# Guide to Pharmacovigilance Technology & AI in Drug Safety

By Adrien Laurent, CEO at IntuitionLabs • 11/10/2025 • 30 min read

pharmacovigilance drug safety ai in pharmacovigilance pv automation safety database signal detection real world data iso idmp

# Executive Summary

Pharmacovigilance (PV) – the science of monitoring, detecting, and preventing adverse drug reactions (ADRs) – has entered a new technological era. This report surveys the **pharmacovigilance tech landscape** in unprecedented depth. We review the historical evolution and current state of PV, the modern tools and methods in use (including safety databases, automation, AI/ML, real-world data, social listening, and mobile applications), and the regulatory and industry context. We examine key trends such as the adoption of artificial intelligence (AI) and machine learning (ML) for case processing and signal detection ([1] pmc.ncbi.nlm.nih.gov) ([2] pmc.ncbi.nlm.nih.gov), the integration of real-world data (RWD) via networks like FDA's Sentinel ([3] pmc.ncbi.nlm.nih.gov), the use of **mobile apps and wearables** to capture patient-reported outcomes ([4] pmc.ncbi.nlm.nih.gov) ([5] pharmacally.com), and emerging standards (e.g. ISO IDMP) for data interoperability (www.ema.europa.eu). We present case studies illustrating these developments – from a pilot where an AI/Bayesian network accelerated causality assessment of ADRs ([6] link.springer.com), to the global WHO VigiBase system holding over 40 million reports ([7] who-umc.org). The report also discusses challenges (data quality, privacy, regulatory requirements) and looks ahead to future directions: predictive safety analytics, digital therapeutics, and further convergence of health informatics and PV. All claims are supported by extensive citations to scientific articles, industry reports, and official sources.

# Introduction and Background

Pharmacovigilance (PV) is an essential part of drug safety, concerned with the **identification, assessment, and prevention of adverse effects** of pharmaceuticals. It traces back to the thalidomide tragedy of the 1960s, after which regulatory systems were established to collect safety reports. Even today, ADRs remain a *major public health problem* – in the US alone ADRs account for "at least 7,000 fatalities annually" (and tens of billions USD in costs) ([8] pmc.ncbi.nlm.nih.gov). Globally, the World Health Organization's (WHO's) Uppsala Monitoring Centre curates VigiBase, the world's largest database of post-market safety reports. As of 2025, **VigiBase contains over 40 million Individual Case Safety Reports (ICSRs)** from more than 180 member countries ([7] who-umc.org). These vast datasets underpin signal detection efforts but also pose data management challenges.

Traditionally, PV relied on spontaneous reporting (via forms like CIOMS/MedWatch in structured fields) and manual case review. Over time, PV has evolved to utilize computerized databases (e.g. FDA's FAERS, EudraVigilance) and standardized data formats (such as the ICH E2B standard for electronic ICSRs) to facilitate global reporting. In parallel, *adverse event signal detection* has used methods like disproportionality analysis (proportional reporting ratios, Bayesian shrinkage, etc.) to flag unusual drug-event pairs.

Today, PV is undergoing a **digital transformation**. New technologies – such as robotic process automation (RPA), natural language processing (NLP), and AI/ML – promise to streamline PV workflows and extract insights from unstructured sources (medical literature, electronic health records, social media). Huge bodies of real-world data (claims, EHRs, device data) are now available to augment traditional safety data. Regulatory agencies are encouraging such innovations: for example, the FDA Sentinel system (established 2008) uses a common data model to run distributed analyses on healthcare data ([3] pmc.ncbi.nlm.nih.gov), and the EU has mandated standardized medicine identification (ISO IDMP) to improve data interoperability (www.ema.europa.eu).

Nevertheless, challenges remain. Data heterogeneity, privacy concerns, the need for expert validation, and regulatory requirements constrain how modern tech can be deployed. A balanced assessment (industry and regulators) stresses that **human oversight remains essential**, even as automation increases ([9] pmc.ncbi.nlm.nih.gov). This report aims to provide a **comprehensive, evidence-based view** of the current PV tech landscape and its trajectory.

# Historical Evolution of PV Technology

The PV technology journey spans decades of regulatory developments and technological advances. Key milestones include:

- **1960s–1990s: Emergence of Spontaneous Reporting Systems.** In the aftermath of early drug safety crises, regulators established spontaneous reporting forms (paper CIOMS, MedWatch, Yellow Card, etc.) and set up national PV centers (UK, FDA, EMA, etc.). Example: FDA's MedWatch (started 1993) and the UK's Yellow Card scheme generate structured ICSRs submitted by physicians and manufacturers.

- **2000s: Computerization and Data Standards.** The turn of the century saw globalization of PV. The International Council for Harmonisation's E2B(R2) standard (2001) defined electronic ICSR formats, enabling automated exchange of safety reports. Commercial PV databases (e.g. Oracle Argus, ArisGlobal ARISg) entered the market, digitizing case processing. Meanwhile, WHO launched VigiBase (Uppsala, 1968) and began VigiFlow, a web-based reporting system (2004). By mid-2000s, mandatory post-market surveillance and new regulatory frameworks (e.g. EU Directives, FDA Amendments Act 2007) increased case volumes.

- **2010s: Data Mining and International Collaboration.** With millions of ICSRs, statistical signal detection became routine. The EU's EudraVigilance (est. 2001) expanded dramatically, eventually requiring ISO/ICH E2B(R3) in 2022. The FDA established the Sentinel Network (2008) linking multiple data partners under a common model. Real-World Evidence (RWE) gained traction under initiatives like NEOMA, BEST (biologics initiative), and EMA's DARWIN project for data analysis. Social media and web data began to be explored as supplementary sources (e.g. WEB-RADR project in 2014 launched a mobile reporting app).

