Drug Discovery Today • Volume 28, Number 9 • September 2023 REVIEWS

Machine-learning-based adverse drug event prediction from observational health data: A review

Jonas Denck [1],*, Elif Ozkirimli [1], Ken Wang [2]

[1] Roche Informatics, F. Hoffmann-La Roche AG, Kaiseraugst, Switzerland
[2] Roche Pharmaceutical Research and Early Development, Roche Innovation Center, Basel, Switzerland

Adverse drug events (ADEs) are responsible for a significant number of hospital admissions and fatalities. Machine learning models have been developed to assess the individual patient risk of having an ADE. In this article, we have reviewed studies addressing the prediction of ADEs in observational health data with machine learning. The field of individualised ADE prediction is rapidly emerging through the increasing availability of additional data modalities (e.g., genetic data, screening data, wearables data) and advanced deep learning models such as transformers. Consequently, personalised adverse drug event predictions are becoming more feasible and tangible.

Keywords: machine learning; adverse drug event; electronic health record; prediction model

Introduction

Adverse drug events (ADEs) are defined as ''an appreciably harmful or unpleasant reaction, resulting from an intervention related to the use of a medicinal product, which predicts hazard from future administration and warrants prevention or specific treatment, or alteration of the dosage regimen, or withdrawal of the product''.[1] In Europe, ADEs are responsible for 0.5%[2] to 12.8%[3] of all hospital admissions and are fatal for up to 0.5% of hospital stays.[4] In addition to the increased morbidities and impact on patient quality of life, ADEs substantially impact society owing to additional treatment and prolonged length of stays in hospital,[5] which increase the costs for the healthcare system.

To prevent an ADE, the subgroup of patients most likely to be susceptible to the adverse effect must be identified first. Then, the treatment protocol and monitoring mitigation strategy can be tailored accordingly.[6] Understanding the patient risk factors for specific drugs is crucial to reducing ADEs, especially for drugs with a narrow therapeutic window.[7] However, clinical trials are often not large enough to report all adverse events. Post-market and real-world data, such as electronic health records (EHRs), electronic medical records (EMRs) or administrative claims, can be used to evaluate risk factors and the incidence of ADEs in a real-world setting. These data are primarily used for digitised and standardised data collection and sharing across healthcare providers. Nevertheless, EHR and EMR data have also been utilised in a variety of secondary use applications, such as patient trajectory modelling,[8] patient risk stratification[9] and more.[10]

Patient risk stratification, for example for ADEs, typically relies on a purely statistical approach but can also be enhanced through machine learning. The main benefit of using machine learning instead of purely statistical approaches is the ability of many machine learning methods to model complex relationships between covariates. Machine learning models can then predict the individualised risk for an ADE. However, utilising the data for training artificial intelligence (AI) and machine learning models also involves several challenges, such as erroneous, noisy or missing data, the modelling of irregular-spaced temporal data with high-dimensional features and the need for big data and appropriate labels.[11]

To assess the challenges and opportunities associated with machine-learning-based ADE risk prediction, we conducted a literature survey to review studies that apply machine learning

* *Corresponding author. Denck, J. (jonas.denck@roche.com)*

1359-6446/© 2023 Elsevier Ltd. All rights reserved.
https://doi.org/10.1016/j.drudis.2023.103715

approaches to the prediction of drug-associated adverse events from observational health data, such as EHR. Our review focuses on the challenges associated with training AI models on observational health data and assesses recommended practices for evaluating ADEs using these data.[12] ADE prediction typically uses a patient baseline period to extract patient features and tries to predict the risk for a patient to have an ADE in the (near) future (Figure 1). This differentiates ADE prediction from ADE detection, which typically tries to identify ADEs within a given period.

The scope of this paper is limited to observational health data because these data are widely available, can provide big dataset sizes and allow modelling the patient trajectory that can be used to identify risk factors for adverse events. Machine learning applications on other data sources (e.g., clinical trial data) might be mentioned but are not extensively surveyed. We follow current recommended practices for evaluating ADEs using electronic health records,[12] base our data abstraction criteria upon existing literature,[13] and expand these with relevant machine learning concepts (temporal data representation, interpretability, uncertainty quantification, calibration).

