and system-organ class distributions serve as micro-surveillance datasets that inform local and sometimes national safety priorities. For example, a prospective study at a private tertiary care teaching hospital documented ADR prevalence and patterns, providing denominators for incidence estimates and highlighting commonly implicated drug classes' data that can feed into regional signal detection priorities and targeted investigations [59]. Similarly, more recent multi-centre analyses of serious ADRs from hospital databases have characterized age, gender, and organ-system patterns that help prioritize safety monitoring [60].

Key lessons: well-designed hospital-based surveillance (prospective or electronic extraction) can act as high-quality RWD for early detection of clinically important ADRs and for validating signals emerging from spontaneous reports [59, 60].

8.3 Consumer-facing tools and programmatic evaluations: ADR alert cards and mobile reporting apps

India has experimented with consumer-oriented interventions to enhance reporting and patient engagement. Evaluation studies have examined the real-world effectiveness of ADR alert cards (patient-held cards indicating suspected ADRs and reporting information) and the PvPI mobile app for ADR reporting. An evaluation of the ADR Alert Card initiative reported measurable interactions between patients and healthcare providers and suggested the cards improved communication about ADRs and facilitated reporting to PvPI [61]. The PvPI mobile app (Android) has been described in open literature as a mechanism to facilitate direct patient and HCP reporting; early assessments highlight feasibility and uptake while noting opportunities for usability and awareness improvements [44].

Key lessons: patient-centric tools increase reach of PV systems and broaden RWD capture (e.g., consumer-reported events), but require parallel investments in awareness, de-duplication workflows, and linkage to clinical records for event validation [61, 44].

8.4 Vaccine safety monitoring HPV vaccine post-licensure and local studies

Vaccines provide a high-public-health priority area for PV and RWE. In India, HPV vaccine introduction and post-licensure monitoring have prompted local safety studies and active surveillance projects. Open-access clinical and post-licensure reports from India have documented tolerability and absence of serious vaccine-related safety signals in hospital-based and trial settings [62, 63]. These Indian studies align with global RWE for HPV vaccines showing favourable safety profiles, and they illustrate how local RWD collection (trial follow-ups, hospital surveillance) can reassure policymakers and the public about vaccine safety.

Key lessons: vaccine RWE in India has been generated through targeted post-licensure studies and clinical trials; scaling such surveillance to national immunization programmes requires integrated systems for near-real-time RWD capture and rapid signal assessment [62, 63].

8.5 Analysis of serious ADRs from multi-centre datasets

Recent open-access analyses using larger hospital or multi-centre datasets have characterized serious ADRs in Indian clinical settings, identifying common culprit drug classes, age distributions, and outcomes. These studies show that aggregated hospital data can be systematically analysed to identify epidemiologic patterns in serious ADRs and to prioritize medicines/classes for regulatory review and targeted pharmaco-epidemiologic studies [60]. Such analyses represent intermediate RWE products not necessarily nation-wide but important for hypothesis generation and for designing registry or cohort studies.

Key lessons: systematic aggregation and analysis of hospital data can approximate denominator-based incidence estimates and provide context to spontaneous report signals [60].

8.6 Registry potential oncology and disease-specific registries as safety platforms

India's disease registries (e.g., National Cancer Registry Programme) are well-established in specific domains and demonstrate the potential of registry data to support safety monitoring for high-priority therapeutic areas [64, 65]. Cancer registries with tumor staging, treatment details, and outcomes can serve as foundations for oncology drug safety surveillance (e.g., monitoring adverse events associated with newer targeted therapies), provided relevant treatment and ADR fields are augmented. Similarly, disease control programmes (e.g., TB, HIV) maintain treatment registries that have been used for pharmacovigilance in programmatic contexts.

Key lessons: disease-specific registries represent ready RWD platforms for focused PV; targeted augmentation (treatment exposure fields, ADR modules) and linkages to hospital EHRs can convert registries into powerful tools for active safety surveillance [64, 65].

9. **Comparative analysis with EU/US RWE frameworks**

Real-World Evidence (RWE) has been adopted into regulatory practice in both the United States and the European Union, but the two jurisdictions differ in infrastructure, operational models, governance approaches, and the current breadth of routine regulatory use. Understanding these differences helps frame realistic, phased strategies for India as it builds national capability.

9.1 High-level regulatory stance and policy

Both the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA) explicitly endorse RWE for certain regulatory purposes, including safety surveillance, label expansions, and post-marketing commitments [1, 2]. FDA's RWE Program and related guidance documents provide operational considerations for using RWD to generate RWE and describe study types and evidentiary standards for regulatory use [66, 1]. EMA has developed an RWE framework and pilots (including DARWIN-EU) that focus on enabling access to federated European data for regulatory evidence generation [2, 1].

A key difference is emphasis and maturity of operational programs. The FDA has operationalized RWE primarily through two pillars: (a) the **Sentinel** system for active safety surveillance (a distributed, claims/EHR network originally evolved from Mini-Sentinel), and (b) programmatic guidance and pilot projects to evaluate RWE for regulatory submissions [18, 66]. The EMA has emphasized creation of the **DARWIN-EU** network as a central service to deliver RWE to support regulatory assessments across EU member states and has run pilots to assess feasibility and methods [19, 1, 2]. Whereas FDA's Sentinel has been operational for over a decade with mature methods for sequential monitoring and distributed analytics, DARWIN-EU is a newer pan-European coordination effort designed to scale federated studies across many national datasets while navigating diverse national rules on data access and privacy [18, 19, 1].

9.2 Data architecture: distributed / federated networks vs centralized repositories

Both jurisdictions use distributed (federated) approaches to respect data custodianship and privacy, but implementation details vary:

- **FDA / Sentinel**: operates predominantly as a distributed network where data partners transform their local data to a common data model (Sentinel Common Data Model), enabling centrally authored analytic programs to run locally and return aggregated results. Sentinel's long experience emphasized governance procedures, common data model harmonization, and rapid-cycle surveillance methods [18, 12].
- **EMA / DARWIN-EU**: similarly uses a federated model where participating data partners standardize to a common data model (often OMOP/other standards) and respond to research requests coordinated

by a central coordination centre [1, 19]. DARWIN-EU's design accommodates the EU's multi-jurisdictional legal environment and diverse healthcare systems; scaling requires orchestration across national data access procedures and ethical approvals [19, 1, 2].

