Results in Engineering 27 (2025) 106002

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

Review article

Predicting adverse drug reactions in oncology: A critical review of machine learning approaches and future directions

Abid Yahya [a], Phatsimo Lobelo [a], Afiya Eram [b,*], Sana Althaf Hussain [c], Irfan Anjum Badruddin [d], Lory Liza D. Bulay-og [e], Dionel O. Albina [e]

[a] Department of Electrical and Communications Systems Engineering, Botswana International University of Science and Technology, Botswana
[b] Department of Conservative Dentistry Endodontics, Manipal College of Dental Sciences, Manipal Academy of Higher Education, Manipal 576104, India
[c] Department of General Surgery, Hamdard Institute of Medical Sciences and Research, New Delhi-110062, India
[d] Department of Mechanical Engineering, College of Engineering, King Khalid University, Abha 61421, Saudi Arabia
[e] Department of Electrical Engineering, University of Science & Technology of Southern Philippines, Cagayan de Oro City, 9000, Philippines

ARTICLE INFO

Keywords:
Adverse drug reactions (ADRs)
Machine learning
Cancer therapy
Pharmacovigilance
Deep learning
Supervised learning
Clinical decision support
Natural language processing (NLP)
Biomedical informatics
Oncology

ABSTRACT

Because of polypharmacy and complicated treatment protocols, adverse drug reactions (ADRs) continue to be a major problem in oncology and frequently lead to serious clinical complications. Recent developments in the use of artificial intelligence (AI) and machine learning (ML) for ADR prediction in anticancer therapy are critically assessed in this review. We go over a variety of methods for utilizing both structured and unstructured clinical data, such as supervised, unsupervised, and deep learning models in addition to natural language processing (NLP) strategies. Strong performance has been demonstrated by ensemble techniques like Random Forest and Gradient Boosting, while deep neural networks allow for sophisticated feature extraction, albeit with interpretability issues. We highlight new integrative techniques based on current literature trends, such as integrating demographic information, treatment history, and physiological signals with CNN-based models and SHAP-based.

Contents

1. Introduction . . . 2
2. Clinical implication and predictive modeling of ADRs in oncology . . . 3
2.1. Strategies for ADR prediction . . . 3
2.1.1. Pharmacovigilance . . . 3
2.2. Supervised machine learning (ML) approaches to ADR prediction . . . 4
2.2.1. Decision trees . . . 4
2.2.2. Support vector machines (SVMs) . . . 4
2.2.3. Random forests . . . 5
2.2.4. Gradient boosting machines (GBMs) . . . 5
2.2.5. AdaBoost (adaptive boosting) . . . 5
2.3. Unsupervised ADR prediction methods . . . 6
2.3.1. K-means clustering . . . 6
2.3.2. Deep learning (DL) in predicting ADRs . . . 6
2.3.3. Convolutional neural networks (CNNs) . . . 6
2.3.4. Artificial neural networks (ANNs) . . . 7
2.4. Natural language processing (NLP) and text mining . . . 7
2.4.1. Text mining and NLP for oncology ADR prediction . . . 7
3. Limitations and challenges of contemporary approaches in ADR prediction . . . 7
3.1. Limitations of traditional pharmacovigilance systems . . . 7

* Corresponding author.
E-mail address: afiya.eram@manipal.edu (A. Eram).

https://doi.org/10.1016/j.rineng.2025.106002
Received 5 May 2025; Received in revised form 14 June 2025; Accepted 26 June 2025
Available online 1 July 2025
2590-1230/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

3.2. Supervised machine learning model limitations . . . 8
3.3. Challenges in unsupervised ADR modeling . . . 8
3.4. Limitations of deep learning for the prediction of oncology ADR . . . 8
3.5. Limitations of NLP-based ADR signal extraction . . . 8
4. Data infrastructure and challenges in ADR predictive modeling for oncology . . . 9
4.1. Gaps in the literature . . . 9
5. Performance comparison of ML algorithms in oncology ADR prediction . . . 10
5.1. Integrated framework proposal . . . 10
5.2. Chemotherapy regimens and cancer type-specific drug profiles . . . 11
6. Discussion . . . 11
7. Conclusion and future directions . . . 12
CRediT authorship contribution statement . . . 12
Declaration of competing interest . . . 13
Acknowledgement . . . 13
Data availability . . . 13
References . . . 13

1. Introduction

The growing utilization of more intricate anticancer regimens has amplified the demand for strong predictive models that have the capability to foresee adverse drug reactions (ADRs) prior to their jeopardization of patient safety or treatment outcomes. This review discusses the role of machine learning (ML) in the prediction of ADRs in oncology, based on the considerable toxicities of chemotherapeutic and targeted therapies. ADRs not only disrupt treatment continuity but can also result in severe morbidity, and hence their early detection and proactive management strategies are crucial. Specifically, we discuss the current status of ML applications—ranging from supervised methods such as decision trees (DT), support vector machines (SVM), and ensemble algorithms; unsupervised techniques such as k-means clustering; deep learning models; and text-mining or natural language processing (NLP) tools—for ADR prediction.

In addition, we discuss pharmacovigilance systems for long-term safety surveillance. The review's emphasis is on identifying methodological advancements, pragmatic limitations, and research gaps in the existing literature. We also highlight the importance of including patient-level features, such as physiological measures (e.g., SpO_2, heart rate) and demographic data, to enhance prediction accuracy and clinical usefulness. Comparing their workflow, effectiveness, and interpretability, we seek to lay the foundations for a more sophisticated, personalized ML platform for ADR prediction in cancer therapy, and thus improved patient outcomes and efficiency of healthcare systems.

The increased clinical complexity of oncology treatment, combined with the high incidence rate of toxicities due to treatment, stresses the urgent need for predictive models capable of predicting cancer patients' ADRs. Because chemotherapy and targeted therapy drugs typically present low therapeutic windows, unexpected ADRs risk undermining the effectiveness of therapy, reducing patient compliance, and causing massive morbidity and health impact. Consequently, machine learning (ML) has emerged as a transformative tool, capable of finding complex, non-linear relationships in high-dimensional biomedical data [1]. Such prospects have led to the development of models that can predict ADRs from data derived from electronic health records, drug pharmacodynamics, patient-specific biomarkers, and treatment history [2–4].

Although initial studies have primarily centered on predictive performance gains—achieved through accuracy, sensitivity, and area under the curve (AUC)—most are deficient in clinical interpretability, reproducibility, and utility for real-world workstreams [5–7]. The more recent literature demonstrates the capacity of a variety of ML methods, from ensemble and semi-supervised methods, to enhance ADR detection. For example, studies by Kim et al. [8], Ho et al. [9], Xiao et al. [10], and Liu et al. [11] have determined that the utilization of real-world data and combining labeled and unlabeled sources increase robustness when dealing with predictive modeling. Data imbalance issues, sparse annotation, and external validation shortage remain ever-present and limit model generalizability across clinical settings. For example, Xiao et al. [10] showed that by learning latent clinical representations, integrating unlabeled electronic health record (EHR) data into a semi-supervised neural network improved model generalizability and captured uncommon or underreported ADR patterns. This is especially helpful in oncology, where labeled datasets are more scarce.

Notably, Ponraj and Charles [12] explored the use of artificial neural networks and random forest classifiers in drug discovery and ADR prediction with promising predictive performance from biological data. Nonetheless, the absence of external clinical validation and limited focus on model transparency lowers the generalizability of their findings to clinical translation. Likewise, Agbonkhese [13] employed classification algorithms for drug response prediction based on prescription patterns without properly addressing overfitting risk or false positives—key concerns in clinical decision support systems. At the same time, Chen [14] confirmed the predictive ability of ML approaches for ADRs but lacked comparative metrics against existing pharmacovigilance practices.

Together, the current body of work validates the promise of ML for ADR prediction enhancement but illustrates significant gaps in interpretability, validation, and integration of multi-modal patient data. The resolution of these gaps is critical to creating robust, personalized models capable of informing oncological practice and improving treatment safety.

Adverse drug reactions (ADRs) are a persistent challenge throughout the drug development pipeline. Although pre-clinical and clinical trials are designed to evaluate the safety and effectiveness of new treatments, they are inherently constrained by their reliance on relatively small, controlled patient populations and restricted follow-up periods [15,16]. These constraints guarantee that many ADRs—particularly those that are rare, delayed, or population-specific—will not be identified until after regulatory approval and broader population use.

To compensate for this limitation, the safety of approved medicines is closely monitored by post-marketing pharmacovigilance systems [17, 18]. These systems are necessary for the identification of ADRs that manifest over time due to cumulative exposure, long-term treatment effects, or interactions with other medications. In some cases, these ADRs may not manifest until many years or decades after a medicine has been marketed [16]. Thus, ongoing surveillance is required both to ensure patient safety and to inform dynamic, data-driven risk mitigation strategies.

Table 1 summarizes anticancer drugs in common use and their associated adverse drug reactions (ADRs). The drugs indicated—cyclophosphamide, oxaliplatin, paclitaxel, carboplatin, capecitabine, cisplatin, and doxorubicin—are utilized across numerous cancer types and have well-established toxicity profiles. By clustering each agent's most commonly reported ADRs, the table emphasizes the clinical importance

Table 1
Summary of Common ADRs in Frequently Used Anticancer Drugs (Adapted from Sunny et al., 2017 [19]).

Anticancer Drug	Common Adverse Drug Reactions (ADRs)
Cyclophosphamide	Fatigue, GI disturbances (nausea, vomiting, diarrhea), hematologic toxicity (neutropenia, anemia), mucositis, anorexia, fever
Oxaliplatin	Hematologic toxicity, GI symptoms, peripheral neuropathy, mucositis, anorexia, fever
Paclitaxel	Fatigue, hematologic effects, neuropathy, GI symptoms, alopecia, rash, muscle/back pain, sinus tachycardia
Carboplatin	Myelosuppression, GI disturbances, alopecia, rash, peripheral neuropathy, myalgia
Capecitabine	GI toxicity, myelosuppression, fatigue, mucositis, skin reactions, fever
Cisplatin	Nausea, vomiting, nephrotoxicity (dysuria), hematologic toxicity, hypokalemia, muscle pain, alopecia, dyspnea
Doxorubicin	Mucositis, fatigue, myelosuppression, alopecia, nausea, cardiac effects (sinus tachycardia), dermatologic changes

Note: This table's adverse drug reactions (ADRs) are taken from Sunny et al. (2017) [19]. To guarantee consistency and readability within the framework of this review, the adaptation entails choosing representative medications that are frequently used in oncology as well as summarizing and rewording ADR descriptions.

of proactive toxicity monitoring and depicts the variety and intensity of side effects that must be dealt with in the course of cancer treatment.

2. Clinical implication and predictive modeling of ADRs in oncology

Adverse drug reactions (ADRs) are a key hindrance to effective cancer therapy, particularly with the introduction of targeted therapies and immunotherapies. Even as they augment oncologic benefits, these novel regimens have been associated with an increased frequency of complex and often severe ADRs [20–22]. Effective pharmacovigilance mechanisms are therefore important to track, report, and manage these toxicities [23–27]. However, the traditional reliance on passive reporting systems will result in underreporting, particularly in patients who have undergone treatment through oral therapies with fewer clinical contact points [28]. This proves the need for active predictive models and data-driven solutions.

Recent research has indicated that predictive modeling can potentially be critical in early detection of ADRs, enhancing patients' quality of life and clinical outcomes. Monestime et al. [28] noted that the integration of real-world evidence into risk assessment tools allowed for the early identification of patients at risk, facilitating timely therapeutic adjustment. Javaid et al. [29] also emphasize the clinical applicability of ML algorithms in tailoring treatment plans through the prediction of individualized ADR risk profiles.

