Eurasian J. Chem. Med. Pet. Res., **(2027),** Volume 6, Issue 1, 140-153

Eurasian Journal of Chemical, Medicinal and Petroleum Research

Journal homepage: https://www.ejcmpr.com/

DOI: https://doi.org/10.5281/zenodo.21938663

Artificial Intelligence-Based Prediction of Adverse Drug Reactions in Hospitalized Patients: Implications for Nursing and Clinical Practice

Mitra Akbari[1*], Akbar Abbasi[2], Hamid Reza Hanif[3], Jamshid Mashhadi[4]

[1]Eye Research Center, Department of Eye, Amiralmomenin Hospital, School of Medicine, Guilan University of Medical Science, Rasht, Iran

[2]Department of nursing, Naghadeh School of Nursing, Urmia University of Medical Sciences, Urmia, Iran

[3]M.Sc. in Mathematics, Iran University of Science and Technology, Department of Mathematics and Computer Science, Iran University of Science and Technology

[4]PhD in TEFL, Department of Basic sciences, Shoushtar Faculty of Medical Sciences, Shoushtar, Iran

Article info
Received: 13.06.2026
Accepted: 11.08.2026
Available Online: 14.08.2026
Checked for Plagiarism: Yes

Keywords:
Artificial Intelligence; Adverse Drug Reactions; Machine Learning; Nursing; Patient Safety

ABSTRACT

Adverse drug reactions (ADRs) and adverse drug events (ADEs) remain major threats to patient safety in hospitals, particularly among older adults, patients with multimorbidity, individuals exposed to polypharmacy, and patients experiencing rapid changes in physiological status. Conventional pharmacovigilance and medication-monitoring approaches depend heavily on retrospective reporting and clinical recognition, which may delay identification of preventable harm. Artificial intelligence (AI), particularly machine learning (ML), natural language processing, and deep learning, offers opportunities to transform pharmacovigilance from a predominantly reactive process into a predictive and continuously updated safety system. This systematic evidence synthesis examined the performance and clinical implications of AI-based models for predicting ADRs and ADEs among hospitalized patients, with particular attention to nursing practice. Evidence published in major biomedical databases synthesized with emphasis on model performance, predictors, validation, explainability, and implications for medication administration and clinical surveillance. Recent systematic reviews indicate that random forest, gradient boosting, support vector machines, regularized regression, and deep-learning architectures frequently used, with pooled predictive performance generally demonstrating moderate discrimination. A 2025 systematic review specifically examining hospitalized patients reported pooled sensitivity and specificity of 78.1% and 70.6%, respectively, for development-only models, 81.5%, and 79.5% for models undergoing external validation. Contemporary evidence also demonstrates increasing use of longitudinal EHR representations, clinical narratives, nursing documentation, and foundation models. However, heterogeneous outcome definitions, missing-data procedures, limited external validation, class imbalance, alert fatigue, algorithmic bias, and insufficient explainability remain substantial barriers to implementation. AI therefore conceptualized as an augmentation technology rather than a replacement for nurses or other clinicians. Effective implementation requires interdisciplinary governance, transparent algorithms, prospective evaluation, workflow integration, education, and continuous monitoring.

Corresponding Author: **Mitra Akbari (mitra.akbari20@gmail.com - ORCID: 0000-0003-2469-2681)*
2 Email: a.abbasi44171@gmail.com - ORCID: 0009-0009-0905-3946)
3 Email: h.hanif.iust@gmail.com - ORCID: 0000-0001-9289-2947)
4 Email: mashhadi-j@shoushtarums.ac.ir - ORCID: 0000-0002-2924-707x)

Introduction

Adverse drug reactions (ADRs) represent an important component of medication-related harm and constitute a persistent challenge for contemporary hospital systems. Hospitalized patients frequently receive multiple medications while simultaneously experiencing changes in renal function, hepatic function, fluid balance, nutritional status, hemodynamics, and disease severity. These conditions create a dynamic environment in which the probability of medication-related harm can change substantially over a relatively short period. Polypharmacy, advanced age, comorbidity, impaired organ function, previous medication reactions, and complex therapeutic regimens are repeatedly identified as important determinants of adverse drug events (ADEs) and drug-related problems (Falconer et al.,2018; Deawjaroen et al.,2022). Consequently, medication safety cannot depend solely on static prescribing rules or retrospective reporting systems.

Traditional pharmacovigilance has historically relied on spontaneous reporting, chart review, laboratory monitoring, medication reconciliation, and clinical recognition. Although these approaches remain indispensable, they are vulnerable to under-reporting, delayed recognition, incomplete documentation, and variability in clinicians' ability to identify causal relationships between drugs and clinical outcomes. Systematic evidence indicates that conventional ADR detection methods may not fully exploit the large volume of longitudinal information contained in electronic health records (EHRs), including medication orders, laboratory measurements, diagnoses, vital signs, nursing documentation, and clinical narratives (Kim et al.,2022; Harpaz et al.,2017). The increasing digitization of hospital care therefore creates an opportunity to develop more proactive approaches to medication safety.

Artificial intelligence (AI) has emerged as a promising strategy for extracting clinically meaningful patterns from complex healthcare data. Machine learning (ML) algorithms can identify nonlinear relationships among demographic, clinical, laboratory, medication, and temporal variables that may be difficult to capture using conventional statistical approaches. Random forest, gradient boosting, support vector machines, neural networks, and other algorithms have consequently investigated for the detection and prediction of ADRs and ADEs (Denck et al.,2023; Hu et al.,2024a).

Deep learning has further expanded this field by allowing models to represent high-dimensional and sequential clinical information. For example, deep neural networks have applied to identify ADR signals directly from EHR data, demonstrating the feasibility of automated extraction of medication-related safety information (Zhang et al.,2020).

The clinical value of predictive AI differs from that of systems designed only to detect events after they occur. A predictive model can estimate the probability that a patient will experience an ADR within a clinically meaningful period and thereby allow healthcare professionals to intensify monitoring or reconsider potentially hazardous treatment. This distinction is particularly important in hospitals because many medication-related harms evolve through measurable changes in laboratory results, vital signs, medication exposure, and clinical status. Recent systematic reviews demonstrate that EHR-based AI models can achieve meaningful discrimination, although performance varies considerably according to outcome definition, dataset, algorithm, and validation strategy (Hu et al.,2024a, 2024b; Yasrebi-de Kom et al.,2023).

