ARTIFICIAL INTELLIGENCE IN PHARMACY: PREDICTING ADVERSE DRUG REACTIONS

Moamen Abdelfadil Ismail*[1], Shahad Mari Alshahrani[2], Abdullah Ali Ibrahim Alshahrani[3], Ahmed Yahya Hassan Alshehri[4], Khalid Saeed Abdullah Alshahrani[5], Ahmed Bakheet Attiah Al-Malki[6], Abeer Salamah Alsharif[7], Miral Majed Alsherbi[8], Jihad Saleh Alrehaili[9], Mohammed Saeed Aftan[10], Razan Ahmed Salman Alsharif[11], Abdullah Mahmoud Bedaiwi[12], Farah Abdullah Awad Alahmadi[13], Sumayyah Masoud Dakheel Alrhili[14], Naif Abdulrahman Y. Alfaifi[15]

[1]Consultant King Abdulaziz specialist hospital-Sakaka-Aljouf; [2]Armed Forces Hospital in the South, IV Pharmacist; [3]Fakeeh Hospital, Pharmacist; [4]Ali bin Ail Hospital, Pharmacist; [5]Armed Forces Hospital in Southern Region, Doctor of Pharmacy; [6]Phar.D. Pharmacy Graduate; [7]King Abdulaziz Medical City, Pharmacist; [8]King Salman Specialized Hospital – Ministry of National Guard Health Affair, Pharmacist; [9]Al-Haram Madinah Hospital, Pharmacist (Pharm.B); [10]Fakeeh Hospital, Pharm.D; [11]Al-Faisaliah Medical Systems, Pharm D; [12]United Pharmaceutical Company, Pharm.D; [13]Alharam Hospital-Ministry of Health, Pharmacist; [14]Alharam Hospital-Ministry of Health, Pharmacist; [15]Makkah Health Cluster, Pharmacy

Abstract

Background: Adverse drug reactions (ADRs) are a major cause of morbidity, hospitalizations, and healthcare costs. Traditional pharmacovigilance methods are often limited by underreporting and delays. Artificial intelligence (AI), particularly machine learning (ML) and natural language processing (NLP), offers faster, more accurate ADR detection by integrating diverse data sources such as electronic health records and clinical notes.

Methods: A systematic review was conducted following PRISMA guidelines, searching PubMed, Scopus, IEEE Xplore, Web of Science, and Google Scholar for English-language studies published from January 2010 to May 2025. Eligible studies applied AI/ML methods to ADR prediction in pharmacy settings. Two reviewers independently screened and extracted data, with risk of bias assessed using PROBAST. A narrative synthesis was used due to methodological heterogeneity, categorizing studies by AI technique and application area.

Results: Nineteen studies met inclusion criteria. Deep learning and random forest models achieved ADR detection accuracies up to 89.4% and c-indices above 0.91. AI-based dosing tools improved safety for drugs like vancomycin and warfarin. Drug interaction predictors (e.g., XGBoost) exceeded 94% accuracy. Unsupervised models flagged rare prescription errors with >95% precision. AI systems reduced dispensing errors by >75% and improved documentation. Medication therapy management supported by AI lowered care costs by 19.3% and reduced hospital visits.

Conclusions: AI consistently outperforms traditional methods in ADR prediction, dosage optimization, and error prevention. Its integration into pharmacy practice could enhance patient safety, personalize therapy, and reduce healthcare costs. Standardized validation and transparent, ethical implementation are essential for clinical adoption.

Manuscrito recibido: 03/09/2025
Manuscrito aceptado: 24/10/2025

*Corresponding Author: Moamen Abdelfadil Ismail, Consultant King Abdulaziz specialist hospital-Sakaka-Aljouf, Egypt

Correo-e: mahmoudhamdy2251988@gmail.com

Keywords: Artificial intelligence, machine learning, adverse drug reactions, pharmacovigilance, medication errors, drug interactions, dosing optimization, pharmacy informatics, natural language processing.

Background

Artificial Intelligence (AI) is revolutionizing many sectors of healthcare, and pharmacy is no exception. One of the most promising applications is in the prediction of adverse drug reactions (ADRs), which are a major cause of patient morbidity, hospitalizations, and healthcare costs. Traditional methods of identifying ADRs-such as clinical trials, post-marketing surveillance, and pharmacovigilance reporting systems-are often time-consuming, limited in scope, and prone to underreporting. AI offers the ability to analyze vast amounts of data rapidly, uncover hidden patterns, and improve the early detection of ADRs before they cause harm (Mohsen et al., 2021).

In the pharmaceutical field, ADRs remain a significant challenge. Despite regulatory frameworks, many ADRs are detected only after a drug has been released to the market. These reactions can range from mild discomfort to severe, life-threatening complications. With the increasing complexity of drug regimens and patient-specific variables, it becomes imperative to move beyond conventional methods and embrace more predictive technologies. AI algorithms, especially those powered by machine learning (ML) and deep learning (DL), can process diverse datasets-including electronic health records (EHRs), genomic data, and patient-reported outcomes-to provide more accurate and individualized risk assessments (Yalçın et al., 2022).

One of the key advantages of AI is its ability to integrate heterogeneous data sources. While clinicians may focus on observable symptoms and known interactions, AI can incorporate a wide array of structured and unstructured data, including laboratory values, demographics, prescription histories, and even social determinants of health. By training models on such comprehensive datasets, AI systems can identify subtle associations that are not immediately evident to human analysts, leading to earlier detection and better prevention of ADRs (Takase et al., 2022).