The federated model adopted by both systems reflects a trade-off: it preserves local control and privacy but requires extensive investment in data standardization, transformation, and governance at each data partner site. In the EU, heterogeneity across member states (coding systems, healthcare delivery models) makes this task more complex than in the U.S., where large national insurers and integrated delivery networks provide large, relatively consistent claims/EHR sources [18, 19, 12].

9.3 Methodological toolkits and analytic approaches

Methodologically, both regions combine disproportionality screening, targeted pharmaco-epidemiologic designs (cohort studies, self-controlled case series), and sequential/near-real-time monitoring. FDA Sentinel has pioneered sequential monitoring and rapid-cycle analysis methods (e.g., maximized sequential probability ratio testing) and operational workflows for signal evaluation [18, 12]. EMA pilots have emphasized multi-pathway evidence generation (regulatory, safety, HTA use cases) and method validation across different data partners [15, 1]. Both agencies advocate transparent study protocols, pre-specified analysis plans, sensitivity analyses, and where possible, replication in multiple data sources to enhance robustness [66, 15]. The EU's DARWIN pilots include demonstration projects to evaluate methods in multi-country contexts and to identify when single-country data suffice versus when pooled/federated analyses are required [19, 1].

9.4 Evidence standards and regulatory acceptance

FDA's guidance documents articulate the evidentiary considerations for accepting RWE to support regulatory decisions, emphasizing data fitness for purpose, study design validity, and transparent reporting [66, 16]. The FDA has accepted RWE in several regulatory decisions (e.g., label changes, post-marketing commitments), often when the RWD source allowed credible exposure and outcome measurement and when studies were conducted with rigorous epidemiologic designs [1, 5]. EMA's RWE experience is growing through structured pilots and case studies; the agency emphasizes not only methods but also governance and multi-stakeholder trust-building [15, 1]. DARWIN-EU aims to provide regulatory-grade evidence supporting assessments across member states, but harmonizing national legal frameworks remains a practical challenge. EMA's RWE framework report highlights that while RWE has supported regulatory assessments, experience to date has been limited and more pilots and transparency are needed [15].

9.5 Governance, privacy and legal frameworks

A major differentiator is the legal and data privacy landscape. The EU's General Data Protection Regulation (GDPR) imposes stringent, harmonized privacy rules that affect cross-border RWD access and require careful legal frameworks for each study [1]. DARWIN-EU designs account for these constraints via federated analyses and careful governance. In the U.S., privacy laws are sectoral (HIPAA) and the Sentinel model relies on business-associate arrangements and de-identification or limited datasets negotiated with partners; nevertheless, privacy concerns remain operational considerations [18, 12].

For India, this comparative lesson is salient: cross-institutional linkage and federated analyses must be architected within robust national data-protection and consent frameworks [37, 54]. The EU and U.S. experience shows that governance and trust frameworks are as important as technical solutions.

9.6 Operational scale, resourcing, and human capital

Sentinel's multi-year maturation demonstrates the resource intensity required site onboarding, data transformation, analytic tool development, and governance structures require sustained funding and skilled data scientists and epidemiologists [18, 12]. EMA's DARWIN-EU likewise requires a coordination centre, trained methodologists, and partner engagement to scale. India must plan for analogous investments: building common data models (e.g., OMOP), funding coordination centres, training pharmacoepidemiologists, and piloting use cases that deliver quick regulatory value (e.g., vaccine safety, high-impact therapeutic classes) [19, 43, 66].

9.7 Lessons for India from EU/US models

1. Start with high-value, feasible pilots use sentinel-style near-real-time monitoring for widely used drugs or national immunization programmes to demonstrate value [1, 18].
2. Adopt a federated model while investing in national standards federated analysis preserves custodianship but depends on common data models and local readiness; India's ABHA and digital health standards can support this evolution [19, 37].
3. Prioritize data fitness and transparency regulatory acceptance hinges on documented data provenance, pre-specified protocols, and replication across sources [15, 66].
4. Invest in governance and legal frameworks early privacy, consent, and data-sharing agreements must be clarified to avoid delays akin to EU cross-border legal negotiations [1].
5. Build human and computational capacity gradually sustained funding for coordination centres, tool libraries, and training will accelerate credible RWE generation [18, 19].

Table 3: Comparison of RWE Frameworks for Pharmacovigilance; US, EU, and India

Policy Dimension	United States (FDA)	European Union (EMA)	India – Current Status & Implications
Regulatory Position on RWE	RWE formally embedded in regulatory decision-making with published guidance and accepted use cases.	RWE formally endorsed through framework documents and coordinated regulatory pilots.	RWE acknowledged but not yet routinely embedded; requires formal policy articulation and guidance.
Core RWE Infrastructure	Sentinel System: mature, distributed network supporting active, near-real-time safety surveillance.	DARWIN-EU: federated EU-wide network delivering regulatory-grade evidence; still scaling.	PvPI primarily relies on spontaneous reporting; national RWE infrastructure remains undeveloped.
Data Architecture Model	Distributed analytics using a common data model with local execution and aggregated outputs.	Federated analyses across national datasets using common data models under central coordination.	Fragmented data landscape; limited interoperability and cross-institutional linkage.
Analytic Capability	Advanced pharmacoepidemiology, sequential monitoring, and validated rapid-cycle analyses.	Comparable analytic toolkit with emphasis on method validation across countries.	Predominantly clinical signal review; quantitative RWE analytics are emerging.
Evidence Acceptance Standards	Clear evidentiary criteria (data fitness, design rigor, transparency); multiple regulatory precedents.	Growing experience through pilots; strong	Evidentiary standards for RWE not clearly defined;

		emphasis on governance and transparency.	regulatory acceptance remains cautious.
Governance & Privacy	Sectoral privacy laws; governance via contractual and technical safeguards.	GDPR-driven, stringent cross-border data protection and study-specific governance.	National data-protection and consent frameworks still evolving; early clarification essential.
Operational Resources	Long-term funding and dedicated epidemiology and data science expertise.	Central coordination centre with skilled methodologists and partner networks.	Limited specialized workforce; capacity building and sustained investment needed.
Key Strategic Lesson for India	Demonstrates feasibility of phased, high-value pilots with sustained investment.	Highlights importance of governance, legal clarity, and federated coordination.	Prioritize sentinel-style pilots, common data standards, early governance, and gradual scaling.