Advances in molecular biology have also helped predict ADR by determining genetic and epigenetic biomarkers—like long non-coding RNAs—linked to drug resistance and toxicity [30–33]. When integrated into predictive models, these molecular factors enable more precise therapy personalization. For example, Imai et al. [34] reported that artificial neural networks (ANNs) achieved 85% accuracy in predicting chemotherapy-induced ADRs upon training on multi-omics data. Bharatiya et al. [35] and Ben Youssef et al. [36] later demonstrated that machine learning could effectively identify at-risk cohorts using complex clinical and genomic information.

Yet, while the majority of models are highly accurate, they typically operate as 'black boxes' with low interpretability—preventing them from being used in clinical practice. Additionally, there are extremely few studies that conduct external validation, and there are numerous of which utilize small, institution-specific datasets. Bellanca et al. [37] highlighted incidence differences by ADR type of cancer and by treatment modality but did not assess generalizability of models between settings or populations. Moreover, performance metrics such as sensitivity, specificity, and false positives are not consistently reported, such that it becomes challenging to make critical comparisons across models.

ADR prediction in oncology has been positively impacted by machine learning and integrating molecular data, there are considerable challenges still present. These range from the importance of transparent and interpretable models to having available diverse and high-quality training sets and stringent external validation. Insights from oncology can provide invaluable lessons for other predictive frameworks across infectious disease, including tuberculosis (TB). Prolonged multidrug

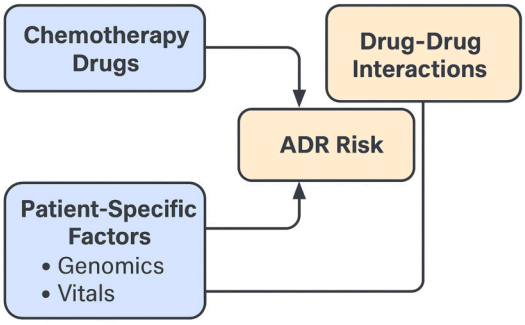


Fig. 1. Conceptual representation of factors influencing ADR prediction in oncology.

treatment-induced ADRs also present a major roadblock to achieving treatment success. As the Fig. 1 shows, the ADR prediction is built upon an interplay between factors that are unique to patients, treatment protocols, and computational model approaches.

Prior to classification, we employ intermediate fusion by merging latent features from CNN/LSTM, NLP, and structured data. During training, the "Optimize Loss, Sensitivity, Specificity" loop is used to match predictions with clinical priorities through threshold tuning or composite objectives.

2.1. Strategies for ADR prediction

2.1.1. Pharmacovigilance

Pharmacovigilance (PV) is one of the pillars of monitoring drug safety post-marketing. The World Health Organization (WHO) defines PV as "the science and activities related to the collection, assessment, and prevention of adverse effects or any other undesired effect produced by a drug" [38]. PV has conventionally relied on spontaneous reporting systems, clinical case reports, and regulatory databases. While these are the foundations of PV, they are presently besieged by their inability to predict proactively ADRs.

One of the inherent limitations of traditional PV is its reactive aspect. Most of the reports come after the occurrence of the undesirable event, when opportunities for prompt intervention are being lost. Underreporting is still a ubiquitous issue—most notably in low-resource settings and among groups with poor access to healthcare or limited digital literacy. Besides, the voluntary reporting aspect introduces a large bias towards well-known or significant ADRs, leaving more subtle or emerging patterns undetected [39,40]. The data generated are generally sparse, lagged, and non-standardized, which diminishes their utility for predictive model construction.

Li et al. [39] gave a comprehensive review of PV systems across Europe, the USA, and China, mentioning lifecycle monitoring in particular, targeting special groups such as children, the elderly, and pregnant women. The authors acknowledged that the systems are generally descriptive, though, and lack predictive with analysis ability. Similarly, Baldo et al. [40] drew attention to PV's significance in oncology but not

operationalizing collected data for future risk stratification or in machine learning initiatives.

Recent discussion has begun examining how PV has evolved towards a yet smarter, adaptive system. Tadge et al. and MedText [41,42] suggest that artificial intelligence (AI) and real-world evidence (RWE) can reframe PV from historic surveillance to prospective prediction of risk. These new methods provide more sensitive signal detection and pattern recognition but introduce new challenges too. The majority of AI-assisted PV tools are black boxes (i.e., opaque), and their accuracy is sensitive to completeness and quality of data. Moreover, privacy regulations, interoperability issues, and the absence of standard model validation frameworks prevent extensive application in clinical settings. In the regulatory, academic, and industrial sectors, a number of AI-assisted pharmacovigilance (PV) systems are presently in pilot or limited use. To help with signal detection from claims and EHR data, for example, the U.S. FDA's *Sentinel System* uses machine learning models within its distributed data network. Although its internal workings are still unknown, the *World Health Organization's VigiBase* has integrated a ranking algorithm called *VigiRank* that uses artificial intelligence (AI) to prioritize case reports based on factors like report completeness and geographical distribution. NLP is used by academic projects like *MedWatcher Social*, which was created with FDA assistance, to extract ADR signals from social media data; however, this is difficult because of noise and ambiguity.

Although they are theoretically attractive, few AI-supported PV models have comprehensively improved ADR detection sensitivity and specificity in more than one therapeutic area and population. For instance, while pilot studies using machine learning for PV have been described with increased signal detection rates, these are frequently lacking in proper benchmarking against conventional methods. Therefore, the clinical usefulness of such tools is as much theoretical as real.

Pharmacovigilance systems establish a necessary but inadequate foundation for ADR monitoring. Their future research should aim for model interpretability, cross-institutional data harmonization, and stringent external validation to enhance their predictive capacity. Without these advances, PV will lag behind addressing the complex, multifaceted nature of ADRs particularly in rapidly evolving fields such as oncology and infectious diseases.

2.2. Supervised machine learning (ML) approaches to ADR prediction

The most important machine learning algorithms that shall be employed in this study for comparison are decision trees, support vector machines (SVM), ensemble methods, deep learning-convolutional neural networks and artificial neural networks along with k-means clustering for unsupervised. Additionally, text-mining and natural language processing (NLP) shall also be employed.

Supervised learning is a type of machine learning course that acquires a mapping function from input feature to predetermined output labels by learning labeled training data [43]. Supervised models rely on external supervision within annotated datasets in a bid to reduce prediction error and improve performance in an iterative process [44,45]. For pharmacovigilance purposes, controlled methods are best suited to forecast ADRs because they can leverage structured data sources e.g., patient demographics, treatment history, and clinical laboratory findings to model risk profiles.

Popular supervised methods applied to ADR modeling include decision trees, support vector machines (SVMs), and ensemble models like Random Forests and Gradient Boosting. Such methodologies have shown robust performance in a variety of biomedical applications, especially when dealing with large, high-featured data sets. SVMs, for example, are particularly adept at recognizing non-linear decision boundaries, whereas ensemble models improve generalizability and reduce overfitting by combining multiple weak learners. Their success in oncology and more generic pharmacovigilance investigations suggests potential application in TB treatment contexts, where ADRs due to multidrug regimens remain a significant clinical headache. There have been several recent studies on the convergence of machine learning and biomedical application. Specifically, recent advancements in nanoparticle-based cancer therapy [46], deep learning for protein structure prediction [47], and machine learning techniques in additive manufacturing [48] highlight the diverse biomedical and healthcare applications of AI.

However, supervised approaches are based on high-quality, labeled data—something that is often not found in many real-world healthcare environments. In addition, model performance generally rests on domain-specific feature design, hyperparameter optimization, and external validation across diverse patient populations. In this way, while the robust methodology of supervised prediction for ADR is considerable, its use in clinical practice rests on rigorous methodological design and verification.

2.2.1. Decision trees

Decision trees are intuitive and interpretable machine learning models that represent decisions as a hierarchy, where each node corresponds to some feature split [43,44,49]. The models explicitly learn decision rules from training data, such that simple classification or regression is enabled. They have also been applied effectively in the prediction of ADR; for instance, Hammann et al. developed a model based on compounds' chemical and structural properties with precision between 78.9% and 90.2% in the prediction of potential ADRs [50–52]. Despite these advantages, decision trees are vulnerable to overfitting especially when used on small or noisy datasets and are sensitive to minor data distributional shifts. These limitations undermine their generalizability in real-world clinical practice. Moreover, while decision trees have demonstrated potential in oncology and general pharmacovigilance, their use in the specific scenario of tuberculosis-related ADR prediction is less researched, representing an avenue of potential future study. There are several reasons for this gap in the literature. First, the majority of TB cases are treated in low- and middle-income countries (LMICs), where there are frequently few digitized clinical datasets and ADR annotations. This restricts the use of machine learning methods that need labeled data. Second, computational ADR modeling has historically received less attention in TB research than infection control, medication resistance, and treatment adherence. Third, multifactorial etiologies involving genetic, immunological, and pharmacokinetic interactions are frequently responsible for TB-related adverse drug reactions (ADRs), including hepatotoxicity, rash, and peripheral neuropathy. This has led to the use of more complex models at the expense of interpretability. However, because of their high transparency and low computational overhead features that are particularly appealing in healthcare settings with limited resources decision trees continue to be a promising but underutilized method.

2.2.2. Support vector machines (SVMs)

SVMs are supervised machine learning algorithms that attempt to identify an optimal hyperplane in a high dimensional feature space that maximally separates data points belonging to different classes [12,44]. The learning procedure involves solving a convex optimization problem to minimize structural risk, thereby enhancing model generalizability. Through the application of kernel functions i.e., radial basis function (RBF), polynomial, or sigmoid—SVMs can effectively transform nonlinearly separable data into higher-dimensional spaces for linear separation [53,54]. These attributes make SVMs ideally suited for biomedical applications involving noisy, sparse, or high-dimensional data such as gene expression matrices or ADR monitoring reports [55]. Yaseen et al. [56] employed SVMs to predict drug-target interactions (DTIs) at accuracy 93.76% and AUC of 0.753 ± 0.006. While this performance was slightly lower than more complex network-based approaches (AUC $= 0.886 \pm 0.010$), the SVM model was a solid and computationally lighter option. While not designed with ADR modeling in mind, this study illustrates the broader relevance of SVMs to pharmacological data analysis.

In oncology pharmacovigilance, where both treatment complexity and patient heterogeneity are present in noisy datasets, SVMs can prove to be especially helpful in ADR prediction since SVMs are resistant to overfitting and can model interactions of high dimensionality.

Having said that, the approach is not without limitations. SVMs possess quadratic training complexity relative to dataset size and are therefore computationally expensive for large-scale ADR monitoring. They also behave reasonably to kernel and hyperparameter tuning (i.e., cost C, kernel parameter γ), and the resulting decision boundaries are often difficult to interpret clinically. To increase clinical usability, ADR-predicting SVMs need to be paired with interpretability enhancing techniques such as SHAP or LIME. With careful tuning and validation, they can offer a compelling balance between accuracy and generalizability—particularly for detecting less frequent but clinically relevant ADRs in cancer treatment datasets. SVMs are classified as "Moderate" in terms of clinical readiness, despite their strong classification accuracy, especially in high-dimensional biomedical datasets. This rating takes into account real-world restrictions on deployment, scalability, and interpretability. To establish trust and guarantee alignment with medical reasoning, clinicians frequently demand models with transparent decision paths; however, SVMs employ high-dimensional hyperplanes and intricate kernel functions that are not naturally explicable. Furthermore, SVM training complexity increases quadratically with dataset size, which can be problematic in clinical settings where large amounts of EHR data are handled. Finally, integration into clinical decision support systems is less intuitive due to their decision logic, which is based on margin maximization rather than interpretable rules. Even though they perform well on benchmark tasks, these factors taken together limit their transfer from research to clinical settings.

2.2.3. Random forests

Random forests are ensemble methods that enhance prediction performance by aggregating the output of many decision trees, each learned from different bootstrap samples with randomly selected feature subsets [57]. This design reduces model variance and overfitting—common issues for single decision trees—by promoting diversity between base learners. The final prediction is realized by majority voting (classification) or averaging (regression), making Random Forests suitable for handling high-dimensional, non-linear biomedical data.

