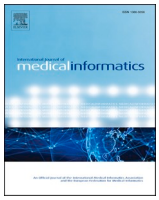
International Journal of Medical Informatics 180 (2023) 105241

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf



Review article

Machine learning in medication prescription: A systematic review

Alexa Iancu, Ines Leb, Hans-Ulrich Prokosch, Wolfgang Rödle *

Chair of Medical Informatics, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Wetterkreuz 15, 91058 Erlangen, Germany

ARTICLE INFO

Keywords:
Machine Learning (ML)
Medication prescription
Electronic medical records (EMR)

ABSTRACT

Background: Medication prescription is a complex process that could benefit from current research and development in machine learning through decision support systems. Particularly pediatricians are forced to prescribe medications "off-label" as children are still underrepresented in clinical studies, which leads to a high risk of an incorrect dose and adverse drug effects.
Methods: PubMed, IEEE Xplore and PROSPERO were searched for relevant studies that developed and evaluated well-performing machine learning algorithms following the PRISMA statement. Quality assessment was conducted in accordance with the IJMEDI checklist. Identified studies were reviewed in detail, including the required variables for predicting the correct dose, especially of pediatric medication prescription.
Results: The search identified 656 studies, of which 64 were reviewed in detail and 36 met the inclusion criteria. According to the IJMEDI checklist, five studies were considered to be of high quality. 19 of the 36 studies dealt with the active substance warfarin. Overall, machine learning algorithms based on decision trees or regression methods performed superior regarding their predictive power than algorithms based on neural networks, support vector machines or other methods. The use of ensemble methods like bagging or boosting generally enhanced the accuracy of the dose predictions. The required input and output variables of the algorithms were considerably heterogeneous and differ strongly among the respective substance.
Conclusions: By using machine learning algorithms, the prescription process could be simplified and dosing correctness could be enhanced. Despite the heterogenous results among the different substances and cases and the lack of pediatric use cases, the identified approaches and required variables can serve as an excellent starting point for further development of algorithms predicting drug doses, particularly for children. Especially the combination of physiologically-based pharmacokinetic models with machine learning algorithms represents a great opportunity to enhance the predictive power and accuracy of the developed algorithms.

1. Introduction

Children cannot easily be seen as "young adults", especially when it comes to medication prescription. Unfortunately, this group of patients is still underrepresented in clinical studies [1], mainly due to the parents' missing knowledge and acceptance [2] and ethical issues [3]. But also, financial reasons regarding the profitability of clinical studies with children are a big obstacle - their implementation is quite expensive and the profit for pharmaceutical companies is rather small as the target group is limited. Additionally, only one third of the pediatric studies approved by the Food and Drug Administration between 2007 and 2014 was completed [4], which leads to ongoing drug approvals without any or with too little information on pediatric drug doses [5]. This situation forces pediatricians to off-label drug use, which makes up about half of the medication prescriptions for children [6]. The danger of an incorrect dose is extremely high, especially for small children and neonates [7], as the drug dose prescription mainly depends on the pediatricians experience and knowledge. Regarding the physiological changes in children's early years of life, many factors need to be considered when choosing the correct dose: pharmacodynamics, pharmacogenetics and pharmacokinetics [8]. Hence, the risk of adverse drug effects is much higher for off-label drug use than for regular prescriptions [9–11].

Many approaches were taken to solve the pediatric dose problem, mainly depending on body weight and body surface area - i.e. the method of Du Bois [12], the Broselow tape [13] and the equation of Stuart-Taylor [14]. However, these approaches seem to be imprecise and outdated [15]. More recent publications considered the individual physiology of children and developed flexible models to calculate the correct dose depending on pharmacokinetic factors like renal functions or serum concentrations [16,17]. Additionally, model-based

* Corresponding author.
E-mail address: wolfgang.roedle@fau.de (W. Rödle).

https://doi.org/10.1016/j.ijmedinf.2023.105241
Received 6 February 2023; Received in revised form 17 September 2023; Accepted 27 September 2023
Available online 20 October 2023

calculations for pediatric drug doses were conducted [18–20] for individual substances.

Although these flexible approaches were more precise and reliable than fixed dose schemes, the results are still missing an all-embracing overview about individual factors of pediatric dose calculations [21].

The recent progress in machine learning (ML) methods and the increasing use of artificial intelligence in medicine leads to the idea to solve the pediatric dosing problem with the help of these methods. And the demand is definitely showing the importance – 11 out of 12 pediatricians expressed interest in an online platform for pediatric dosing support [22]. However, since many questions remain unsolved, e.g. regarding the best working algorithm, input variables and output values, the objective of this review is to get an overview about existing ML methods for predictions in medication dosing. We also aim to evaluate the different approaches concerning their applicability in pediatrics. To have an overall quality assessment of the included literature we used the IJMEDI checklist.