Studies overview

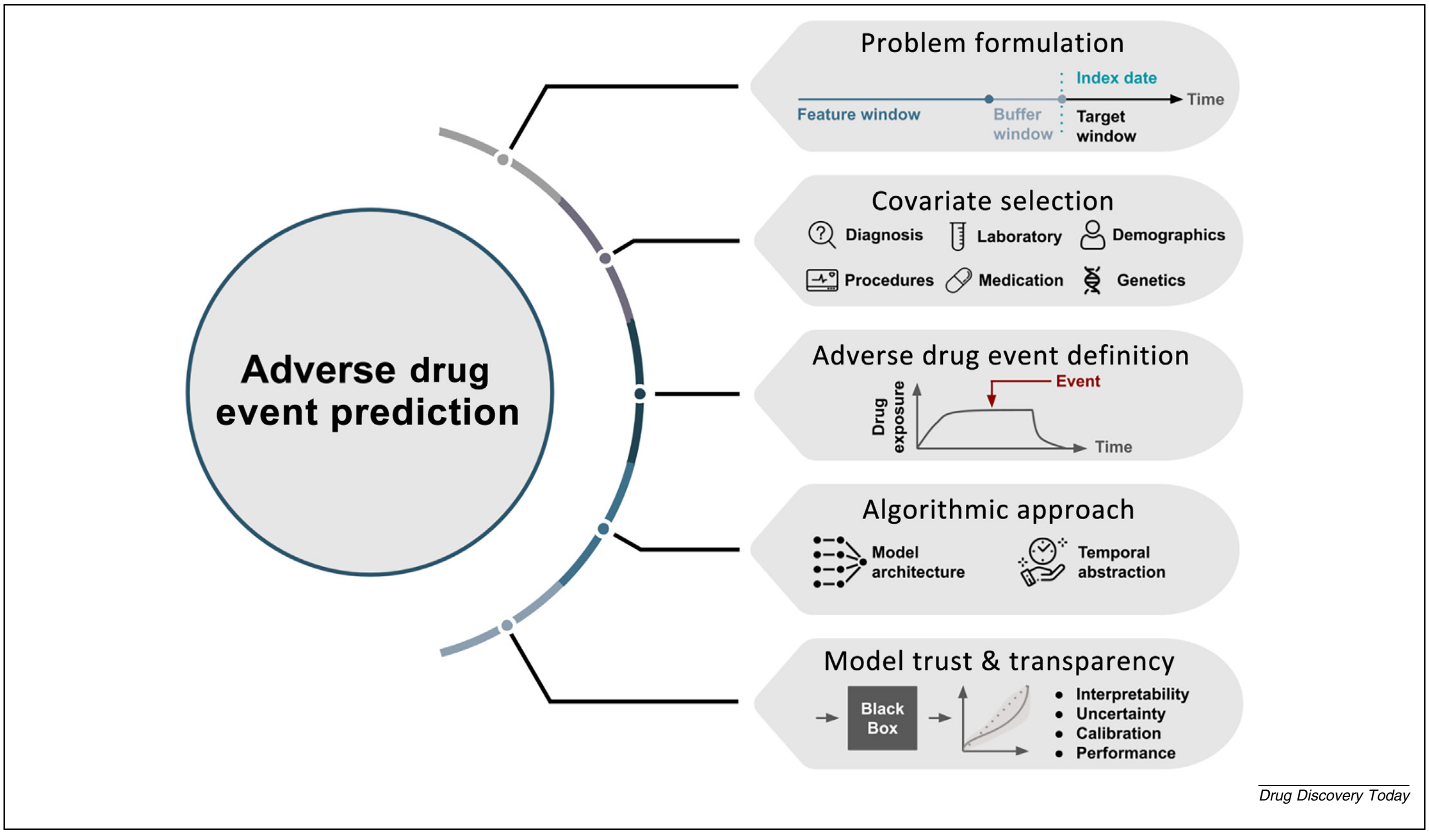
From 5444 unique records that were identified through the literature search (for details on the literature search, see Supplementary material online) 41 articles met the target criteria and were included in the review. Articles were excluded if:

- their data input did not include EHR, EMR or claims data;
- no machine learning approach was utilised;
- ADE detection or a general health outcome prediction tasks rather than ADE prediction were tackled;
- or if the study was not an original research article.

The studies evaluated different adverse events, whereas 31 [76%] out of 41 focused on specific adverse events or affected body regions, and 10 [24%] out of 41 were adverse event agnostic (Table 1).[14–54] Most studies focused on antineoplastic agents [8 (20%) out of 41] or did not specify the medication class of interest or were medication agnostic [8 (20%) out of 41]. Other captured criteria and information are evaluated more thoroughly in the following sections. Although we present and discuss important design decisions and issues associated with EHR data in the following sections, the same points apply to other types of observational health data, such as EMR or administrative claims data.

Covariate selection

Machine learning models can only make meaningful predictions if the input covariates and their representation describe the



Drug Discovery Today

FIGURE 1

Graphical overview of important components of an adverse drug event (ADE) machine learning prediction pipeline. Depending on the problem formulation, feature and target time windows must be selected appropriately. Covariate selection and robust ADE definition are crucial for training the right machine learning model. The surveyed studies differed significantly in their underlying machine learning model architecture and how they abstracted the temporal component of electronic health record (EHR) data. To assess model trust and transparency, we also investigated how the surveyed studies approached model interpretability, uncertainty quantification and model calibration.

TABLE 1

Overview of included studies.

Refs	Publication year	Dataset size	ADE	Drug class	Model type	AUC-ROC
14	2014	1746	Multiple	Multiple	XGBoost	–
15	2016	1558	Anaemia	Anti-anaemic preparations	Neural network	–
16	2016	1 200 000	Multiple	Multiple	Random forest	0.69–0.98
17	2017	592	Nephrotoxicity	Antibacterials for systemic use	Decision tree	–
18	2017	608	Venous thromboembolism	Antineoplastic agents	Kernel machine learning	0.72
19	2017	60 534	Acute kidney injury	Multiple	Random forest	0.67–0.77
20	2018	53 126	Hypoxaemia	Anaesthetics	XGBoost	0.92
21	2018	1 250 825	Unspecified	Multiple	Random forest	0.70–0.95
22	2019	560 057	Overdose	Analgesics	Neural network	0.91
23	2019	1042	Acute liver failure	Analgesics	Support vector machine	0.73
24	2019	17 446	Heart failure	Multiple	XGBoost	0.91
25	2019	396	Neutropenia	Antivirals for systemic use	Decision tree	–
26	2019	402	Nephrotoxicity	Antibacterials for systemic use	Decision tree	–
27	2019	1 200 000	Multiple	Multiple	Random forest	–
28	2019	38 780	Hypoglycaemia	Drugs used in diabetes	Random forest	0.9
29	2020	50 397	Major depressive disorder	Beta-blocking agents	Logistic regression	0.74
30	2020	1271	β-lactam hypersensitivity	Antibacterials for systemic use	Neural network	0.94
31	2020	254	Hepatic injury	Antimycobacterials	Neural network	0.90
32	2020	1 314 646	Multiple	Multiple	Recurrent neural network	0.77
33	2020	4960	Arrhythmia, heart failure, myocarditis, pericardial disease	Antineoplastic agents	XGBoost	0.65
34	2020	1141	Nephrotoxicity	Antibacterials for systemic use	Neural network	0.83
35	2020	222	Hypotension	Anaesthetics	Random forest	0.84
36	2020	80 768	Overdose	Analgesics	Neural network	0.70–0.76
37	2021	267	Nausea	Analgesics	Decision tree	–
38	2021	291 560	Hospitalisation and death	Multiple	Graph neural network	0.9
39	2021	353	Hepatotoxicity	Antineoplastic agents	Logistic regression	0.61–0.65
40	2021	5 544 150	Anaphylaxis, agranulocytosis	Vaccines	Random forest	0.70–0.77
41	2021	214 676	Death, heart failure, myocardial infarction, stroke	Drugs used in diabetes	Recurrent neural network	0.79–0.81
42	2021	5644	Unspecified	Psycholeptics	Transformer	0.84
43	2021	1326	Therapeutic failure (ICU transfer or death), stroke, major bleeding	Antithrombotic agents	XGBoost	0.95
44	2021	290	Overdose	Analgesics	XGBoost	0.89
45	2021	392 979	Multiple	Analgesics	XGBoost	0.87
46	2021	–	Acute coronary syndrome	Anti-inflammatory and antirheumatic products	XGBoost	0.72
47	2021	138	Drug retention rate	Immunosuppressants	Decision tree	–
48	2021	68 889	Acute coronary syndrome and death	Anti-inflammatory and antirheumatic products	XGBoost	0.72–0.84
49	2022	36 030	Multiple	Antineoplastic agents	XGBoost	0.82
50	2022	935	Multiple	Antineoplastic agents	Neural network	0.83
51	2022	499	Multiple	Antineoplastic agents	Gradient boosting decision tree	0.83
52	2022	118	Acute kidney injury	Antineoplastic agents	Neural network	0.9
53	2022	1616	Acute kidney injury	Antineoplastic agents	Support vector machine	0.83
54	2022	241	Thromboembolism	Antithrombotic agents	XGBoost	0.81