Random Forests have been employed in oncology pharmacovigilance to predict complicated drug-event interactions. As an example, Ietswaart et al. [58] developed a Random Forest classifier to predict ADRs in various drugs with high generalization performance on a test collection of drugs left out from the training phase. While specific metrics were not entirely described, the study highlighted the algorithm's ability to identify significant, not before interpreted drug-ADR associations—strongly indirectly suggesting great potential for applied usage in pharmacovigilance.

In spite of these benefits, Random Forests have their drawbacks. As the trees grow in number, the model becomes computationally demanding, especially at inference—challenging integration into real-time clinical pipelines. Also, interpretability is still limited: while feature importance scores may provide some insight, the underlying decision paths are challenging to follow through, and thus clinical trust and transparency are impacted.

In general, Random Forests tightly balance predictive accuracy against stability, particularly in the case of high-dimensional biomedical data. Their ability to detect complex drug-ADR interactions renders them a likely choice for pharmacovigilance platforms in oncology, provided concerns regarding computational cost and interpretability are dealt with effectively.

2.2.4. Gradient boosting machines (GBMs)

GBMs are ensemble learning algorithms that construct sequentially an orderly series of shallow decision trees, with each tree individually trained to predict the residuals of the preceding tree. Through gradient descent minimization of a differentiable loss function, GBMs make increasingly more accurate predictions at each step, enabling solid performance on challenging, non-linear classification tasks [59]. Compared to variance-reducing bagging methods like Random Forests, GBMs are optimized for model bias in a sequential manner and generally yield improved accuracy at the cost of increased computational effort and parameter tuning.

Although Malik et al. [59] applied GBM to predict pan evaporation in environmental science, the underlying architecture is highly adaptable for biomedical applications. In pharmacovigilance oncology, GBMs have shown promise in ADR prediction by being able to manage heterogeneous, feature-rich data such as drug dosing patterns, treatment regimens, and lab results without sacrificing robustness to missing values and nonlinear interactions.

For instance, in Zhao et al.'s research, a GBM-based ADR predictive model from electronic health records was proposed. The model achieved an AUC of 0.89 and outperformed logistic regression as well as traditional decision tree methods. This indicates the power of the model to pick up subtle risk signals in the presence of high-dimensional clinical data.

While strong predictors, GBMs are not without flaws. With ease, it will overfit if allowed to execute in its native, unregularized state, so it needs to be seriously regularized and optimized regarding the learning rate, tree depth, and estimators' hyperparameters. And, as with most ensemble learners, GBMs are not interpretable, at least not directly, a significant hurdle to clinical deployment.

In cancer, GBMs are best suited to mimic multifactorial ADR risks, particularly when integrated with patient-specific and treatment-related variables. However, for clinical usefulness, their performance must be validated across different patient cohorts, and efforts must be made to enhance model explainability through techniques such as SHAP or feature attribution frameworks.

2.2.5. AdaBoost (adaptive boosting)

AdaBoost is a comparatively ancient but extremely influential boosting algorithm, designed to transform a collection of weak learners, typically shallow decision trees, into a high-performing ensemble classifier. It trains base classifiers sequentially and dynamically rescales sample weights: more weight is assigned to misclassified instances in subsequent rounds, so the model can concentrate more on harder cases to classify.

Hu et al. [60] applied AdaBoost to forecast ADR in a current investigation and achieved superior performance to other ensemble methods, including Gradient Boosting and Bagging. The model achieved area under the curve (AUC) of 0.861, F1-score of 0.837, and precision of 0.812, indicating the capacity of AdaBoost to handle imbalanced pharmacovigilance datasets with high predictive accuracy.

While robust, AdaBoost is noisy data- and outlier-sensitive due to the algorithm's strong dependence on challenging-to-classify instances, potentially mislabeled or irrelevant samples. Additionally, the iterative process of training increases computational expenses, especially for high-volume biomedical datasets such as electronic health records (EHRs) or omics-based drug response data. Interpretability is also a problem since the final ensemble can conceal the contribution of an individual feature, where relevant features may not have their effect transparent.

AdaBoost holds potential for the optimization of risk stratification models of oncology pharmacovigilance, in which ADRs are infrequent and heterogeneous. Its ability to emphasize under-represented ADR patterns can be particularly valuable for early detection and intervention. However, model calibration and the selection of domain-specific features must be accomplished with caution to avoid overfitting and guarantee clinical applicability. Attempts to integrate AdaBoost into explainability pipelines and external validation procedures will be crucial for translating it to practical cancer care settings. Table 2 enumerates key

Table 2
Comparison of Supervised Machine Learning Models for ADR Prediction.

Model	Accuracy	Interpretability	Complexity	Clinical Readiness
Decision Tree	Moderate–High	High	Low	Good
SVM	High	Moderate	High	Moderate
Random Forest	High	Moderate	Moderate	Good
GBM	Very High	Low	High	Moderate
AdaBoost	High	Moderate	High	Moderate

characteristics of well-known supervised machine learning models applied for ADR prediction, compared with respect to accuracy,

Both dataset-specific limitations and the intrinsic characteristics of the models themselves are to blame for the generalization limitations seen with supervised machine learning models. When it comes to data, a lot of supervised models are trained on limited cohorts that are unique to particular cancer types, organizations, or geographic populations. This results in overfitting and decreased external validity when used in different contexts.

2.3. Unsupervised ADR prediction methods

Unsupervised learning involves the process of model training on unlabeled data so that the algorithm identifies patterns, groupings, or structure automatically without pre-existing output labels [44,61,62]. In the case of predicting adverse drug reactions (ADRs), unsupervised methods are highly beneficial for exploratory data analysis along with hypothesis generation. Such techniques are able to reveal latent structures in clinical datasets—e.g., clustering patients by ADR profiles, grouping drugs by similar toxicity profiles, or discovering subgroups likely to exhibit differential responses to treatment regimens. Unlike supervised techniques, they are applicable where data are scarce and do not rely on expert-labeled outcomes. This makes them especially valuable for early-stage oncology pharmacovigilance research or underreported ADR cases.

2.3.1. K-means clustering

K-means is a well-known unsupervised method for partitioning data into k non-overlapping clusters based on feature similarity. It starts by initializing cluster centroids and updating them iteratively to minimize intra-cluster variance [44,63,64]. K-means clusters patients or drugs with similar profiles effectively by assigning all data points to the nearest centroid.

In oncology, Ahmadian et al. [65] applied K-means to electronic health records (EHRs) for clustering patients with similar treatment response and ADR profiles. Their work demonstrated the potential for clustering to identify abnormal patterns of adverse events that would be worthy of further investigation, e.g., in situations involving off-label drug use or multi-agent chemotherapy regimens.

Despite its advantages, K-means has known limitations. It assumes convex, isotropic clusters and requires the number of clusters k to be prespecified, neither of which may be reconcilable with the underlying structure of clinical data. One of the most important steps in using k-means, especially in pharmacovigilance settings, is estimating the ideal number of clusters (k). The *elbow method* is a popular heuristic technique that plots the within-cluster sum of squares (WCSS) to determine the point at which adding more clusters no longer significantly reduces variance. Despite its widespread use, this approach may be arbitrary. Higher average silhouette scores indicate better cluster cohesion and separation. *Silhouette analysis* offers a more robust metric by assessing how well each point fits within its assigned cluster in comparison to others. Applications like text-derived feature space clustering and ADR signal grouping have shown success with this. The algorithm is also sensitive to the initialization of centroids and can converge on local minima, giving non-reproducible results between runs.

Nonetheless, K-means remains useful for hypothesis generation and exploratory data analysis in ADR research. In cancer pharmacovigilance, it can be utilized to define subpopulations at risk for delayed-onset or rare toxicities as a precursor to more complex modeling. Its usefulness can be enhanced by pairing dimensionality reduction techniques such as Principal Component Analysis (PCA) or t-SNE to enable cluster separation in high-dimensional space.

2.3.2. Deep learning (DL) in predicting ADRs

DL is a subset of machine learning derived from neural networks that learn to operate on vast amounts of data and comprehend unstructured data as well as generate general observations without the requirement for manual feature extraction [66]. DL finds extensive use in object and speech recognition. DL is replicating the human brain's learning process. Shuang Li et al. [67] presented a deep learning framework aimed at ADR prediction in new drugs from a binary network-based message passing through the neural network integrating drug expression to all ADR and its feature representation. It was found to assist pharmacists and HCPs to comprehend potential dangers of drug reaction and to correct underreporting issues. Jhih-Yuan Huang et al. [68] also proposed a deep learning approach using the Bidirectional Encoder Representations from Transformers (BERT) model with a new batch-wise adaptive approach to enhance the predictive ability, which was pre-trained on a big corpus. They used the BERT-design for language encoding to tackle the diversity of ADRs-related terms and used the pre-trained model to tackle the imbalanced dataset.

2.3.3. Convolutional neural networks (CNNs)

CNNs are deep learning-based neural networks that were originally developed for processing image data [69]. The general structure of a CNN includes three major layers: convolutional layers to detect local patterns, pooling layers to reduce dimensionality, and fully connected layers to produce the output. Although CNNs have historically been image-focused, some recent variants have used them with structured biomedical information by converting electronic health records (EHRs), drug-target interaction matrices, or genomic sequences into grid-structured representations suitable for convolutional operations.

Li et al. [70] employed a CNN model to predict ADRs and obtained improved performance over other baseline models. The CNN attained an accuracy rate of 91.4% and AUC of 0.88, representing strong prediction capability and credibility in a well-organized drug-response dataset.

In oncology, CNNs are applicable to high-throughput molecular data (e.g., proteomics, gene expression) or structured EHR embeddings to find subtle patterns for drug toxicity. Their spatial hierarchy-learning ability and capacity to represent complex, non-linear interactions place CNNs at the forefront as a technology to detect ADR in multidimensional cancer data sets.

However, CNNs are not limitation-proof. They tend to require enormous amounts of tagged training data in order to not overfit and are computationally costly to train. Additionally, their black-box nature limits interpretability—a highly desirable aspect when it comes to clinical decision-making. Efforts to improve transparency using saliency maps or gradient-based interpretation methods are on-going but have not yet reached the stage of being routinized in ADR modeling workflows.

Despite such challenges, CNNs have tremendous potential when applied to structured clinical and omics data, especially in the field of oncology, where the unification of multi-modal information becomes crucial for personalized therapy and minimization of drug-induced injury.

2.3.4. Artificial neural networks (ANNs)

ANNs are computational models inspired by the structure and function of the biological brain. They consist of layers of interconnected nodes—or "neurons"—that encode input data as weighted summations and non-linear activation functions to generate output predictions [71]. ANNs are particularly good at modeling complex, non-linear relationships in high-dimensional data and have been widely applied in biomedical applications for classification, regression, and pattern recognition problems.

Through iterative training using algorithms like backpropagation and stochastic gradient descent, ANNs adjust their internal parameters based on error feedback to minimize prediction loss and optimize performance [72]. This ability to learn enables them to generalize from historical data, making them susceptible to adverse drug reaction (ADR) prediction after training on large-scale electronic health records (EHRs), pharmacogenomic profiles, or drug usage trajectories.

For example, Zhang et al. [73] demonstrated the application of ANNs in automating medical image analysis, significantly reducing the time and subjectivity of traditional evaluation protocols. Although the study was imaging-based, the principles highlight the ANN's proficiency in feature extraction and pattern recognition—capabilities that can be extended to ADR prediction in oncology or tuberculosis, where clinical information can be framed and coded for supervised learning.

However, ANNs also suffer from significant drawbacks. ANNs require enormous quantities of labeled training data to avoid overfitting and may be susceptible to hyperparameter sensitivity, such as learning rate, network depth, and regularization. In addition, their decisions are generally not interpretable, which is challenging for clinical interpretability and trust. Hence, it is extremely crucial to couple ANN predictions with explainability techniques (e.g., LIME, SHAP) for real-world deployment in the context of pharmacovigilance.

Despite these challenges, ANNs are a highly successful modeling approach to ADR detection when combined with domain-specific feature engineering and adequate validation practices.