2. Materials and methods

In order to satisfy the multiple complex requirements in pediatric medication prescription, extensive literature research was conducted to carry out this systematic review. Publications were searched in *PROSPERO, PubMed* and *IEEE Xplore* and screened after analysis following the guidelines on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [23]. Due to the small number of publications dealing with ML algorithms for pediatric medication prescription, publications outside the field of pediatrics were also included in the review.

We applied the following keywords to the database search: “machine learning” combined with “dosing/dose/dosage” or “drug dose/dosing” or “dose/dosage prediction” (and “pediatrics/children”) to identify suitable publications. Publications fulfilling the following criteria were included in the review:

i) presentation and analysis of a specified ML algorithm for calculating/predicting medication doses,
ii) evaluation of the used algorithm by calculating the error rate or any other parameter,
iii) naming the origin and dimension of the designated data set,
iv) describing the used input variables and output values of the algorithm and their role in the dose calculation.

Publications that miss information about any of the criteria above cannot be included in the review since these factors are crucial for classifying the algorithms according their eligibility for pediatric dose prediction.

We followed the PRISMA flow (see Fig. 1) for exclusion of publications and according to the eligibility criteria described above. All steps were performed at least by two authors and ambiguous decisions were discussed by all authors until consensus.

The remaining publications after completing the PRISMA flow were included in the review. These publications were examined in detail where special attention was put to the data items shown in Table 1. The publications were categorized into groups according to the extracted data items. This categorization simplifies the statistical calculations for the evaluation and makes correlations between individual factors more obvious.

As the pediatric medication prescription makes use of rather small drug doses, the performance of the algorithm in small dose ranges is fundamental for its future use and opportunities in this field. Moreover, the therapeutic window for drug prescriptions is even smaller for children than for adults, which makes an exact calculation, e.g. low variance and a low error rate, necessary. In order to find the algorithms that fulfill these requirements in the best way, special attention during the evaluation process was given to the performance of the respective algorithms for small dose ranges. The properties of the particular algorithms working best and worst for this scenario were compares with each other with the aim of finding individual factors correlating with the performance.

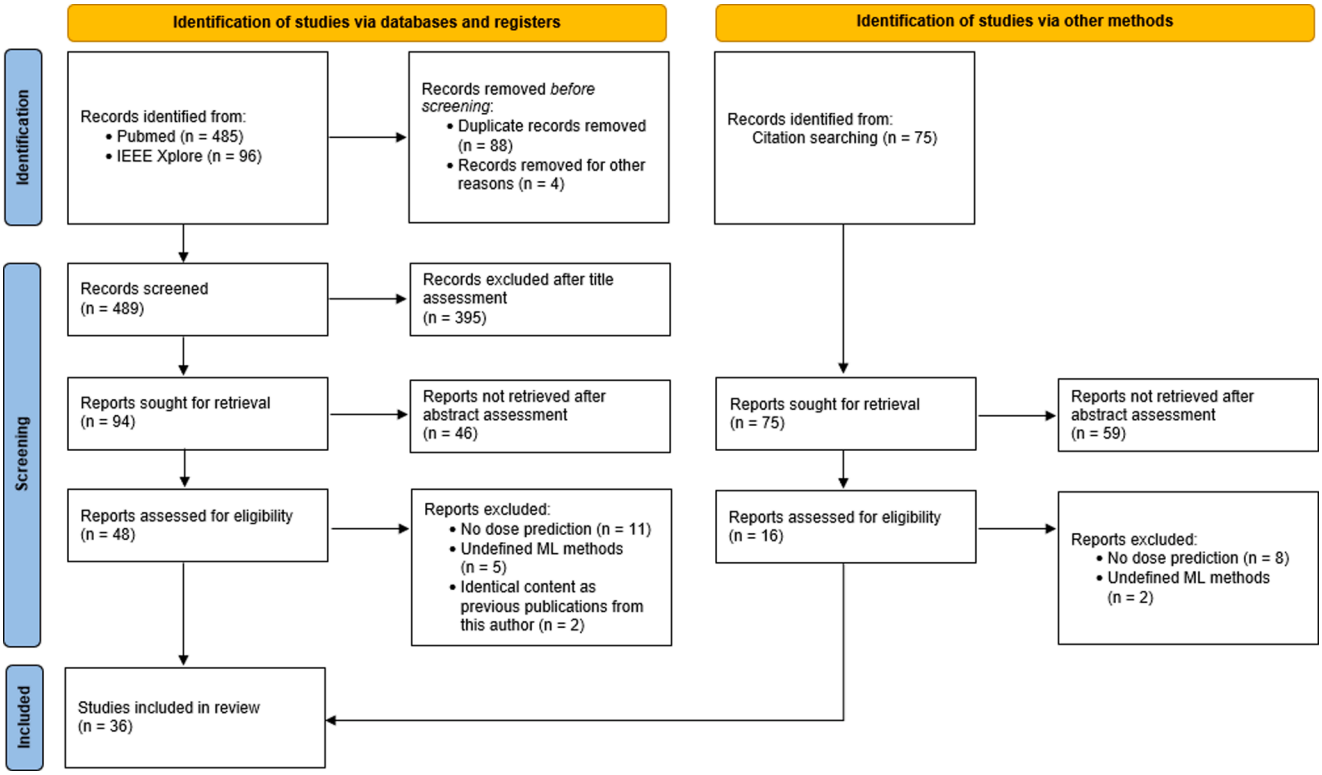


Fig. 1. PRISMA flow diagram.