10. Operational, Regulatory, and Ethical Challenges in India

Implementation of Real-World Evidence (RWE)-driven pharmacovigilance (PV) in India faces interrelated operational, regulatory, legal, and ethical challenges. These challenges span data quality and interoperability, governance and legal frameworks, workforce and analytic capacity, public trust and consent, and equity considerations. Below we synthesize these issues and reference illustrative sources and guidance.

10.1 Data Quality, Fragmentation and Interoperability

A foundational operational barrier is heterogeneous data capture across the Indian health system. Many care settings still rely on paper records or locally developed hospital information systems with inconsistent coding practices; even where Electronic Health Records (EHRs) exist, variation in structure, use of terminologies (ICD, SNOMED, RxNorm/Anatomical Therapeutic Chemical codes), and missing data (outpatient prescriptions, over-the-counter medicines) limit fitness-for-purpose of EHRs as PV RWD [33, 67, 68]. Claims databases are fragmented because a substantial proportion of care is out-of-pocket and not captured in insurer datasets, limiting denominator estimation at national scale [36, 8]. Interoperability is a critical technical and governance challenge. India's digital health push (Ayushman Bharat Digital Mission / ABHA) provides promising building blocks unique health identifiers and APIs but large-scale interoperability requires wide provider uptake, standardized data models (e.g., OMOP, FHIR), and incentives for data custodians to harmonize and expose data for authorized PV use [69, 37]. Without coordinated adoption of standards, federated analytics or pooling will remain technically and operationally difficult [18, 19].

10.2 Data Governance, Privacy, and Legal Uncertainty

Robust legal and policy frameworks are essential for responsible RWD use. India does not yet have a fully enacted, consolidated personal data protection statute equivalent to GDPR; earlier drafts of a Personal Data Protection Bill (PDPA) have been debated and revised, creating some uncertainty among data custodians about lawful data sharing, de-identification standards, and re-use for secondary research and surveillance [70]. Meanwhile, NDHM/ABDM has developed consent and data-sharing architectures intended to operationalize user consent and data portability, but practical implementation and legal harmonization across states, public programmes and private providers is still evolving [69, 37]. This legal uncertainty complicates cross-institutional linkages, federated analyses, and secondary use of clinical data for PV. Data custodians (hospitals, insurers, registries) may be reluctant to participate in federated systems without clear legal assurances and governance mechanisms that balance public health utility with individuals' privacy rights [1, 70].

10.3 Ethical Considerations: Consent, Transparency and Public Trust

RWE-driven PV raises several ethical questions: when and how to obtain consent for secondary use of routine clinical data; how to ensure transparency about data use and governance; how to manage re-identification risks; and how to ensure equitable benefits from surveillance activities [54, 71, 72]. WHO and CIOMS guidance emphasize that public health surveillance and PV have legitimate public-interest justifications for secondary data use, but they also highlight requirements for minimal intrusion, adequate safeguards, and proportional governance [71, 72]. In the Indian context, obtaining explicit, informed consent at scale for PV use is operationally challenging; therefore, many programmes rely on legal provisions for public health surveillance or on broad consent models. To retain public trust, PV systems must adopt transparent communication, easy opt-out mechanisms where appropriate, and strong technical de-identification and governance safeguards [37, 54, 71].

10.4 Analytical Validity, Bias and Confounding

RWD often lack standardized outcome definitions, consistent exposure ascertainment, and may suffer from confounding by indication, channeling bias, and other systematic biases [21, 33]. Using RWE for signal detection demands careful study design (target trial emulation, self-controlled designs), pre-specification, and sensitivity analyses to address potential biases [21, 66]. In India, variable coding practices, incomplete lab data, and missing outpatient medication records exacerbate these methodological challenges and increase the risk of spurious signals or missed signals if analytic approaches do not explicitly account for such limitations [33, 35].

10.5 Workforce, Technical Capacity and Resource Constraints

High-quality RWE generation and advanced analytics (AI/ML) require skilled epidemiologists, data scientists, informaticians, and software engineers, as well as sustained computational resources [10, 18]. While India has centers of excellence and a growing digital health workforce, distribution is uneven; many regional centres, smaller hospitals, and state public health agencies lack personnel trained in pharmacoepidemiology or in implementation of common data models and federated analytics [22, 18]. Building capacity will require academic-industry-government partnerships, training curricula, and investment in infrastructure.

10.6 Equity, Accessibility and Representation

There is a risk that early RWE systems built on private hospital EHRs or insurance claims will over-represent urban, higher-income populations and under-represent rural, poorer, and marginalized groups who access care in under-digitized settings [8, 36]. This differential representation can produce biased estimates of incidence and risk and perpetuate inequities in safety surveillance. Ethical PV practice requires deliberate strategies to ensure inclusive data capture e.g., engaging public hospitals, primary health centres, and national programme registries to build representative evidence for the diverse Indian population [64, 43].

10.7 Operationalizing Regulatory Use: Standards and Decision Workflows

For RWE to inform regulatory action, regulators need documented standards for data quality, study design, pre-registration of analysis protocols, and procedures for rapid review of RWD-based findings [66, 15]. While CDSCO and PvPI have guidance on PV obligations [25], operational guidance on integrating RWE (e.g., acceptance criteria, minimum data standards, analytic specs) is still under development. India can adapt international practice (FDA/EMA case examples and DARWIN/Sentinel playbooks) while tailoring thresholds and governance for local context [66, 1].

11. Recommendations for Strengthening RWE-Based Pharmacovigilance in India

Strengthening India's capability to generate and use Real-World Evidence (RWE) for pharmacovigilance (PV) requires coordinated action across technical infrastructure, governance, regulatory processes, and workforce development. The following recommendations are informed by global best practices (FDA, EMA, CIOMS), India-specific analyses (PvPI, digital health programmes), and open-access literature.