The drug class is derived from the Anatomical Therapeutic Chemical (ATC)[61] 2nd level code of the drug(s) of interest. The model type refers to the best-performing machine learning model used in the study. The AUC-ROC was the most reported performance metric and is rounded to two decimals in this table. If multiple experiments were evaluated, we reported the lowest and highest AUC-ROC among these experiments.

important relationships and dependencies to the output. The studies mainly used demographic data, diagnoses, procedures, laboratory values and medication information as covariates. Typically, EHR data are encoded through different medical vocabularies. For diagnosis, the most used coding system was the International Classification of Diseases (ICD) [21 (51%) out of 41]. In 17 [42%] out of the 41 studies that used diagnosis codes, the diagnosis coding system or the process of diagnosis information retrieval was not specified. Moreover, the Medical Dictionary for Regulatory Activities (MedDRA),[55] Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT),[56] the WHO Adverse Reaction Terminology[57] or the Observational Medical Outcomes Partnership (OMOP) common data model[58] were used in a minority of the reviewed studies.

Some studies utilise only a single type of covariates, for example diagnosis codes[59] or lab values.[27] However, the use of multiple covariates beyond clinical features is anticipated to yield better models. Genetics can be pivotal in patient adverse and therapeutic responses to medications, and including omics-related features can enhance the prediction.[31,52] Furthermore, integrating clinical- and mechanistic-model-derived features can increase the predictive performance of developed models.[59]

Adverse drug events and negative controls

Identifying the right ADE and defining appropriate negative controls is crucial for the correct training of machine learning models. Only a well-labelled dataset enables the model to build an accurate understanding of the task that is represented by the input features and the associated labels. However, identifying ADEs is difficult owing to noisy, erroneous or missing data and coding inaccuracy,[60,61] different coding practices between institutions,[61] the granularity of coding systems and difficulties in establishing drug causality for adverse events.[62]

TABLE 2

Overview of ADE and negative control definitions, separated by used covariates.

Covariate used	ADE definition	Negative control definition
Diagnosis and procedures	Code occurrence	Code occurrence
	• Occurrence of target code(s)[16,22,28,29,32,33,36,40,41,43,45,46,49,54]	No occurrence of any of the target codes[25,41]
	• Multiple occurrences of the target code[25]	No occurrence of any of the target codes but similar code occurrence, i.e., ICD code from the same chapter (first three characters)[16,27]
	CTCAE	CTCAE
	• Grade > 0[38]	Grade = 0[38]
	Naranjo scale	Naranjo scale
	• Score > 0[59]	Score = 0[59]
	• Score > 1[82]	Score ≤ 1[82]
Clinical notes and manual assessment	Automated text extraction[29]	
	Consensus manual assessment[15]	
	CTCAE	CTCAE
	• Grade > 0[51]	Grade = 0[51]
	• Grades 3–5[52]	Grade < 3[52]
Laboratory tests	Thresholds[16,17,20,26,28,35]	
	Skin tests and drug provocation tests[31]	
	CTCAE	CTCAE
	• Grade > 0[40,52]	Grade = 0[40,52]
	• Referring to CTCAE thresholds for classifying neutropenia[26]	
	CIOMS[32]	
	KDIGO[20,53]	
	West Haven Criteria	West Haven Criteria
	• Grade 3 or 4[24]	• Grade 1 or 2[24]
	Vancomycin therapeutic guidelines[35]	
Medication	Reduction in dose[22]	No dose reduction[22]
	Retention rate[48]	
Other	Death registry[49]	
	Hospitalisations and deaths[39]	
	Autopsy data[45]	

The ADE definition varies significantly across the surveyed literature, depending on the ADE of interest, the available covariates and the research focus. Some studies only specified the definition of the ADE, for example through the occurrence of a particular target diagnosis code. In these cases, we assume that the absence of the criteria defines the negative control.

- Council for International Organizations of Medical Science (CIOMS).[79]
- Common Terminology Criteria for Adverse Events (CTCAE).
- Kidney Disease Improving Global Outcomes (KDIGO)-based modifications of the Acute Kidney Injury Network and Risk, Injury, Failure, Loss, and End-Stage Kidney classification criteria.
- Vancomycin Therapeutic Guidelines: 2009 vancomycin consensus statement of the Infectious Diseases Society of America.[80]
- West Haven Criteria for hepatic encephalopathy.[81]

Table 2 summarises the ADE and negative control definitions used in the studies. ADEs are mainly defined based on diagnosis and procedure codes, (manually) retrieved from clinical notes or defined through laboratory results. In many cases, ADEs and negative controls are defined specifically for the studies. Alternatively, validated criteria for defining adverse events were used, such as the Common Terminology Criteria for Adverse Events (CTCAE).[37,50]