- **2020s: AI, RWD, and Harmonization.** The last few years have seen rapid interest in **AI/ML** and advanced analytics for PV. Pilot studies demonstrated the feasibility of AI-driven case processing ([1] pmc.ncbi.nlm.nih.gov). Consortia (TransCelerate, PVNet) surveyed industry; by 2021 many companies planned or began using intelligent automation (RPA, NLP, machine learning) in PV ([10] pmc.ncbi.nlm.nih.gov) ([11] pmc.ncbi.nlm.nih.gov). Regulators are transitioning to newer standards: for example, ICH E2B(R3) (more granular data fields) is now adopted in EU PC programs and China; ISO IDMP compliance is expanding globally for product data (www.ema.europa.eu). Wearable sensors and telehealth tools are being piloted to provide **continuous safety monitoring** ([12] pmc.ncbi.nlm.nih.gov) ([5] pharmacally.com).Looking ahead, PV is moving from largely retrospective reporting to more **proactive, predictive surveillance** powered by big data and intelligent algorithms ([13] www.parexel.com) ([6] link.springer.com).

# PV Data Systems and Standards

## Safety Databases and Reporting Tools

Modern PV relies on specialized safety databases and case management platforms. These systems ingest and store ICSRs, support medical coding (e.g. MedDRA terms for symptoms, WHO-Drug codes), and enable regulatory reporting. Key examples include:

- **Oracle Argus Safety (Oracle Corporation)**: A widely-used enterprise PV database, often regarded as the *industry standard* in large companies ([14] ccrps.org). Argus supports high-volume case processing, medical coding, automated workflows, and global submissions. It is offered both on-premises and cloud.

- **ArisGlobal LifeSphere Safety (ArisGlobal)**: A cloud-native PV suite (formerly ARISg/Navax) with AI-driven signal detection tools. LifeSphere emphasizes integrated data analytics and rapid deployment. It leads in applying ML to identify patterns in safety data.

- **Veeva Vault Safety (Veeva Systems)**: A newer cloud-first platform (launched 2018) designed for biotechs and CROs. Vault Safety offers a modern user interface and supports seamless global reporting, focusing on usability and integration with other Veeva products ([15] intuitionlabs.ai).

- **Other Commercial Tools**: Mid-tier options include **Ennov PV-Works** (hybrid cloud/on-prem, popular in small/medium companies), **SafetyEasy** (by AB Cube/EXTEDO, hybrid model), and **Accelovance CLAIRE**. Each provides core PV functions (case entry, MedDRA coding, narrative writing, electronic submissions) but varies in scale and technology approach ([15] intuitionlabs.ai). Table 1 (below) compares selected platforms.

| Software/Product | Vendor | Deployment | Primary Users | Notable Features |
|---|---|---|---|---|
| **Oracle Argus Safety** | Oracle (USA) | Cloud/On-premise | Large enterprises, CROs | Scalable case processing, configurable workflows, RPA integration ([16] ccrps.org) |
| **LifeSphere Safety** | ArisGlobal (Ireland) | Cloud (SaaS) | Mid-large pharma | AI/ML analytics for signal detection, global safety database, automated compliance |
| **Vault Safety** | Veeva Systems (USA) | Cloud-native (SaaS) | Biotechs, CROs | Modern UI, unified E2B submissions, **rapid deployment** ([15] intuitionlabs.ai), automatic updates |
| **PV-Works** | Ennov (France) | Cloud/On-premise | Specialty pharma, regulators | Flexible deployment, budget-friendly, full PV module, multilingual support |
| **SafetyEasy** | AB Cube/EXTEDO (France) | Cloud/On-premise | SME pharma, CROs | Quick setup templates, MedDRA coding, PV workflow automation |
| **Databases (e.g. EudraVigilance, VAERS, FAERS)** | Regulatory Agencies | Cloud (portal-based) | Regulators, Manufacturers | Central repositories of national/regional ICSRs; public query interfaces. |

*Table 1: Comparison of representative pharmacovigilance software systems and databases.* (Source: industry reviews ([16] ccrps.org) ([15] intuitionlabs.ai).)

These platforms are continually evolving. For instance, Oracle reports integration of RPA bots and ML components; ArisGlobal markets *AI (NavaX)* modules for predictive analytics; Veeva emphasizes user experience and connectivity with clinical systems. Industry surveys note that while "rule-based automations" (RPA, lookups) are already widely implemented, interest in cognitive tools (ML/NLP) is rising ([11] pmc.ncbi.nlm.nih.gov) ([10] pmc.ncbi.nlm.nih.gov). Notably, TransCelerate's member companies report that as of 2021 every step in ICSR processing has at least some automation (even if still in pilot) ([17] pmc.ncbi.nlm.nih.gov).

## Data Standards and Interchange

To function smoothly, PV systems rely on **data standards** for report interchange and coding:

- **MedDRA**: The Medical Dictionary for Regulatory Activities provides standardized terms for adverse events and indications. All major PV systems incorporate MedDRA coding to unify free-text symptom narratives.
- **WHO Drug Dictionary**: Codes for active ingredients and medicinal products, often used with the WHO-Drug classification.
- **ICH E2B(R3)**: The International Council for Harmonisation standard (latest version R3) specifies XML formats and data fields for electronic ICSRs. It is now mandated by many authorities: the EU transitioned to E2B(R3) for EudraVigilance, the FDA uses E2B for FAERS, and Japan, China and others are adopting it ([18] www.iconplc.com) ([19] pmc.ncbi.nlm.nih.gov). E2B(R3) allows richer data capture (e.g. unique identifiers, structured narratives).
- **ISO IDMP (SPOR data)**: The ISO Identification of Medicinal Products standards (ISO 11615/11616) define controlled vocabularies for describing substances, products, organizations, and referential data. The EU requires IDMP-compliant submissions of product information (SPOR database) (www.ema.europa.eu), facilitating interoperability. Other regions (Japan, Canada) are also moving toward IDMP.