11.1 Build National RWD Infrastructure Using Standardized Data Models

A major foundational requirement is strengthening India's digital health infrastructure and harmonizing data capture across healthcare providers.

Adopt and promote common data models (CDMs) such as the Observational Medical Outcomes Partnership (OMOP) framework for hospital systems and registry programmes, enabling structured RWD across sites to be analyzed consistently [73, 74, 75]. Early pilots under the Ayushman Bharat Digital Mission (ABDM) and ABHA framework can incorporate CDM-aligned templates into provider-facing systems [37].

Benefits:

- Enables federated analytics (similar to FDA Sentinel and EMA DARWIN-EU) [18, 1].
- Ensures interoperability between hospitals, registries, insurers, and PvPI systems [73].

Actions:

- Develop national implementation guidelines for OMOP/FHIR in PV use-cases [37, 73].
- Incentivize adoption through accreditation standards or reimbursement-linked incentives.

11.2 Strengthen EHR Adoption and Integration With PvPI

Scale-up EHR adoption in secondary and tertiary care facilities using ABDM-compliant systems to ensure structured data capture of medication exposures, diagnoses, laboratory values, and outcomes [35, 37].

Integrate EHR systems with PvPI to allow automated extraction of suspected ADR data, improving completeness and timeliness of ICSRs. Hospitals with functional EHRs could pilot semi-automated signal monitoring pipelines using rule-based triggers or ML-enhanced screening [33, 10].

Actions:

- Develop PvPI EHR-ADR integration API guidelines.
- Pilot hospital-based active surveillance modules (e.g., for hepatotoxicity or antimicrobial ADRs).

11.3 Leverage Disease Registries for Condition-Focused Safety Monitoring

India's established registries especially in oncology (NCRP), HIV/TB programmes, and maternal/child health represent fertile ground for registry-based pharmacovigilance [64, 65, 43]. Enhance registries with medication exposure fields, ADR modules, and standardized follow-up intervals.

Use-cases:

- Monitor targeted cancer therapies using NCRP-linked datasets [64].
- Track ADRs in national TB/HIV programmes, where treatment is protocolized and follow-up is routine [43].

Benefits:

- Enables population-specific active safety surveillance with high data quality.

11.4 Develop National Frameworks for Federated Analytics

Given India's diversity and fragmented data sources, a federated analytics model similar to FDA Sentinel or EMA DARWIN-EU should be adopted.

Federated networks allow data to remain with custodians while standardized queries are executed locally [18, 1], addressing privacy and governance concerns.

Actions:

- Establish an India-specific sentinel network linking major insurers, select public hospitals, national registries, and PvPI.
- Create governance structures for secure local execution of analytic code with aggregate outputs returned to a central coordination centre [69, 72].

Benefits:

- Avoids transfer of sensitive patient data.
- Scales efficiently with India's large provider ecosystem.

11.5 Strengthen Data Governance, Privacy, and Ethical Frameworks

India must enact and operationalize robust data-governance rules to support lawful, ethical RWD use for PV.

Priority areas:

1. Clear legal basis for secondary use of health data for public health surveillance [70, 71].
2. Consent architecture aligned with ABDM's digital consent framework [37].
3. De-identification standards for data shared with PvPI or federated networks [72].
4. Transparency mechanisms including periodic public communication on how health data are used for PV [71].

Outcome: strengthens public trust, reduces hesitancy among data custodians, and aligns India with global privacy norms.

11.6 Capacity Building: Pharmacoepidemiology, Data Science, and AI/ML Skills

A sustained effort is needed to train professionals in pharmacoepidemiology, data science, informatics, and regulatory science.

Actions:

- Establish national training programmes (virtual and in-person) in collaboration with universities and WHO collaborating centres [72].
- Create PvPI Centres of Excellence for RWE and AI/ML in PV.
- Incorporate OMOP/FHIR, ML methods, and advanced causal inference training into curricula [53, 73].

Benefits:

- Reduces dependence on external vendors.
- Creates long-term national expertise for regulatory-grade RWE.

11.7 Implement AI/ML in High-Value, Controlled Pilots

AI/ML holds significant potential but must be introduced cautiously [46, 54].

Recommended pilot areas:

- Automated ICSR coding and duplicate detection [50].
- NLP-based ADR extraction from EHRs in select tertiary hospitals [34, 52].
- Hybrid pipelines combining disproportionality analysis with anomaly detection [48].

Governance:

- Follow WHO ethics guidance [47].
- Record provenance of training data, performance metrics, and model-update policies [49].

11.8 Modernize Regulatory Frameworks for RWE

CDSCO's PV and MAH guidance documents should be expanded into a comprehensive RWE regulatory framework modeled partly on FDA/EMA RWE programs [66, 1].

Recommendations:

- Publish an "Indian RWE Framework for Regulatory Decision-Making."
- Specify minimum data-quality standards, acceptable study designs, and criteria for RWE in label updates or post-marketing studies.
- Enable RWE in conditional approvals, risk-minimization plans, and renewal decisions.

11.9 Improve Public and Provider Engagement

Enhancing ADR reporting rates remains essential, even within RWE ecosystems.

Actions:

- Nationwide scale-up of QR-code reporting and patient-facing reporting channels [31, 32].
- Incentive-based reporting schemes for hospitals and clinicians.
- Public awareness campaigns tied to major health programmes (TB, HIV, oncology).

Outcome: better numerator data for PV, improved representativeness, and early detection of emerging safety issues.

11.10 Promote Collaboration Among Stakeholders

India's RWE ecosystem must be co-created by regulators, PvPI, academic institutions, insurers, hospitals, and health-tech companies.

Actions:

- Establish a National RWE–PV Consortium to set standards and coordinate pilots.
- Align data standards with global initiatives (OMOP, FHIR) to support international research.
- Encourage public–private partnerships to accelerate technical innovation while maintaining data governance oversight [1].