Data quality issues due to erroneous data can be mitigated by not relying on single medical (e.g., diagnosis) codes but by using multiple criteria to define an ADE or using validated algorithms to increase confidence in ADE labelling. ADE mislabelling can be a major concern, especially if no validated criteria or algorithm has been used to define ADEs and negative controls.[12] If no validated criteria are used, it is essential to ensure that the granularity of the coding system is sufficient to identify the adverse event of interest. For example, many drugs are known to increase the risk for a specific type of cardiac arrhythmias, torsades de pointes (TdP), for which a dedicated ICD-10 diagnosis code was only introduced in October 2022. Before this dedicated code was introduced, it was difficult to identify TdP precisely.[63]

Whereas the general prediction of a health outcome is already difficult, establishing drug association with the incidence of the outcome complicates it further. For drug association, appropriate drug exposure or drug era windows must be defined. Different ways to define such are available, typically based on prescribing or dispensing records, whereas using dispensing records is preferred.[12] Moreover, guidelines and validated algorithms are available to assess the likelihood of drug association with an adverse event, such as the Naranjo scale,[64] which is used by some studies.[58,64] Outcome definition is typically focused on the adverse event. However, defining the negative control can be difficult with EHR data. Often the absence of an event is used to define a negative instance. Owing to the potential underreporting of ADEs, this might be insufficient to determine class affiliation. Consequently, a surrogate event could be a better fit to define the negative control.[27]

Algorithmic approach

Machine learning model

A wide range of machine learning algorithms is used for ADE predictions, with random forest or boosted trees [19 (46%) out of 41] and neural networks [8 (20%) out of 41] being the most used algorithms. Although the main benefit of deep learning approaches is the ability of automatic feature engineering, gradient-boosted trees (including XGBoost) often outperform deep learning approaches on tabular data and, consequently, are a good choice for many machine learning tasks.[65,66] However, conventional machine learning approaches, including

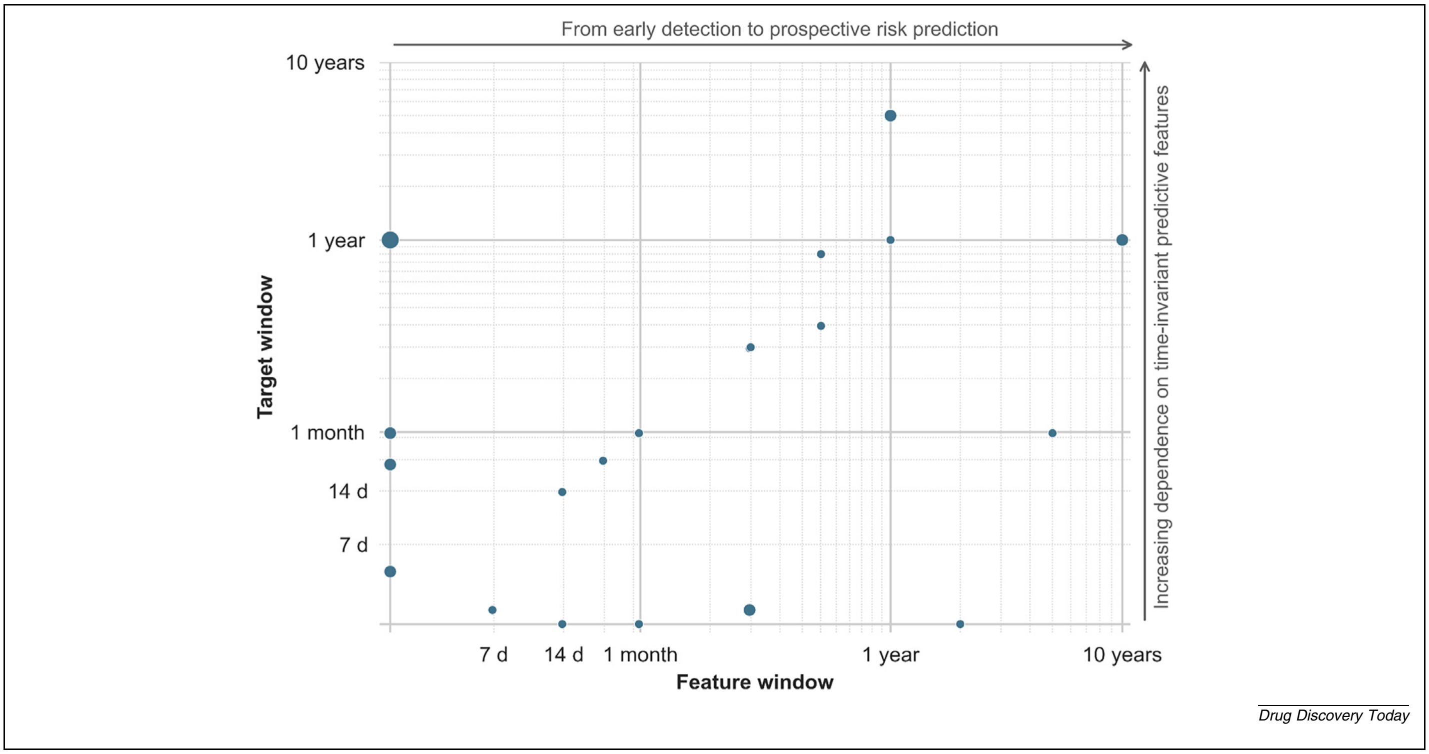


FIGURE 2

Scatter plot showing feature and target window lengths in the surveyed studies, with the marker size corresponding to the number of studies. A zero value indicates that the study did not report the window size. The majority of studies used window lengths within 3 months. Six studies used window lengths longer than 1 year.[28,36,41,45,46,48] The studies not shown in this plot did not specify the window lengths used for feature extraction or target period.

simple multi-layer perceptrons (neural networks), neglect the temporality of the EHR data and can disregard useful information for the prediction.