Table 1
Data Extraction Items.

Item	Value	Explanation
Method(s)	Applied Machine Learning Method(s)	Which machine learning method(s) was/were used in particular to generate the output value?
Dataset	Size and Origin of the used test and training datasets	Where did the data items for training and test come from and how many were used?
Substance (s)	Name(s) of the active substance(s) the algorithm was developed for	Which active substance(s) is/are addressed by the algorithm?
Input	Input variables for the algorithm	Which variables are needed as input for the prediction?
Output	Output value predicted by the algorithm	Which output value does the algorithm give as result (e.g. daily dose/weekly dose/dose rate change, etc.)?
Error	Value of the prediction error and the calculation	Which error value is given, how is it calculated and what is the respective value?
Application	If applicable: the field of application	Which fields of application are covered by the algorithm (e.g. substances, patient group [age, race, diagnoses], etc.) and which limitations are mentioned?

In order to have a standardized evaluation of the studies' quality in the field of artificial intelligence in medical applications, the IJMEDI checklist [24] was used. With this checklist, developers can evaluate their own applications according to the necessary criteria and researchers can assess the quality of existing AI studies. The summarized score for the checklists' 30 questions is 50 points, while the questions consist of 20 high-priority items (40 points) and 10 low-priority items (10 points). In this systematic review, the studies examined were divided into low (0–25), medium (25.5–39.5), and high (40–50) quality according to their respective scores.

3. Results

Overall, we identified n = 656 publications from our literature research sources; *Pubmed* (n = 485), *IEEE Xplore* (n = 96), citation searching (n = 75). The process of evaluating the eligibility is shown in the PRISMA flow diagram in Fig. 1.

N = 36 publications were included in this review [25–60], the majority of them (n = 19) dealing with warfarin, the remaining ones dealing with vancomycin (n = 3), tacrolimus (n = 2), heparin (n = 2), thyroxine (T4) (n = 2) and dofetilide, lamotrigine, propofol and fentanyl, remifentanil, insulin, digoxin, opioids and substances catalyzed by cytochrome P450 (each n = 1). N = 21 of the identified publications are comparative analyses, which means that they are comparing multiple different ML algorithms with each other with respect to their individual use case. These analyses especially represent great value for this review as they might give hints for well-performing algorithms in direct comparison.

Regarding the respective ML methods, Decision Trees were used most often (n = 19) as well as Regression Analysis (n = 18), Support Vector Machines (SVM) (n = 17) and Neural Networks (NN) (n = 15). N = 7 publications used the k-Nearest Neighbor method (k-NN), whereas other methods like Bayesian Analysis or Relevance Vector Machines (RVM) were only used in individual cases. Besides, the majority of the approaches made use of Ensemble Methods (n = 19), i.e. Bagging, Boosting or Voting.

Looking at the datasets used by the different studies, it can be seen that there is a wide variety of the sizes and the types of input data. Many studies (at least partially) made use of already existing datasets like the IWPC dataset [61] or the MIMIC-III dataset [62] (n = 12). On average, the studies in this review collected data from 1766 patients (median: 650). What stands out is the size of the dataset used by Gu et al. [33], which includes more than 15,000 patients. In contrast to that, Gonzalez-Cava et al. [31] included only 15 patients.

The amount of input values also differs among the different studies: some included less than 10 input features (e.g. Li et al. [41]) and others gathered more than 25 features (e.g. Sharabiani et al. [51]) to predict the respective output values. A promising result is that 34 of the 36 identified publications were successful in terms of having at least one satisfying approach for their respective use case. Only the remaining n = 2 publications were not capable of finding an algorithm that can predict the outcome value in a similar or better way than existing guidelines or experts. The results for each individual substance are shown in Table 2, where the respective substances are linked to their best-performing algorithm(s) including their error rates, the strongest predictors and the outcome values.

Apart from that, it can easily be seen that the strongest predictors apart from age, gender and weight differ enormously among the different substances. Output values are either daily or weekly doses or, i.e. for T4, an individual value that is related to the pharmacodynamics of the respective substance.

For the most frequent methods the results are shown in Table 3. In general, algorithms based on Decision Trees and/or Regression Methods are most likely to be the best performing algorithm (in 7 out of 15 comparative analyses, respectively). Algorithms using Ensemble methods perform above average as well, being incorporated in the best

Table 2
Results classified by active substances (abbreviations can be found in Supplementary Material).