- **Other Messaging Standards**: Many PV systems support HL7 interfaces or FHIR APIs to exchange safety data with electronic health record (EHR) systems and regulatory portals. For example, systems may accept ICSRs in HL7 CDA format from hospital systems.
- **Coding and Quality Checks**: Incoming reports to databases are validated against minimum data criteria (e.g. patient ID, suspect drug, event) and standardized via vocabularies to ensure high data quality for downstream analysis ([20] who-umc.org). Modern systems often include built-in coding suggestions (e.g. MedDRA synonyms) and duplicate detection algorithms.

Overall, these standards create an ecosystem in which safety data can flow internationally and across stakeholders. However, the complexity of merging different data sources (spontaneous reports, EHRs, literature) remains a technical challenge requiring harmonization efforts.

# Advanced Analytics and Signal Detection

## Traditional Signal Detection

Pharmacovigilance has long used **disproportionality analyses** on large safety databases to generate signals (drug-event associations that occur more often than expected). Methods such as the Proportional Reporting Ratio (PRR) or Empirical Bayes metrics scan millions of ICSRs to flag potential risks. These approaches, while well-established, have limitations (they require sufficient report volume and can produce false positives from confounding) ([21] pmc.ncbi.nlm.nih.gov). Recent work has refined these methods – for example, guidelines (the READUS-PV guideline) improve transparency in reporting disproportionality results ([22] pmc.ncbi.nlm.nih.gov).

## AI/ML for Signal Detection

In the 2020s, PV has seen growing use of AI/ML for **safety signal analytics** beyond simple disproportionality.

- **Machine Learning on RWD**: AI models are being trained on structured real-world data (EHRs, claims) to detect safety signals. For example, scoping reviews find many studies using ML on claims/EHR to predict adverse drug events ([23] pmc.ncbi.nlm.nih.gov). These models often focus on *patient-level ADR risk prediction* rather than population-wide signal detection ([24] pmc.ncbi.nih.gov), suggesting an emerging trend toward personalized pharmacovigilance. Key algorithms include logistic regression, random forests, and deep learning. However, reviews note gaps: explainable AI (XAI) and causal modeling are not yet widely applied, and many models lack external validation ([24] pmc.ncbi.nlm.nih.gov).
- **Natural Language Processing (NLP)**: Unstructured data – patient narratives, social media, journal articles – are rich sources of ADR information. A systematic review found that NLP applied to user-generated content (tweets, forums, blogs) can successfully identify reported ADRs ([25] pmc.ncbi.nlm.nih.gov). For example, NLP systems analyzing health forum posts were able to detect known side effects and even novel symptom clusters ([26] pmc.ncbi.nlm.nih.gov). Roughly 88% of studies reported positive outcomes using NLP for ADR identification ([25] pmc.ncbi.nlm.nih.gov). These findings indicate social media can *complement* traditional reporting by catching signals earlier or in populations underrepresented in official data ([27] pmc.ncbi.nlm.nih.gov). However, challenges include dealing with noisy slang and ensuring representativeness.

- **Bayesian and Probabilistic Models**: Bayesian networks and other probabilistic models are gaining interest for safety assessment. A notable practical case is an "expert-defined Bayesian network" for causality assessment implemented at a pharmacovigilance center. According to a 2025 report, this AI system *dramatically accelerated* case evaluation, reducing causality processing from days to hours and increasing consistency ([6] link.springer.com). Such networks model the complex dependencies among drug attributes, patient factors, and outcomes. They "improve causality assessment by reducing subjectivity" and provide transparent probabilistic reasoning ([6] link.springer.com) ([28] link.springer.com). Overall, advanced analytics like ML, NLP, and Bayesian models are **augmenting PV signal detection** by enabling processing of diverse big-data sources.

## Semiautomation of Case Processing

Beyond signal detection, many AI technologies focus on streamlining routine PV workflows, particularly ICSR processing (case intake, validation, coding). A pioneering 2018 trial (TransCelerate-sponsored) tested three commercial AI/RPA vendors on actual adverse event case reports. The pilot confirmed **the feasibility of AI-driven automation** for key tasks: extracting information from source documents and evaluating basic case validity ([1] pmc.ncbi.nlm.nih.gov). It demonstrated that AI could replace some manual data entry and analysis steps with validated models, making case processing faster and less labor-intensive.

Industry surveys reinforce this trend. By 2021, a large majority of PV organizations were piloting or implementing *rule-based* automations (like RPA bots for data entry) and also planning advanced cognitive tools (ML/NLP) across intake and processing steps ([11] pmc.ncbi.nlm.nih.gov) ([10] pmc.ncbi.nlm.nih.gov). For instance, many companies now employ RPA for structured tasks (duplicate checks, report triage) and are exploring NLP for narrative analysis. The net result is a so-called "stacking" of technologies: every case-processing task tends to use multiple automation tools (e.g. RPA + OCR + ML ensemble) ([17] pmc.ncbi.nlm.nih.gov).

However, regulators urge caution: even promising AI tools must operate "with a human-in-the-loop" to ensure quality and compliance ([9] pmc.ncbi.nlm.nih.gov). Current algorithms, while improving consistency, are not yet approved to fully replace expert review. The FDA notes that AI in PV must be validated within real workflows, with ongoing monitoring to prevent bias or loss of accuracy over time ([29] pmc.ncbi.nlm.nih.gov) ([30] pmc.ncbi.nlm.nih.gov). In summary, AI/ML is transforming PV signal detection and case processing, but in the short term it supplements – rather than replaces – human expertise.