Temporal representation

The temporality of EHR data with irregular time intervals, varying patient trajectories and a high degree of sparsity increases the complexity of the time series modelling for the machine learning task. Temporal data can be represented through different aggregate statistics (e.g., mean, minimum, maximum value within a specific period) for a time-varying feature. However, any aggregation statistics neglect the natural temporality of the data. Thus, sequence models that can incorporate the sequential nature of EHR data are anticipated to improve performance. Two studies used a sequential or temporal representation through RNNs.[32,41] Three studies incorporated temporal information through weighted feature aggregation,[16] piecewise and symbolic aggregation[28] or exponentially decaying temporal weights.[20] The remaining studies [35 (85%) out of 41] used aggregated features, features at baseline or a single time point, or did not specify how the EHR data are abstracted.

In addition to the temporal representation method used, the selection of an appropriate baseline period (i.e., feature window) as well as the follow-up period (i.e., target window) is essential to capture all necessary covariates and the ADE (Figure 1). The feature and target window sizes reported in the literature are shown in Figure 2. The remaining studies did not report the window size but often stated only that features were extracted upon, for example hospital admission or study start. A buffer window was only used in two studies.[19,45]

Large target windows infer that the predictive features are time-invariant or only change with very slow dynamics. For instance, Lo-Ciganic *et al.* developed a machine learning model to assess the risk of opioid overdose with a 5-year target window.[22] It was found that demographic (i.e., time-invariant) features are among the most predictive ones. Shorter feature windows (with small or no buffer windows) are more related to early prediction tasks because early-onset features can be used for prediction, for example daily predicting the risk for acute liver failures based on a 7-day feature window.[23] If the feature window increases more chronic risk factors can be considered (e.g., chronic heart failure or diabetic neuropathy) for the risk prediction of hypoglycemia.[28]

Appropriate window sizes are crucial to record all relevant risk factors or covariates (feature window or baseline period), ensure that ADEs are captured (target window or follow-up period) and avoid data leakage (through a buffer window or other measures). These design choices should be informed by the context of use. Target windows can be selected based on the pharmacokinetics property of the drug of interest (e.g., half-life) and an additional washout period.[67,68] Information leakage, for example when highly correlated covariates co-occur with the ADE rather than causing the adverse event,[16] can be avoided by selecting an appropriate buffer window or masking covariates during training that are highly correlated with the ADE. A buffer window is particularly important if long-lasting risk factors are identified rather than short-term correlations.

Model trust and transparency

We also evaluated which studies addressed model interpretability, uncertainty quantification and calibrated model outputs. If implemented correctly, interpretable machine learning models, uncertainty quantification of the model predictions and calibrated models increase the robustness and trustworthiness of machine learning approaches. This is essential for mitigating concerns related to opaqueness, potential bias, accountability and responsibility. This, consequently, is also essential for successful deployment in clinical trials or clinical practice.

Interpretability

Interpretability can be achieved through model-based (e.g., decision trees) or post hoc interpretability (e.g., statistical feature importance assessment, prediction-level interpretation).[69] These methods enable quantifying the contribution of individual features for individual prediction, which is useful for quality assurance and protects against label leakage and spurious correlations.[69] We evaluated whether an interpretable model has been used, a well-known interpretability measure or feature importance evaluation (such as SHAP,[70] LIME[71] or Gini importance for random forests) was evaluated or the studies discuss the topic of interpretability in any other manner. The majority of studies [35 (85%) out of 41] used an interpretability measure.

Uncertainty estimation

The ability to assess the prediction's certainty is crucial for establishing trust in machine learning systems. Overconfidence is a common problem in machine learning. The ability to say 'the model is not sure in its prediction' will boost trust in machine learning models supporting medical decisions. This includes the estimate of uncertainty of individual predictions and overall performance uncertainty.[69] However, no study evaluated the uncertainty of their model predictions.

Calibration

To manage the over- and/or under-confidence of trained machine learning models, the scores of the model should be calibrated.[71,72] This is especially important when the model's output is used for clinical risk stratification, where treatment might be selected based on certain risk thresholds.[73] Three studies[23,30,46] reported calibrated models.

Performance

Reporting the appropriate performance metrics is an essential step toward a trusted machine learning model. The mainly reported performance metric was AUC-ROC [33 (80%) of 41], ranging from 0.61[40] to 0.978.[16] Although AUC-ROC was most often reported, it does not provide a comprehensive view of the actual performance on its own because it focuses on the majority class in a binary classification problem. Often, the minority class (i.e., the ADE cohort) is of more interest. Therefore, additional performance metrics, such as precision, recall and AUC-PR, provide useful information for an in-depth performance assessment.

Future directions and opportunities

The numerous machine learning models that have already been developed demonstrate the significant progress made in this field. It becomes increasingly evident that individualised ADE prediction emerges as an important component of precision medicine. For the future, we anticipate several key enablers that could catalyse the successful development and deployment of machine learning models for ADE prediction.

Data

Whereas EHR data offer great insights into the patient trajectory, there are many possibilities to enhance the feature space and get a more complete picture of the patient. Additional valuable data sources are for example omics data, data from wearables, screening data and self-reported data. They enhance the feature space and can be used to define the ADE and negative controls more confidently. In some cases, defining a true negative control cohort can be challenging owing to the underreporting problem of ADEs. Real-time patient data from wearables can help to increase confidence and reduce the uncertainty of identifying negative cohorts. Moreover, drug-specific risk can be quantified via pharmacokinetics/pharmacodynamics (PK/PD) modelling. Thus, combining mechanistic models with a machine learning approach could further increase the predictive performance of trained models.

Methods

Recently, transformer models have been proposed for disease or health outcome predictions from EHR data.[74,75] They leverage several concepts from the natural language processing (NLP) domain that enhance performance, because EHR data is sequential data that can be modelled similarly to natural languages.[74] The patient pathways can be modelled as a sequence of tokens, representing for example diagnosis, procedure codes or drug prescriptions. Owing to the large amount of EHR data, large foundation transformer models can be pretrained and are anticipated to create more generalisable models.[75,76] This can be important because patient journeys are complex,[10] the feature space is very high-dimensional[77] and it requires large amounts of data to train robust machine learning models. Pretrained models can be fine-tuned on specific ADEs and patient cohorts or used in a zero-shot learning setting.[77]

The temporal component of EHR data can be modelled by leveraging the positional embedding used in transformers,[74,75] which allows true temporal encoding and sampling of the irregular time intervals. Moreover, the attention mechanism[78] can enhance prediction interpretability. Furthermore, instead of building ADE-specific machine learning models trained on hand-selected comorbidities and features, an alternative approach can be building health outcome or disease prediction models with drugs as covariates. Establishing drug association is inherently difficult, even more so in EHR data, because causality assessments such as the Narjano scale were developed for clinical trials rather than clinical practice. Therefore, medication could be simply treated as another covariate and drug causality assessment could be removed. This might lead to a larger dataset and increase the generalisability of a trained model.