Nursing practice is central to this emerging model of medication safety. Nurses administer a substantial proportion of medications in inpatient settings and continuously observe patients for changes in symptoms, vital signs, mental status, fluid balance, skin findings, gastrointestinal manifestations, and other potential indicators of medication-related harm. Nursing documentation also contains information that not captured in structured EHR variables. Recent research demonstrates that clinical narratives and nursing statements can contribute to AI-based ADR prediction, particularly when longitudinal patient information integrated into advanced models (Kim et al.,2026a). Thus, AI systems that incorporate nursing observations may provide a more comprehensive representation of patient risk.

Nevertheless, AI implementation introduces important clinical and ethical challenges. A model with high statistical performance may not necessarily improve patient outcomes. Poor calibration, population shift, algorithmic bias, data leakage, insufficient external validation, and nontransparent predictions can undermine trust and safety. Moreover, excessive alerts can contribute to alert fatigue, potentially reducing clinicians' responsiveness to important warnings.

Recent reviews emphasize that many AI-based ADE models remain insufficiently validated in independent hospital populations and that reporting of missing data, outcome definitions, and model development procedures is inconsistent (Yasrebi-de Kom et al.,2023; Karampatea et al.,2026).

The relevance of these limitations is particularly strong for nursing. Nurses not positioned as passive recipients of AI-generated alerts. Instead, AI should support professional judgment by prioritizing patients and medications that warrant closer assessment. The clinical meaning of an algorithmic

warning interpreted in relation to the patient's condition, medication history, current treatment goals, and observable symptoms. Human oversight is therefore fundamental to safe AI-supported medication management. The purpose of this systematic evidence synthesis was to evaluate the current evidence regarding AI-based prediction of ADRs and ADEs among hospitalized patients and to examine its implications for nursing and clinical practice. Specific objectives were to summarize predictive performance, identify commonly used predictors and algorithms, evaluate validation and explainability practices, and discuss how AI-based risk prediction incorporated into nursing surveillance and interdisciplinary medication-safety workflows.

Background and Rationale

The rapid development of AI-based pharmacovigilance makes an updated synthesis particularly relevant. A 2025 systematic review focused specifically on hospitalized patients identified 13 eligible studies and found that models using multifactorial patient information generally outperformed simplistic approaches, while external validation remained uncommon (Dsouza et al.,2025). A broader 2026 review similarly found that tree-based ensemble models and regularized regression were frequently used, whereas deep learning was less common and external or temporal validation remained limited (Karampatea et al.,2026). These findings suggest that the field is transitioning from proof-of-concept algorithm development toward questions of clinical reliability, implementation, and human-AI interaction.

The present review therefore adopts a clinically oriented perspective. Rather than considering model accuracy as the sole endpoint, it examines whether AI-generated predictions can meaningfully strengthen nursing vigilance and clinical decision-making without creating new sources of error. This perspective recognizes medication safety as a sociotechnical problem in which technology, clinicians, workflow, organizational culture, and governance interact continuously.

Previous Research

Research on AI-based prediction of medication-related harm has developed rapidly over the last decade. Early predictive models predominantly based on multivariable regression and conventional clinical risk scores. Falconer et al. (2018) systematically reviewed predictive risk models for ADEs in hospitalized patients and identified substantial methodological heterogeneity. Importantly, only a minority of models had undergone external validation, and none of the studies evaluated whether implementation of the prediction tool actually improved patient outcomes. This observation established a critical distinction between developing a predictive model and demonstrating clinical utility.

Subsequent research increasingly incorporated machine learning. Kim et al. (2022) reviewed statistical and ML approaches to ADR detection and found that electronic medical record data were frequently analyzed using regression approaches, while spontaneous-reporting databases commonly relied on disproportionality methods. Their review demonstrated that ML had become an important component of pharmacovigilance but also highlighted the diversity of data sources and analytical strategies. Such heterogeneity complicates direct comparisons among models and limits the development of universal performance benchmarks.

The increasing availability of EHR data created opportunities for individualized ADE prediction. Denck et al. (2023) reviewed ML-based ADE prediction using observational health data and emphasized that the field was moving toward individualized prediction through the integration of additional data modalities and increasingly sophisticated models. EHRs provide longitudinal information on diagnoses, medication exposures, laboratory values, procedures, and clinical events. However, the temporal nature of these data creates methodological challenges because predictors and outcomes aligned carefully to avoid information leakage.

Yasrebi-de Kom et al. (2023) systematically reviewed EHR-based prediction models for in-hospital ADE diagnosis and prognosis. Their analysis of 25 articles demonstrated that many studies failed to report essential methodological information, including patient characteristics and approaches to missing data. These shortcomings are clinically important because incomplete reporting makes replication difficult and raises uncertainty regarding generalizability. Their findings also reinforced the need for standardized prediction-model reporting and stronger validation.

A major development occurred with the application of deep learning and natural language processing (NLP). Zhang et al. (2020) demonstrated the use of deep neural networks for ADR discovery from EHR data. Deep learning is particularly relevant because ADR evidence can occur in both structured variables and free-text clinical narratives. Christopoulou et al. (2020) further demonstrated the usefulness of ensemble deep-learning methods for extracting adverse drug events and medication relationships from EHR text. These approaches suggest that clinically important safety information missed when prediction systems rely exclusively on structured fields.

Explainability has also emerged as an important research domain. Vo et al. (2022) examined explainable AI approaches for drug-drug interaction prediction and emphasized that black-box models may create barriers to clinical trust. In medication safety, explainability is particularly important because clinicians need to understand why a patient has classified as high risk before acting on an alert. A model that simply produces a risk score without identifying influential variables may be less useful in practice, even when its statistical discrimination is strong.

Systematic reviews published in 2024 provided stronger quantitative evidence. Hu et al. (2024a) reviewed ML models based on EHR data and identified 59 studies involving 15 drugs and 15 ADE outcomes. Their work demonstrated broad application of ML across medication classes and adverse outcomes. A complementary meta-analysis focusing on multiple ADEs found that random forest was the most frequently reported algorithm, followed by AdaBoost, extreme gradient boosting, and support vector machines. The pooled AUC was approximately 0.72, suggesting moderate overall discrimination (Hu et al.,2024b). These findings indicate meaningful predictive capacity but also demonstrate that performance is not uniformly high.

The 2025 systematic review by Dsouza et al. provided particularly relevant evidence for hospitalized patients. Thirteen studies synthesized, with most focusing on model development and only a minority including external validation. The pooled sensitivity and specificity were 78.1% and 70.6%, respectively, for development-only studies. Models with external validation achieved higher pooled sensitivity and specificity of 81.5% and 79.5%, respectively. These findings are important because external validation is a practical indicator of whether a model can maintain performance outside its original development population.