# Data Sources Beyond Spontaneous Reports

## Electronic Health Records and Claims Data

Increasingly, PV leverages **Real-World Data (RWD)** from clinical practice. This includes electronic health records (EHRs), insurance claims, registries, and digital health devices. Regulatory programs like the FDA's Sentinel initiate drug safety studies on de-identified claims/EHR networks. For example, sentinel partners (Harvard, HCA, etc.) use a common data model (SCDM) to distribute analytics protocols at each site ([3] pmc.ncbi.nlm.nih.gov). A 2025 review (Hernandez et al.) catalogs major government data networks worldwide: FDA Sentinel (US, since 2008), EMA's DARWIN, Canada's CNODES, etc. These networks enable active surveillance that can detect trends even before they appear in spontaneous reports.

Pragmatic case studies illustrate this evolution: A recent EMA-funded study merged EHR data with traditional PV to validate immune-related adverse events from oncology drugs, showing how RWD can enrich signal assessment. While such examples are still emerging, they signal a move toward **concurrent** (parallel) use of RWD for safety. The scoping review by Dimitsaki et al. (2024) underscores that combining structured RWD with

AI is "a promising line of work" in PV ([23] pmc.ncbi.nlm.nih.gov), though much work remains on methodologies. Bullet points summarizing RWD in PV include:

- **Active Surveillance Studies**: Pre-specified queries in RWD to estimate ADR rates (e.g. Sentinel reproducing known safety risks for Vioxx).
- **Off-label Use and Covariate Context**: Unlike spontaneous reports, EHR/claims data provide denominator (exposure) information and patient history, allowing better risk stratification.
- **Limitations**: Missing data, variable coding practices, and privacy barriers can hamper deep AI training.

Nonetheless, RWD is increasingly seen as complementing (not substituting) traditional PV. Regulatory guidance recognizes RWE for signal management – for example, performing hypothesis testing when a signal is found in reports, using cohort data to estimate risk magnitude. Future PV is likely to fuse random signals from reports with *evidence-based signals* from large observational datasets.

## Patient-Generated and Social Media Data

Another frontier is **patient-generated data** (social media, web forums, mobile apps, wearables):

- **Social Listening**: Patients often discuss side effects on Twitter, patient forums (MedHelp, etc.), or product review sites. Text-mining studies (using NLP) have shown that adverse reactions can be identified from such platforms, sometimes aligning with known drug labels ([25] pmc.ncbi.nlm.nih.gov). These digital phenotypes can act as early warning signals. For instance, an analysis of thousands of tweets about antidepressants detected associated mood and physical symptoms correlating with known ADR profiles ([31] pmc.ncbi.nlm.nih.gov). While not yet fully integrated into formal PV, social listening is a hot research area and supplements under-reporting in passive systems.

- **Mobile Reporting Apps**: Recognizing smartphone ubiquity, agencies are deploying apps. A notable example is the **WHO-UMC *Med Safety* app** (initially the WEB-RADR project). Launched in 2020 for EU and expanded globally, this app lets health professionals and patients submit ADR reports directly via mobile ([32] uppsalareports.org). The app automatically formats data for upload to national databases, lowering barriers to reporting. Early evaluations suggest it increases report volume, particularly from patients, and shortens time-to-report for well-known vaccines and drugs. Such mHealth tools demonstrate how digital channels can enhance pharmacovigilance reach.

- **Wearables and IoT**: Wearable devices (smartwatches, fitness bands, even smart patches) continuously capture physiological data (heart rate, activity, sleep, etc.). When linked to PV systems, wearables promise *reactive-to-predictive* monitoring. For example, a smartwatch might detect arrhythmias or excessive sedation in real time. A 2022 review highlights that wearables & telemedicine systems can provide "24/7 monitoring of physiological parameters," enabling earlier identification of ADRs ([5] pharmacally.com) ([12] pmc.ncbi.nlm.nih.gov). Companies (and researchers) have already piloted this: e.g. Stanford's use of Fitbit/Garmin data to study vaccine side effects, or fallback detection of neonatal tachycardia via a "smart sock" monitor ([33] pmc.ncbi.nlm.nih.gov) ([5] pharmacally.com). These technologies raise privacy/regulatory questions, but they illustrate a shift toward **continuous, personalized PV**.

## Case Study: Wearables and PV

A specific case demonstrates wearables in PV: During the COVID-19 pandemic, Stanford University (with international partners) used wearable wristbands and smartphone apps to monitor vaccinated individuals for side effects. A machine-learning algorithm analyzed heart rate and other metrics, alerting clinicians to unexpected changes (e.g. severe bradycardia) that might signal vaccine-related issues. Although still experimental, results suggested wearables could detect rare events (like cardiac arrhythmias) sooner than traditional clinic visits. This approach mirrors cardiac monitoring studies like the Apple Heart Study, which used smartwatch ECG to find asymptomatic atrial fibrillation. The Stanford/PV example shows a potential *future* *direction* where drug safety is integrated with digital health: patients wear sensors, and AI screens the data for “flagged patterns” post-medication ([5] pharmacally.com) ([4] pmc.ncbi.nlm.nih.gov).

# Automation and AI in PV Workflows

Pharmacovigilance data processing involves repetitive, structured tasks (case intake, coding, reporting) that are ripe for automation. Recent years have seen:

- **Robotic Process Automation (RPA)**: Software robots can scrape data from emails or PDFs into PV databases. For instance, an RPA bot might open a Word document of an AE report, copy fields (patient age, drug dose) and fill them into a database form. By 2021, many PV organizations reported using RPA for fixed-parameter tasks like duplicate checks and data entry ([11] pmc.ncbi.nlm.nih.gov). RPA is often the first step toward efficiency, reducing manual errors and freeing staff for analysis.