Applications

Models for individualised ADE prediction can be used to inform patient inclusion and exclusion criteria in clinical trials, especially if the model is drug-agnostic or drug-related features are abstracted. Moreover, in the scope of clinical decision support tools, the models can be used for an individualised benefit–risk analysis, inform treatment selection and risk monitoring or mitigation plans. Risk factors can be analysed through trained prediction models, for example through SHAP value analysis.[20] Individual risk prediction can improve therapy adherence and treatment outcome, enable systematic monitoring, reduce the overall ADE rate and therefore can reduce costs for the healthcare system in the future.

Concluding remarks

Machine learning is widely applied for ADE prediction with observational health data. In this review, we have surveyed recent literature and summarised the main findings from ADE prediction from observational health data through machine learning. There are many important design decisions for framing the right problem, capturing the important covariates and evaluating the performance of the machine learning model properly. We presented key considerations for machine learning approaches for ADE prediction, which are anticipated to help future researchers to design better machine learning models.

However, our review does not answer all technical and methodological questions. Patient cohort identification (inclusion and exclusion criteria) was not evaluated but can introduce significant biases to the machine learning model. Moreover, a meta-analysis of performance comparison between the approaches was not feasible owing to the heterogeneity of the dataset, ADEs and negative control definitions and the reported performance metrics. As we pointed out, disease or health outcome prediction models[43,74,75] can probably be applied to the task of ADE prediction but were not systematically surveyed, potentially neglecting relevant literature. We believe that new machine learning approaches, especially for modelling sequential or temporal data (such as transformers), facilitate the development of new applications for predicting ADEs. These models are anticipated to have a great role in clinical prediction tasks in the future, creating the conditions to make personalised healthcare more feasible and tangible.

Conflict of Interest

All authors are employed by F. Hoffmann-La Roche AG.

Data availability

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.drudis.2023.103715.