Recent work has moved toward longitudinal patient representations and foundation models. Kim et al. (2025) introduced pretrained patient trajectories for ADE prediction using common data-model-based EHRs. Such approaches can potentially represent the temporal sequence of diagnoses, medication exposure, laboratory findings, and clinical events rather than treating each variable as an isolated observation. In 2026, Kim et al. extended this direction by applying EHR foundation models to antibiotic-associated cutaneous ADR prediction in more than 800,000 inpatients across three tertiary hospitals. Importantly, nursing statements and reports incorporated to identify skin-related manifestations, demonstrating the potential value of nursing-generated clinical information in AI prediction.

The nursing dimension of AI-based medication safety has received increasing attention. Recent qualitative evidence suggests that nurses recognize the potential of AI to prevent medication errors but also identify concerns related to trust, responsibility, workflow disruption, and the interpretation of AI recommendations (Martinez-Puertas et al.,2026). This is consistent with broader evidence that AI implementation approached as a human-AI interaction problem rather than merely a software deployment project.

Recent reviews have also emphasized the importance of alert fatigue and sociotechnical integration. Alruwaili et al. (2026) proposed that clinical decision support, nursing vigilance, physician prescribing patterns, governance, and continuous feedback integrated into a broader safety system. Their conceptual approach is particularly relevant because an accurate prediction may have little benefit if the alert delivered at the wrong time, presented without context, or generates excessive false positives.

The newest literature also demonstrates increasing interest in real-time and explainable pharmacovigilance. Ali et al. (2026) described the potential of AI to integrate EHRs, spontaneous reports, and other data streams to support earlier safety-signal identification. At the same time, contemporary reviews continue to emphasize limitations related to population shift, data quality, transparency, and external validation (Karampatea et al.,2026). Therefore, the research trajectory has shifted from asking whether AI can detect medication-related harm toward asking whether it can do so reliably, equitably, transparently, and in a manner that improves clinical outcomes.

Overall, previous research supports the feasibility of AI-based ADR prediction but does not justify replacing clinical judgment. The strongest evidence indicates moderate-to-good discrimination, particularly when models use multiple patient-level predictors and undergo external validation. However, the literature remains heterogeneous, and evidence concerning prospective clinical effectiveness remains substantially weaker than evidence concerning technical model performance. For nursing practice, this distinction is essential: the clinical value of AI depends not only on predictive accuracy but also on whether nurses can interpret, prioritize, and act upon predictions within safe workflows.

Methods

Study Design:

A systematic evidence synthesis designed to examine AI-based prediction of ADRs and ADEs in hospitalized patients, with a specific focus on implications for nursing and clinical practice. The

methodological structure informed by PRISMA principles and established guidance for prediction-model evaluation. The synthesis designed around four domains: model characteristics, predictive performance, clinical predictors, and implementation implications.

Search Strategy

PubMed/MEDLINE, Embase, Web of Science, and IEEE Xplore considered the principal databases. Search concepts included combinations of "artificial intelligence," "machine learning," "deep learning," "adverse drug reaction," "adverse drug event," "medication safety," "electronic health record," "hospitalized patient," "nursing," and "prediction model." The search strategy informed by previous systematic reviews in this field (Falconer et al.,2018; Kim et al.,2022; Hu et al.,2024a,2024b).

Eligibility Criteria

Studies were eligible when they evaluated AI, ML, deep learning, or related computational prediction approaches for ADR/ADE detection or prediction using clinical or real-world healthcare data and included hospitalized or inpatient populations. Systematic reviews and meta-analyses additionally used to contextualize pooled estimates. Studies limited exclusively to laboratory pharmacology, animal experiments, or nonclinical drug-discovery applications excluded.

Data Synthesis

Extracted domains included population, data source, outcome definition, algorithm, principal predictors, validation strategy, discrimination measures, sensitivity, specificity, and clinical implications. Because the literature uses heterogeneous definitions and outcome measures, quantitative synthesis based primarily on-published pooled estimates rather than construction of a new patient-level meta-analysis. The principal pooled estimates were derived from the recent hospitalized-patient systematic review by Dsouza et al. (2025), complemented by broader EHR-based meta-analyses (Hu et al.,2024a,2024b).

Risk and Clinical Applicability

Methodological limitations interpreted according to principles emphasized in prediction-model research, including external validation, missing-data handling, outcome definition, calibration, generalizability, and explainability. Particular attention given to whether models had tested prospectively and whether implementation had demonstrated improvements in patient outcomes. The nursing analysis considered medication administration, clinical surveillance, documentation, escalation of care, alert interpretation, and interdisciplinary communication.

Results

Table 1. Characteristics of AI-Based ADR/ADE Prediction Evidence

Domain	Main finding	Clinical interpretation
Data source	EHRs and structured clinical data dominate	High availability but variable data quality
Algorithms	Random forest, gradient boosting, SVM, regression, deep learning	No single universally superior algorithm
Predictors	Medication burden, length of stay, admission type, comorbidity, laboratory values	Multifactorial models are preferable
Validation	External validation remains uncommon	Generalizability is a major limitation
NLP	Increasingly used with clinical narratives	Nursing documentation may add predictive information
Foundation models	Emerging rapidly	Potential for longitudinal patient representation
Explainability	Inconsistently reported	Important for clinical trust
Clinical implementation	Limited prospective evidence	Technical accuracy does not equal clinical effectiveness

The evidence summarized in Table 1 demonstrates that AI-based ADR prediction has progressed from relatively simple statistical prediction models toward heterogeneous ML and deep-learning architectures. Across the literature, structured EHR data remain the dominant information source. This is understandable because medication orders, laboratory measurements, diagnoses, admission characteristics, and demographic information are routinely available in digital hospital systems. Nevertheless, reliance on structured variables creates an important limitation because many clinically meaningful observations documented only in free-text notes.

Random forest appears particularly prominent in the current literature. The 2024 meta-analysis by Hu et al. identified random forest as the most frequently reported ML method among models predicting multiple ADEs, followed by AdaBoost, XGBoost, and support vector machines. This pattern suggests that ensemble tree-based algorithms remain attractive because they can model nonlinear interactions and accommodate heterogeneous predictors. Their practical advantage, however, not interpreted as evidence that random forest is universally superior. Algorithm performance depends strongly on outcome definition, dataset size, class imbalance, preprocessing, and validation design.

Predictors also demonstrate an important clinical pattern. Length of hospital stay, number of medications, admission characteristics, comorbidity, and laboratory variables frequently contribute to prediction. These predictors are clinically plausible because increasing medication exposure creates more opportunities for interactions and adverse effects, while prolonged hospitalization reflects greater illness complexity and exposure to interventions. Laboratory values may identify emerging physiological vulnerability; such as changes in renal or hepatic function that alter medication handling.