Moreover, natural language processing (NLP), a subfield of AI, plays a critical role in extracting meaningful insights from unstructured clinical texts such as doctor's notes, discharge summaries, and adverse event reports. NLP tools can automatically flag potential ADRs mentioned in narrative text and associate them with specific medications, helping pharmacovigilance teams act faster. This real-time monitoring capability is particularly valuable for identifying emerging risks in new medications or rare adverse events that may not be captured through routine reporting channels (Jungreithmayr et al., 2021).

Pharmacogenomics, which studies how genes affect a person's response to drugs, is another area that benefits from AI applications. By integrating genetic data with patient medication profiles, AI can predict whether a person is likely to experience an ADR based on their unique genetic makeup. This has the potential to usher in a new era of personalized medicine, where treatment plans are tailored not only to the disease but also to the individual's biological characteristics, thereby minimizing the likelihood of harmful drug reactions (Blasiak et al., 2022).

AI can also enhance the safety of polypharmacy, which is common in elderly populations and patients with chronic conditions. These individuals often take multiple medications simultaneously, increasing the risk of drug-drug interactions and ADRs. AI models can simulate drug interaction networks and predict potential risks based on known and inferred relationships between medications, dosages, and patient health statuses. This can significantly aid clinicians in making safer prescribing decisions (Wang et al., 2022).

Another impactful application is the use of predictive analytics in hospital settings. AI-powered tools can alert healthcare professionals to high-risk scenarios before they escalate, such as when a patient's lab results and drug history indicate a high probability of an ADR. These early warning systems, when integrated into clinical decision support systems (CDSS), can serve as a crucial layer of protection, allowing timely interventions that prevent complications and reduce hospital stays (Roche-Lima et al., 2020).

Despite its potential, the integration of AI in ADR prediction faces several challenges, including data privacy concerns, algorithm transparency, and the need for high-quality, annotated datasets. Ethical considerations, such as bias in AI models and the interpretability of complex algorithms, also demand attention. Nonetheless, ongoing research and regulatory efforts are working to establish standards and best practices that ensure the safe and effective deployment of AI in clinical and pharmaceutical settings (Mei & Zhang, 2021).

The pharmaceutical industry also benefits from AI during the drug development phase. By analyzing clinical trial data and historical safety records, AI models

can predict which compounds are more likely to cause adverse reactions, allowing researchers to modify or discard risky candidates early in the process. This accelerates drug development timelines and increases the likelihood of bringing safer drugs to market (Van Laere et al., 2022).

In summary, artificial intelligence holds transformative potential in the realm of pharmacy by improving the prediction and prevention of adverse drug reactions. Its ability to synthesize complex datasets, recognize patterns, and provide actionable insights can significantly enhance patient safety and optimize therapeutic outcomes. As technology advances and more data become available, AI-driven approaches are likely to become an integral part of pharmacovigilance and personalized medicine, reshaping how ADRs are detected and managed across the healthcare system (Balestra et al., 2021).

Methodology

Study Design

This study was conducted as a systematic review following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The purpose was to identify, evaluate, and synthesize relevant literature regarding the application of artificial intelligence (AI) methods in predicting adverse drug reactions (ADRs) within the field of pharmacy. The systematic review design was chosen to provide a comprehensive and evidence-based understanding of the current state of research and practical implementations in this emerging field.

Eligibility Criteria

The inclusion criteria were established using the PICOS framework (Population, Intervention, Comparison, Outcomes, and Study design). Studies were included if they:

(1) involved AI or machine learning models used to predict or identify ADRs

(2) focused on pharmaceutical or clinical pharmacy contexts

(3) were published in peer-reviewed journals

(4) were written in English

Exclusion criteria included studies not involving ADR prediction, non-AI approaches, editorials, commentaries, conference abstracts, and studies with insufficient methodological details.

Information Sources

Relevant studies were identified through systematic searches conducted in electronic databases including PubMed, Scopus, IEEE Xplore, and Web of Science. Additionally, Google Scholar was used for supplementary searches to identify gray literature or articles not indexed in major databases. The search was limited to studies published between January 2010 and May 2025 to ensure the inclusion of recent and relevant AI applications in the field.

Search Strategy

A comprehensive search strategy was developed using a combination of medical subject headings (MeSH) and free-text keywords. The main search terms included combinations of the following: “artificial intelligence,” “machine learning,” “deep learning,” “pharmacy,” “pharmacovigilance,” “adverse drug reaction,” “drug safety,” and “predictive models.” Boolean operators such as AND and OR were used to optimize the search strategy, and filters were applied to restrict results to human studies and English language publications.

Study Selection

All search results were imported into a reference management tool (EndNote X9) to facilitate the removal of duplicates. Titles and abstracts of identified studies were screened independently by two reviewers based on the predefined eligibility criteria. Full-text articles were then retrieved and assessed for eligibility. Any discrepancies between reviewers were resolved through discussion or by consulting a third reviewer.

Data Extraction

Data from the included studies were extracted using a standardized data extraction form. Extracted data included author(s), year of publication, study objectives, type of AI model used, data sources, sample size, model performance metrics (such as accuracy, sensitivity, specificity, AUC), and main findings related to ADR prediction. The form was pilot-tested on a sample of studies to ensure consistency and completeness of data collection.