12. Future Directions

As India builds capacity for Real-World Evidence (RWE)–driven pharmacovigilance (PV), several future directions can accelerate impact and ensure sustainability. These directions focus on technological architectures, methodological advances, governance innovation, workforce development, and international collaboration. The following strategic avenues aim to translate near-term pilots into a mature, regulator-grade RWE ecosystem that supports robust signal detection and public health action.

12.1 Federated, Privacy-Preserving Analytics at Scale

A federated analytics model where data custodians retain local control while standardized analytic queries run locally and return aggregate results will remain central to large-scale PV in India. The Sentinel and DARWIN-EU experiences demonstrate the feasibility and governance advantages of federated networks, while highlighting the importance of common data models (CDMs) and robust local capacity to implement them [18, 19, 73]. In India, the ABHA/ABDM architecture combined with OMOP/FHIR mapping is a practical route to federated analytics that balance privacy, legal constraints, and utility [37, 73] .

Enhancements to federated systems include:

- **Privacy-enhancing technologies (PETs):** Techniques such as secure multiparty computation, homomorphic encryption, and differential privacy can enable aggregate analytics with provable privacy guarantees; combining PETs with federated learning can allow model training without centralizing raw data [49, 54]. Pilots using PETs for specific PV use cases (e.g., vaccine safety) can build trust while proving technical feasibility.
- **Standardized query libraries and containerized analytic packages:** Reusable analytic modules (for cohort definitions, sequential monitoring, SCCS analyses) that run identically at each node reduce variability and accelerate evidence generation [18, 73].

12.2 National Reference Implementations and Modular Pilots

Rather than attempting nationwide scale immediately, India should pursue reference implementations and modular pilots that deliver regulatory value quickly:

- **Immunization safety pilot:** Leverage existing immunization programmes and ABHA to run near-real-time safety monitoring for nationally administered vaccines a high-value, politically salient use-case that demonstrates public benefit.
- **High-priority therapeutic pilots:** Target widely used therapeutic classes (e.g., cardiovascular drugs, anti-infectives) where hospital EHR clusters exist, enabling cohort and self-controlled studies to validate signal detection pipelines.
- **Registry-anchored pilots:** Augment NCRP or disease-programme registries with exposure fields and ADR modules to pilot registry-based active surveillance for oncology or TB/HIV therapeutics.

Pilots should be pre-registered with clear success metrics (timeliness, sensitivity, specificity, regulatory actionability) and operate under explicit governance agreements [66, 19].

12.3 Hybrid Methods and Causal Inference Advances

Future PV systems should adopt hybrid detection pipelines that combine:

- Disproportionality screening on spontaneous reports (for sensitivity),
- EHR/claims sequential monitoring (for incidence and temporal patterns), and
- Target trial emulation / advanced causal inference for robust estimation of effect sizes and confounding control [21, 18].

Investment in standardized phenotyping, validated outcome algorithms (lab thresholds, composite endpoints), and cross-source replication will increase the credibility of RWE signals presented to regulators [21, 66].

12.4 Responsible, Explainable AI/ML Integrated into Workflows

AI/ML will continue to mature as an enabler rather than a substitute for epidemiologic design. Future directions include:

- Explainable models for triage and prioritization that surface interpretable reasons for flagging a case or pattern [54].
- Continuous learning systems with documented model-update procedures and real-world performance monitoring [49, 55].
- Indian-context annotated corpora (multi-language) and shared gold standards to improve NLP and ML performance on local clinical narratives [53].

Regulatory sandboxes under PvPI/CDSCO can pilot explainable ML models for non-critical tasks (e.g., coding, de-duplication) and progressively expand to higher-impact use cases once performance and governance are proven [49, 55].

12.5 Strengthened Governance, Consent and Public Engagement

Future systems must embed trust and transparency:

- Operational consent frameworks within ABHA that allow citizens to grant limited, auditable research/monitoring access to their de-identified health data for PV purposes, with clear opt-out mechanisms [37].
- Transparent public dashboards (aggregate findings, how data are used, privacy safeguards) and regular public communications on PV outcomes to build confidence.
- Ethics oversight aligned with WHO/CIOMS recommendations to govern secondary use and ML deployments [71, 72].

12.6 Workforce Ecosystem and Research-Regulatory Pathways

Sustained capacity building includes:

- Formal training pathways (certificates, fellowships) in pharmacoepidemiology, RWE methods, and health data engineering.
- Centers of excellence that provide analytic services, share tools and annotated datasets, and mentor regional PV units [74].
- Clear regulatory pathways (pre-submission consultation, data-quality checklists) to help sponsors and investigators design RWE studies likely to meet regulatory evidentiary thresholds [66].

12.7 International Collaboration and Interoperability

India benefits from global collaboration:

- Aligning data models and phenotypes with international standards (OMOP, MedDRA, SNOMED) to enable cross-country replication and meta-analysis [73].
- Participating in multinational RWE consortia and federated studies to detect rare events that transcend national sample sizes, particularly for novel therapeutics and vaccines [19, 18].
- Sharing methodology and governance lessons through WHO/CIOMS fora to accelerate mutual learning [3, 71].

12.8 Sustainable Financing and Incentives

Finally, achieving mature RWE-PV requires sustainable funding models:

- Public investment to establish national coordination centres and support public hospital participation.
- Incentives for private providers and insurers (data-access agreements, reimbursement linkages, or accreditation benefits) to contribute standardized data.
- Public–private partnerships for building tooling (analytics libraries, data transformation services) under clear governance to avoid vendor lock-in while leveraging private sector capacity.