Active substances	Methods studied	Best performing method(s)	Lowest error (MAE)	Input values	Output value
Warfarin (n = 19)	NN, MLP, RF(R), SVR, MARS, BART, CART, k-NN, MLR, …	RF, SVR, MARS + Bagging/ Voting	0.21 $\frac{mg}{day}$ [35] 4.39 $\frac{mg}{week}$ [41] accuracy: 82 % [24]	age, gender, weight, diagnoses, VKORC1 & CYP2C9, amiodarone use, aspirin, target INR value	Weekly dose or daily dose
Vancomycin (n = 3)	CART, M5, XGBoost, SVR	CART, M5, XGBoost	9.14 $\frac{mg}{l}$ and 6.41 $\frac{mg}{l}$ for peak and through concentrations [34] accuracy: 70 % [26]	age, gender, weight, Scr, eGFR, medications, dosing intervals	Daily dose or peak and through concentrations
Heparin (n = 2)	RF, SVM, XGBoost, AdaBoost, NN	NN	F1 Score: 87 % [26]	age, gender, weight, AST/ALT ratio, Scr, therapeutic dose	therapeutic effect of the given dose / aPTT after 4–6 h
Tacrolimus (n = 2)	MLR, NN, RT, BRT, SVR, RFR, LAR, BART, MARS, LR	RT, Lasso + LR	0.72 $\frac{mg}{day}$ [53] accuracy: 52 % [53]	age, gender, weight, organ type, CYP3A5, lab values, diagnoses, medications	Daily dose or T1 level after 36–48 h
T4 (n = 2)	DT, SVM, NN, RF, OLSR, LAR, Poisson Regression, Gamma Regression, RR	DT, Poisson Regression	13 $\frac{\mu g}{day}$ [28] accuracy: 75 % [31]	age, gender, weight, TSH value, dietary supplements	Daily dose or daily dose adjustment

Table 3
Results classified by ML methods (abbreviations can be found in Supplementary Material); bold marked approaches show the best performance in a comparative analysis; study quality is indicated by the colour of the respective cells (green = high quality, yellow = medium quality, red = low quality).

Method	occurrence in comparative analyses (best performed)	active substances	lowest error (MAE) / accuracy	output value
Decision Trees (n=20)	n=15 (n=7)	warfarin, vancomycin, lamotrigine, tacrolimus, propofol/fentanyl, remifentanil, heparin, T4, digoxin, substances catalyzed by cytochrome P450	**0.21 $\frac{mg}{day}$ [37]**	**daily dose**
			82% [43]	
			4.49 $\frac{mg}{week}$ [41]	weekly dose
			13 µg [28]	**dose adjustment (4-6 weeks after initial dose)**
			83.9% [36]	**dose correctness**
			R^2: 0.388 [27]	single dose
			F1: 73.8% [52]	aPTT
			9.14 $\frac{mg}{l}$ [34]	**peak / through concentrations**
Regression methods (n=17)	n=15 (n=7)	warfarin, tacrolimus, propofol/fentanyl, remifentanil, T4, digoxin	**4.39 $\frac{mg}{week}$ [41]**	**weekly dose**
			0.21 $\frac{mg}{day}$ [35]	**daily dose**
			R^2: 0.501 [27]	**single dose**
			75% [31]	dose adjustments
			MSE: 0.6 $\frac{ug}{l}$ [46]	**T1 level**
			64.8% [36]	dose correctness
SVM (n=17)	n=15 (n=2)	warfarin, tacrolimus, propofol/fentanyl, remifentanil, opioids, vancomycin, heparin, T4	4.53 $\frac{mg}{week}$ [41]	weekly dose
			0.21 $\frac{mg}{day}$ [35]	**daily dose**
			R^2: 0.49 [27]	single dose
			85% [31]	dose adjustments
			F1: 76.2% [52]	aPTT
			9.8 $\frac{mg}{l}$ [34]	peak / through concentrations
NN (n=15)	n=13 (n=1)	warfarin, tacrolimus,	4.71 $\frac{mg}{week}$ [41]	weekly dose
			0.23 $\frac{mg}{day}$ [35]	daily dose
		remifentanil, heparin, T4, digoxin,	80.9% [36]	dose correctness
			81% [31]	dose adjustments
			F1: 87.3% [52]	**aPTT**
			92% [42]	**therapeutic effect**
k-NN (n=7)	n=6 (n=1)	warfarin, digoxin, propofol/fentanyl, remifentanil, insulin	6.15 $\frac{mg}{week}$ [48]	weekly dose
			0.22 $\frac{mg}{day}$ [35]	daily dose
			R^2: 0.424 [27]	single dose
			88% [31]	**dose adjustments**
			60.2% [36]	dose correctness
Ensemble methods (n=19)	n=16 (n=7)	warfarin, tacrolimus, vancomycin, digoxin, propofol/fentanyl, remifentanil, heparin, T4, substances catalyzed by cytochrome P450	4.52 $\frac{mg}{week}$ [41]	weekly dose
			0.21 $\frac{mg}{day}$ [35]	**daily dose**
			R^2: 0.488 [27]	single dose
			83% [29]	dose adjustments
			F1: 78.85% [50]	aPTT
			83.9% [36]	**dose correctness**

Table 4
Results of the quality assessment of the 36 publications according to the IJMEDI checklist; study quality is indicated by the colour of the respective cells (green = high quality, yellow = medium quality, red = low quality).