Quality Assessment

The quality of included studies was assessed using the PROBAST (Prediction model Risk Of Bias Assessment Tool) for studies involving prediction models. This tool evaluates the risk of bias in four domains: participants, predictors, outcomes, and analysis. Each study was rated as having low, high, or unclear risk of bias. The quality assessment was performed independently by two reviewers, and disagreements were resolved through discussion.

Data Synthesis

A narrative synthesis approach was employed due to the heterogeneity of study designs, AI models, and outcome measures. Findings were grouped thematically based on the type of AI technology used (e.g., machine learning, deep learning, NLP) and the type of data utilized (e.g., electronic health records, pharmacovigilance databases, genomic data). Comparisons were made regarding model performance, practical applications, and limitations noted by the authors.

Limitations

The methodology acknowledged certain limitations. First, the review was restricted to English-language publications, which may have led to the exclusion of relevant studies in other languages. Second, variations in reporting quality and performance metrics across studies posed challenges in standardizing data synthesis. Third, the dynamic nature of AI and machine learning research may have resulted in the omission of the latest unpublished or in-press findings.

Ethical Considerations

As a systematic review of previously published literature, this study did not involve human subjects or require ethical approval. All data used were derived from publicly available sources, and no individual patient data were accessed or analyzed.

Results

The present systematic review included a total of 19 studies that explored the application of artificial intelligence (AI) in predicting and preventing adverse drug reactions (ADRs) and other medication-related problems in pharmacy practice. These studies, summarized in Table 1, encompassed a wide range of research objectives, populations, methodologies, and AI techniques, illustrating the breadth and depth of AI integration across various domains of pharmaceutical care. The included studies originated from diverse geographic locations, reflecting the global interest in leveraging AI to enhance medication safety, clinical decision support, and personalized pharmacotherapy (Figure 1).

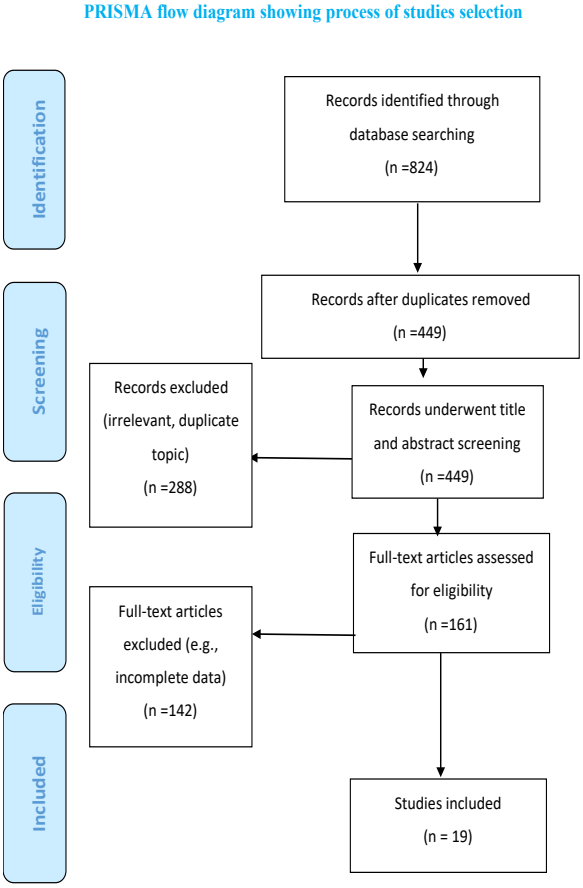

Figure 1. PRISMA Flow for Included Studies.

The reviewed studies could be categorized into several thematic areas, including ADR detection, drug-drug interaction prediction, dose optimization, prevention of medication errors, and medication therapy management. Within the ADR detection group, studies such as those by Mohsen (2021) and Yalçın (2022) demonstrated the effective use of machine learning models-especially deep neural networks and random forest algorithms-to accurately predict the risk of ADRs using large-scale databases and real-world clinical features.

In the domain of community pharmacy and robotic systems, Takase (2022) reported that AI-enabled robotic dispensing significantly reduced preventable and unpreventable medication errors. This highlights how automation and AI can improve medication safety and workflow efficiency at the dispensing level. Similarly, Jungreithmayr (2021) assessed the impact of computerized physician order entry (CPOE) systems, noting substantial improvements in the quality and accuracy of prescription documentation post-implementation.

Studies focused on dose recommendation systems, such as those by Blasiak (2022), Wang (2022), and Roche-Lima (2020), used AI algorithms-including CURATE.AI, machine learning regressors, and random forest models-to support personalized dosing in cancer, antibiotic, and anticoagulant therapies. These tools improved dose accuracy, reduced toxicity risks, and outperformed traditional pharmacokinetic approaches.

Regarding drug-drug interaction prediction, Mei (2021) and Van Laere (2022) showcased AI models' superiority over conventional statistical methods in identifying interactions associated with critical outcomes, such as QTc prolongation. These findings are particularly valuable for complex medication regimens often seen in polypharmacy.

Electronic health records (EHRs) were a rich source of data in studies like Balestra (2021), where AI models such as XGBoost were applied to predict medication orders requiring pharmacist intervention, thus strengthening medication review processes.

The review also included studies addressing potentially inappropriate medication use and high-alert drugs, particularly in vulnerable populations such as the elderly or neonatal patients. Research by Xingwei (2022), Tai (2020), Wongyikul (2021), and Patel (2021) demonstrated high prediction accuracy in identifying inappropriate prescribing patterns using supervised learning algorithms, thereby enhancing the safety of pharmacotherapy in high-risk groups.