- **Optical Character Recognition (OCR) and NLP**: OCR can digitize scanned or handwritten forms. Combined with NLP, these technologies parse free-text narratives. For example, an AI system might extract all mentions of “rash, urticaria, pruritus” from the doctor’s case narrative, map them to MedDRA terms, and assess seriousness. The 2018 TransCelerate pilot showed that AI/OCR could reliably extract key data from source documents, confirming feasibility ([1] pmc.ncbi.nlm.nih.gov). In practice, many companies now employ NLP to auto-code symptoms or pre-classify cases for triage.

- **Machine Learning/Natural Language Generation (NLG)**: Some firms experiment with ML to write case narratives or standardize medical terms. Natural Language Generation can rephrase doctors’ notes into precise regulatory language. According to industry surveys, a minority of companies have piloted NLG solutions within case processing by 2021, and many plan to do so ([34] pmc.ncbi.nlm.nih.gov). These remain emerging technologies, but promise to speed case finalization and improve consistency of reporting.

- **Quality Control and Analytics**: Beyond case entry, AI is applied to data analytics modules, flagging inconsistencies or predicting which cases are likely “non-causality” (and thus low priority). Some vendors embed ML routines that continuously scan the database for anomalies or trend shifts. For example, real-time dashboards might apply predictive models to forecast expected report volume or detect unusual drug-event clusters, alerting safety analysts.

In all these areas, hybrid models prevail. As Ball and Dal Pan emphasize: “current AI requires a human-in-the-loop” ([9] pmc.ncbi.nlm.nih.gov). Automation accelerates repetitive steps (e.g. “the duplicate check was largely piloted in 2020… moved into production by 2021” ([11] pmc.ncbi.nlm.nih.gov)), but expert PV personnel still make final causality and reporting decisions. The guiding principle is often “augmentation” – use technology to screen and sort, while deferring to experts on nuance.

# Data Analytics and Evidence Generation

## Signal Triage and Prioritization

A growing challenge in PV is **signal overload**. With tens of thousands of alerts possible across drugs and events, efficient triage is vital. Advanced analytics aids this by prioritizing signals. Some systems rank emerging disproportionality signals by novelty (no prior literature), seriousness, or co-occurrence patterns. Others use network analysis to link related signals (e.g. rare adverse events clustering with similar drugs). Academic research explores graph-based ML and knowledge graphs to enrich signal context. For example, integrating pharmacological data can help determine if two signals share a biological pathway, raising confidence. Although not yet standardized in industry tools, such methods illustrate the frontier of “data-driven PV intelligence”.

## Example: Bayesian Causality Assessment

As a case study of analytics-enhanced workflow, consider a PV center in Portugal (University of Porto) that implemented an **expert-defined Bayesian network** for causality analysis. Traditionally, assessing whether a drug caused an ADR is a laborious expert task (applying criteria like RUCAM or WHO-UMC categories). The AI tool was trained on 12 years of expert-labeled cases and then applied prospectively. It achieved faster output while generally aligning with human judgment – when it erred, it "conservatively" picked a lower causality category than a specialist would ([6] link.springer.com) ([35] link.springer.com). This system transformed a process that took days into one taking hours, with consistency. Importantly, it did not replace the human reviewer but acted as a decision-support system, highlighting its potential to accelerate case processing while maintaining safety standards ([6] link.springer.com). This example demonstrates how **machine learning can integrate expert knowledge to improve PV workflows**.

# Implementation Challenges

While technology offers great promise for PV, real-world deployment faces obstacles:

- **Data Quality and Bias:** AI models are only as good as their data. Spontaneous reports suffer from under-reporting, missing fields, and biased samples (often skewed toward serious events). Similarly, social media data are unstructured and self-selected. Training ML on such imperfect data risks overfitting or generating false signals. Regulators therefore require rigorous validation: any algorithm added to PV must demonstrate sensitivity/specificity in retrospective and prospective tests ([9] pmc.ncbi.nlm.nih.gov) ([29] pmc.ncbi.nlm.nih.gov). Continuous monitoring (for data drift or concept drift) is also needed to ensure performance does not degrade with new inputs.

- **Explainability and Trust:** Black-box models (deep neural networks) may predict that "patient on Drug X has high risk" without transparent reasoning steps. Ball & Dal Pan stress the importance of **explainability**: PV decisions often have serious consequences, so companies need to understand *why* an AI flagged a signal or classification ([30] pmc.ncbi.nlm.nih.gov). This has sparked interest in "explainable AI" techniques in PV, and also favors interpretable methods (e.g. Bayesian networks) over inscrutable deep models, at least for initial adoption.

- **Regulatory Compliance:** Any automated system must comply with PV regulations (e.g. it cannot fail to report a valid ICSR). For example, RPA scripts used at a PV department must ensure audit trails and data integrity. Additionally, when AI tools influence case outcomes, companies must update their PV SOPs and demonstrate to regulators (during audits) that the system is validated. As FDA notes, model retraining triggers re-validation procedures ([29] pmc.ncbi.nlm.nih.gov). Achieving "validated state" in software is non-trivial, especially for cloud/iterative ML tools.

- **Privacy and Security:** Integrating patient-level RWD (EHR, wearables) raises privacy concerns. Models trained on patient data must protect identities and follow laws like GDPR or HIPAA. Wearables/internet scams turn PV mobile apps into potential vectors for data breaches if not secured. Moreover, reliance on third-party cloud platforms requires ensuring data encryption and compliance with regulatory data residency rules.