	Problem Understanding (10)	Data Understanding (6)	Data Preparation (8)	Modeling (6)	Validation (12)	Deployment (8)	Total (50)
Altay [25]	2	3	8	6	1	0.5	**20.5**
Alzubiedi [26]	10	4	8	5	3	1	**31**
Asai [27]	9	0	6	6	3.5	1	**25.5**
Chen [28]	5.5	2	6	6	6.5	0.5	**26.5**
Cosgun [29]	9	0	8	5	5.5	0.5	**28**
Coulet [30]	10	1	8	6	9.5	3	**37.5**
Gonzalez-Cava [31]	7	1	8	5	2.5	1.5	**25**
Grossi [32]	10	5	7	6	11	1	**40**
Gu [33]	10	5	7	6	11	3.5	**42.5**
Hu [34]	6	3	7	6	4.5	1	**27.5**
Hu [35]	3.5	2	4	6	7	0.5	**23**
Hu [36]	9	2	7	6	9.5	1.5	**35**
Huang [37]	9	3	7	6	8	5	**38**
Imai [38]	5.5	3	6	6	2	1.5	**24**
Klein [39]	10	5	8	6	9	2.5	**40.5**
Levy [40]	5	5	4	6	9	0	**29**
Li [41]	7	3	5	6	3	0.5	**24.5**
Li [42]	9.5	4	8	6	11	3	**41.5**
Liu [43]	6.5	2	8	5	3	1	**25.5**
Liu [45]	10	3	5	5	4	0	**27**
Ma [44]	9.5	5	5	6	6	2	**33.5**
Min [46]	9	3	7	3	5	2	**29**
Olesen [47]	8	2	2	5	2	0	**19**
Roche-Lima [48]	9.5	2	4	4	4.5	0.5	**24.5**
Schellemann [49]	7.5	4	6	4	4	0	**25.5**
Sharabiani [50]	6.5	2	6	4	3	0	**21.5**
Sharabiani [51]	9	5	6	6	8	1	**35**
Su [52]	9.5	4	8	6	6	1.5	**35**
Tang [53]	10	4	6	6	5	1	**32**
Tao [54]	5	2	8	6	9	2	**32**
Tao [55]	9	3	8	6	12	3.5	**41.5**
Truda [56]	5.5	6	5	6	9	3	**34.5**
Tyler [57]	10	6	3	6	8	4.5	**37.5**
van Nguyen [58]	8	4	6	6	10.5	2	**36.5**
Zaborek [59]	8	5	4	6	4	1	**28**
Zhu [60]	7	6	5	6	9.5	1	**34.5**