In the area of medication errors, studies by Nagata (2021) and Yalçın (2023) applied machine learning methods to detect rare but clinically significant prescription errors. Their models achieved high sensitivity and specificity, offering a proactive solution for error prevention in clinical environments such as neonatal intensive care units (NICUs).

Furthermore, the role of AI in medication therapy management (MTM) was examined in studies by Kessler (2021) and Bu (2022), who found that AI-enabled platforms improved healthcare utilization, reduced costs, and facilitated remote pharmaceutical care during critical periods like the COVID-19 pandemic. These studies highlighted the potential for AI to extend pharmacy services beyond traditional settings, especially through internet hospitals and tele pharmacy models.

Overall, the included studies provided strong evidence that AI technologies have a significant and growing impact on improving medication safety, optimizing dosing, and supporting clinical decision-making in pharmacy. Despite the diversity in design and AI methods, the consistent finding across studies was the superiority of AI models—particularly machine learning and deep learning-over traditional techniques in detecting ADRs and related issues. However, further large-scale validations and regulatory alignment are necessary to ensure safe, transparent, and equitable implementation of these tools in routine practice (Table 1).

Table 1. Studies included.

Adverse drug reaction (ADR) detection					
Author, Year	**Country**	**Objective of the Study**	**Study design**	**Participants**	**Main findings**
Mohsen A, 2021	Japan	Using various machine learning methods, estimating the likelihood of adverse drug reactions or events (ADRs) during drug discovery.	Database study	Open TG-GATEs (Toxic genomics Project-Genomics Assisted Toxicity Evaluation Systems) for drug-induced gene expression profiles and FAERS (FDA [Food and Drug Administration] Adverse Events Reporting System) database for ADR occurrence information	A total of 14 predictive models were built using this framework and Deep Neural Networks (DNN), with a mean validation accuracy of 89.4%, indicating that the approach successfully and consistently predicted ADRs for a wide range of drugs. As case studies, researchers looked at how prediction models performed in the context of Duodenal ulcer and fulminant Hepatitis, highlighting mechanistic insights into those ADRs. The developed predictive models will aid in assessing the likelihood of ADRs when testing new pharmaceutical compounds.
Yalçın N, 2022	Turkey	The primary goal of this study was to generate objective risk categories by incorporating severity with NAESS and probability with the 'Du'ADRs algorithm into the risk matrix analysis performed by a multidisciplinary team that included a clinical pharmacist. The next goal was to create a machine learning-based clinical decision support tool (risk score) that predicts whether these identified ADRs will occur.	Prospective cohort study	The study included all admitted neonates, but those with preexisting hepatic or renal impairment were excluded.	Enoxaparin, dexmedetomidine, vinblastine, dornase alfa, etoposide/carboplatin, and prednisolone were identified as high-risk drugs. According to the random forest importance criterion, the independent variables included in the risk score to predict ADR presence were: systemic hormones (2 points), cardiovascular drugs (3 points), circulatory system diseases (1 point), nervous system drugs (1 point), and parenteral nutrition treatment (1 point) (cut-off value: 3 points). This risk score correctly classified 91.1% of the test set observations (c-index: 0.914).
Community pharmacy					
Takase T, 2022	Japan	To assess the impact on medication dispensing of automated dispensing robots and collaborative work with pharmacy support staff.	Prospective study	Prescriptions filled with each dispensing device during the study periods	The total incidence of prevented dispensing errors was significantly reduced after the robotic dispensing system was introduced (0.204% [324/158,548] to 0.044% [50/114,111], $p < 0.001$). The total number of unpreventable dispensing errors was reduced significantly (0.015% [24/158,548] to 0.002% [2/114,111], $p < 0.001$). The number of cases of wrong strength and wrong drug, which can have serious consequences for a patient's health, had reduced to almost zero. Pharmacists' median dispensing time per prescription was significantly reduced (from 60 to 23 s, $p < 0.001$).