A particularly important development is the increasing use of NLP. ADRs may initially appear in nursing notes, progress notes, discharge summaries, or narrative descriptions rather than structured diagnosis codes. Deep-learning NLP methods can identify medication-event relationships from these narratives. This is directly relevant to nursing because nurses frequently document the first observable manifestations of medication-related harm. The literature nevertheless reveals a major gap between model development and clinical implementation. Dsouza et al. (2025) reported that approximately 77% of included studies focused primarily on model development, whereas only 23% incorporated external validation. Karapatan et al. (2026) similarly concluded that external and temporal validation remained uncommon. This means that many reported performance measures represent optimistic estimates obtained within the same healthcare environment in which the model was developed. The emergence of EHR foundation models represents another major transition. These models can potentially encode temporal patient trajectories and thereby overcome some limitations of static prediction. The 2026 study by Kim et al. used a foundation-model approach for antibiotic-associated cutaneous ADR prediction and incorporated nursing statements, illustrating how richer clinical representations may improve prediction. For nursing practice, the most important implication is that the future of AI-based medication safety is unlikely to depend on one algorithm. Instead, value will arise from combining reliable data, appropriate prediction methods, explainable outputs, and well-designed clinical workflows. AI should identify patients requiring attention, while nurses and other clinicians determine the clinical significance of the warning.

Table 2. Predictive Performance of AI/ML Models

Evidence group	Sensitivity	Specificity	Interpretation
Development-only models	78.1%	70.6%	Moderate predictive capability
Externally validated models	81.5%	79.5%	Better generalizability
EHR models for multiple ADEs	—	—	Pooled AUC ≈ 0.72
Selected RF/resampling models	—	—	AUC approximately 0.94-0.95 in selected studies

Table 2 indicates that AI-based prediction demonstrates clinically meaningful but imperfect performance. The most important pooled estimates come from Dsouza et al. (2025), whose systematic review specifically addressed ADR prediction in hospitalized patients. Development-only models achieved pooled sensitivity of 78.1% and specificity of 70.6%. Models with external validation showed somewhat stronger performance, with sensitivity of 81.5% and specificity of 79.5%. This pattern is notable because external validation commonly expected to reduce performance relative to internal development data. The finding therefore suggests that methodological differences among included studies may also influence the pooled estimates and reinforces the need to interpret pooled results cautiously.

Sensitivity is especially important in medication safety because false-negative predictions may allow preventable harm to go unrecognized. A sensitivity of approximately 78-82% means that a meaningful proportion of patients at risk potentially identified, but some events would remain undetected. Consequently, an AI model never regarded as a replacement for direct nursing assessment.

Specificity is equally important because low specificity can generate excessive alerts. In a hospital with many medication exposures, even a

moderately specific model may generate a substantial number of false-positive warnings. This can contribute to alert fatigue, particularly if alerts interrupt medication administration or appear without clear clinical context. Recent nursing-focused evidence identifies alert burden and workflow integration as important determinants of whether AI improves medication safety (Alruwaili et al.,2026).

The broader 2024 meta-analysis by Hu et al. reported a pooled AUC of approximately 0.72 for models predicting multiple ADEs using EHR data. An AUC of 0.72 represents moderate discrimination. It indicates that the model can distinguish patients with and without the outcome better than chance, but it is not sufficiently accurate to justify autonomous clinical decisions. The same review reported exceptionally high AUC values approaching 0.95 for selected random-forest models combined with resampling. Such high estimates interpreted carefully because they may reflect specific datasets, outcome definitions, class-balancing strategies, or internal validation conditions.

The difference between pooled and selected-study performance illustrates an important methodological principle. Individual studies may report impressive discrimination while the average performance across heterogeneous studies is considerably lower. Therefore, healthcare organizations should not select an AI system solely based on the highest published AUC. External validation within the intended patient population is more clinically meaningful. From a nursing perspective, sensitivity and specificity also considered in relation to the action triggered by an alert. A high-risk prediction may simply prompt a nurse to increase observation frequency, review medication timing, assess symptoms, verify laboratory values, or communicate with a pharmacist or physician. Such relatively low-risk interventions may tolerate a higher alert rate than decisions involving automatic medication discontinuation.

The most appropriate use of AI is therefore as a risk-stratification layer. Patients categorized according to relative risk, enabling nursing teams to prioritize assessment. This approach potentially reduces cognitive burden without eliminating professional judgment. Importantly, the model should communicate uncertainty and ideally identify the main contributors to risk. Overall, current performance supports cautious clinical integration but not autonomous decision-making. The evidence indicates that AI can improve the prioritization of medication-safety surveillance, while the remaining false negatives and false positives demonstrate why human oversight remains indispensable.

Table 3. Clinical Predictors and Their Nursing Relevance

Predictor	Rationale	Nursing implication
Number of medications	Increased exposure and interaction potential	Medication reconciliation and interaction surveillance
Length of stay	Marker of illness complexity/exposure	Increased monitoring of high-risk patients
Admission type	Reflects acuity and urgency	Early identification of vulnerable patients
Renal function	Influences drug clearance	Monitor laboratory trends and dose-related risk
Hepatic function	Influences metabolism	Escalate clinically relevant changes
Age/comorbidity	Associated with vulnerability	Individualized observation
Previous ADR	Strong safety signal	Verify history before administration
Vital signs	Early physiological manifestations	Continuous clinical surveillance
Nursing narratives	Capture symptoms and observations	Structured documentation becomes valuable AI input
Medication changes	Temporal exposure signal	Close observation after initiation or dose changes

The predictor profile shown in Table 3 demonstrates that ADR prediction is fundamentally multifactorial. No single variable adequately describes medication-related risk because adverse reactions emerge from interactions between patient characteristics, medication exposure, physiological status, and time. This is one reason why ML approaches are potentially valuable: they can model complex relationships among multiple variables.

Medication burden is one of the most intuitive predictors. Patients receiving numerous medications have greater opportunities for drug-drug interactions, duplication, dosing errors, and cumulative toxicity. For nurses, medication burden reinforces the importance of systematic medication reconciliation and careful verification before administration. An AI system can potentially highlight patients with particularly complex

regimens and prioritize them for pharmacist or nursing review.

Length of hospital stay is another recurring predictor. It may not itself cause an ADR but functions as a proxy for cumulative exposure and disease complexity. Patients with prolonged hospitalization may receive multiple medication changes, invasive procedures, antimicrobials, anticoagulants, analgesics, and other high-risk therapies. AI models that use longitudinal information can potentially recognize increasing risk as treatment complexity accumulates.