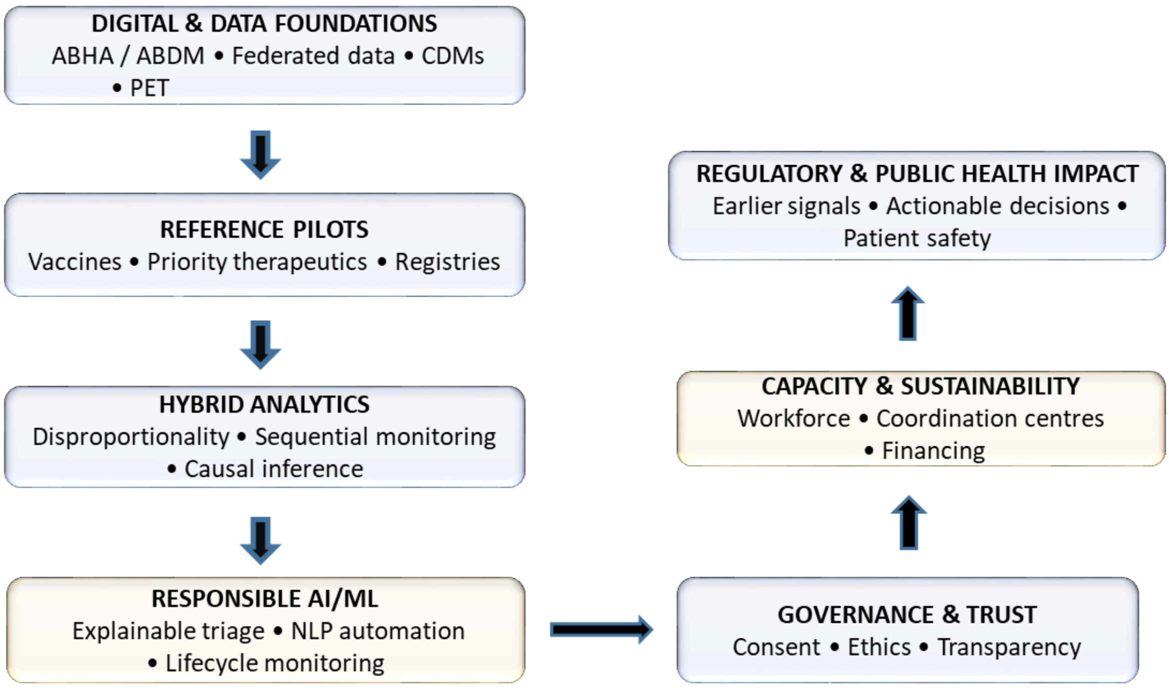

Figure 2: Roadmap for a Regulator-Grade RWE-Driven Pharmacovigilance System in India

13. CONCLUSION

Real-World Evidence (RWE) has emerged globally as a critical complement to traditional pharmacovigilance (PV) systems, enabling earlier, more accurate, and more contextually grounded detection of adverse drug reactions (ADRs). As the U.S. FDA, EMA, and international bodies such as CIOMS and WHO increasingly integrate RWE into regulatory decision-making, India is uniquely positioned to leverage its vast and diverse patient population, emerging digital health infrastructure, and evolving PV ecosystem to build a next-generation drug safety surveillance model (1, 2, 3, 4, 18, 19).

While India's Pharmacovigilance Programme (PvPI) has significantly strengthened spontaneous reporting and national signal management processes, the next decade must focus on transitioning from a predominantly passive system to a hybrid model that integrates structured real-world data (RWD) sources including EHRs, insurance claims, registries, and digital patient-reported outcomes. The Ayushman Bharat Digital Mission (ABDM) and ABHA framework provide foundational digital health identifiers, interoperability standards, and consent architectures essential for scaling RWD access and federation across institutions [37, 69].

Global experiences demonstrate that RWE-driven PV requires not only technology but also governance, legal clarity, validated analytic methods, and sustained capacity building. India must adopt federated analytics models, promote national common data standards (e.g., OMOP/FHIR), strengthen evidence requirements for regulatory RWE submissions, and invest in training pharmacoepidemiologists, data scientists, and ML engineers [66, 1, 73]. AI/ML tools offer substantial potential but must be developed and deployed responsibly, with transparency, human oversight, and ethical safeguards [49, 54].

The future of Indian pharmacovigilance lies in building a scalable, privacy-preserving, federated RWE ecosystem capable of near–real-time safety monitoring, robust signal validation, and rapid regulatory action. By aligning national efforts with global frameworks, harmonizing data standards, and embedding trust and public engagement, India can become a leader in RWE-driven drug safety science. This transformation promises not only improved patient safety but also stronger regulatory decision-making and enhanced public health outcomes.

14. REFERENCES

1. U.S. Food and Drug Administration. Real-World Evidence. Silver Spring (MD): FDA; 2025 Sep 22
2. European Medicines Agency. Real-world evidence (RWE) in regulatory decision-making. Amsterdam: EMA; 2025
3. Council for International Organizations of Medical Sciences. Real-World Data and Real-World Evidence in Regulatory Decision-Making. Geneva: CIOMS; 2021
4. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W. What is real-world data? A review of definitions based on literature and stakeholder interviews. Value in health. 2017 Jul 1;20(7):858-65.
5. Corrigan-Curay J, Sacks L, Woodcock J. Real-world evidence and real-world data for evaluating drug safety and effectiveness. Jama. 2018 Sep 4;320(9):867-8.
6. Dang A, Vallish BN. Real world evidence: an Indian perspective. Perspectives in Clinical Research. 2016 Oct 1;7(4):156-60.
7. Bhatt A. Conducting real-world evidence studies in India. Perspectives in Clinical Research. 2019 Apr 1;10(2):51-6.
8. Indian Pharmacopoeia Commission. Annual performance report: Adverse Drug Reaction Monitoring Centres under PvPI. Ghaziabad: IPC; 2023
9. National Digital Health Mission. ABHA Health ID Portal. New Delhi: Ministry of Health & Family Welfare
10. Beam AL, Kohane IS. Big data and machine learning in health care. Jama. 2018 Apr 3;319(13):1317-8.
11. Sendak MP, D'Arcy J, Kashyap S, Gao M, Nichols M, Corey K, Ratliff W, Balu S. A path for translation of machine learning products into healthcare delivery. EMJ Innov. 2020 Jan 27;10:19-00172.
12. World Health Organization. The importance of pharmacovigilance – safety monitoring of medicinal products. Geneva: WHO; 2002
13. Hazell L, Shakir SA. Under-reporting of adverse drug reactions. Drug safety. 2006 May;29(5):385-96.