2.4. Natural language processing (NLP) and text mining

2.4.1. Text mining and NLP for oncology ADR prediction

Text mining and Natural Language Processing (NLP) have become vital pharmacovigilance tools with a particular application in enhancing adverse drug reaction (ADR) prediction in oncology. These technologies facilitate the extraction of clinically useful information from large amounts of unstructured data sources such as EHRs, physician notes, and patient narratives normally inaccessible to traditional structured-data models [74]. Non-image inputs, like genomic sequences or EHRs, are usually converted into grid-structured representations so that CNNs can process them. Building a temporal matrix with rows representing clinical features (such as diagnoses, prescriptions, and lab results) and columns denoting discretized time intervals is a popular method for working with EHR data. A numerical value, an embedding vector, or a binary presence indicator could be present in every cell. Multivariate pattern learning can be facilitated by grouping feature types into distinct channels, similar to the RGB channels in images. By converting free-text data to structured formats, NLP allows the systematic extraction of drug-event relationships, comorbidities, treatment regimens, and temporal relationships that are crucial for ADR risk assessment.

NLP has been especially useful in oncology, where critical diagnostic and prognostic data is often locked in narrative reports. For instance, Datta et al. [75] developed a framework to extract cancer-related biomarkers and prognostic variables from free-text surgical pathology reports, allowing for more personalized treatment stratification. Similarly, Wang et al. [76] utilized an NLP system for extracting disease stage, histological subtype, and treatment modalities from clinical notes of patients with lung cancer with high precision and recall rates increasing the granularity of patient phenotyping downstream for ADR modeling.

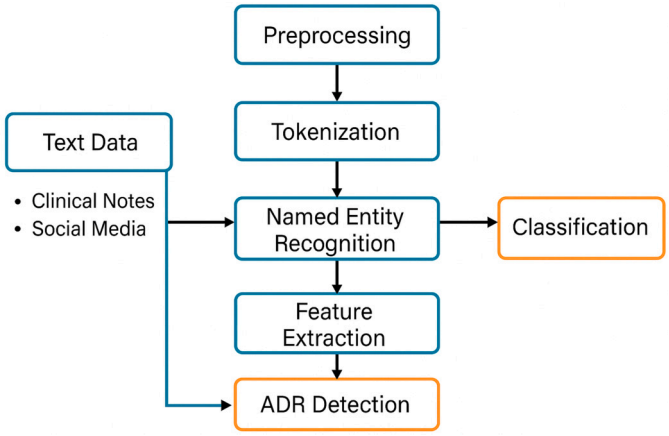


Fig. 2. Illustrative pipeline for ADR extraction from clinical notes and social media using NLP.

Moreover, Houssein et al. [77] emphasized the necessity of NLP for the extraction of pharmacological patterns and ADR events from EHRs in complex treatment regimes like chemotherapy, where drug effects and patient responses are extremely heterogeneous. This capability is essential for populating machine learning models with dense, real-world clinical features that would otherwise not be reflected in structured databases.

Yet, text mining approaches are not without limitations. NLP models must contend with ambiguous language, non-standardized clinical terms, and errors introduced during the documentation process. Additionally, generalizability to other institutions is typically limited due to variations in EHR system architectures and reporting conventions. Model robustness is therefore assured only by large annotated corpora and regular retraining on in-domain texts.

Despite these challenges, the integration of NLP and ADR prediction models holds great promise especially in oncology, where the richness of clinical narratives offers untapped potential for early risk detection and personalized pharmacovigilance. Fig. 2 illustrates natural language processing that enables ADR signal extraction from unstructured sources through a pipeline that includes multiple steps of preprocessing, named entity recognition, and classification.

3. Limitations and challenges of contemporary approaches in ADR prediction

3.1. Limitations of traditional pharmacovigilance systems

Traditional PV systems, while necessary for post-marketing drug safety surveillance, are hampered by severe challenges that downplay their predictive capability. Chief among these is the long-standing issue of under-reporting—stemming from the voluntary nature of patient and healthcare professional ADR reporting—that leads to biased and incomplete data [16,78]. The absence of obligatory reporting procedures contributes to delayed signal detection and truncates the external validity of risk assessment.

Moreover, current PV infrastructures are resource-intensive and reactive rather than proactive. Reporting, validating, and analyzing cases are associated with high administrative expenses, yet these activities barely result in timely ADR risk reduction. This inefficiency points to the inadequacy of conventional surveillance models under dynamic clinical settings, particularly in oncology, where treatment complexity demands faster, more precise information.

Another significant limitation is pharmacovigilance databases' heterogeneity and fragmentation. Unstandardized terminologies, reporting standards, and data formats are barriers to platform integration and constrained the scalability of machine learning application. Moreover,

traditional PV systems lack capabilities for the identification of rare, delayed, or polypharmacy-related ADRs—events commonly encountered in oncology due to multidrug regimens and lengthy treatment durations [79–82].

These limitations highlight the need to augment pharmacovigilance with predictive analytics that have the capability to leverage large, heterogeneous data. The latter must integrate structured clinical data with real-world evidence in order to evolve from passive monitoring to proactive risk stratification.

3.2. Supervised machine learning model limitations

Decision trees, random forests, support vector machines (SVMs), AdaBoost, and gradient boosting machines (GBMs) supervised learning algorithms are the foundation of the majority of adverse drug reaction (ADR) prediction models. Supervised learning algorithms make predictions on new examples by relying on large quantities of labeled training data to learn useful patterns. In oncology pharmacovigilance, however, high-quality, accurately labeled ADR datasets remain in short supply [83]. This dearth of data is a powerful barrier to model validity and leads to overfitting, where models perform outstandingly well on training data but cannot generalize to novel clinical settings [84,85].

Interpretability likewise seems to be a critical constraint. Decision trees, though usually interpretable, are less so with increased tree depth—particularly where complicated treatment regimens, drug-drug interactions, or dynamic clinical statuses are involved [52]. Ensemble methods like random forests and GBMs boost predictive power at the expense of interpretability, limiting their utility where interpretable decisions must be made in order to enable clinical adoption.

Support vector machines (SVMs), although well suited to high-dimensional feature spaces, are very sensitive to the presence and distribution of labeled data. They are also influenced by kernel selection and hyperparameter tuning, and their decision-making processes are mostly transparent [86].

Boosting algorithms such as AdaBoost enhance classification accuracy, particularly for imbalanced data. Nevertheless, their sensitivity to noisy and misclassified instances can warp decision boundaries, especially when outliers overwhelm the training set [87]. Likewise, GBMs, although robust, are computationally intensive and need careful parameter tuning (e.g., learning rate, tree depth) to prevent overfitting and instability [87].

In combination, these limitations highlight the challenge of model complexity vs. interpretability, data dependence vs. generalizability—questions that must be seriously balanced in using supervised models to forecast ADR in oncology.

3.3. Challenges in unsupervised ADR modeling

Unsupervised learning techniques, including clustering algorithms such as K-means, offer an alternative approach to exploring adverse drug reaction (ADR) patterns when labeled data is not available. These techniques are helpful for uncovering latent patterns, patient subgroups, or novel drug-event associations in large clinical datasets. Their application in oncology pharmacovigilance, however, presents considerable challenges.

K-means clustering, although widely used due to its simplicity and scalability, requires the specification of the number of clusters (k) beforehand—a serious limitation when the underlying data distribution is unknown or highly heterogeneous [86,88,89]. Specification of an incorrect value of k can result in underfitting (too few clusters) or overfitting (too many clusters), distorting the interpretation of ADR patterns and potentially misleading downstream clinical conclusions.

Furthermore, K-means assumes that the clusters are convex, isotropic, and homogeneous, which is rarely the case in heterogeneous biomedical data. In cancer, patient therapeutic responses are determined by many overlapping variables such as genetics, comorbidities, and drug combinations—attributes that tend to defy the assumptions of standard clustering models.

In addition, K-means is also initial position sensitive in terms of centroid and can get trapped in local optima in the case of high-dimensional or sparse data. This kind of instability makes outcomes less reproducible, particularly where clustering is being used for patient stratification or risk characterization.

Unsupervised approaches need to be blended with dimension reduction approaches (e.g., PCA, t-SNE) and measures of cluster validity to improve robustness. While these enhancements notwithstanding, because of the lack of any explicit labeling and limited clinical interpretability, they are most suitable for exploratory and not confirmatory ADR prediction.

3.4. Limitations of deep learning for the prediction of oncology ADR

Deep learning (DL) models, e.g., convolutional neural networks (CNNs) and artificial neural networks (ANNs), have shown immense potential to handle complex, non-linear relationships among biomedical data. However, there are a few major limitations of their application for predicting adverse drug reaction (ADR) in the context of oncology.

Maybe the most critical issue is algorithmic bias and injustice. Such systems are extremely sensitive to the composition of their training set. If the data are skewed or homogeneous—i.e., capture some groups over-representatively and underrepresent or omit others—DL systems can amplify or even exacerbate health inequities in ADR risk prediction [90]. This is especially an issue in cancer treatment, where heterogeneity among patients in genetics and comorbidities is a big player in drug response.

CNNs, which are used in imaging and structured data transformation, require extensive annotated datasets in order to learn generalizable features. Their reliance on extensive, diverse input data increases data preprocessing and annotation time and resource usage. Overfitting remains a threat despite the use of regularization methods [91].

Artificial Neural Networks (ANNs), while more tolerant of input structure, suffer from limitations in continuous learning. ANNs are susceptible to catastrophic forgetting—where new data overwrite old information—so it is challenging to add new pharmacovigilance signals over time. ANNs also offer no guaranteed convergence to an optimal solution and can become stuck in local minima, especially when trained on noisy or high-dimensional ADR datasets [92].

Finally, interpretability of DL models remains a fundamental challenge. Unlike decision trees or linear models, DL architectures are black boxes, offering limited transparency into feature contributions or decision logic. This transparency limitation restricts their uptake in clinical workflows where explainability is critical.

DL models must be paired with robust data governance, fairness audits, and explainable AI techniques (e.g., SHAP, attention mechanisms) in order to enhance their usability in the prediction of oncology ADRs. Without these practices, scalability and credibility in pharmacovigilance are constrained.

3.5. Limitations of NLP-based ADR signal extraction

Text mining and Natural Language Processing (NLP) hold promising potential for identifying adverse drug reactions (ADRs) in unstructured sources such as electronic health records (EHRs), clinical narratives, and social media. The two methods are valuable in unearthing signals, which structured sets are prone to overlook. Their value is, however, restricted in the use in oncology pharmacovigilance by several critical limitations.

One of the key challenges is language inconsistency and ambiguity in unstructured data. Clinical notes include abbreviations, misspelling, and idiosyncratic language, while social media data include informal grammar, sarcasm, and noise. Such linguistic heterogeneity makes it difficult to accurately extract ADR-related information [74,93–96]. NLP

Table 3
Summary of Limitations in ADR Prediction Approaches.

Category	Key Limitations	Cited Sources
Pharmacovigilance	Underreporting, voluntary bias, slow update cycles	Pilipiec et al. (2022), Costa et al. (2023)
Supervised ML	Dependent on large, well-labeled datasets; overfitting; limited generalization	Sarker et al. (2015), Imran et al. (2022)
Unsupervised ML	Difficulty defining clusters; choice of K; poor clinical interpretability	Yuan et al. (2019), Ikotun et al. (2023)
Deep Learning	High data requirement, risk of overfitting, black-box nature	Krichen et al. (2023), Ahmed et al. (2023)
NLP/Text Mining	Noisy and unstructured input, contextual ambiguity, domain adaptation issues	Kumar et al. (2016), Khan et al. (2020)

methods can read the context incorrectly—e.g., confusing negated statements ("no signs of nausea") with positive ADR events—or unable to distinguish between hypothetical, past, and present references to adverse events.