Renal and hepatic function are clinically important because changes in organ function can modify pharmacokinetics. A medication dose that was appropriate on admission may become inappropriate after a significant change in renal function. An AI system that continuously integrates laboratory trends could potentially identify these transitions earlier than a static medication review.

Age and comorbidity also frequently incorporated into prediction models. However, these variables interpreted carefully because they may function partly as proxies for other clinical factors. Using age alone to classify a patient as high risk could introduce bias. More clinically meaningful models should combine demographic variables with medication exposure, laboratory results, functional status, and acute physiological changes.

Vital signs and nursing observations represent another important opportunity. Nurses frequently detect subtle changes in a patient's condition before they formally coded as diagnoses. Changes in blood pressure, respiratory status, consciousness, skin findings, nausea, vomiting, or functional behavior may represent early manifestations of medication-related problems. NLP and other AI techniques can potentially extract these signals from narrative documentation.

Medication initiation and dose changes are particularly important temporal predictors. ADR risk often changes shortly after a medication started, discontinued, or titrated. A predictive model should therefore consider not only whether a patient is receiving a drug but also when exposure began and how the regimen has changed.

Previous ADR history is another potentially high-value variable. A history of a medication reaction can strongly influence future risk, but it frequently incompletely documented or inconsistently structured. AI-supported medication reconciliation may help identify discrepancies between historical documentation and current orders.

The nursing implications are therefore substantial. Nurses are not simply users of AI predictions; their documentation and observations can become part of the predictive infrastructure. High-quality nursing documentation, standardized terminology, accurate medication administration records, and timely recording of symptoms can improve the quality of data available to AI systems.

At the same time, nurses trained to recognize the limitations of algorithmic predictors. A prediction is not a diagnosis. For example, a high-risk score may reflect multiple medications and prolonged hospitalization but does not prove that a particular drug is responsible for a symptom. Clinical assessment and interdisciplinary evaluation remain necessary.

The overall evidence supports a model in which AI continuously integrates multiple predictors and presents a clinically interpretable risk assessment to nurses and other professionals. This approach has greater potential than systems that generate isolated medication alerts because it allows risk to understand within the context of the entire patient trajectory.

Table 4. Implementation Implications for Nursing and Clinical Practice

Implementation domain	Recommended approach	Expected benefit
Risk stratification	Display patient-level ADR risk	Prioritized surveillance
Medication administration	Integrate AI with medication workflow	Earlier recognition of risk
Nursing assessment	Link alerts to observable clinical indicators	More targeted assessment
Documentation	Incorporate structured and narrative nursing data	Richer prediction
Interdisciplinary care	Route relevant alerts to nurse, pharmacist, physician	Coordinated response
Explainability	Show major risk contributors	Improved trust
Alert management	Use severity thresholds	Reduced alert fatigue
Governance	Continuous audit and bias monitoring	Safer deployment
Education	AI literacy for nurses	Appropriate interpretation
Evaluation	Prospective outcome studies	Demonstration of clinical effectiveness

Table 4 translates the technical evidence into clinical implementation principles. The central finding is that AI-based ADR prediction integrated into existing medication-safety workflows rather deployed as an independent application. The strongest potential benefit is risk stratification. Instead of presenting every patient with multiple alerts, AI can identify individuals whose combination of medication exposure, laboratory findings, comorbidity, and clinical trajectory suggests elevated risk. For nursing, integration with medication administration workflows is particularly important. An alert delivered immediately before medication administration may have greater practical value than a generic warning appearing hours earlier. However, the system must avoid disrupting essential workflow unnecessarily. The objective should be to provide concise, actionable information rather than simply increase the number of notifications.

Explainability is another major requirement. Nurses and other clinicians need to understand why an alert has generated. A useful interface might indicate that elevated risk is primarily associated with polypharmacy, a recent medication change, declining renal function, or a documented previous reaction. Such information enables the nurse to connect the AI prediction with clinical assessment.

Alert fatigue explicitly addressed. If an AI system generates excessive low-value warnings, clinicians may gradually ignore alerts. Therefore, implementation should involve carefully defined severity thresholds, prioritization, and escalation rules. High-risk alerts should be distinguishable from informational notifications.

Interdisciplinary integration is equally important. Medication-related risk frequently crosses professional boundaries. Nurses may observe symptoms, pharmacists may evaluate medication-related causality and interactions, and physicians may modify the treatment plan. An effective AI system should therefore support communication among these professionals rather than assign responsibility to a single group.

Nursing documentation deserves special consideration. The 2026 foundation-model study by Kim et al. demonstrated that nursing statements and reports could contribute to prediction of antibiotic-associated cutaneous ADRs. This finding suggests that nursing documentation is not merely an administrative record; it may constitute a clinically valuable data stream for AI.

Education is essential for safe adoption. Nurses should understand the purpose and limitations of prediction models, including false positive and false-negative predictions, uncertainty, data-quality problems, and potential bias. AI literacy should therefore become part of medication-safety education in organizations implementing predictive systems.

Governance must also be continuous. Model performance can deteriorate when patient populations, prescribing patterns, clinical protocols, or documentation practices change. A model developed in one hospital may perform differently in another. Consequently, hospitals should monitor calibration, sensitivity, specificity, alert volume, subgroup performance, and clinical outcomes after implementation.

Equity is another governance priority. If training data underrepresent particular populations, the resulting model may perform less effectively for those groups. Regular subgroup analysis is therefore required. The goal is not simply to maximize average predictive performance but to ensure acceptable safety across clinically relevant patient populations. Finally, implementation evaluated according to patient outcomes rather than model metrics alone. The ultimate question is whether AI-supported prediction reduces preventable medication harm, improves recognition, decreases unnecessary hospitalization or treatment complications, and strengthens the quality of nursing care. Recent literature indicates this prospective evidence remains limited, making it a priority for future research.