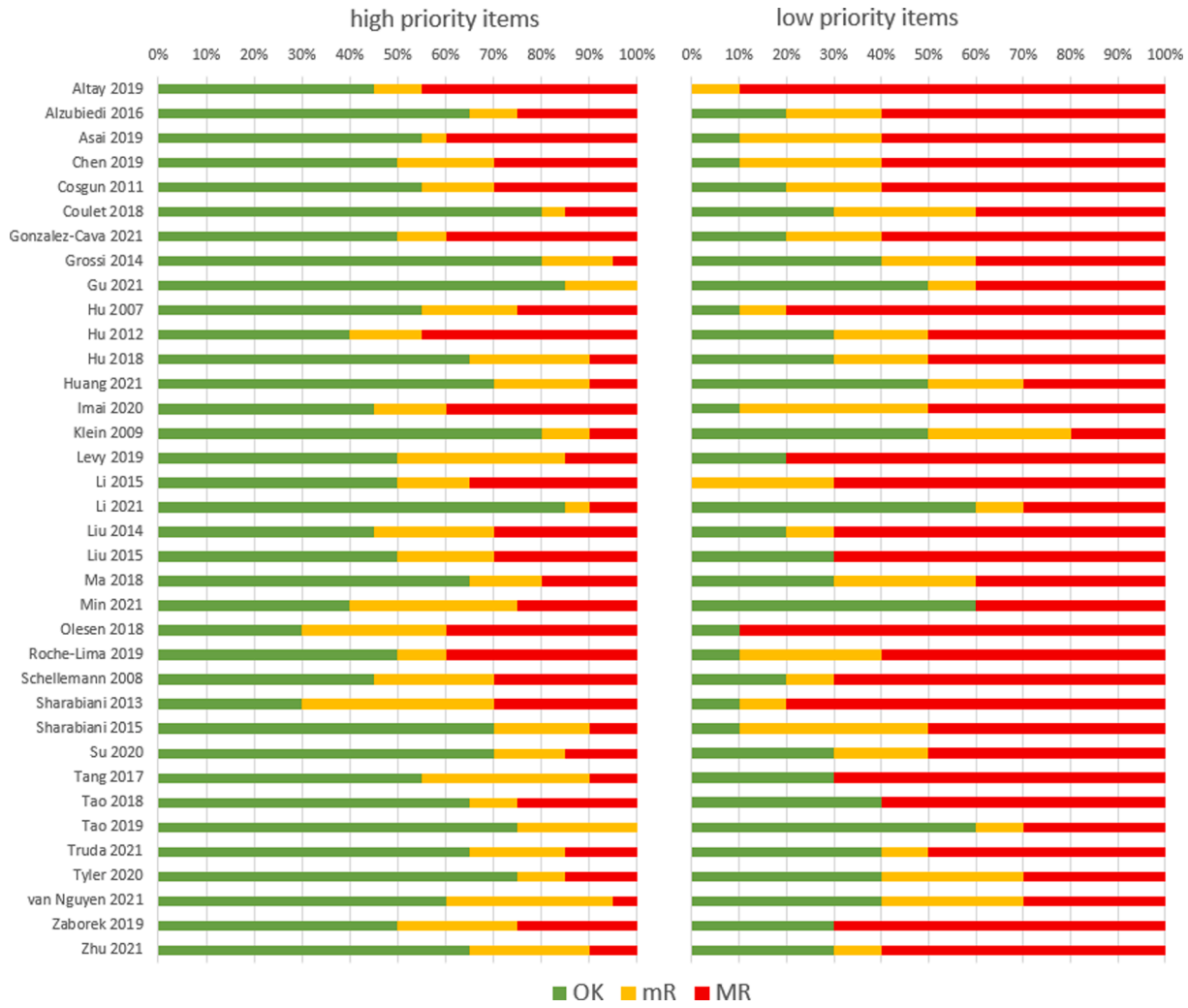


Fig. 2. Proportion of the results in the high- and low-priority items (OK = adequately addressed; mR = sufficient but improvable; MR = inadequately addressed).

algorithm in 7 out of 16 comparative analyses. On the contrary, NN (in 1 out of 13) or SVM (in 2 out of 15) are unlikely to be the best performing algorithms in a comparative analysis.The Supplementary File (Appendix C) details the results of the quality assessment. Table 4 summarises the scores of each category and the total score of each study. The average score of the IJMEDI checklist was 30.9 (range: 19 to 42.5). Most of the studies were of a medium quality. We identified five studies [32,33,39,42,55], that were of a high quality, while eight were of a low quality.

The results of the high-quality studies show that most of them used a very large or at least above average dataset (see above). Accordingly, most of the low-quality studies used a dataset of a size significantly below the average. Low and intermediate quality studies were also more likely to achieve either very good [35,41] or very poor results [25,47]. Fig. 2 shows the quantitative results of the checklist, ordered by publications and priority of items.

4. Discussion

The results of this review show that the performance of an algorithm mainly depends on its use case. Most of the reviewed publications developed very individual and specialized algorithms and used variables depending on the respective substances and datasets. For instance, calculating the correct dose for vancomycin requires completely different parameters (e.g. the glomerular filtration rate as a strong predictor and the peak and through concentrations as output variables) than for any other substance. As vancomycin is an antibiotic, it is necessary to maintain a steady concentration level to reach the desired effect. Even small inconsistencies or variances in the concentration level may lead to adverse side effects or even no effects at all. Moreover, Vancomycin overdoses are suspected to be nephrotoxic and should be avoided at all times [63,64]. In contrast to that, the crucial steps for the correct warfarin dosing are genotyping (Vitamin K epOxide Reductase Complex subunit 1 [VKORC1] and Cytochrome P450 family 2 subfamily C member 9 [CYP2C9]) and calculating the target International Normalized Ratio (INR) value. The research for warfarin made huge progress during the last years and the establishment of the International Warfarin Pharmacogenetics Consortium (IWPC) [61] in 2008 provided important knowledge about detailed warfarin medication data in combination with pharmacogenetic and pharmacodynamic variables. Therefore, the information on warfarin dosing is by far more developed than for other substances. By using the data from IWPC, it shall be possible to predict completely individualized dose values by incorporating pharmacogenetic information. Especially for warfarin, genetic data seems to be the most important predictor for the correct dose and was used in 12 out of 19 publications dealing with warfarin – especially in the five high-quality studies. Nevertheless, it is still discussed if pharmacogenetic data really improves the predictive accuracy. Some studies found that clinical data alone offers as much accuracy as the combination of clinical and genetic data [65–67], but these results are not necessarily transferable to long-term observations [68]. For warfarin, researchers aim to develop a fully patient-individualized

approach to predict the prescription and dose using genetic data [61] as these predictors differ strongly among different patient groups, e.g. age, race and prescription of other medications (see Appendix A, column "Application" for detailed information).