In addition, jargon specific to oncology in the field makes it more challenging. Identification of ADRs for targeted agents or immunotherapies requires awareness of clinical subtleties and drug-specific lexicons, for which general-purpose NLP tools are poorly trained. Errors in named entity recognition, relation extraction, and sentiment analysis can lead to false positives or missed ADR signals.

Besides, NLP models trained on the data of a single institution are not generalizable. Variations in documentation practices, EHR formats, and local medical vocabularies limit model portability across healthcare systems, inhibit large-scale deployment, and erode model stability.

Effective NLP-based ADR discovery systems must utilize state-of-the-art techniques such as contextual embeddings (e.g., BERT), domain ontologies, and error correction based on rules for overcoming these drawbacks. External validation against multi-institutional corpora and harmonization against expert-curated dictionaries are equally important to ensure trust in clinical use. In NLP pipeline, domain adaptation is essential, especially when transferring ADR prediction models between healthcare institutions or between oncology sub-specialties (e.g., hematologic vs. solid tumors). Variability in terminology, abbreviations, coding conventions, and clinical note structure can cause distributional shifts that impair model performance. Table 3 summarizes key limitations concerning various ADR prediction approaches, tagging methodological and practical issues identified in the literature.

4. Data infrastructure and challenges in ADR predictive modeling for oncology

Developing accurate and generalizable models for the prediction of adverse drug reactions (ADRs) in oncology depends heavily on the availability and quality of underlying data. Among the most prominent challenges is the small size, completeness, and standardization of datasets that are used to train artificial intelligence (AI) and machine learning (ML) models. As the complexity of models increases—especially with the application of deep learning—does the demand for large, diverse, and well-annotated data. However, the majority of existing ADR datasets are small-scale, domain-specific, or lack standard data formats because of their heterogeneous nature.

Preclinical cancer models, in vitro and in vivo, have emerged as valuable sources of data for pharmacogenomic insights [58,97–99]. While these models can be complementary to clinical datasets, they generally possess limited translational fidelity and need to be meticulously harmonized when combined with real-world patient data. Combination of preclinical findings with real-world evidence (RWE) data i.e., electronic health records (EHRs), cancer registries, and pharmacovigilance databases—has been shown to improve ADR signal detection [100,101].

The complexity of cancer treatment itself presents another daunting challenge. Polypharmacy, regimen switching, and comorbidities generate multidimensional noise that obscures clean signal detection [102–104]. As a result, ML model performance in real-world deployment often lags despite promising results in controlled experimental settings [105,106]. This divide underscores the necessity of developing robust data pipelines that can handle missing values, temporal incoherence, and heterogeneous outcome definitions [107,108].

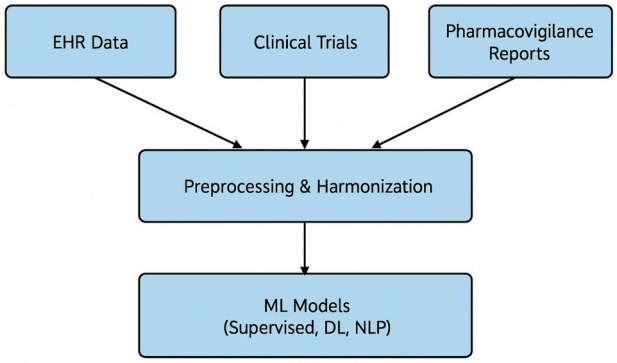


Fig. 3. Data flow for ADR prediction in oncology.

Recent developments have embraced large-scale AI-driven systems capable of mining EHRs and clinical trial data for ADR associations [52,109,110]. More significantly, graph-based architectures like the Tumor-Biomarker Knowledge Graph have been demonstrated to infer implicit relations between toxicity profiles, patients, and drugs [111,112]. These approaches herald a promising future for multimodal, multi-source data integration for ADR prediction.

Nonetheless, underlying data challenges persist. Many datasets remain incomplete, sparsely annotated, or biased towards ADRs of severe nature due to underreporting [113,114]. Additionally, ADR association records are often noisy, inconsistently coded, or lack precise temporal resolution—causal inference being difficult with these [115]. To overcome these challenges, future endeavors must invest in data standardization efforts, incentivized reporting mechanisms, and federated learning methods enabling cross-institutional model training without patient privacy trade-offs.

4.1. Gaps in the literature

Despite growing interest in machine learning (ML) for the prediction of adverse drug reactions (ADR), certain important gaps in the literature remain to be completely addressed:

1. **Remote Monitoring and Telemedicine:** How can ML-based ADR prediction systems be adapted to support real-time monitoring and remote delivery of healthcare services—particularly in resource-scarce settings such as rural areas or developing countries, where access to trained healthcare providers is limited?
2. **Enhancing Robustness of Supervised Learning:** In what ways can existing predictive models—particularly those based on supervised learning methods like decision trees and support vector machines—be enhanced to mitigate the vulnerabilities introduced by small, noisy, or incorrectly labeled datasets?
3. **Multimodal Personalization:** How do ML models integrate heterogeneous, patient-specific features such as real-time physiological signals (heart rate, SpO_2, body temperature), demographics, and genomic profiles to offer personalized, context-aware ADR predictions that evolve dynamically with differing health states?

Addressing these gaps is the path toward moving from static ADR models to adaptive, patient-tailored prediction systems. This study con-



Fig. 4. Heatmap showing comparative performance and applicability of machine learning models for ADR prediction.

tributes to the literature by preprocessing a multimodal patient dataset comprising physiological indicators (heart rate, SpO_2, body temperature) and demographic factors (age, gender, height, weight). The proposed framework applies ensemble learning and deep neural networks to mitigate challenges posed by sparsely labeled data. The model aims to make very personalized Androgen Deprivation Therapy (ADT) risk predictions and enhance safety across many oncology treatment scenarios through the integration of real-time signals and individual-level characteristics. As Fig. 3 illustrates, ADR prediction models rely on harmonized data from EHRs, clinical trials, and pharmacovigilance sources pooled across preprocessing pipelines. Particularly important for ADR prediction in oncology are physiological indicators like heart rate variability (HRV), peripheral oxygen saturation (SpO_2), body temperature, blood pressure, respiratory rate, and electrocardiogram (ECG) readings. Particularly in reaction to medications like doxorubicin, bleomycin, or VEGF inhibitors, these metrics can act as early markers for common adverse drug reactions (ADRs) like cardiotoxicity, respiratory distress, infection-induced fever, and hypertensive episodes. Wearable biosensors, such as fingertip pulse oximeters, adhesive patches, and smartwatches, are making it more and more feasible to collect these signals practically. Such signals can be temporally aligned with structured clinical records from electronic health records (EHRs) and streamed to cloud-based platforms when they are continuously monitored. This integration improves the model's sensitivity to acute or subtle ADR patterns and supports real-time risk stratification.

5. Performance comparison of ML algorithms in oncology ADR prediction

A critical comparative evaluation of machine learning models applied in ADR prediction highlights each technique's strengths, limitations, and performance trade-offs. Fig. 4 summarizes key performance metrics and practical attributes for commonly used models, derived from experimental studies and literature benchmarks.

As seen from the table, Random Forest and Gradient Boosting Machines have the best overall accuracy and AUC and are therefore most suitable for structured, high-dimensional ADR datasets [52,58]. Convolutional Neural Networks (CNNs) are also suitable but require a lot more training data and computational resources [70,91]. Support Vector Machines (SVMs) have moderate accuracy but have high training times and lower interpretability [44,56]. Because of its clearer feature importance and simpler structure, RF has a higher interpretability score than GBM. Workflow compatibility, robustness, efficiency, and interpretability are all combined to create the Clinical Applicability score.

Notably, simpler models such as AdaBoost may be comparable in imbalanced data but are prone to noise and very sensitive to parameter tuning [60,87]. The lack of transparency of most models (excluding decision trees) hinders their straightforward clinical application, and thus explainability tools become critical [72,90].

5.1. Integrated framework proposal

This study suggests an enhanced hybrid multimodal framework intended for reliable ADR prediction in oncology patients in response to reviewer comments and limitations noted in the literature. The framework, depicted in Fig. 5, incorporates a variety of data sources, such as unstructured clinical notes, imaging data, physiological signals (e.g., SpO_2, HR, BP), and structured clinical records. Specialized modules process each data stream: CNNs process image data, transformer-based NLP models process text analytics, and LSTM-based encoders process time-series physiological signals. Early feature-level fusion is used in the core hybridization strategy, which is followed by a deep ensemble predictor that combines the advantages of model-agnostic interpretability (e.g., SHAP or attention weights) with data-driven deep learning.

The architecture begins by ingesting multimodal patient data, including structured clinical features (e.g., age, weight, treatment history) and real-time physiological signals (e.g., SpO_2, heart rate, temperature). Preprocessing and normalization of the data streams are performed by feature engineering modules to be input into a Convolutional Neural Network (CNN) that learns hierarchical representations of the patient profile. The embeddings are then input into a Random Forest (RF) classifier that has been trained on known ADR outcomes.

This architecture combines the strength of CNNs in identifying complex patterns and RF's resistance to noise and overfitting. In order to address the unavoidable interpretability challenge, the model consists of post-hoc explainability layers with SHAP values, allowing clinicians to view which features have the strongest contribution to each ADR prediction.

Besides, the system can support deployment in telemedicine environments. Its modularity allows for the incorporation of remote monitoring instruments for real-time ADR risk scoring even in low-resource settings. Incorporating deep learning, ensemble modeling, and explainable AI,

Table 4
Common Chemotherapeutic Regimens by Cancer Type [116].

Chemotherapeutic Agents	Associated Cancer Types
Cyclophosphamide + Doxorubicin + Vincristine + Prednisone	Non-Hodgkin's Lymphoma (NHL)
Paclitaxel + Cisplatin	Cervical Cancer
Cisplatin + Cyclophosphamide; Liposomal Doxorubicin; Paclitaxel + Carboplatin	Ovarian Cancer
Capecitabine + Oxaliplatin + 5-FU + Leucovorin	Rectal Cancer
Oxaliplatin + 5-FU + Leucovorin	Colon Cancer
Etoposide + Actinomycin-D + Methotrexate; Cyclophosphamide + Vincristine	Choriocarcinoma
Cisplatin + 5-FU; Paclitaxel + Cisplatin	Esophageal Cancer
Paclitaxel + Cisplatin; Gemcitabine + Cisplatin	Lung Cancer
Capecitabine + Docetaxel + Cisplatin + 5-FU	Stomach Cancer
Docetaxel + Carboplatin; Cyclophosphamide + Adriamycin + 5-FU	Breast Cancer
Cyclophosphamide + Adriamycin + Vincristine	Ewing's Sarcoma

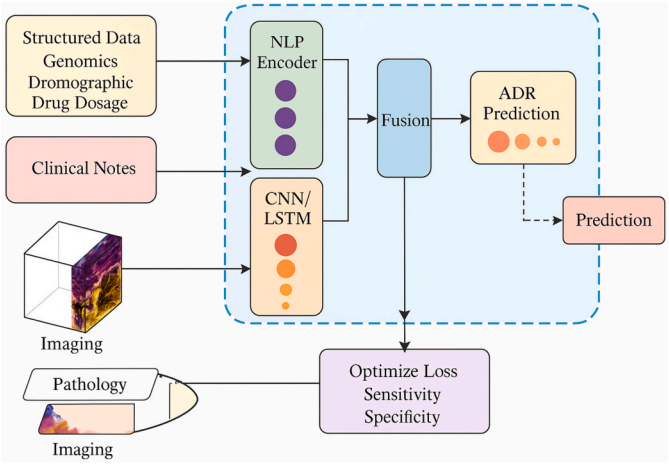


Fig. 5. Overview of machine learning framework integrating clinical, physiological, textual, and imaging data for ADR prediction in oncology patients.

the framework aims to provide a clinically meaningful and scalable solution to ADR prediction in oncology.

In order to complement the model-comparative analysis, Fig. 6 presents a graphical summary of the primary dimensions that have been compared between the machine learning models. They are accuracy, interpretability, training time, real-time applicability, and clinical application—each measured on a 0 to 1 scale for comparison.