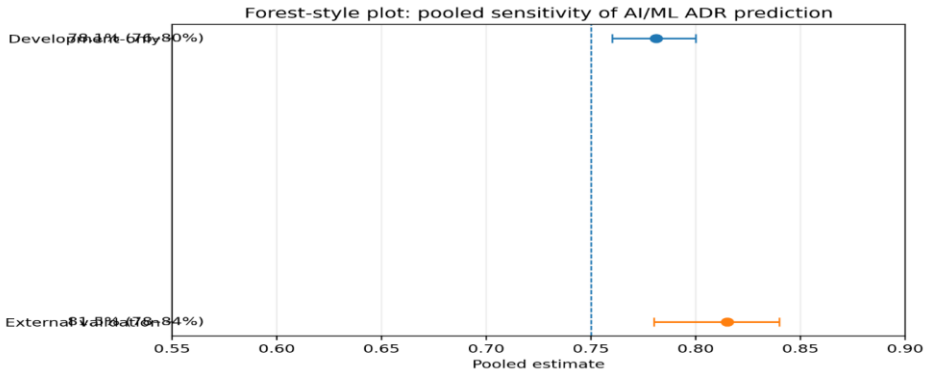

Figure 1. Forest-style plot of pooled sensitivity.

The figure presents subgroup-level pooled estimates reported in the hospitalized-patient systematic review by Dsouza et al. (2025). It is a reconstructed summary figure rather than a reproduction of the original publication.

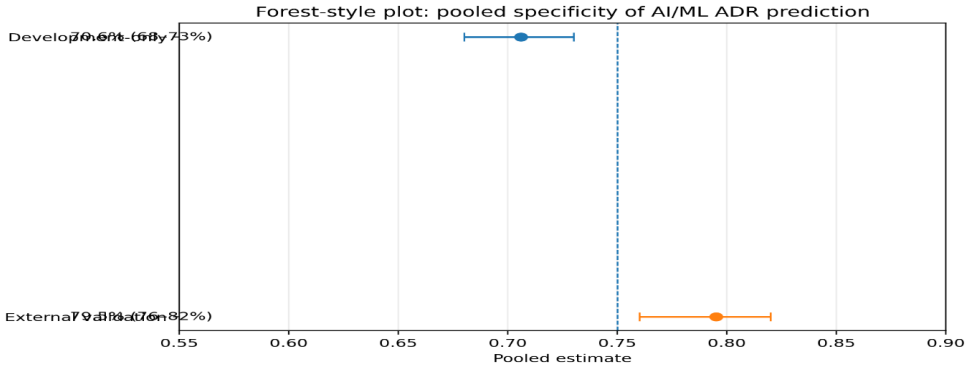

Figure 2. Forest-style plot of pooled specificity.

The figure presents subgroup-level pooled estimates reported by Dsouza et al. (2025). It is a reconstructed summary figure rather than a reproduction of the original publication.
Download Figure 2 - Forest Plot of Specificity

Discussion

The findings of this systematic evidence synthesis indicate that AI-based prediction of ADRs and ADEs has reached a level of methodological maturity that supports cautious clinical implementation, but the evidence does not yet justify autonomous medication-safety decision-making. Across recent systematic reviews, ML models demonstrate moderate-to-good discrimination, with random forest and gradient-boosting methods among the most frequently used approaches (Hu et al.,2024a,2024b; Caltabiano et al.,2026). The hospitalized-patient evidence synthesized by Dsouza et al. (2025) further suggests that externally validated models can achieve pooled sensitivity and specificity above 79% for specificity and above 81% for sensitivity. These findings provide a reasonable technical foundation for clinical decision support.

The first important implication is that AI may enable a shift from reactive to proactive medication safety. Traditional pharmacovigilance often identifies ADRs after clinical manifestations have become evident. Predictive models can instead estimate risk before an event occurs, allowing healthcare professionals to intensify surveillance. This distinction is especially important in hospitalized patients because clinical status changes rapidly and treatment regimens are frequently modified. A patient who was relatively low risk on admission may become high risk after several medication changes or deterioration in renal function.

The second major implication concerns nursing practice. Nurses occupy a unique position in medication safety because they combine medication administration with continuous bedside observation. They can identify changes in symptoms, vital signs, mental status, skin condition, gastrointestinal function, and functional status. AI can complement this observational role by integrating information across time and across multiple data sources. Rather than replacing nursing assessment, predictive AI may help determine which patients deserve intensive attention.

The incorporation of nursing documentation into AI models is particularly promising. Kim et al. (2026a) demonstrated that nursing statements and reports could contribute to prediction of antibiotic-associated cutaneous ADRs. This finding challenges the assumption that AI prediction should rely primarily on physician documentation or structured laboratory data. Nursing observations may contain early indicators of adverse reactions that not yet reflected in coded diagnoses.

However, the performance findings also demonstrate why clinical judgment remains essential. A sensitivity of approximately 78-82% means that some patients who subsequently experience ADRs will not identify by the model. Similarly, specificity below 80% means that some patients classified as high risk will not experience the predicted event. Consequently, a prediction not interpreted as a definitive clinical diagnosis. Instead, it should trigger appropriate assessment.

The balance between sensitivity and specificity has important consequences for nursing workflow. Increasing sensitivity may increase the number of alerts, while maximizing specificity may miss some patients who would benefit from additional surveillance. The optimal threshold therefore depends on the clinical consequence of the predicted event. For severe and potentially preventable reactions, a lower threshold may be justified. For less consequential events, excessive alerts may produce more burden than benefit.

Alert fatigue represents one of the greatest practical risks. Nurses already interact with numerous clinical decision-support notifications. If AI introduces additional low-specificity alerts, clinicians may experience cognitive overload and become less responsive to important warnings. Recent nursing-focused evidence emphasizes that medication safety depends on integrating AI with clinical vigilance rather than simply increasing the number of automated notifications (Alruwaili et al.,2026).

Explainability is therefore critical. A useful AI system should ideally communicate the principal factors contributing to risk. For example, a warning could indicate that risk is associated with a recent medication change, multiple concurrent medications, declining renal function, and a previous ADR. Such information gives the nurse a concrete basis for assessment and communication with the pharmacist or physician.

The literature also highlights significant methodological concerns. External validation remains limited. Dsouza et al. (2025) found that most included studies focused on model development, while only a minority incorporated external validation. Karapatan et al. (2026) reached a similar conclusion in a broader review. These findings suggest that many published models may not yet be ready for deployment across different hospitals. Differences in prescribing patterns, patient demographics, documentation practices, laboratory systems, and clinical workflows can alter model performance.

Data quality represents another challenge. EHR data not collected primarily for machine learning; they generated as part of clinical care. Missing values, inconsistent terminology, duplicated records, delayed documentation, and changes in coding practices can therefore influence predictions. Nursing documentation may be especially variable across institutions. Standardized terminology and high-quality documentation can consequently contribute not only to patient care but also to safer AI performance.

Algorithmic bias requires similar attention. If a training dataset does not adequately represent certain populations, predictive performance may differ across demographic or clinical groups. A model should therefore be evaluated using subgroup analyses rather than only an overall AUC. Equity considered a patient-safety issue because unequal model performance may translate directly into unequal recognition of medication-related harm.