14. Uppsala Monitoring Centre. Signal management at UMC: managing safety signals Emerging tools and best practices.
15. U.S. Food and Drug Administration. FDA Adverse Event Reporting System (FAERS). Silver Spring (MD): U.S. Food and Drug Administration.
16. Bate A, Evans SJ. Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiology and drug safety. 2009 Jun;18(6):427-36.
17. European Medicines Agency. Signal management. Amsterdam: EMA
18. U.S. Food and Drug Administration. Sentinel Initiative. Silver Spring (MD
19. European Medicines Agency. DARWIN-EU Data Network. Amsterdam: EMA
20. European Medicines Agency. *Benefit-risk methodology project: Work package 2 report: Applicability of current tools and processes for regulatory benefit-risk assessment*. London: European Medicines Agency; 2010 Aug 31
21. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. American journal of epidemiology. 2016 Apr 15;183(8):758-64.
22. Kalaiselvan V, Thota P, Singh GN. Pharmacovigilance Programme of India: Recent developments and future perspectives. Indian journal of pharmacology. 2016 Nov 1;48(6):624-8.
23. Lihite RJ, Lahkar M. An update on the Pharmacovigilance Programme of India. Frontiers in pharmacology. 2015 Sep 22;6:194.
24. Indian Pharmacopoeia Commission, Pharmacovigilance Programme of India (PvPI). *Annual Performance Report of PvPI 2023-2024*. Ghaziabad: Indian Pharmacopoeia Commission.
25. Central Drugs Standard Control Organization (CDSCO), Directorate General of Health Services, Ministry of Health & Family Welfare, Government of India. *Guidance Document for Marketing Authorization Holders*. New Delhi: CDSCO; 2017.
26. Indian Pharmacopoeia Commission, Pharmacovigilance Programme of India (PvPI). ADR Monitoring Centres list. Ghaziabad; 2023.
27. Agrawal V, Shrivastava TP, Adusumilli PK, Vivekanandan K, Thota P, Bhushan S. Pivotal role of Pharmacovigilance Programme of India in containment of antimicrobial resistance in India. Perspectives in clinical research. 2019 Jul 1;10(3):140-4.
28. Central Drugs Standard Control Organization (CDSCO), Directorate General of Health Services, Ministry of Health & Family Welfare, Government of India. *CDSCO – Home*. New Delhi: CDSCO
29. Kalaiselvan V, Srivastava S, Singh A, Gupta SK. Pharmacovigilance in India: present scenario and future challenges. Drug safety. 2019 Mar 7;42(3):339-46.
30. Times of India. *Scan to report side effects: Government orders mandatory QR codes at all pharmacies*. 2025 Dec 3
31. Economic Times. *CDSCO directs pharmacies to display QR code and toll-free number for reporting adverse drug reactions*. 2025 Dec 3
32. Times of India. *Chemists across India will soon have new QR code and toll-free number, here's what it will mean for buyers*. 2025 Dec 4.
33. Davis SE, Zabotka L, Desai RJ, Wang SV, Maro JC, Coughlin K, Hernández-Muñoz JJ, Stojanovic D, Shah NH, Smith JC. Use of electronic health record data for drug safety signal identification: a scoping review. Drug safety. 2023 Aug;46(8):725-42.
34. Haerian K, Varn D, Vaidya S, Ena L, Chase HS, Friedman C. Detection of pharmacovigilance-related adverse events using electronic health records and automated methods. Clinical Pharmacology & Therapeutics. 2012 Aug;92(2):228-34.
35. Srivastava SK. Adoption of electronic health records: a roadmap for India. Healthcare informatics research. 2016 Oct 31;22(4):261-9.
36. Makady A, Ten Ham R, de Boer A, Hillege H, Klungel O, Goettsch W. Policies for use of real-world data in health technology assessment (HTA): a comparative study of six HTA agencies. Value in Health. 2017 Apr 1;20(4):520-32.

37. National Health Authority, Government of India. *Ayushman Bharat Digital Mission*. New Delhi: National Health Authority: 2023.
38. National Health Authority, Government of India. *Health ID – Ayushman Bharat Digital Mission*. New Delhi: National Health Authority.
39. Gandhi AP, Soundappan K. Perception towards electronic health records & uptake of digital health IDs among the urban residents in northern India: A mixed methods study. The Indian Journal of Medical Research. 2024 Sep 30;160(1):51.
40. Behera P, Patro BK. Population based cancer registry of India–the challenges and opportunities. Asian Pacific journal of cancer prevention: APJCP. 2018;19(10):2885.
41. Indian Council of Medical Research – National Centre for Disease Informatics and Research (ICMR-NCDIR). *Report of National Cancer Registry Programme 2020*. Bengaluru: ICMR-NCDIR; 2020.
42. Mathur P, Sathishkumar K, Chaturvedi M, Das P, Sudarshan KL, Santhappan S, Nallasamy V, John A, Narasimhan S, Roselind FS, Icmr-Ncdir-Ncrp Investigator Group. Cancer statistics, 2020: report from national cancer registry programme, India. JCO global oncology. 2020 Jul;6:1063-75.
43. Directorate General of Health Services, Ministry of Health & Family Welfare, Government of India. *National Tuberculosis Elimination Programme*. New Delhi: Government of India.
44. Prakash J, Joshi K, Malik D, Mishra O, Sachan A, Kumar B, Bhushan S, Kalaiselvan V, Singh GN. “ADR PvPI” Android mobile app: Report adverse drug reaction at any time anywhere in India. Indian journal of pharmacology. 2019 Jul 1;51(4):236-42.
45. Indian Pharmacopoeia Commission, Pharmacovigilance Programme of India (PvPI). *PvPI Newsletter, Volume 5, Issue 12, 2015* . Ghaziabad: Indian Pharmacopoeia Commission; 2015
46. Pilipiec P, Liwicki M, Bota A. Using machine learning for pharmacovigilance: a systematic review. Pharmaceutics. 2022 Jan 23;14(2):266.
47. Murphy RM, Klopotowska JE, de Keizer NF, Jager KJ, Leopold JH, Dongelmans DA, Abu-Hanna A, Schut MC. Adverse drug event detection using natural language processing: A scoping review of supervised learning methods. Plos one. 2023 Jan 3;18(1):e0279842.
48. Desai MK. Artificial intelligence in pharmacovigilance–Opportunities and challenges. Perspectives in Clinical Research. 2024 Jul 1;15(3):116-21.
49. U.S. Food and Drug Administration (FDA). *Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD) Discussion paper and request for feedback* . Silver Spring, MD: U.S. Food and Drug Administration; 2019
50. Kassekert R, Grabowski N, Lorenz D, Schaffer C, Kempf D, Roy P, Kjoersvik O, Saldana G, ElShal S. Industry perspective on artificial intelligence/machine learning in pharmacovigilance. Drug Safety. 2022 May;45(5):439-48.
51. Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. Journal of the American Medical Informatics Association. 2005 Jul 1;12(4):448-57.
52. Luo Y, Thompson WK, Herr TM, Zeng Z, Berendsen MA, Jonnalagadda SR, Carson MB, Starren J. Natural language processing for EHR-based pharmacovigilance: a structured review. Drug safety. 2017 Nov;40(11):1075-89.
53. Golder S, Xu D, O’Connor K, Wang Y, Batra M, Hernandez GG. Leveraging natural language processing and machine learning methods for adverse drug event detection in electronic health/medical records: a scoping review. Drug Safety. 2025 Apr;48(4):321-37.
54. World Health Organization. Ethics and governance of artificial intelligence for health: large multi-modal models. WHO guidance. World Health Organization; 2024 Jan 18.
55. U.S. Food and Drug Administration. *FDA releases Artificial Intelligence/Machine Learning Action Plan*. Silver Spring (MD): U.S. Food and Drug Administration; 12 Jan 2021.
56. Indian Pharmacopoeia Commission. *Pharmacovigilance Programme of India (PvPI) Performance Report 2019-2020* [Internet]. Ghaziabad: Indian Pharmacopoeia Commission, Ministry of Health & Family Welfare, Government of India; 2019-20.