Fig. 6 suggests Random Forest and Gradient Boosting Machines take the middle path of being highly precise with moderate explainability, so they are strong contenders for structured ADR datasets. CNNs are very precise but have heavy computational loads and low transparency. Support Vector Machines (SVMs) and AdaBoost are stable across moderate datasets but lack scalability and deployment in real-time. This comparative narrative emphasizes the importance of aligning model selection with clinical objectives.

5.2. Chemotherapy regimens and cancer type-specific drug profiles

More insight into drug-cancer pairings is required in order to place ADR prediction into perspective in oncology. Cancers with representative chemotherapy regimens are given in Table 4 and are taken from Wahlang et al. [116]. Key adverse drug reactions (ADRs) by regimen are summarized in Table 4; however, [116] also discusses aggregated data on frequency and severity (e.g., CTCAE grading). Only high-level observations are made here to prevent overgeneralization because these trends differ depending on the patient population, dosage, and clinical setting.

Abbreviations: Ca = Cancer, NHL = Non-Hodgkin's Lymphoma.

These regimens are frequently implicated in severe ADRs, as reflected in Table 4, and highlight the importance of integrating regimen-specific features into machine learning frameworks. Mapping specific drug combinations to ADR incidence can enhance feature representation in supervised learning models and inform clinical risk stratification.

Along with systemic cancers, patients receiving therapies for oral cancers, most notably squamous cell carcinoma, also have severe adverse drug reactions (ADRs) with chemoradiotherapy. Among the most common ADRs dental in nature are mucositis, xerostomia, osteoradionecrosis, and increased risk of infection, which all significantly affect quality of life and therapy compliance. ML can be used in dental oncology to forecast oral adverse drug reactions (ADRs) like osteonecrosis, xerostomia, and mucositis. In order to facilitate early detection and individualized care, these models may incorporate oral imaging, dental records, and systemic treatment data. A summary of the ML models used in ADR prediction based on important deployment-relevant parameters is given in Table 5.

6. Discussion

This critique summarizes a decade of machine learning (ML)-based adverse drug reaction (ADR) prediction research in oncology, highlighting methodological strengths but also translational challenges. Multiple supervised models such as Random Forests and Gradient Boosting Machines have repeatedly proved to be top-performing across structured datasets [52,58]. On the other hand, advanced architectures like CNNs and ANNs have robust feature learning ability but come with poor interpretability and heavy computing costs [70,91]. Techniques like SHAP and LIME are evolving into very important tools in explaining black-box models in a more human-centered manner [72,90]. Although attributing predictions to input features improves interpretability, SHAP's clinical usefulness in ADR contexts is still restricted. Although it offers local (per-patient) explanations, clinicians may find it cognitively complex and lack temporal awareness. Additionally, it makes the assumption that features are independent, which might not be true in clinical settings that are correlated. Therefore, without integration into more intuitive or domain-aware interpretability frameworks, SHAP by itself might not provide fully actionable insights, even though it is helpful.

While good at prediction, one too frequent limitation throughout literature has been a lack of high-quality, multi-institutional, large-sized datasets with structured clinical data, physiological vital sign information, and pharmacogenomic biomarkers. Flaws of prevailing practices on annotated data cause external model generalizability as well as cross-validation challenges as well [113,115]. Throughout real field-deployment scenarios, ML algorithms fail in performing equally as good because alterations take place within data distributions and lacking due to their not exhaustive recording of related health records [109,110].

Furthermore, regulatory and ethical barriers are in the way of easy clinical integration. ADR prediction tools must be compliant with data privacy (e.g., GDPR, HIPAA), equitable across all demographic subgroups, and rigorously audited post-deployment. All of these points are overlooked in technical reports and must be urgently given priority so as not to deploy ADR prediction tools that are merely precise but also dangerous and unfair to use.

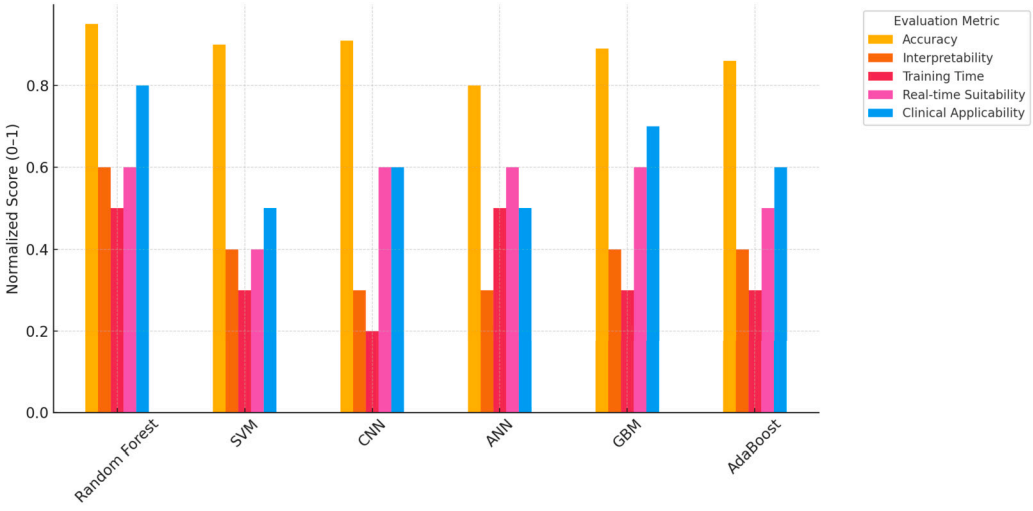


Fig. 6. Bar chart comparing key attributes of machine learning models for ADR prediction.

Table 5
Comparison of ML Models in ADR Prediction: Accuracy, Interpretability, and Clinical Readiness.

Model	Accuracy Range	Interpretability (0–1)	Clinical Readiness[†]
Decision Tree (DT)	70–85%	0.9	Good
Random Forest (RF)	78–92%	0.6	Good
Gradient Boosting (GBM)	80–94%	0.4	Moderate
Support Vector Machine (SVM)	75–90%	0.3	Moderate
Logistic Regression (LR)	68–85%	0.8	Good
Naïve Bayes (NB)	60–78%	0.7	Moderate
Convolutional Neural Network (CNN)	82–95%	0.2	Low
Recurrent Neural Network (LSTM/GRU)	80–94%	0.3	Low
Transformer-based (BERT variants)	85–96%	0.3	Moderate

The optimal way forward is hybrid architectures that fuse real-time physiologic information with multimodal static inputs (e.g., treatment history, demographics). These models can adaptively personalize risk assessments and support applications like remote monitoring, especially in telemedicine or resource-poor settings. Their embedding in clinical workflows and scalability, however, remains an open research area.

In conclusion, while the ML context for ADR prediction is rich in algorithmic innovation, transforming these successes into clinical utility rests on prioritizing data quality, interpretability, regulatory consistency, and integration with infrastructure. Next-generation research must focus on federated learning, multimodal data harmonization, and co-design of tools with clinicians to span the gap between model development and bedside use. Federated learning makes it possible to model ADR across institutions without sharing data, but it has drawbacks such as heterogeneous data, unbalanced infrastructure, privacy risks from gradient leakage, and regulatory obstacles. Oncology settings require customized FL strategies.

7. Conclusion and future directions

This review has presented a comprehensive overview of machine learning (ML) techniques as they have been applied to adverse drug reaction (ADR) prediction in oncology. With the heightened use of complex chemotherapeutic regimens and targeted therapies comes the need for robust, personalized models with the ability to predict ADRs and avert harm. While classical supervised models such as decision trees and support vector machines provide interpretability and sufficient accuracy, ensemble methods and deep learning models gain more predictive power at the cost of transparency and computational effort.

We highlighted the emergence of hybrid models that leverage multimodal data sources—including patient demographics, clinical history, physiological signals, and genomic profiles—charting a course toward more dynamic and personalized ADR prediction systems. In addition, explainability tools such as SHAP and LIME are bridging the interpretability gap, making complex models more clinically implementable.

Despite these advances, several challenges persist. Data fragmentation, class imbalance, and lack of standardized annotations continue to limit generalizability and reproducibility. Furthermore, the absence of real-world validation and disregard for regulatory, ethical, and infrastructural constraints are barriers to deployment in clinical practice.

Looking forward, future research must prioritize the development of large, heterogeneous, and multimodal datasets, the implementation of federated and privacy-preserving learning pipelines, and the development of clinician-in-the-loop systems to translate predictive outputs into real-world decision-making. Continued focus on explainability, fairness, and cross-institutional validation will be paramount to translate these models from experimental studies to bedside applications. Expanding predictive models to other similar fields, i.e., dental oncology, may have the added benefits in integrated care pathways In this way, machine learning can be extended beyond prediction to proactive and equitable clinical risk management in oncology. Ongoing efforts are starting to address the lack of data in oncology ADR prediction. Data standardization and federated analytics are encouraged by platforms such as OHDSI (through OMOP CDM) and EU consortia (like IMI-PREFER). These initiatives set the foundation for interoperable, machine learning-ready datasets, even though no single resource is specifically focused on oncology-ADR.

CRediT authorship contribution statement

Abid Yahya: Writing – original draft, Formal analysis, Writing – review & editing, Methodology, Conceptualization, Investigation. **Phatsimo Lobelo:** Formal analysis, Writing – review & editing. **Afiya Eram:** Investigation, Funding acquisition, Writing – review & editing. **Sana Al-**

thaf Hussain: Writing – review & editing, Methodology. **Irfan Anjum Badruddin:** Methodology, Funding acquisition, Writing – review & editing. **Lory Liza D. Bulay-og:** Project administration, Resources. **Dionel O. Albina:** Validation, Writing – review & editing, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the results and interpretations reported in this study.

Acknowledgement

The authors acknowledge the support provided by Manipal Academy of Higher Education, Manipal in carrying out this work. The author I.A. Badruddin extends his appreciation to the Deanship of Research and Graduate Studies at King Khalid University for supporting through Large Research Project under RGP.2/460/46.

Data availability

No data was used for the research described in the article.