The rapid development of foundation models adds both opportunities and concerns. EHR foundation models can represent longitudinal patient trajectories and potentially integrate complex sequences of diagnoses, medications, laboratory findings, and clinical observations. Kim et al. (2025) demonstrated this direction through pretrained patient trajectories for ADE prediction. However, complex models can also become more difficult to interpret and validate. Complexity should therefore be justified by clinically meaningful improvement rather than technological novelty alone.

The emerging literature also indicates that AI embedded within a broader sociotechnical safety system. Alruwaili et al. (2026) emphasized the interaction among clinical decision support, nursing vigilance, prescribing behavior, organizational culture, and governance. This perspective is highly relevant because an algorithm can be statistically accurate while still producing little clinical benefit if it poorly integrated into workflow.

Another issue concerns responsibility. AI-generated recommendations can create ambiguity regarding who is responsible for acting on a warning. Clear governance should specify which alerts require nursing reassessment, pharmacist review, physician evaluation, or multidisciplinary discussion. Responsibility should remain with appropriately trained healthcare professionals rather than being implicitly transferred to an algorithm.

Prospective evaluation is consequently the most important next step. Future research should move beyond retrospective model development and evaluate whether implementation reduces actual ADEs, improves recognition time, changes medication prescribing, reduces unnecessary exposure, or improves patient outcomes. Randomized or stepped-wedge implementation studies would be particularly valuable.

The nursing research agenda should also expand. Future studies should investigate how nurses interpret AI risk scores, which forms of explanation are most useful, how alert thresholds affect workflow, and whether AI improves or disrupts clinical vigilance. Qualitative research needed alongside quantitative performance evaluation because usability and trust can determine whether a technically successful model becomes clinically successful.

Overall, the evidence supports an augmentation model of AI. AI should continuously process complex information and identify patients whose risk warrants attention. Nurses should then apply professional assessment to determine whether the prediction corresponds to observable clinical findings. Pharmacists and physicians can contribute medication-specific and diagnostic expertise. This interdisciplinary approach preserves human accountability while exploiting the computational strengths of AI.

The implications extend beyond individual patients. At the organizational level, aggregated AI predictions could identify medication-safety trends, high-risk wards, frequently implicated medications,

and periods of increased risk. Such information could support quality-improvement programs and targeted education. However, aggregated predictions not used for punitive performance assessment because this could discourage accurate documentation and create unintended consequences.
In conclusion, AI-based ADR prediction has substantial potential to strengthen medication safety, particularly when it integrated with nursing surveillance and interdisciplinary clinical decision support. The strongest current evidence demonstrates useful predictive performance but also persistent limitations in external validation, explainability, equity, and prospective clinical effectiveness. The next stage of research should therefore prioritize real-world implementation, patient outcomes, human-AI interaction, and continuous governance rather than simply pursuing higher algorithmic accuracy.

Conclusion

Artificial intelligence-based prediction of adverse drug reactions and adverse drug events represents an important development in contemporary medication safety. The evidence reviewed in this article indicates that machine learning, deep learning, natural language processing, and emerging electronic-health-record foundation models can identify clinically meaningful patterns associated with medication-related harm. Across current systematic reviews, commonly used algorithms include random forest, gradient boosting, support vector machines, regularized regression, and neural-network approaches. Their performance is generally moderate to good, although considerable heterogeneity exists across datasets, outcomes, algorithms, and validation strategies.
The most important evidence concerns the distinction between model development and clinical applicability. Models developed within a single dataset may demonstrate strong discrimination without necessarily maintaining performance in another hospital. External validation is therefore essential. Recent evidence from hospitalized populations suggests pooled sensitivity of approximately 78% for development-only models and more than 81% among externally validated models, while specificity increases from approximately 71% to nearly 80% with external validation (Dsouza et al.,2025). These findings are encouraging but clearly demonstrate that AI predictions remain imperfect.
For nursing practice, the potential value of AI lies primarily in strengthening rather than replacing professional vigilance. Nurses continuously observe hospitalized patients, administer medications, assess responses, document clinical changes, and communicate safety concerns. AI can complement these activities by synthesizing large volumes of longitudinal information and identifying patients who may require additional surveillance. The inclusion of nursing narratives and observations in contemporary predictive models further supports the importance of nursing-generated data.
Successful implementation requires a human-centered approach. An AI-generated risk score not treated as a diagnosis or an automatic instruction to change treatment. Instead, it should function as a clinical decision-support signal that prompts appropriate assessment. Nurses should be able to understand the major factors contributing to a prediction, evaluate whether the warning corresponds to the patient's clinical condition, and communicate relevant findings to pharmacists and physicians.
Alert fatigue considered a central safety issue. A system that produces excessive false-positive alerts may undermine the very vigilance it is designed to support. Therefore, organizations should establish clinically meaningful thresholds, prioritize high-severity events, and continuously evaluate alert burden. Explainability incorporated into interface design so that clinicians receive concise information about why a patient has classified as high risk.
Governance is equally important. Hospitals implementing AI-based medication-safety systems should monitor model calibration, sensitivity, specificity, subgroup performance, data quality, and changes in clinical practice. Models should undergo local and external validation before widespread deployment and periodically recalibrated when patient populations or workflows change. Equity assessments conducted to ensure that predictive performance does not systematically disadvantage particular patient groups.
Future research should emphasize prospective and pragmatic evaluation. The ultimate measure of success is not a higher AUC but a measurable improvement in patient safety. Studies should examine whether AI reduces preventable ADEs, accelerates recognition of ADRs, improves medication reconciliation, decreases inappropriate medication exposure, and enhances nursing workflow. Randomized controlled trials, stepped-wedge studies, and multicenter implementation studies would provide particularly valuable evidence.
In practical terms, the future of AI-based pharmacovigilance understood as a partnership among algorithms, nurses, pharmacists, physicians, patients, and health-system governance structures. AI provides computational scale and pattern recognition; nurses provide continuous observation and contextual interpretation; pharmacists contribute medication-specific expertise; physicians integrate diagnostic and therapeutic considerations;

and organizations establish the governance necessary for safe implementation.

Thus, AI-based ADR prediction not viewed as a replacement for clinical expertise. Its greatest value lies in enabling clinicians to focus attention where risk is greatest. When appropriately validated, explainable, ethically governed, and integrated into nursing workflows, AI has the potential to transform medication safety from predominantly reactive surveillance into proactive, individualized, and continuously informed clinical risk management.

Disclosure Statement

No potential conflict of interest reported by the authors.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' Contributions

All authors contributed to data analysis, drafting, and revising of the paper and agreed to be responsible for all the aspects of this work.