References

1. Edwards IR, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. *Lancet*. 2000;356:1255–1259.
2. Thuermann PA et al. Detection of adverse drug reactions in a neurological department: comparison between intensified surveillance and a computer-assisted approach. *Drug Saf*. 2002;25:713–724.
3. Alexopoulou A et al. Adverse drug reactions as a cause of hospital admissions: a 6-month experience in a single center in Greece. *Eur J Intern Med*. 2008;19:505–510.
4. Bouvy JC, De Bruin ML, Koopmanschap MA. Epidemiology of adverse drug reactions in Europe: a review of recent observational studies. *Drug Saf*. 2015;38:437–453.
5. Amelung S et al. Association of preventable adverse drug events with inpatients' length of stay-A propensity-matched cohort study. *Int J Clin Pract*. 2017;71. https://doi.org/10.1111/ijcp.12990.
6. Coleman JJ, Pontefract SK. Adverse drug reactions. *Clin Med*. 2016;16:481–485.
7. Peck RW. Precision dosing: an industry perspective. *Clin Pharmacol Ther*. 2021;109:47–50.
8. Allam A, Feuerriegel S, Rebhan M, Krauthammer M. Analyzing patient trajectories with artificial intelligence. *J Med Internet Res*. 2021;23:e29812.
9. Davies MR et al. Use of patient health records to quantify drug-related pro-arrhythmic risk. *Cell Rep Med*. 2020;1 100076.
10. Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE J Biomed Health Inform*. 2018;22:1589–1604.
11. Tayefi M et al. Challenges and opportunities beyond structured data in analysis of electronic health records. *Wiley Interdiscip Rev Comput Stat*. 2021;13. https://doi.org/10.1002/wics.1549.
12. Ng DQ et al. Current and recommended practices for evaluating adverse drug events using electronic health records: a systematic review. *J Am Coll Clin Pharm*. 2021;4:1457–1468.
13. Syrowatka A et al. Key use cases for artificial intelligence to reduce the frequency of adverse drug events: a scoping review. *Lancet Digit Health*. 2022;4:e137–e148.
14. Pandit AA, Dubey SA. A comprehensive review on Adverse Drug Reactions (ADRs) Detection and Prediction Models. *2021 13th International Conference on Computational Intelligence and Communication Networks (CICN)*. 2021. doi:10.1109/cicn51697.2021.9574639
15. Yu Z et al. Predicting adverse drug events in Chinese pediatric inpatients with the associated risk factors: a machine learning study. *Front Pharmacol*. 2021;12 659099.
16. Barbieri C et al. Performance of a predictive model for long-term hemoglobin response to darbepoetin and iron administration in a large cohort of hemodialysis patients. *PLoS One*. 2016;11:e0148938.
17. Zhao J, Henriksson A. Learning temporal weights of clinical events using variable importance. *BMC Med Inform Decis Mak*. 2016;16:71.
18. Imai S, Yamada T, Kasashi K, Kobayashi M, Iseki K. Usefulness of a decision tree model for the analysis of adverse drug reactions: evaluation of a risk prediction model of vancomycin-associated nephrotoxicity constructed using a data mining procedure. *J Eval Clin Pract*. 2017;23:1240–1246.
19. Ferroni P, Zanzotto FM, Scarpato N, Riondino S, Guadagni F, Roselli M. Validation of a machine learning approach for venous thromboembolism risk prediction in oncology. *Dis Markers*. 2017;2017:8781379.
20. Cheng P, Waitman LR, Hu Y, Liu M. Predicting inpatient acute kidney injury over different time horizons: how early and accurate?. *AMIA Annu Symp Proc*. 2017;2017:565–574.
21. Lundberg SM et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat Biomed Eng*. 2018;2:749–760.
22. Coulet A, Shah NH, Wack M, Chawki MB, Jay N, Dumontier M. Predicting the need for a reduced drug dose, at first prescription. *Sci Rep*. 2018;8:15558.
23. Lo-Ciganic WH et al. Evaluation of machine-learning algorithms for predicting opioid overdose risk among medicare beneficiaries with opioid prescriptions. *JAMA Netw Open*. 2019;2:e190968.
24. Speiser JL, Karvellas CJ, Wolf BJ, Chung D, Koch DG, Durkalski VL. Predicting daily outcomes in acetaminophen-induced acute liver failure patients with machine learning techniques. *Comput Methods Programs Biomed*. 2019;175:111–120.
25. Yang X et al. Identifying cancer patients at risk for heart failure using machine learning methods. *AMIA Annu Symp Proc*. 2019;2019:933–941.
26. Imai S, Yamada T, Kasashi K, Ishiguro N, Kobayashi M, Iseki K. Construction of a flow chart-like risk prediction model of ganciclovir-induced neutropaenia including severity grade: a data mining approach using decision tree. *J Clin Pharm Ther*. 2019;44:726–734.
27. Imai S, Yamada T, Kasashi K, Niinuma Y, Kobayashi M, Iseki K. Construction of a risk prediction model of vancomycin-associated nephrotoxicity to be used at the time of initial therapeutic drug monitoring: a data mining analysis using a decision tree model. *J Eval Clin Pract*. 2019;25:163–170.
28. Bagattini F, Karlsson I, Rebane J, Papapetrou P. A classification framework for exploiting sparse multi-variate temporal features with application to adverse drug event detection in medical records. *BMC Med Inform Decis Mak*. 2019;19:1–20.
29. Li X et al. Predictive modeling of hypoglycemia for clinical decision support in evaluating outpatients with diabetes mellitus. *Curr Med Res Opin*. 2019;35:1885–1891.
30. Jin S et al. Prediction of major depressive disorder following beta-blocker therapy in patients with cardiovascular diseases. *J Pers Med*. 2020;10. https://doi.org/10.3390/jpm10040288.
31. Moreno EM et al. Usefulness of an artificial neural network in the prediction of β-lactam allergy. *J Allergy Clin Immunol Pract*. 2020;8:2974–2982.e1.
32. Lai NH et al. Comparison of the predictive outcomes for anti-tuberculosis drug-induced hepatotoxicity by different machine learning techniques. *Comput Methods Programs Biomed*. 2020;188 105307.
33. Rebane J, Samsten I, Papapetrou P. Exploiting complex medical data with interpretable deep learning for adverse drug event prediction. *Artif Intell Med*. 2020;109 101942.
34. Heilbroner SP et al. Predicting cardiac adverse events in patients receiving immune checkpoint inhibitors: a machine learning approach. *J Immunother Cancer*. 2021;9. https://doi.org/10.1136/jitc-2021-002545.
35. Imai S et al. Validation of the usefulness of artificial neural networks for risk prediction of adverse drug reactions used for individual patients in clinical practice. *PLoS One*. 2020;15:e0236789.
36. Kang AR et al. Development of a prediction model for hypotension after induction of anesthesia using machine learning. *PLoS One*. 2020;15:e0231172.
37. Hastings JS, Howison M, Inman SE. Predicting high-risk opioid prescriptions before they are given. *Proc Natl Acad Sci U S A*. 2020;117:1917–1923.
38. Kumai M et al. Construction of a risk prediction model of extended release oxycodone tablet-induced nausea and clarification of predictive factors. *Biol Pharm Bull*. 2021;44:593–598.
39. Anastopoulos IN, Herczeg CK, Davis KN, Dixit AC. Multi-drug featurization and deep learning improve patient-specific predictions of adverse events. *Int J Environ Res Public Health*. 2021;18. https://doi.org/10.3390/ijerph18052600.
40. Kim JS, Han JM, Cho YS, Choi KH, Gwak HS. Machine learning approaches to predict hepatotoxicity risk in patients receiving nilotinib. *Molecules*. 2021;26. https://doi.org/10.3390/molecules26113300.
41. Kim Y et al. Machine learning approach for active vaccine safety monitoring. *J Korean Med Sci*. 2021;36:e198.
42. Longato E, Fadini GP, Sparacino G, Avogaro A, Tramontan L, Di Camillo B. A deep learning approach to predict diabetes' cardiovascular complications from administrative claims. *IEEE J Biomed Health Inform*. 2021;25:3608–3617.
43. Wang Z, Poon J, Wang S, Sun S, Poon S. A novel method for clinical risk prediction with low-quality data. *Artif Intell Med*. 2021;114 102052.
44. Falsetti L et al. Risk prediction of clinical adverse outcomes with machine learning in a cohort of critically ill patients with atrial fibrillation. *Sci Rep*. 2021;11:18925.
45. Lo-Ciganic WH et al. Integrating human services and criminal justice data with claims data to predict risk of opioid overdose among Medicaid beneficiaries: a machine-learning approach. *PLoS One*. 2021;16:e0248360.
46. Sharma V, Kulkarni V, Eurich DT, Kumar L, Samanani S. Safe opioid prescribing: a prognostic machine learning approach to predicting 30-day risk after an opioid dispensation in Alberta, Canada. *BMJ Open*. 2021;11:e043964.
47. Ward IR, Wang L, Lu J, Bennamoun M, Dwivedi G, Sanfilippo FM. Explainable artificial intelligence for pharmacovigilance: what features are important when predicting adverse outcomes?. *Comput Methods Programs Biomed*. 2021;212 106415.
48. García-Dorta A et al. Association of gender, diagnosis, and obesity with retention rate of secukinumab in spondyloarthropathies: results form a multicenter real-world study. *Front Med*. 2021;8 815881.
49. Lu J et al. Machine learning risk prediction model for acute coronary syndrome and death from use of non-steroidal anti-inflammatory drugs in administrative data. *Sci Rep*. 2021;11:18314.