Regarding the input variables, a large variety among the reviewed publications could be observed which can be explained by the completely different factors that need to be considered for the different substances. For some, one might apply genetic variables (e.g. VKORC1 and CYP2C9 for warfarin), for others pharmacokinetic and pharmacodynamic variables (e.g. the glomerular filtration rate). Additional factors about medication (e.g. dosing times, intervals, and interaction with other active substances) and laboratory values (e.g. serum creatinine) can also be important.

It was suggested in former publications to generally include genetic data in medical records [69] and to use the entire available data in research and clinical care [70–72]. Nevertheless, it has to be considered that this data collection process would imply a significant increase of work load for clinicians and more examinations and time consumption for patients, which does not seem feasible, especially for children. Existing physiologically-based pharmacokinetic (PBPK) models could be the solution to fill the data gaps in a fast and simple way. Many models exist which developed calculations to simulate pharmacokinetic variables in children's bodies, e.g. for the clearance rate of cytochrome metabolized substances [73], the clearance rate of midazolam [74] or even for predicting the clearance for multiple substances (including warfarin and vancomycin) simultaneously. [75]. The use of such PBPK models seems to be helpful for finding the correct dose in pediatric use cases [76] as they are able to avoid overdosing by considering the individual physiological conditions [77]. Consequently, it seems reasonable to include these models in the prediction using ML methods to overcome missing data items and increase the predictive power. The results of the models may serve as a further information source for the following ML algorithm. However, further research on the PBPK models is needed and they cannot replace clinical experts dosing decisions for now [78].

Although most of the publications chose weekly or daily doses as output values, some use cases made individual outputs necessary, e.g. drug concentration values for vancomycin, recommendations for changes in the current dosing scheme for lamotrigine, the correctness of the current digoxin dose or the adjustment of the daily dose of T4. The output value appears to be as individual as the respective substances and application areas. The best performing algorithm also differs among the different output values: for concentration values and dose correctness predictions approaches using Random Forest (RF), Classification and Regression Trees (CART), Extreme Gradient Boosting (XGBoost) or M5 are most successful. For weekly and daily dose predictions, as in warfarin, the best methods appeared to be RF, Support Vector Regression (SVR) and MARS in combination with Bagging or Voting (see Table 3).

In general, it could be observed that the majority of algorithms needs a large amount of demographic, clinical, and substance-individualized data to be able to perform well. But a comparison between the different approaches remains difficult. Many publications do not offer detailed information on their dataset and how they curated the data. Especially the eight low-quality studies were lacking crucial information about the data extraction and preparation process. Moreover, none of the publications defined the term "data from n patients" - which could mean the complete data from every single of these n patients, a single examination of n patients, single hospital stays of n patients, a single laboratory value of n patients or even a mix of these. The difference in data quality and redundancy of single values collected for a single study versus a complete medical record of a patient is enormous and should not be compared to each other. Hence, the data quality, which could not be evaluated in this review, is much more important than the given size of the dataset, because a larger dataset does not equal higher quality data. This leads to the complicated situation, where an algorithm trained with a larger dataset is not necessarily more accurate than an algorithm trained with a smaller dataset [79], even when larger datasets usually lead to an increased predictive power [80]. Additionally, the problem of "Overfitting" is present in very large datasets, where redundant data leads to a decreasing performance of the algorithm [81]. Especially Adaptive Boosting and Logistic Regression are sensitive to Overfitting [82]. To avoid this, dimension reduction methods can be used [82] as it was done e.g. by Grossi et al. [32] (one of the 5 identified high-quality studies).

Furthermore, not only the different datasets are challenging to compare to each other, but also the algorithms itself. Although the publications stated clearly, which algorithm they were using, a large scope of differences can occur even when using the same algorithm. For example, Sharabiani et al. [50] and Alzubiedi [26] both used a Neural Network to predict the optimal dose of Warfarin for African-American patients and nevertheless, both have achieved very different results (MAE of 20.2 [50] versus MAE of 10.9 [26]). This could be because the stated methods, e.g. "Support Vector Machine" or "Neural Network" do not make any guidelines on the actual implementation and the chosen parameters of the algorithm itself. As the algorithms were mainly developed by the researchers themselves, the performance can strongly depend on the experience and knowledge about the methods of the respective researcher. The best results in this review were achieved by algorithms based on regression methods or Decision Trees [35,41,48], which are both relatively simple methods and easier to implement than other approaches. Neural Networks, in contrast, could be more complex to understand and to develop, which could be another reason why they performed poorly in comparison to easier methods.