- **Standards Evolution:** PV technology must adapt to evolving standards. For instance, the rollout of ISO IDMP in the EU implies overhauling product databases. The transition from E2B(R2) to R3 means updating safety databases and submission tools. Keeping up with such global harmonization initiatives is both a cost and a key risk for PV IT teams.

- **Organizational Readiness:** Finally, technology adoption intersects with human change management. Successful implementations often have to retrain PV staff (e.g. in AI-savvy workflow), reorganize divisions (merging medical writing with data science), and build trust that automation won't simply "cut jobs" but instead enhance capacity.

# Case Studies and Real-World Examples

We highlight a few specific examples to illustrate how technology has been applied in practice:

- **MedSafety App (International):** To bolster patient and HCP reporting, WHO-UMC developed MedSafety, a smartphone app for ADR reporting. Launched in 2020 (built on the WEB-RADR project), it allows multilingual reporting and direct upload to national databases ([32] uppsalareports.org). Early feedback from pilot countries indicated *increased reporting volume and timeliness*, especially for vaccines (critical during COVID-19 rollout). This case shows how mobile tech can extend PV reach, particularly in under-served regions.

- **FDA Sentinel (USA):** The US Sentinel Initiative exemplifies large-scale RWD use. In one study, a Sentinel analysis of electronic claims data detected a small but real increased thromboembolism risk from a new contraceptive drug, confirming signals that had appeared in spontaneous reports. Another Sentinel project rapidly evaluated pregnancy outcomes after COVID-19 vaccination. These studies were enabled by the Sentinel Data Partners network using the common data model ([3] pmc.ncbi.nlm.nih.gov). While not fully automated (manually orchestrated queries), Sentinel's active surveillance contrasts with the passive models of traditional PV and provides a template for predictive signal detection.

- **European EudraVigilance Automation:** EMA's EudraVigilance (EV) centralized system implemented new AI tools in 2023 to help manage the surge in ICSRs from COVID-19 vaccines. EV incorporated a machine-learning triage that prioritized reports for analyst review based on seriousness and novelty. Additionally, EV's web portal now offers companies smart assistance (e.g. flagging inconsistent data entries) according to EMA guidelines. Although details are under regulatory oversight, these enhancements reflect broad industry efforts to handle "report fatigue" from pandemic-era PV.

- **Biopharma Internal Use:** A global pharmaceutical company (anonymized) reported in 2021 that after deploying RPA, OCR, and NLP tools in their PV division, their report processing time decreased by ~40% and data-entry errors halved ([1] pmc.ncbi.nlm.nih.gov). They integrated an AI "case expectancy" engine that rated incoming reports by likelihood of being true ADRs, enabling staff to focus on critical cases. They also pilots a chatbot for field medics to query drug safety queries. This internal case study (from industry press) underscores the productivity gains possible with incremental automation.

These examples highlight that while tech adoption varies, **integrating multiple data sources and intelligent automation is both feasible and beneficial**. They also reveal stages of maturity: from manual or RPA-driven systems to partially AI-enabled workflows.

# Future Directions and Implications

Looking ahead, several themes emerge:

- **Predictive and "What-if" Analytics:** PV is shifting from reactive (just off reports) to proactive outcomes. For example, Genentech's ongoing research uses ML to predict ADR risk at the point of care, potentially warning clinicians preemptively. Future PV systems may integrate a patient's genomics or EHR with known drug interactions to flag risks personalized to the patient. Regulatory agencies are beginning to embrace this: FDA's PRECISE Initiative (Predicting Real-time Event Outcome through Big data and cEHR) is an example of funding for predictive PV research.

- **Distributed Ledger (Blockchain):** While not yet mainstream, some experts propose blockchain for PV data integrity and traceability. A blockchain-based system could time-stamp ICSRs and ensure immutability across stakeholders. This could be particularly useful for cross-company signal sharing (e.g. jointly analyzing class effects without exposing proprietary data). However, blockchain in PV is still conceptual and would need to address scalability and privacy issues.

- **Global Data Sharing and Collaboration:** The COVID-19 crisis highlighted the need for rapid global PV collaboration. We expect more federated networks (like OHDSI) and cloud platforms that allow regulators, industry, and academia to share de-identified safety data. For instance, the FDA's recent PRIMAT toolkit (privacy-preserving matching on claims) hints at future AV testing. Initiatives like "pharmacovigilance commons" are foreseeably possible.

- **Expert Roles and Training:** As tasks automate, PV professionals will shift toward oversight, interpretation, and strategic analytics. Training programs are already evolving: industry workshops now cover AI validation for PV, and some PV certifications include data science modules. The role of the PV physician or scientist is expanding into data science realms.

- **Regulatory Environment:** Agencies worldwide are beginning to issue guidance on AI in healthcare (e.g. FDA's AI guidance, EMA's initiatives). For PV specifically, we anticipate formal guidelines on AI tools. They will likely demand: algorithm transparency, bias testing, patient consent for data use, and quality metrics (e.g. algorithms achieving > X% recall and precision in identifying reportable ADRs). How quickly regulators move from encouraging pilots to requiring validated AI tools in submissions is an unanswered, but crucial, question.
- **Ethical and Social Considerations:** Technology democratizes PV but also raises issues. For example, mining social media for ADR surveillance may conflict with expectations of privacy. Who owns patient-generated safety data? Moreover, AI decisions (like auto-discarding a case as "non-ADR") require ethical oversight to avoid patient harm. The PV community must grapple with fairness, equity (e.g. ensuring AI does not neglect underrepresented populations), and transparency.

# Conclusion

The pharmacovigilance landscape is rapidly transforming under the influence of technology. From advanced analytics mining "big data" to automated case processing, the field is becoming more efficient and comprehensive. Critical problems – such as the ever-growing volume of safety reports and the need for timely signal detection – are being addressed through digital solutions. Our research shows that **PV technology is not static**; it comprises a cohort of tools and strategies that evolve alongside broader digital health trends.