References

[1] L. Zhou, S. Pan, J. Wang, A.V. Vasilakos, Machine learning on big data: opportunities and challenges, Neurocomputing 237 (2017) 350–361.
[2] A. Aliper, S. Plis, A. Artemov, A. Ulloa, P. Mamoshina, A. Zhavoronkov, Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data, Mol. Pharm. 13 (7) (2016) 2524–2530.
[3] S. Jamal, S. Goyal, A. Shanker, A. Grover, Predicting neurological adverse drug reactions based on biological, chemical and phenotypic properties of drugs using machine learning models, Sci. Rep. 7 (1) (2017) 872.
[4] S. Jamal, W. Ali, P. Nagpal, S. Grover, A. Grover, Computational models for the prediction of adverse cardiovascular drug reactions, J. Transl. Med. 17 (2019) 1–13.
[5] I. Spasic, G. Nenadic, et al., Clinical text data in machine learning: systematic review, JMIR Med. Inform. 8 (3) (2020) e17984.
[6] A.N. Richter, T.M. Khoshgoftaar, A review of statistical and machine learning methods for modeling cancer risk using structured clinical data, Artif. Intell. Med. 90 (2018) 1–14.
[7] O.Y. Al-Jarrah, P.D. Yoo, S. Muhaidat, G.K. Karagiannidis, K. Taha, Efficient machine learning for big data: a review, Big Data Res. 2 (3) (2015) 87–93.
[8] H.R. Kim, M. Sung, J.A. Park, K. Jeong, H.H. Kim, S. Lee, Y.R. Park, Analyzing adverse drug reaction using statistical and machine learning methods: a systematic review, Medicine 101 (25) (2022) e29387.
[9] T.-B. Ho, L. Le, D.T. Thai, S. Taewijit, Data-driven approach to detect and predict adverse drug reactions, Curr. Pharm. Des. 22 (23) (2016) 3498–3526.
[10] C. Xiao, Y. Li, I.M. Baytas, J. Zhou, F. Wang, An mcem framework for drug safety signal detection and combination from heterogeneous real world evidence, Sci. Rep. 8 (1) (2018) 1806.
[11] N. Liu, C.-B. Chen, S. Kumara, Semi-supervised learning algorithm for identifying high-priority drug–drug interactions through adverse event reports, IEEE J. Biomed. Health Inform. 24 (1) (2019) 57–68.
[12] T.E. Ponraj, J. Charles, Investigation of single and multiple mutations prediction using binary classification approach, Intell. Autom. Soft Comput. 36 (1) (2023).
[13] C. Agbonkhese, H.A. Soriyan, K. Mosaku, The efficiency of machine learning algorithms in the prediction of drug reactions in clinical settings, Indian J. Comput. Sci. Eng. 14 (6) (Dec. 2023) 891–901.
[14] A.W. Chen, Predicting adverse drug reaction outcomes with machine learning, Int. J. Commun. Med. Public Health 5 (3) (2018) 901.
[15] C. Ngufor, J. Wojtusiak, J. Pathak, A systematic prediction of adverse drug reactions using pre-clinical drug characteristics and spontaneous reports, in: 2015 International Conference on Healthcare Informatics, IEEE, 2015, pp. 76–81.
[16] P. Pilipiec, M. Liwicki, A. Bota, Using machine learning for pharmacovigilance: a systematic review, Pharmaceutics 14 (2) (2022) 266.
[17] M. Alomar, A.M. Tawfiq, N. Hassan, S. Palaian, Post marketing surveillance of suspected adverse drug reactions through spontaneous reporting: current status, challenges and the future, Ther. Adv. Drug Saf. 11 (2020) 2042098620938595.
[18] G. Gonzalez-Hernandez, M. Krallinger, M. Muñoz, R. Rodriguez-Esteban, Ö. Uzuner, L. Hirschman, Challenges and opportunities for mining adverse drug reactions: perspectives from pharma, regulatory agencies, healthcare providers and consumers, Database 2022 (2022) baac071.
[19] S. Sunny, A. Thampi, N. Shetty, S.K. Babasahib, C. Chacko, et al., Assessment of adverse effects of most commonly prescribed anticancer drugs in a tertiary care teaching hospital, Indian J. Pharm. Pract. 10 (4) (2017).
[20] J.D. Martin, H. Cabral, T. Stylianopoulos, R.K. Jain, Improving cancer immunotherapy using nanomedicines: progress, opportunities and challenges, Nat. Rev., Clin. Oncol. 17 (4) (2020) 251–266.
[21] S. Zhu, T. Zhang, L. Zheng, H. Liu, W. Song, D. Liu, Z. Li, C.-x. Pan, Combination strategies to maximize the benefits of cancer immunotherapy, Int. J. Pediatr. Hematol./oncol. 14 (1) (2021) 156.
[22] T.A. Yap, E.E. Parkes, W. Peng, J.T. Moyers, M.A. Curran, H.A. Tawbi, Development of immunotherapy combination strategies in cancer, Cancer Discov. 11 (6) (2021) 1368–1397.
[23] P. Inácio, A. Cavaco, M. Airaksinen, The value of patient reporting to the pharmacovigilance system: a systematic review, Br. J. Clin. Pharmacol. 83 (2) (2017) 227–246.
[24] J. Pant, H. Marwah, R. Singh, S. Hazra, An overview of world-wide master key for drug safety monitoring (pharmacovigilance) and its role in India, J. Pharamacovigil 9 (2021) 325.
[25] V.P. Londhe, R. Chanshetti, S. Dhole, Pharmacovigilance: past, present and future, Asian J. Pharm. Clin. Res. 14 (2) (2024) 175–182.
[26] A. Santoro, G. Genov, A. Spooner, J. Raine, P. Arlett, Promoting and protecting public health: how the European Union pharmacovigilance system works, Drug Safety 40 (2017) 855–869.
[27] P. Agrawal, V. Kushwaha, S. Siddiqui, N.F. Khan, M. Tripathi, History of pharmacovigilance and drug safety, World J. Pharm. Res. 12 (2022) 575–586.
[28] S. Monestime, R. Page, W.M. Jordan, S. Aryal, Prevalence and predictors of patients reporting adverse drug reactions to health care providers during oral targeted cancer treatment, J. Am. Pharm. Assoc. 61 (1) (2021) 53–59.
[29] M. Javaid, A. Haleem, R.P. Singh, R. Suman, S. Rab, Significance of machine learning in healthcare: features, pillars and applications, Int. J. Intell. Netw. 3 (2022) 58–73.
[30] X. Zhou, X. Ao, Z. Jia, Y. Li, S. Kuang, C. Du, J. Zhang, J. Wang, Y. Liu, Non-coding rna in cancer drug resistance: underlying mechanisms and clinical applications, Front. Oncol. 12 (2022) 951864.
[31] V. Das, J. Kalita, M. Pal, Predictive and prognostic biomarkers in colorectal cancer: a systematic review of recent advances and challenges, Biomed. Pharmacother. 87 (2017) 8–19.
[32] E. Koncina, S. Haan, S. Rauh, E. Letellier, Prognostic and predictive molecular biomarkers for colorectal cancer: updates and challenges, Cancers 12 (2) (2020) 319.
[33] G. Goel, Molecular characterization and biomarker identification in colorectal cancer: toward realization of the precision medicine dream, Cancer Manag. Res. (2018) 5895–5908.
[34] S. Imai, Y. Takekuma, H. Kashiwagi, T. Miyai, M. Kobayashi, K. Iseki, M. Sugawara, Validation of the usefulness of artificial neural networks for risk prediction of adverse drug reactions used for individual patients in clinical practice, PLoS ONE 15 (7) (2020) e0236789.
[35] J.P. Bharadiya, Machine learning and ai in business intelligence: trends and opportunities, Int. J. Comput. 48 (1) (2023) 123–134.
[36] D. Ben-Israel, W.B. Jacobs, S. Casha, S. Lang, W.H.A. Ryu, M. de Lotbiniere-Bassett, D.W. Cadotte, The impact of machine learning on patient care: a systematic review, Artif. Intell. Med. 103 (2020) 101785.
[37] C.M. Bellanca, E. Augello, A.F. Cantone, R. Di Mauro, G.A. Attaguile, V. Di Giovanni, G.A. Condorelli, G. Di Benedetto, G. Cantarella, R. Bernardini, Insight into risk factors, pharmacogenetics/genomics, and management of adverse drug reactions in elderly: a narrative review, Pharmaceuticals 16 (11) (2023) 1542.
[38] W.H. Organization, Pharmacovigilance, https://www.who.int/teams/regulation-prequalification/regulation-and-safety/pharmacovigilance, 2024. (Accessed 12 October 2024).
[39] Y. Li, Y. Wu, T. Jiang, H. Xing, J. Xu, C. Li, Y. Liu, Opportunities and challenges of pharmacovigilance in special populations: a narrative review of the literature, Ther. Adv. Drug Saf. 14 (2023) 20420986231200746.
[40] P. Baldo, G. Fornasier, L. Ciolfi, I. Sartor, S. Francescon, Pharmacovigilance in oncology, Int. J. Clin. Pharm. 40 (2018) 832–841.
[41] S. Tadge, M. Gambhire, A. Kulkarni, A Narrative Review on Adverse Drug Reactions Reporting: Current Practices and Challenges, 2023.
[42] M. Publications, A Narrative Review on Adverse Drug Reactions Reporting: Current Practices and Challenges, MedText Publications, 2023.
[43] V. Nasteski, An overview of the supervised machine learning methods, Horizons B 4 (51–62) (2017) 56.
[44] B. Mahesh, Machine learning algorithms-a review, Int. J. Sci. Res. 9 (1) (2020) 381–386.
[45] M. Praveena, V. Jaiganesh, A literature review on supervised machine learning algorithms and boosting process, Int. J. Comput. Appl. 169 (8) (2017) 32–35.
[46] S. Chauhan, E. Marcharla, S. Hariharan, S. Vinayagam, T. Sundaram, S. Priya, Swamynathan, Insights into the molecular mechanisms and nanoparticle-based therapies for gastric cancer: a review, Results Eng. 25 (2025) 101324.
[47] U. Vignesh, R. Parvathi, K. Gokul Ram, Ensemble deep learning model for protein secondary structure prediction using nlp metrics and explainable ai, Results Eng. 24 (2024) 101687.
[48] S. Inayathullah, R. Buddala, Review of machine learning applications in additive manufacturing, Results Eng. 25 (2025) 103676.