57. Thota P, Thota A, Medhi B, Sidhu S, Kumar P, Selvan VK, Singh GN. Drug safety alerts of pharmacovigilance programme of India: A scope for targeted spontaneous reporting in India. Perspectives in clinical research. 2018 Jan 1;9(1):51-5.
58. Mulchandani R, Kakkar AK. Reporting of adverse drug reactions in India: A review of the current scenario, obstacles and possible solutions. International Journal of Risk & Safety in Medicine. 2019 Feb;30(1):33-44.
59. Sriram S, Ghasemi A, Ramasamy R, Devi M, Balasubramanian R, Ravi TK, Sabzghabaee AM. Prevalence of adverse drug reactions at a private tertiary care hospital in south India. Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences. 2011 Jan;16(1):16.
60. Singh P, Verma S, Vaishnav Y, Joshi U, Agrawal M. Probing the in-depth analysis of Serious Adverse Drug Reactions in a tertiary care hospital of Central India. Exploratory Research in Clinical and Social Pharmacy. 2025 Mar 1;17:100579.
61. Mulkalwar S, Chitale S, Dandage P, Bapat S, Tilak AV, Patil S, Tilak A. Evaluation of the Effectiveness of the Adverse Drug Reaction Alert Card System in Preventing the Recurrence of Adverse Drug Reactions. Cureus. 2024 Jul 16;16(7).
62. Sharma H, Anil K, Parekh S, Pujari P, Shewale S, Desai S, Madhusudhan RL, Patel J, Eswaraiah A, Rao H, Gairola S. A phase-I, open label clinical trial to assess the safety & tolerability of qHPV vaccine manufactured by Serum Institute of India Pvt. Ltd. in adults. Vaccine: X. 2023 Aug 1;14:100313.
63. Sankaranarayanan R, Bhatla N, Basu P. Current global status & impact of human papillomavirus vaccination: Implications for India. Indian Journal of Medical Research. 2016 Aug 1;144(2):169-80.
64. National Centre for Disease Informatics and Research (NCDIR), Indian Council of Medical Research (ICMR). *Report of National Cancer Registry Programme, Bengaluru, India 2020* [Internet]. Bengaluru: ICMR-NCDIR; 2020
65. Behera P, Patro BK. Population based cancer registry of India–the challenges and opportunities. Asian Pacific journal of cancer prevention: APJCP. 2018;19(10):2885.
66. 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. 2023
67. Norén GN, Donegan K, Muñoz MA, Alshammari TM, Pratt N, Candore G, Morales D, Rijnbeek P, Bate A, Postigo R, Toh S. Recommendations to Enable Broader Use of Real-World Evidence to Inform Decision-Making Throughout Pharmacovigilance Signal Management. Pharmacoepidemiology and Drug Safety. 2025 Oct;34(10):e70231.
68. Yuan L, Rahman M, Concato J. Comparison of two assessments of real-world data and real-world evidence for regulatory decision-making. Clinical and Translational Science. 2024 Jan;17(1):e13702.
69. Government of India, Ministry of Health and Family Welfare. *National Digital Health Blueprint (NDHB)* . New Delhi: Ministry of Health and Family Welfare; 2019.
70. PRS Legislative Research. *The Personal Data Protection Bill, 2019*. New Delhi: PRS India; 2019.
71. World Health Organization. *WHO Guidelines on Ethical Issues in Public Health Surveillance*. Geneva: World Health Organization; 2017.
72. Council for International Organizations of Medical Sciences (CIOMS). *International Ethical Guidelines for Health-related Research Involving Humans* . 4th ed. Geneva: Council for International Organizations of Medical Sciences; 2016.
73. Observational Health Data Sciences and Informatics (OHDSI) Collaborative. *The Book of OHDSI*: OHDSI; 2019
74. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, Suchard MA, Park RW, Wong IC, Rijnbeek PR, Van Der Lei J. Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. Studies in health technology and informatics. 2015;216:574.
75. Akehurst R, Murphy LA, Solà-Morales O, Cunningham D, Mestre-Ferrandiz J, de Pouvourville G. Using real-world data in the health technology assessment of pharmaceuticals: strengths, difficulties, and a pragmatic way forward. Value in Health. 2023 Apr 1;26(4):11-9.