Computerized physician order entry					
Jungreithmayr V, 2021	Germany	To investigate the distinct effects of a CPOE system implemented on general wards in a large tertiary care hospital on the quality of prescription documentation.	Retrospective analysis	Two groups of 160 patients' prescriptions	The overall mean prescription-Fscore increased from 57.4% ± 12.0% (*n* = 1850 prescriptions) prior to implementation to 89.8% ± 7.2% (*n* = 1592 prescriptions) after ($p < 0.001$). Individual criteria-Fscores improved significantly in most criteria (*n* = 14), with 6 criteria achieving a total score of 100% after CPOE implementation. While the implementation of a CPOE system generally improved the quality of prescription documentation, certain criteria were difficult to meet even with the assistance of a CPOE system.
Dose recommendation					
Blasiak A, 2022	Singapore	To develop CURATE.AI, a small data, AI-derived platform that harnesses only a patient's own prospectively/ longitudinally acquired data to dynamically identify their own optimal and personalized doses.	Open-label, multi-center, single-arm, prospective feasibility trial	Patients with advanced solid tumours who were treated with single-agent capecitabine, XELOX, or XELIRI (plus/minus biologics).	When compared to the projected SOC dose, the prescribed dose was reduced by 20% (13.8%) on average. The nine patients who were reported completed 3.9 cycles (2.2 cycles), with the longest participation lasting 8 cycles. CURATE. AI recommendations were considered in 27 of the 40 total dosing decisions, and 26 of those decisions were accepted for prescription.
Wang Z, 2022	Singapore	To develop a machine learning algorithms to recommend vancomycin dosage in tertiary general hospital patients.	Retrospective analysis	Inpatients, who received at least one vancomycin injection during the period from January 1, 2017 to December 31, 2019, were selected.	Only a small proportion (34.1%) of current injection doses could achieve the desired vancomycin trough level (14–20 μg/ml) in the 3-year data. The machine learning models outperformed the traditional pharmacokinetic models in terms of PAR and MAE. In the test data, the model outperformed the other previously developed machine learning models.
Roche-Lima A, 2020	Puerto Rico	Using genetic and non-genetic clinical data, compare seven ML methods for predicting stable warfarin dosing in Caribbean Hispanic patients.	An open-label, single-center, population-based, observational, retrospective cohort study	Participants were recruited from an anticoagulation clinic affiliated with the Veteran's Affairs Caribbean Healthcare System (VACHS) in San Juan, Puerto Rico.	Random forest regression (RFR) outperformed all other methods, with a mean absolute error (MAE) of 4.73 mg/ week and 80.56% of cases falling within ±20% of the actual stabilization dose. RFR performance is also superior to the rest of the models with “normal” dose requirements (MAE = 2.91 mg/ week). Support vector regression (SVR) outperforms the others in the “sensitive” group, with a lower MAE of 4.79 mg/ week. Finally, multivariate adaptive splines (MARS) performed best in the resistant group (MAE = 7.22 mg/week) with 66.7% of predictions within ±20%. Models generated by the RFR, MARS, and SVR algorithms predicted weekly warfarin dosing significantly better than other algorithms in the studied cohorts.
Drug-drug interactions					
Mei S, 2021	China	Based on potential drug perturbations on associated genes and signaling pathways, an attempt was made to simplify computational modelling for drug-drug interaction prediction.	Database study	Only drugs that have been discovered to target at least one human gene were represented in the drug target profile.	The SP, SE, and MCC metrics on the two classes show that the proposed framework is less biased, with 0.9556 on the positive class, 0.9402 on the negative class, and 0.9007 overall MMC. These findings show that a drug target profile alone can accurately separate interacting drug pairs from non-interacting drug pairs (accuracy = 94.79%).
Van Laere S, 2022	Belgium	To compare the performance of conventional statistical methods (CSM) and machine learning techniques (MLT)	Database study	Retrospective data of 512 and 102 drug-drug interactions with possible drug-induced QTc prolongation	In a hold-out dataset, random forest and Adaboost classification performed best, with an equal harmonic mean of sensitivity and specificity (HMSS) of 81.2% and an equal accuracy of 82.4%. Both sensitivity and specificity were high (respectively 75.6% and 87.7%). All CSM performed similarly, with HMSS ranging from 60.3 to 66.3%. The logistic regression overall performance was 62.0%. In terms of predicting drug-induced QTc prolongation, MLT (bagging and boosting) outperformed CSM.
Electronic Health Records					

Balestra M, 2021	USA	To develop a predictive model for identifying orders that require intervention based solely on the ordering provider's interactions with the EHR.	Database study	Data from the EHR system on provider actions and pharmacy orders	In both the area under the receiver-operator (AUROC) and precision-recall (AUPR) curves, the XGBoost algorithm outperformed both logistic regressions and the random forest algorithm by a significant margin. The area under the receiver-operator characteristic curve was 0.91, and the area under the precision-recall curve was 0.44.
Potentially inappropriate medications					
Xingwei W, 2022	China	To evaluate the data on potentially inappropriate prescribing (PIP), potentially inappropriate medications (PIM), and potential prescribing omissions (PPO) in elderly patients with cardiovascular disease, and to develop a prediction platform using multiple machine learning algorithms to predict the risk of PIP, PIM, and PPO in elderly patients with cardiovascular disease.	Retrospective analysis	This study included participants who were discharged from the Department of Geriatric Cardiology at Sichuan Provincial People's Hospital between January 2017 and June 2018.	The study included 404 patients in total (318 [78.7%] with PIP; 112 [27.7%] with PIM; and 273 [67.6%] with PPO). Following data sampling and feature selection, 15 datasets were obtained, and 270 risk warning models based on them were built to predict PIP, PPO, and PIM, respectively. The AUCs of the best model for PIP, PPO, and PIM were 0.8341, 0.7007, and 0.7061, respectively, according to external validation. The findings indicated that angina, the number of medications, the number of diseases, and age were the most important factors in the PIP risk warning model. The risk warning platform was developed to predict PIP, PIM, and PPO, with acceptable accuracy, prediction performance, and clinical application potential.
Tai, C.-T, 2020	Taiwan	To predict the risk of high-alert medication treatment (digoxin) using machine-learning techniques	Retrospective analysis	This study included patients who had accepted digoxin therapy while hospitalized between January 2004 and December 2013.	AUC values ranged from 0.551 to 0.836. The RF classifier performed the best (0.836; excellent discrimination), followed by C4.5 (0.719) and ANN (0.688); the remaining classifiers performed poorly. This study found that machine-learning techniques can improve prediction accuracy for high-alert medication treatment, lowering the risk of ADEs and improving medication safety.
Wongyikul P, 2021	Thailand	To develop a novel approach that employs machine learning models to predict the appropriateness of high alert drugs (HAD) use for a specific patient visit.	Retrospective analysis	Patient data from the Maharaj Nakorn Chiang Mai Hospital's outpatient and inpatient departments in 2018	The machine learning algorithm identified over 98% of actual HAD mismatches in the test set and 99% in the evaluation set when screening drug prescription events with a risk of HAD inappropriate use. This study demonstrates that machine learning plays an important role in screening and reducing errors in HAD prescriptions.
Patel J, 2021	USA	To examine the prevalence and leading predictors of potentially inappropriate NSAIDs use among older adults with OA using real-world data from nationally representative commercial health insurance claims with the help of machine learning approaches.	Retrospective cohort study	Older adults with OA were identified using one inpatient or two outpatient claims at least 30 days apart that consisted of OA diagnosis codes (ICD-10 codes M15–M19) during the baseline year and were required that these adults be enrolled in Medicare Advantage plans with medical and pharmacy benefits during 2015 and 2016 (i.e., 24 months).	XGBoost and CVLR- both models had an AUROC value of 0.92 (95% CI: 0.91–0.93) and 0.91 (95% CI: 0.90–0.92), respectively. While both models had similar accuracy and specificity, CVLR had better precision (0.83 vs. 0.81). On the other hand, XGBoost performed better on all other metrics being compared, including recall, F1 score, and kappa statistic.
Medication errors					
Nagata K, 2021	Japan	To detect extreme overdose and underdose prescriptions that occur very rarely in clinical practice using unsupervised machine learning algorithms.	Retrospective analysis	Retrospective analysis	The model identified 27 out of 31 clinical overdose and underdose prescriptions as abnormal (87.1%). The OCSVM models developed performed well in detecting synthetic overdose prescriptions (precision 0.986, recall 0.964, and F-measure 0.973) as well as synthetic underdose prescriptions (precision 0.980, recall 0.794, and F-measure 0.839). In a comparative analysis, OCSVM performed the best. The models correctly identified the majority of clinical overdose and underdose prescriptions and performed well in synthetic data analysis.