50. Li C, Chen L, Chou C, Ngorsuraches S, Qian J. Using machine learning approaches to predict short-term risk of cardiotoxicity among patients with colorectal cancer after starting fluoropyrimidine-based chemotherapy. *Cardiovasc Toxicol*. 2022;22:130–140.
51. On J, Park HA, Yoo S. Development of a prediction models for chemotherapy-induced adverse drug reactions: a retrospective observational study using electronic health records. *Eur J Oncol Nurs*. 2022;56 102066.
52. Zhang N et al. A risk-factor model for antineoplastic drug-induced serious adverse events in cancer inpatients: a retrospective study based on the global trigger tool and machine learning. *Front Pharmacol*. 2022;13 896104.
53. Huang SH et al. How platinum-induced nephrotoxicity occurs? Machine learning prediction in non-small cell lung cancer patients. *Comput Methods Programs Biomed*. 2022;221 106839.
54. Yu X, Wu R, Ji Y, Huang M, Feng Z. Identifying patients at risk of acute kidney injury among patients receiving immune checkpoint inhibitors: a machine learning approach. *Diagnostics (Basel)*. 2022;12. https://doi.org/10.3390/diagnostics12123157.
55. Dai MF et al. Warfarin anticoagulation management during the COVID-19 pandemic: the role of internet clinic and machine learning. *Front Pharmacol*. 2022;13 933156.
56. MedDRA. Accessed 20 July 2022. https://www.meddra.org/.
57. Uppsala Monitoring Centre. WHO-ART legacy service. Accessed 20 July 2022. https://who-umc.org/vigibase/vigibase-services/who-art/.
58. FitzHenry F et al. Creating a common data model for comparative effectiveness with the observational medical outcomes partnership. *Appl Clin Inform*. 2015;6:536–547.
59. McMaster C, Liew D, Keith C, Aminian P, Frauman A. A machine-learning algorithm to optimise automated adverse drug reaction detection from clinical coding. *Drug Saf*. 2019;42:721–725.
60. Green CA, Perrin NA, Janoff SL, Campbell CI, Chilcoat HD, Coplan PM. Assessing the accuracy of opioid overdose and poisoning codes in diagnostic information from electronic health records, claims data, and death records. *Pharmacoepidemiol Drug Saf*. 2017;26:509–517.
61. Cozzolino F et al. A diagnostic accuracy study validating cardiovascular ICD-9-CM codes in healthcare administrative databases. The Umbria Data-Value Project. *PLoS One*. 2019;14:e0218919.
62. Peng M et al. Coding reliability and agreement of International Classification of Disease, 10 revision (ICD-10) codes in emergency department data. *Int J Popul Data Sci*. 2018;3:445.
63. Vallano A et al. Obstacles and solutions for spontaneous reporting of adverse drug reactions in the hospital. *Br J Clin Pharmacol*. 2005;60:653–658.
64. Naranjo CA et al. A method for estimating the probability of adverse drug reactions. *Clin Pharmacol Ther*. 1981;30:239–245.
65. Hung CY, Lin CH, Chang CS, Li JL, Lee CC. Predicting gastrointestinal bleeding events from multimodal in-hospital electronic health records using deep fusion networks. *Conf Proc IEEE Eng Med Biol Soc*. 2019;2019:2447–2450.
66. Shwartz-Ziv R, Armon A. Tabular data: Deep learning is not all you need. *Inf Fusion*. 2022;81:84–90.
67. Agoram BM, Martin SW, van der Graaf PH. The role of mechanism-based pharmacokinetic-pharmacodynamic (PK-PD) modelling in translational research of biologics. *Drug Discov Today*. 2007;12:1018–1024.
68. Zou H, Banerjee P, Leung SSY, Yan X. Application of pharmacokinetic-pharmacodynamic modeling in drug delivery: development and challenges. *Front Pharmacol*. 2020;11:997.
69. Murdoch WJ, Singh C, Kumbier K, Abbasi-Asl R, Yu B. Definitions, methods, and applications in interpretable machine learning. *Proc Natl Acad Sci U S A*. 2019;116:22071–22080.
70. Lundberg SM, Lee SI. A unified approach to interpreting model predictions Accessed 5 August 2022. *Adv Neural Inf Process Syst*. 2017;30 https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf.
71. Ribeiro MT, Singh S, Guestrin C. Why should I trust you?. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM; 2016. doi:10.1145/2939672.2939778.
72. Fernández A, García S, Galar M, Prati RC, Krawczyk B, Herrera F. *Learning from Imbalanced Data Sets*. Springer International Publishing; 2018.
73. Steele AJ, Denaxas SC, Shah AD, Hemingway H, Luscombe NM. Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease. *PLoS One*. 2018;13:e0202344.
74. Rasmy L, Xiang Y, Xie Z, Tao C, Zhi D. Med-BERT: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *NPJ Digit Med*. 2021;4:86.
75. Li Y et al. BEHRT: transformer for electronic health records. *Sci Rep*. 2020;10. https://doi.org/10.1038/s41598-020-62922-y.
76. Krishnan R, Rajpurkar P, Topol EJ. Self-supervised learning in medicine and healthcare. *Nat Biomed Eng*. 2022;6:1346–1352.
77. Berisha V et al. Digital medicine and the curse of dimensionality. *NPJ Digit Med*. 2021;4:153.
78. Vaswani A, et al., Attention Is All You Need, June 2017. https://doi.org/10.48550/arXiv.1706.03762.
79. Bénichou C. Criteria of drug-induced liver disorders. *J Hepatol*. 1990;11:272–276.
80. Rybak MJ et al. Vancomycin Therapeutic Guidelines: A Summary of Consensus Recommendations from the Infectious Diseases Society of America, the American Society of Health-System Pharmacists, and the Society of Infectious Diseases Pharmacists. *Clin Infect Dis*. 2009;49:325–327.
81. Vilstrup H et al. Hepatic encephalopathy in chronic liver disease: 2014 Practice Guideline by the American Association for the Study of Liver Diseases and the European Association for the Study of the Liver. *Hepatology*. 2014;60:715–735.
82. Sperandio da Silva GM et al. A clinical adverse drug reaction prediction model for patients with chagas disease treated with benznidazole. *Antimicrob Agents Chemother*. 2014;58:6371–6377.