[49] B. Gupta, A. Rawat, A. Jain, A. Arora, N. Dhami, Analysis of various decision tree algorithms for classification in data mining, Int. J. Comput. Appl. 163 (8) (2017) 15–19.
[50] F. Hammann, J. Drewe, Decision tree models for data mining in hit discovery, Expert Opin. Drug Discov. 7 (4) (2012) 341–352.
[51] J. Wei, L. Cheng, P. Han, Y. Zhu, W. Huang, Decision tree-based data stratification method for the minimization of the masking effect in adverse drug reaction signal detection, Appl. Sci. 11 (23) (2021) 11380.
[52] X.-W. Wu, J.-Y. Zhang, H. Chang, X.-W. Song, Y.-L. Wen, E.-W. Long, R.-S. Tong, Develop an adr prediction system of Chinese herbal injections containing panax notoginseng saponin: a nested case–control study using machine learning, BMJ Open 12 (9) (2022) e061457.
[53] L. Jia, R. Yarlagadda, C.C. Reed, Structure based thermostability prediction models for protein single point mutations with machine learning tools, PLoS ONE 10 (9) (2015) e0138022.
[54] F. Osisanwo, J. Akinsola, O. Awodele, J. Hinmikaiye, O. Olakanmi, J. Akinjobi, et al., Supervised machine learning algorithms: classification and comparison, Int. J. Comput. Trends Technol. 48 (3) (2017) 128–138.
[55] D. Pandey, K. Niwaria, B. Chourasia, Machine learning algorithms: a review, Mach. Learn. 6 (02) (2019).
[56] B.T. Yaseen, Drug target interaction prediction using support vector machine (svm), in: 2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), IEEE, 2023, pp. 1–4.
[57] S.J. Rigatti, Random forest, J. Insur. Med. 47 (1) (2017) 31–39.
[58] R. Ietswaart, S. Arat, A.X. Chen, S. Farahmand, B. Kim, W. DuMouchel, D. Armstrong, A. Fekete, J.J. Sutherland, L. Urban, Machine learning guided association of adverse drug reactions with in vitro target-based pharmacology, eBioMedicine 57 (2020).
[59] A. Malik, M.K. Saggi, S. Rehman, H. Sajjad, S. Inyurt, A.S. Bhatia, A.A. Farooque, A.Y. Oudah, Z.M. Yaseen, Deep learning versus gradient boosting machine for pan evaporation prediction, Eng. Appl. Comput. Fluid Mech. 16 (1) (2022) 570–587.
[60] Q. Hu, X. Li, et al., Machine Learning to Predict Adverse Drug Reaction or Event Based on Electronic Health Records: a Systematic Review and Meta-Analysis, 2024.
[61] M. Usama, J. Qadir, A. Raza, H. Arif, K.-L.A. Yau, Y. Elkhatib, A. Hussain, A. Al-Fuqaha, Unsupervised machine learning for networking: techniques, applications and research challenges, IEEE Access 7 (2019) 65579–65615.
[62] S. Naeem, A. Ali, S. Anam, M.M. Ahmed, An unsupervised machine learning algorithms: comprehensive review, Int. J. Comput. Digit. Syst. (2023).
[63] D. Xu, Y. Tian, A comprehensive survey of clustering algorithms, Ann. Data Sci. 2 (2015) 165–193.
[64] S. Firdaus, M.A. Uddin, A survey on clustering algorithms and complexity analysis, Int. J. Comput. Sci. Issues 12 (2) (2015) 62.
[65] S. Ahmadian, A. Epasto, R. Kumar, M. Mahdian, Clustering without over-representation, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, pp. 267–275.
[66] Amazon Web Services, "What is deep learning? - deep learning explained", https://aws.amazon.com/what-is/deep-learning/. (Accessed 3 October 2024).
[67] S. Li, L. Zhang, L. Wang, J. Ji, J. He, X. Zheng, L. Cao, K. Li, Bimpadr: a deep learning framework for predicting adverse drug reactions in new drugs, Molecules 29 (8) (2024) 1784.
[68] J.-Y. Huang, W.-P. Lee, K.-D. Lee, Predicting adverse drug reactions from social media posts: data balance, feature selection and deep learning, Healthcare 10 (4) (2022) 618, MDPI.
[69] J. Cardete, Convolutional Neural Networks: A Comprehensive Guide, Feb. 2024. (Accessed 03 October 2024).
[70] Y. Li, W. Zhao, B. Dang, X. Yan, M. Gao, W. Wang, M. Xiao, Research on adverse drug reaction prediction model combining knowledge graph embedding and deep learning, in: 2024 4th International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), IEEE, 2024, pp. 322–329.
[71] V. Ventures, Artificial neural networks definition, explanation, and use cases, https://www.vationventures.com/glossary/artificial-neural-networks-definition-explanation-and-use-cases, 2024. (Accessed 3 October 2024).
[72] S. Wang, J. Di, D. Wang, X. Dai, Y. Hua, X. Gao, A. Zheng, J. Gao, State-of-the-art review of artificial neural networks to predict, characterize and optimize pharmaceutical formulation, Pharmaceutics 14 (1) (January 2022) 183.
[73] J. Zhang, C. Li, Y. Yin, J. Zhang, M. Grzegorzek, Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer, Artif. Intell. Rev. 56 (2) (2023) 1013–1070.
[74] A. Akilan, Text mining: challenges and future directions, in: 2015 2nd International Conference on Electronics and Communication Systems (ICECS), IEEE, 2015, pp. 1679–1684.
[75] S. Datta, E.V. Bernstam, K. Roberts, A frame semantic overview of nlp-based information extraction for cancer-related ehr notes, J. Biomed. Inform. 100 (2019) 103301.
[76] L. Wang, L. Luo, Y. Wang, J.A. Wampfler, P. Yang, H. Liu, Information extraction for populating lung cancer clinical research data, in: 2019 IEEE International Conference on Healthcare Informatics (ICHI), IEEE, 2019, pp. 1–2.
[77] E.H. Houssein, R.E. Mohamed, A.A. Ali, Machine learning techniques for biomedical natural language processing: a comprehensive review, IEEE Access 9 (2021) 140628–140653.
[78] D. Pappa, L.K. Stergioulas, Harnessing social media data for pharmacovigilance: a review of current state of the art, challenges and future directions, Int. J. Data Sci. Anal. 8 (2) (2019) 113–135.
[79] M. Alwhaibi, N.A. Al Aloola, Healthcare students' knowledge, attitude and perception of pharmacovigilance: a systematic review, PLoS ONE 15 (5) (2020) e0233393.
[80] M. Desai, Pharmacovigilance and spontaneous adverse drug reaction reporting: challenges and opportunities, Perspect. Clin. Res. 13 (4) (2022) 177–179.
[81] R. Adisa, T.I. Omitogun, Awareness, knowledge, attitude and practice of adverse drug reaction reporting among health workers and patients in selected primary healthcare centres in Ibadan, Southwestern Nigeria, BMC Health Serv. Res. 19 (2019) 1–14.
[82] C. Costa, P. Abeijon, D.A. Rodrigues, A. Figueiras, M.T. Herdeiro, C. Torre, Factors associated with underreporting of adverse drug reactions by patients: a systematic review, Int. J. Clin. Pharm. 45 (6) (2023) 1349–1358.
[83] A. Sarker, R. Ginn, A. Nikfarjam, K. O'Connor, K. Smith, S. Jayaraman, T. Upadhaya, G. Gonzalez, Utilizing social media data for pharmacovigilance: a review, J. Biomed. Inform. 54 (2015) 202–212.
[84] B.S. Kaas-Hansen, S. Gentile, A. Caioli, S.E. Andersen, Exploratory pharmacovigilance with machine learning in big patient data: a focused scoping review, Basic Clin. Pharmacol. Toxicol. 132 (3) (2023) 233–241.
[85] M. Imran, A. Bhatti, D.M. King, M. Lerch, J. Dietrich, G. Doron, K. Manlik, Supervised machine learning-based decision support for signal validation classification, Drug Safety 45 (5) (2022) 583–596.
[86] A.A. Pandit, S.A. Dubey, A comprehensive review on adverse drug reactions (adrs) detection and prediction models, in: 2021 13th International Conference on Computational Intelligence and Communication Networks (CICN), 2021, pp. 123–127.
[87] A.A. Khan, O. Chaudhari, R. Chandra, A review of ensemble learning and data augmentation models for class imbalanced problems: combination, implementation and evaluation, Expert Syst. Appl. (2023) 122778.
[88] C. Yuan, H. Yang, Research on k-value selection method of k-means clustering algorithm, J - Multidiscipl. Sci. J. 2 (2) (2019) 226–235.
[89] A.M. Ikotun, A.E. Ezugwu, L. Abualigah, B. Abuhaija, J. Heming, K-means clustering algorithms: a comprehensive review, variants analysis, and advances in the era of big data, Inf. Sci. 622 (2023) 178–210.
[90] K. Sharifani, M. Amini, Machine learning and deep learning: a review of methods and applications, World Inf. Technol. Eng. J. 10 (07) (2023) 3897–3904.
[91] M. Krichen, Convolutional neural networks: a survey, Computers 12 (8) (2023) 151.
[92] S.F. Ahmed, M.S.B. Alam, M. Hassan, M.R. Rozbu, T. Ishtiak, N. Rafa, M. Mofijur, A. Shawkat Ali, A.H. Gandomi, Deep learning modelling techniques: current progress, applications, advantages, and challenges, Artif. Intell. Rev. 56 (11) (2023) 13521–13617.
[93] B.S. Kumar, V. Ravi, A survey of the applications of text mining in financial domain, Knowl.-Based Syst. 114 (2016) 128–147.
[94] A. O'Mara-Eves, J. Thomas, J. McNaught, M. Miwa, S. Ananiadou, Using text mining for study identification in systematic reviews: a systematic review of current approaches, Syst. Rev. 4 (2015) 1–22.
[95] D. Khurana, A. Koli, K. Khatter, S. Singh, Natural language processing: state of the art, current trends and challenges, Multimed. Tools Appl. 82 (3) (2023) 3713–3744.
[96] M. Khan, M.S. Khan, Y. Alharbi, Text mining challenges and applications—a comprehensive review, Int. J. Comput. Sci. Netw. Secur. 20 (12) (2020) 138.
[97] N. Niu, L. Wang, In vitro human cell line models to predict clinical response to anticancer drugs, Pharmacogenomics J. 16 (3) (2015) 273–285.
[98] L. Pompili, M. Porru, C. Caruso, A. Biroccio, C. Leonetti, Patient-derived xenografts: a relevant preclinical model for drug development, J. Exp. Clin. Cancer Res. 35 (2016) 1–8.
[99] A. Kalamara, L. Tobalina, J. Saez-Rodriguez, How to find the right drug for each patient? Advances and challenges in pharmacogenomics, Curr. Opin. Syst. Biol. 10 (2018) 53–62.
[100] C.L. Bennett, B. Schooley, M.A. Taylor, B.J. Witherspoon, A. Godwin, J. Vemula, H.C. Ausdenmoore, O. Sartor, Y.T. Yang, J.O. Armitage, et al., Caveat medicus: clinician experiences in publishing reports of serious oncology-associated adverse drug reactions, PLoS ONE 14 (7) (2019) e0219521.
[101] F.C. Barcelos, G.C.d. Matos, M.J.S.d. Silva, F.A.B.d. Silva, E.d.C. Lima, Suspected adverse drug reactions related to breast cancer chemotherapy: disproportionality analysis of the Brazilian spontaneous reporting system, Front. Pharmacol. 10 (2019) 498.
[102] A. Mueller-Schoell, S.L. Groenland, O. Scherf-Clavel, M. van Dyk, W. Huisinga, R. Michelet, U. Jaehde, N. Steeghs, A.D. Huitema, C. Kloft, Therapeutic drug monitoring of oral targeted antineoplastic drugs, Eur. J. Clin. Pharmacol. 77 (2021) 441–464.
[103] M. Barillet, V. Prevost, F. Joly, B. Clarisse, Oral antineoplastic agents: how do we care about adherence?, Br. J. Clin. Pharmacol. 80 (6) (2015) 1289–1302.
[104] D. Chopra, H. Rehan, V. Sharma, R. Mishra, Chemotherapy-induced adverse drug reactions in oncology patients: a prospective observational survey, Indian J. Med. Paediatr. Oncol. 37 (01) (2016) 42–46.
[105] A.H. Lavan, P. Gallagher, Predicting risk of adverse drug reactions in older adults, Ther. Adv. Drug Saf. 7 (1) (2016) 11–22.
[106] R. Osanlou, L. Walker, D.A. Hughes, G. Burnside, M. Pirmohamed, Adverse drug reactions, multimorbidity and polypharmacy: a prospective analysis of 1 month of medical admissions, BMJ Open 12 (7) (2022) e055551.

[107] R. Masumshah, R. Aghdam, C. Eslahchi, A neural network-based method for polypharmacy side effects prediction, BMC Bioinform. 22 (2021) 1–17.
[108] N. Parameswaran Nair, L. Chalmers, G.M. Peterson, B.J. Bereznicki, R.L. Castelino, L.R. Bereznicki, Hospitalization in older patients due to adverse drug reactions–the need for a prediction tool, in: Clinical Interventions in Aging, 2016, pp. 497–505.
[109] A.J. Steele, S.C. Denaxas, A.D. Shah, H. Hemingway, N.M. Luscombe, Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease, PLoS ONE 13 (8) (2018) e0202344.
[110] C. Xiao, E. Choi, J. Sun, Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review, J. Am. Med. Inform. Assoc. 25 (10) (2018) 1419–1428.
[111] M. Wang, X. Ma, J. Si, H. Tang, H. Wang, T. Li, W. Ouyang, L. Gong, Y. Tang, X. He, et al., Adverse drug reaction discovery using a tumor-biomarker knowledge graph, Front. Genet. 11 (2021) 625659.
[112] H.-Y. Wang, C.-H. Chen, S. Shi, C.-R. Chung, Y.-H. Wen, M.-H. Wu, M.S. Lebowitz, J. Zhou, J.-J. Lu, Improving multi-tumor biomarker health check-up tests with machine learning algorithms, Cancers 12 (6) (2020) 1442.
[113] Y. Tan, Y. Hu, X. Liu, Z. Yin, X.-w. Chen, M. Liu, Improving drug safety: from adverse drug reaction knowledge discovery to clinical implementation, Methods 110 (2016) 14–25.
[114] D.M. Bean, H. Wu, E. Iqbal, O. Dzahini, Z.M. Ibrahim, M. Broadbent, R. Stewart, R.J. Dobson, Knowledge graph prediction of unknown adverse drug reactions and validation in electronic health records, Sci. Rep. 7 (1) (2017) 16416.
[115] J.H. Scholl, F.P. van Hunsel, E. Hak, E.P. van Puijenbroek, A prediction model-based algorithm for computer-assisted database screening of adverse drug reactions in the Netherlands, Pharmacoepidemiol. Drug Saf. 27 (2) (2018) 199–205.
[116] J.B. Wahlang, P.D. Laishram, D.K. Brahma, C. Sarkar, J. Lahon, B.S. Nongkynrih, Adverse drug reactions due to cancer chemotherapy in a tertiary care teaching hospital, Ther. Adv. Drug Saf. 8 (2) (2017) 61–66.