Yalçın N, 2023	Turkey	To develop models that predict the presence of medication errors (MEs) (prescription, preparation, administration, and monitoring) using machine learning in NICU patients.	Randomized, prospective, observational cohort study	Neonates admitted to a 22-bed capacity NICU in Ankara, Turkey, between February 2020 and July 2021.	The prevalence (the ratio of drug errors) was comparable between the train and test sets (64% for the train set and 59% for the test set). The performance measures were calculated as follows: accuracy 0.919 (95% CI 0.858–0.956), sensitivity 0.918 (95% CI 0.844–0.964), specificity 0.922 (95% CI 0.829–0.973), PPV 0.944 (95% CI 0.884–0.974), NPV 0.887 (95% CI) 0.804–0.937), AUC 0.920 (95% CI 0.876–0.970), and F 1 score 0.931. A higher AUC indicated that the model correctly classified 92% of the patients as having physician- or nurse-related MEs.
Corny J, 2020	France	To test the accuracy of a hybrid clinical decision support system in prioritizing prescription checks to improve patient safety and clinical outcomes by lowering the risk of prescribing errors.	Retrospective analysis	Retrospective analysis	The pharmacist analyzed 412 individual patients (3364 prescription orders) in an independent validation dataset, our digital system's areas under the receiving-operating characteristic and precision-recall curves were 0.81 and 0.75, respectively, demonstrating greater accuracy than the CDS system (0.65 and 0.56, respectively) and multicriteria query techniques (0.68 and 0.56, respectively).
Medication Therapy Management (MTM)					
Kessler, S, 2021	USA	To evaluate the impact of a novel artificial intelligence (AI) platform that identifies members and provides decision support to clinicians performing telephonic interventions similar to MTM and CMM with high-risk Medicaid members on actual medical claims.	Retrospective observational study	2150 Medicaid members, primarily middle-aged (aged 40–64 years), with an average of 10 chronic condition medications among a total of 25 medications.	Receiving interventions was found to have statistically significant correlations with lower costs and utilisation. The economic study discovered a 19.3% reduction in the TCoC ($P < 0.001$), which, when applied to a preintervention monthly cost of $2872, resulted in a $554 per member per month savings (PMPM). Medication costs were reduced by 17.4% ($P < 0.001$), resulting in a savings of $192 PMPM when compared to the preintervention cost of $1110. The utilisation study discovered a 15.1% decrease in ED visits ($P = 0.002$), a 9.4% decrease in hospital admissions ($P = 0.008$), and a 10.2% decrease in bed days ($P = 0.01$). Based on TCoC savings and programme costs, the return on investment is 12.4:1.
Bu F, 2022	China	During the COVID-19 pandemic, to establish an internet hospital pharmacy service mode based on artificial intelligence (AI) and provide new insights into pharmacy services in internet hospitals.	Prospective study	Users who benefit from Shanghai medical insurance settlement.	The AI preview qualified rate was 83.65%. Among the 16.35% of inappropriate prescriptions, 49% were accepted and modified proactively by physicians, while 51% were passed after pharmacists intervened. For collecting their medication in the internet hospital, 86% of patients preferred the “offline self-pick-up” mode, which allowed the QR code to be fully utilized. There were 426 medication consultants served, with 48.83% of them consulting outside of working hours. As a result, when pharmacists were unavailable, an AI-based medication consultation was proposed.