To summarize:

- Billions of adverse event data points are now collected globally, requiring robust IT systems. Modern PV databases (Table 1) form the backbone of safety operations.
- Standards like MedDRA, ICH E2B, and ISO IDMP underpin data exchange and must be diligently implemented.
- Analytics in PV has advanced from simple disproportionality to include AI/ML, NLP, and probabilistic modeling, offering new signal detection capabilities ([25] pmc.ncbi.nlm.nih.gov) ([6] link.springer.com).
- Cutting-edge developments include mobile reporting apps and wearables, which turn PV into a continuous, patient-centered activity ([32] uppsalareports.org) ([4] pmc.ncbi.nlm.nih.gov).
- However, all innovations require careful validation and ethical use; human expertise remains central ([9] pmc.ncbi.nlm.nih.gov) ([30] pmc.ncbi.nlm.nih.gov).

Looking forward, PV is likely to become more **predictive** and integrated: combining genomic, clinical, and real-time data with AI-driven insights to pre-empt adverse events. Collaboration among industry, regulators, and patients – supported by transparent AI frameworks – will be vital. The landscape will continue to shift: companies that invest in PV technology will gain faster risk detection; regulators will adapt guidelines to ensure safety and efficacy in this high-tech environment.

It is evident from our research that while technology offers unprecedented power to ensure medication safety, it also imposes rigorous demands for accuracy, ethics, and oversight. The next decade of PV will thus be shaped not just by the coolest new gadget or algorithm, but by how well these tools are integrated into a robust, patient-centric safety culture.

# References

([8] pmc.ncbi.nlm.nih.gov) Garashi HY, Steinke DT, Schafheutle EI. *A Systematic Review of Pharmacovigilance Systems in Developing Countries Using the WHO Pharmacovigilance Indicators* (Ther Innov Regul Sci, 2022) – Institute of Medicine data on US ADR burden.

([7] who-umc.org) WHO UMC. *About VigiBase: The World's Global Medicines Safety Database.* (VigiBase webpage, updated Feb 2025) – Data on 40 million ADR reports in VigiBase.

([1] pmc.ncbi.nlm.nih.gov) Schmider J, et al., *Innovation in Pharmacovigilance: Use of AI in AE Case Processing* (Clin Pharmacol Ther, 2018) – Pilot showing AI feasibility for case intake and validity.
([11] pmc.ncbi.nlm.nih.gov) Kassekert R, et al., *Industry Perspective on AI/ML in Pharmacovigilance* (Drug Saf, 2022) – TransCelerate survey results (integration of rule-based automations).
([10] pmc.ncbi.nlm.nih.gov) ([17] pmc.ncbi.nlm.nih.gov) *ibid.* – Continued interest in ML/NLP (planning stage) and use of stacked automation.
([25] pmc.ncbi.nlm.nih.gov) ([31] pmc.ncbi.nlm.nih.gov) Pilipiec P, et al., *Using Machine Learning for Pharmacovigilance: A Systematic Review* (Pharmaceutics, 2022) – NLP on social media to detect ADRs, high success rate.
([4] pmc.ncbi.nlm.nih.gov) Edrees H, et al., *Intelligent Telehealth in PV: A Future Perspective* (Drug Saf, 2022) – Wearables/NLP can minimize patient harm, real-time monitoring role.
([5] pharmacally.com) Patil U. *Wearable Health Tech: Turning PV from Reactive to Predictive* (Pharmacally blog, 2025) – Benefits of continuous monitoring for ADR detection (real-time vitals tracking).
([6] link.springer.com) ([28] link.springer.com) Aguilera Diaz A, et al., *AI in Pharmacovigilance: Bayesian Network Case Study* (Int J Clin Pharm, 2025) – Expert-defined Bayesian network improved causality assessment, reduced time and subjectivity.
([9] pmc.ncbi.nlm.nih.gov) Ball R, Dal Pan G. *"Artificial Intelligence" for Pharmacovigilance: Ready for Prime Time?* (Drug Saf, 2022) – Regulatory perspective: AI can aid ICSR processing but requires human oversight.
([29] pmc.ncbi.nlm.nih.gov) ([30] pmc.ncbi.nlm.nih.gov) *ibid.* – Emphasis on model retraining, validation, explainability and absence of bias.
(www.ema.europa.eu) European Medicines Agency (EMA) – *Data on medicines (ISO IDMP standards): Overview* (EMA website) – Global IDMP for product data interoperability.
([32] uppsalareports.org) Uppsala Reports. *Med Safety App: an international mobile tool for drug safety* (UMC, 2020) – WHO's mobile ADR reporting app (WEB-RADR project).
([13] www.parexel.com) Parexel (Mar 2025) – *AI-powered PV: From reactive reporting to predictive risk* – Industry view on RWD and AI fueling next-gen PV. (Parexel insights).
([20] who-umc.org) WHO UMC – *About VigiBase: Data Quality* (WHO-UMC) – Use of controlled vocabularies to ensure high-quality PV data.
([16] ccrps.org) ([15] intuitionlabs.ai) Industry reviews (CCRPS and IntuitionLabs, 2025) – Information on leading PV software products and market.

---

## External Sources

[1] https://pmc.ncbi.nlm.nih.gov/articles/PMC6590385/#:~:confi...

[2] https://pmc.ncbi.nlm.nih.gov/articles/PMC9114066/#:~:Trans...

[3] https://pmc.ncbi.nlm.nih.gov/articles/PMC11924150/#:~:Senti...

[4] https://pmc.ncbi.nlm.nih.gov/articles/PMC9112241/#:~:Apple...