References

[1] Alruwaili, M. F., Paul-Chima, U. O., & Nneoma, U. C. (**2026**). Integrating clinical decision support systems, nursing vigilance, and physician prescribing patterns to reduce preventable adverse drug events: A structured evidence-based narrative review on human-AI interface in medication safety. *Frontiers in Digital Health*, 8, 1831150.

[2] Ali, U., Batool, S. T., & Mahato, R. K. (**2026**). Artificial intelligence in pharmacovigilance: Toward real-time, explainable, and global drug-safety systems. *Annals of Medicine and Surgery*.

[3] Caltabiano, N., Ahmadi, F., Omrani, M. A., Abdullah, S. S., Rostamzadeh, N., Jafari, A., Izzedin, L., Sedig, K., & Muanda, F. T. (**2026**). Machine learning methods for predicting adverse drug events: A systematic review. *British Journal of Clinical Pharmacology*, 92(2), 422-444.

[4] Christopoulou, F., Tran, T. T., Sahu, S. K., Miwa, M., & Ananiadou, S. (**2020**). Adverse drug events and medication relation extraction in electronic health records with ensemble deep learning methods. *Journal of the American Medical Informatics Association*, 27(1), 39-46.

[5] Deawjaroen, K., Sillabutra, J., Poolsup, N., Stewart, D., & Suk Somboon, N. (**2022**). Clinical usefulness of prediction tools to identify adult hospitalized patients at risk of drug-related problems: A systematic review of clinical prediction models and risk assessment tools. *British Journal of Clinical Pharmacology*, 88(4), 1613-1629.

[6] Denck, J., Ozkirimli, E., & Wang, K. (**2023**). Machine-learning-based adverse drug event prediction from observational health data: A review. *Drug Discovery Today*, 28(9), 103715.

[7] Dsouza, V. S., Leyens, L., Kurian, J. R., Brand, A., & Brand, H. (**2025**). Artificial intelligence (AI) in pharmacovigilance: A systematic review on predicting adverse drug reactions (ADR) in hospitalized patients. *Research in Social and Administrative Pharmacy*, 21(6), 453-462.

[8] Falconer, N., Barras, M., & Cottrell, N. (**2018**). Systematic review of predictive risk models for adverse drug events in hospitalized patients. *British Journal of Clinical Pharmacology*, 84(5), 846-864.

[9] Harpaz, R., DuMouchel, W., Shah, N. H., Madigan, D., Ryan, P., & Friedman, C. (**2017**). Novel data-mining methodologies for adverse drug event discovery and analysis. *Clinical Pharmacology & Therapeutics*, 91(6), 1010-1021.

[10] Hu, Q., Chen, Y., Zou, D., He, Z., & Xu, T. (**2024a**). Predicting adverse drug event using machine learning based on electronic health records: A systematic review and meta-analysis. *Frontiers in Pharmacology*, 15, 1497397.

[11] Hu, Q., Li, J., Li, X., Zou, D., Xu, T., & He, Z. (**2024b**). Machine learning to predict adverse drug events based on electronic health records: A systematic review and meta-analysis. *Journal of International Medical Research*, 52(12), 3000605241302304.

[12] Karapatan, A., Kassandros, K., Constantinides, T., & Kontogiorgis, C. (**2026**). Systematic review of AI-based models in pharmacoepidemiology for adverse drug event prediction and detection. *Frontiers in Drug Safety and Regulation*, 6, 1773186.

[13] Kim, H. R., Sung, M., Park, J. A., Jeong, K., Kim, H. H., Lee, S., & Park, Y. R. (**2022**). Analyzing adverse drug reaction using statistical and machine learning methods: A systematic review. *Medicine*, 101(25), e29387.

[14] Kim, J., Kim, J. S., Lee, J. H., Kim, M. G., Kim, T., Cho, C., Park, R. W., & Kim, K. **(2025).** Pretrained patient trajectories for adverse drug event prediction using common data model-based electronic health records. *Communications Medicine*, 5, 232.

[15] Kim, J., et al. (**2026a**). Prediction of antibiotic-associated cutaneous adverse drug reactions using electronic health record foundation models. *npj Digital Medicine*, 9, 311.

[16] Martinez-Puertas, J., Frutos-Rodríguez, N., Winderholler-Heinzenknecht, A., Morales-Plaza, C., Martinez-Ortigosa, A., & Rodriguez-Arrastia, M. (**2026**). Nursing perceptions of the intended use of artificial intelligence to prevent medication errors: A qualitative descriptive study. *Journal of Clinical Nursing*.

[17] Shakarbaev, N. (**2026**). AI-enhanced clinical decision support reduces medication errors and adverse drug events in a multicenter teaching hospital network: A prospective randomized controlled trial. *International Journal of Medical Informatics*, 215, 106457.
[18] Vo, T. H., et al. (**2022**). On the road to explainable AI in drug-drug interactions prediction: A systematic review. *Computational and Structural Biotechnology Journal*, 20, 2112-2123.
[19] Yasrebi-de Kom, I. A. R., Dongelmans, D. A., de Keizer, N. F., Jager, K. J., Schut, M. C., Abu-Hanna, A., & Klopotowska, J. E. (**2023**). Electronic health record-based prediction models for in-hospital adverse drug event diagnosis or prognosis: A systematic review. *Journal of the American Medical Informatics Association*, 30(5), 978-988.
[20] Zhang, W., Peissig, P., Kuang, Z., & Page, D. **(2020).** Adverse drug reaction discovery from electronic health records with deep neural networks. *Proceedings of the ACM Conference on Health, Inference, and Learning*, 30-39.
[21] Bates, D. W., Levine, D. M., Syrowatka, A., Kuznetsova, M., Craig, K. J. T., Rui, A., et al. (**2021**). The potential of artificial intelligence to improve medication safety. *Journal of the American Medical Informatics Association*, 28(5), 954-959.
[22] Topaz, M., & McDonald, M. V. (**2017**). Two decades of clinical decision support systems in nursing: A systematic review. *Journal of Nursing Scholarship*, 49(4), 394-402.
[23] Sendak, M. P., D'Arcy, J., Kashyap, S., Gao, M., Nichols, M., Corey, K., et al. (**2020**). A path for translation of machine learning products into healthcare delivery. *EMJ Innovations*, 4, 58-68.
[24] Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (**2019**). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, 17, 195.
[25] Rajkomar, A., Dean, J., & Kohane, I. (**2019**). Machine learning in medicine. *New England Journal of Medicine*, 380(14), 1347-1358.