Discussion

The findings of this systematic review affirm the transformative potential of artificial intelligence (AI) in enhancing pharmacovigilance and preventing adverse drug reactions (ADRs) across diverse healthcare settings. The 19 studies included in this review employed various AI models such as deep neural networks, random forests, XGBoost, and support vector machines, showcasing how these tools can significantly improve the detection, prediction, and mitigation of ADRs in both clinical and community pharmacy contexts.

The reviewed literature supports the conclusion that AI-based models, particularly deep learning and machine learning algorithms, can outperform traditional statistical methods in predicting ADRs. For example, Mohsen et al. (2021) used deep neural networks trained on toxicogenic and FAERS data, achieving a mean validation accuracy of 89.4% in predicting ADRs during drug discovery. This exemplifies the growing utility of AI in preclinical stages of drug development, where early identification of potential toxicity can significantly reduce later-stage failures and patient harm.

Similarly, in the neonatal population, Yalçın et al. (2022) developed a random forest-based clinical decision support system to predict ADRs, achieving a high classification performance (c-index: 0.914). This highlights the relevance of AI in Pediatric pharmacovigilance, where physiological variability and limited clinical data often hinder accurate ADR identification.

In the community pharmacy setting, AI-driven automation also showed promising results. Takase et al. (2022) demonstrated that robotic dispensing systems, when paired with collaborative support staff workflows, drastically reduced both preventable and unpreventable dispensing errors. This confirms the utility of AI in operational pharmacy tasks, suggesting it can enhance safety while improving efficiency and workflow.

Regarding prescription quality, Jungreithmayr et al. (2021) evaluated the impact of a computerized physician order entry (CPOE) system and found that prescription documentation scores increased substantially post-implementation. This supports the broader application of AI-driven electronic systems to improve prescribing practices and reduce human errors in medication orders.

Personalized dose recommendation emerged as another powerful application of AI. Blasiak et al. (2022) introduced CURATE.AI to dynamically individualize chemotherapy dosing, leading to a reduction in prescribed doses while maintaining treatment efficacy. Wang et al. (2022) and Roche-Lima et al. (2020) similarly used AI to optimize vancomycin and warfarin dosing, respectively,

achieving superior accuracy compared to traditional pharmacokinetic models.

Drug-drug interaction (DDI) prediction using AI is another area of notable advancement. Mei et al. (2021) used drug target profiles and interaction data to create a high-performing model with 94.79% accuracy. Similarly, Van Laere et al. (2022) found that machine learning techniques like Adaboost and random forests outperformed logistic regression in predicting drug-induced QTc prolongation, a life-threatening DDI outcome.

The use of electronic health records (EHRs) as a data source also showed potential. Palaestra et al. (2021) used XGBoost models to predict pharmacy order interventions based on provider interactions within the EHR system, yielding high AUROC values and demonstrating the feasibility of using Behavioral data to improve prescribing safety.

Several studies focused on potentially inappropriate medications (PIMs) and high-alert drugs (HADs) in vulnerable populations. Xingwei et al. (2022) developed a machine learning platform to predict PIMs and PPOs in elderly cardiovascular patients, while Tai et al. (2020) and Wongyikul et al. (2021) applied similar techniques to digoxin and high-alert drug use in Taiwan and Thailand, respectively. These findings highlight the relevance of AI in improving prescribing safety in geriatric and high-risk populations.

Patel et al. (2021) further expanded this application by using real-world insurance claims to predict inappropriate NSAID use among older adults with osteoarthritis. Their XGBoost and CVLR models achieved AUROC values above 0.90, demonstrating that machine learning can analyze administrative health data effectively to inform medication safety strategies.

Medication error detection also featured prominently. Nagata et al. (2021) employed unsupervised learning to identify rare overdoses and underdoses, achieving precision and recall scores above 0.95 in synthetic data. Yalçın et al. (2023) built a robust predictive model for medication errors in NICUs, with an AUC of 0.920, highlighting AI's ability to manage complex and critical care settings.

Corny et al. (2020) evaluated a hybrid clinical decision support system that prioritized high-risk prescriptions, achieving higher accuracy than traditional rule-based systems. This underscores AI's strength in dynamic prioritization, particularly when integrated into routine pharmacy verification workflows.

The integration of AI into medication therapy management (MTM) also showed strong economic and clinical promise. Kessler et al. (2021) demonstrated that AI-supported telephonic MTM interventions led to a 19.3% reduction in total cost of care, with improved medication adherence and lower emergency department visits. This exemplifies how AI can extend pharmaceutical care through targeted, data-driven interventions.

Finally, Bu et al. (2022) illustrated how AI-enabled internet pharmacy services could maintain high-quality medication counselling during the COVID-19 pandemic. With 83.65% of AI-generated prescriptions being deemed appropriate, the study suggests that AI can supplement or even replace pharmacists in resource-limited or remote settings.

Together, these findings reinforce the conclusion that AI technologies, when rigorously developed and appropriately applied, can significantly improve drug safety, optimize therapy, reduce healthcare costs, and expand access to pharmaceutical services. However, the adoption of AI tools must be accompanied by regulatory oversight, standardized validation protocols, and clinician training to address ethical and technical challenges such as algorithm bias, data privacy, and model transparency.