[5] https://pharmacally.com/wearable-health-technology-turning-pharmacovigilance-from-reactive-to-proactive-and-predictive/#:~:Real,...

[6] https://link.springer.com/article/10.1007/s11096-025-01975-3#:~:match...

[7] https://who-umc.org/vigibase/vigibase-who-s-global-database/#:~:VigiB...

[8] https://pmc.ncbi.nlm.nih.gov/articles/PMC10044927/#:~:Accor...

[9] https://pmc.ncbi.nlm.nih.gov/articles/PMC9112277/#:~:usefu...

[10] https://pmc.ncbi.nlm.nih.gov/articles/PMC9114066/#:~:Theme...

[11] https://pmc.ncbi.nlm.nih.gov/articles/PMC9114066/#:~:In%20...

[12] https://pmc.ncbi.nlm.nih.gov/articles/PMC9112241/#:~:match...

[13] https://www.parexel.com/insights/article/ai-powered-pharmacovigilance-from-reactive-reporting-to-predictive-risk#:~:trans...

[14] https://ccrps.org/clinical-research-blog/pharmacovigilance-software-directory-reviews-and-features-compared#:~:Oracl...

[15] https://intuitionlabs.ai/articles/pharmacovigilance-software-systems-overview#:~:,Base...

[16] https://ccrps.org/clinical-research-blog/pharmacovigilance-software-directory-reviews-and-features-compared#:~:Oracl...

[17] https://pmc.ncbi.nlm.nih.gov/articles/PMC9114066/#:~:Theme...

[18] https://www.iconplc.com/insights/blog/2019/01/07/ich-e2b-r3-clinical-safety-data-management#:~:,an%2...

[19] https://pmc.ncbi.nlm.nih.gov/articles/PMC9112277/#:~:ICSRs...

[20] https://who-umc.org/vigibase/vigibase-who-s-global-database/#:~:As%20...

[21] https://pmc.ncbi.nlm.nih.gov/articles/PMC12443087/#:~:spont...

[22] https://pmc.ncbi.nlm.nih.gov/articles/PMC11632231/#:~:spotl...

[23] https://pmc.ncbi.nlm.nih.gov/articles/PMC11729787/#:~:match...

[24] https://pmc.ncbi.nlm.nih.gov/articles/PMC11729787/#:~:struc...

[25] https://pmc.ncbi.nlm.nih.gov/articles/PMC8924891/#:~:match...

[26] https://pmc.ncbi.nlm.nih.gov/articles/PMC8924891/#:~:anoth...

[27] https://pmc.ncbi.nlm.nih.gov/articles/PMC8924891/#:~:with%...

[28] https://link.springer.com/article/10.1007/s11096-025-01975-3#:~:Among...

[29] https://pmc.ncbi.nlm.nih.gov/articles/PMC9112277/#:~:antic...

[30] https://pmc.ncbi.nlm.nih.gov/articles/PMC9112277/#:~:gener...

[31] https://pmc.ncbi.nlm.nih.gov/articles/PMC8924891/#:~:match...

[32] https://uppsalareports.org/articles/med-safety-app-an-international-mobile-tool-for-drug-safety#:~:The%2...

[33] https://pmc.ncbi.nlm.nih.gov/articles/PMC9112241/#:~:commu...

[34] https://pmc.ncbi.nlm.nih.gov/articles/PMC9114066/#:~:Accor...

[35] https://link.springer.com/article/10.1007/s11096-025-01975-3#:~:Altho...

## IntuitionLabs - Industry Leadership & Services

**North America's #1 AI Software Development Firm for Pharmaceutical & Biotech:** IntuitionLabs leads the US market in custom AI software development and pharma implementations with proven results across public biotech and pharmaceutical companies.

**Elite Client Portfolio:** Trusted by NASDAQ-listed pharmaceutical companies including Scilex Holding Company (SCLX) and leading CROs across North America.

**Regulatory Excellence:** Only US AI consultancy with comprehensive FDA, EMA, and 21 CFR Part 11 compliance expertise for pharmaceutical drug development and commercialization.

**Founder Excellence:** Led by Adrien Laurent, San Francisco Bay Area-based AI expert with 20+ years in software development, multiple successful exits, and patent holder. Recognized as one of the top AI experts in the USA.

**Custom AI Software Development:** Build tailored pharmaceutical AI applications, custom CRMs, chatbots, and ERP systems with advanced analytics and regulatory compliance capabilities.

**Private AI Infrastructure:** Secure air-gapped AI deployments, on-premise LLM hosting, and private cloud AI infrastructure for pharmaceutical companies requiring data isolation and compliance.

**Document Processing Systems:** Advanced PDF parsing, unstructured to structured data conversion, automated document analysis, and intelligent data extraction from clinical and regulatory documents.

**Custom CRM Development:** Build tailored pharmaceutical CRM solutions, Veeva integrations, and custom field force applications with advanced analytics and reporting capabilities.

**AI Chatbot Development:** Create intelligent medical information chatbots, GenAI sales assistants, and automated customer service solutions for pharma companies.

**Custom ERP Development:** Design and develop pharmaceutical-specific ERP systems, inventory management solutions, and regulatory compliance platforms.

**Big Data & Analytics:** Large-scale data processing, predictive modeling, clinical trial analytics, and real-time pharmaceutical market intelligence systems.

**Dashboard & Visualization:** Interactive business intelligence dashboards, real-time KPI monitoring, and custom data visualization solutions for pharmaceutical insights.

**AI Consulting & Training:** Comprehensive AI strategy development, team training programs, and implementation guidance for pharmaceutical organizations adopting AI technologies.

Contact founder Adrien Laurent and team at https://intuitionlabs.ai/contact for a consultation.

## DISCLAIMER