Conclusion

This systematic review demonstrates that artificial intelligence holds immense promise for predicting adverse drug reactions and improving medication safety in pharmacy practice. The included studies showcased diverse AI applications, from dose optimization and DDI prediction to MTM and robotic dispensing, all yielding strong performance outcomes. As AI continues to evolve, its integration into pharmaceutical systems can enhance clinical decision-making, reduce healthcare costs, and promote personalized medicine—provided that its implementation is guided by robust validation, transparency, and regulatory compliance.

References

1. Mohsen A., Tripathi L.P., Mizuguchi K. Deep learning prediction of adverse drug reactions in drug discovery using open TG–GATEs and FAERS databases. Front Drug Discov. 2021 Oct 27; 1:768792.
2. Yalçın N., Kaşıkcı M., Çelik H.T., et al. An artificial intelligence approach to support detection of neonatal adverse drug reactions based on severity and Probality scores: a new risk score as web-tool. Children. 2022 Dec;9(12):1826. doi: 10.3390/children9121826.
3. Takase T., Masumoto N., Shibatani N., et al. Evaluating the safety and efficiency of robotic dispensing systems. J Pharm Health Care Sci. 2022 Oct 1;8(1):24. doi: 10.1186/s40780-022-00255-w.
4. Jungreithmayr V., Meid A.D., Haefeli W.E., Seidling H.M. The impact of a computerized physician order entry system implementation on 20 different criteria of medication documentation—a before-and-after study. BMC Med Inform Decis Mak. 2021 Dec;21(1):1–2. doi: 10.1186/s12911-021-01607-6.
5. Blasiak Agata, Truong Anh, Jeit Wen, et al. PRECISE CURATE.AI: A prospective feasibility trial to dynamically modulate personalized chemotherapy dose with artificial intelligence. J Clin Oncol. 2022;40(16_suppl):1574.
6. Wang Z., Ong C.L., Fu Z. AI models to assist vancomycin dosage titration. Front Pharmacol. 2022 Feb 8; 13:801928. doi: 10.3389/fphar.2022.801928.
7. Roche-Lima A., Roman-Santiago A., Feliu-Maldonado R., et al. Machine learning algorithm for predicting warfarin dose in caribbean hispanics using pharmacogenetic data. Front Pharmacol. 2020; 10:1550. doi: 10.3389/fphar.2019.01550. Published 2020 Jan 22.
8. Mei S., Zhang K. A machine learning framework for predicting drug–drug interactions. Sci Rep. 2021 Sep 2;11(1):17619. doi: 10.1038/s41598-021-97193-8.
9. Van Laere S., Muylle K.M., Dupont A.G., Cornu P. Machine learning techniques outperform conventional statistical methods in the prediction of high risk QTc prolongation related to a drug-drug interaction. J Med Syst. 2022 Nov 23;46(12):100. doi: 10.1007/s10916-022-01890-4.
10. Balestra M., Chen J., Iturrate E., Aphinyanaphongs Y., Nov O. Predicting inpatient pharmacy order interventions using provider action data. JAMIA Open. 2021;4(3) doi: 10.1093/jamiaopen/ooab083. Published 2021 Oct 5.
11. Xingwei W., Huan C., Mengting L., et al. A machine learning-based risk warning platform for potentially inappropriate prescriptions for elderly patients with cardiovascular disease. Front Pharmacol. 2022; 13:804566. doi: 10.3389/fphar.2022.804566. Published 2022 Aug 11.
12. Tai C.-T., Sue K.-L., Hu Y.-H. Machine learning in high-alert medication treatment: a study on the cardiovascular drug. Appl Sci. 2020;10(17):5798.
13. Wongyikul P., Thongyot N., Tantrakoolcharoen P., Seephueng P., Khumrin P. High alert drugs screening using gradient boosting classifier. Sci Rep. 2021;11(1):20132. doi: 10.1038/s41598-021-99505-4. Published 2021 Oct 11.
14. Patel J., Ladani A., Sambamoorthi N., et al. A machine learning approach to identify predictors of potentially inappropriate non-steroidal anti-inflammatory drugs (NSAIDs) use in older adults with osteoarthritis. Int J Environ Res Public Health. 2020;18 doi: 10.3390/ijerph18010155.
15. Nagata K., Tsuji T., Suetsugu K., et al. Detection of overdose and underdose prescriptions—an unsupervised machine learning approach. PloS One. 2021 Nov 19;16(11) doi: 10.1371/journal.pone.0260315.
16. Yalçın N., Kaşıkcı M., Çelik H.T., et al. Development and validation of a machine learning-based detection system to improve precision screening for medication errors in the neonatal intensive care unit. Front Pharmacol. 2023 Apr 14;14:1151560. doi: 10.3389/fphar.2023.1151560.
17. Corny J., Rajkumar A., Martin O., et al. A machine learning-based clinical decision support system to identify prescriptions with a high risk of medication error. J Am Med Inform Assoc. 2020;27(11):1688–1694. doi: 10.1093/jamia/ocaa154.
18. Kessler S., Desai M., McConnell W., et al. Economic and utilization outcomes of medication management at a large Medicaid plan with disease management pharmacists using a novel artificial intelligence platform from 2018 to 2019: a retrospective observational study using regression methods. J Manag Care Spec Pharm. 2021;27(9):1186–1196. doi: 10.18553/jmcp.2021.21036.
19. Bu F., Sun H., Li L., et al. Artificial intelligence-based internet hospital pharmacy services in China: Perspective based on a case study. Front Pharmacol. 2022;13 doi: 10.3389/fphar.2022.1027808. Published 2022 Nov 9.