Furthermore, the quality of the individual studies should be considered when comparing them. Although we did not find that higher quality studies lead to better results in general, it stands to the reason that these studies hold a greater value when it comes to subsequent research studies in that field and should be the preferred source of information.

These challenges regarding the comparison of the input datasets and used ML methods cannot easily be solved, what makes a clear comparison and evaluation of the given approaches difficult. For example, regarding the three publications dealing with vancomycin, it appears that all of them are using different ways to evaluate their algorithms – Huang et al. [37] use the percentage of predictions within 20 % of the dosing interval, Imai et al. [38] use the peak and through concentrations and Hu et al. [34] calculate the MAE.

As hardly any publications regarding the application of ML methods in the pediatric medication field matched the eligibility criteria (n = 1), this review focused on the application of such algorithms in the general medication process. But the results cannot be transferred directly to the pediatric use cases. The physiological differences between children's and adults' bodies are a crucial factor when comparing the dose prediction algorithms. The ability to metabolize substances is a process that does not evolve in a linear way. Since ever, researchers try to model the metabolization processes of youths, children and neonates in complex calculations and studies. An example is the development of the glomerular filtration rate, which represents an important input variable for various algorithms in this review. It indicates the required time for metabolizing substances in the kidney, which is quite low for neonates but achieves the adults' level during the first year of life [8]. In contrast to that, the pH value of the gastric acid is higher in neonates than in adults and decreases slowly during the first years of life [8]. All of these processes, especially the non-linear and complex ones, should be represented by a well-developed model and considered by the ML algorithm in order to achieve a precise dose prediction in pediatrics. It becomes clear that this algorithm cannot easily be derived from an adult's algorithm as there are more requirements that need to be considered.

Hence, it was found, that some algorithms perform well in general and in medium to high dose ranges, but poorly in lower dose ranges (see [32,41,48,27]). However, these small dose ranges are crucial for the

dose prediction for children and should be observed in detail. For instance, the comparative analyses of Li et al. [41] regarding warfarin stated an overall minimal error of 4.39 $\frac{mg}{week}$ using Multivariate Linear Regression (MLR), while the same algorithm leads to overdoses in 89 % for the low dose groups. Moreover, Liu et al. [45] achieved the best overall results using Multivariate Adaptive Regression Splines (MARS) but for small doses an algorithm based on Bayesian Additive Regression Tree (BART) clearly outperformed the other approaches. These observations lead to the conclusion, that it could be necessary to consider a clear distinction between ML approaches for adults and children.

To the best of our knowledge, this systematic review is the first one using the IJMEDI checklist for quality assessment of AI applications in the area of medication prescription. The checklist serves as a helpful tool when it comes to comparing different AI publications qualitatively and is easy to be used. The evaluation of the studies using the IJMEDI checklist shows that most AI studies in this area provide little to no explanation of the data set used (mean: 3.3 of 6) and the deployment (mean: 1.5 of 5). Only the medical reasons for the development (mean: 7.9 of 10) and the developed algorithms (mean: 5.6 of 6) are presented in detail. The results of the AI models are explained, but there is insufficient information on a validation process (mean: 6.4 of 12). This leads a medical professional to partially blindly trust the results and the algorithms. Due to the different ways of explanations, it is noticeable that the authors of the studies did not use a standardized checklist for their development. By using such a checklist in the field of AI developing, the authors could have identified and clarified open questions and gaps in their explanations.

Besides, the 36 studies included in this review are mostly not satisfying the requirements stated in the IJMEDI checklist [24], which is why only 5 studies were considered of a high quality. As the lack of information is especially crucial when it comes to the used dataset and the deployment, the authors strongly suggest to consider these aspects in further work.

This systematic review confirmed the demand for a standardized procedure in AI developing in medicine, especially to ensure comparability of quality and to provide developers a checklist to control their work for completeness.

5. Conclusion

This review shows that a well-trained ML algorithm is able to predict the correct dose in various use cases. But it also highlights the importance of a large amount of high-quality data to reach this goal. The available data items should cover a wide range of variables: demographic, clinical, genetic, pharmacokinetic, etc. Due to the complexity of this topic there is currently no ML system that is able to fulfil the requirements of different substances or even the special demands in a pediatric application area.

The overview identifies algorithms based on regression methods, decision trees and ensemble methods as the top performers in dosing prediction. However, the complexity is in the detail and there is currently too little research for a general statement. The respective method as well as input and output variables need to be tailored to the respective use cases, patient groups, substances and other requirements. PBPK models may serve as additional information source to simplify the data collection process. Finally, we have found that ML algorithms are able to support physicians in the medication prescription process, even though much research and development work is still needed in this area.

Summary table

What was already known on the topic:

- There is an increasing use of machine learning in clinical context in various disciplines.
- A vast amount of publications exists on the application of machine learning methods in the field of medication prescription. Most of them are not taking the special requirements in pediatrics into account.

What this study added to our knowledge:

- This systematic review is the first of its kind to provide a broad overview of machine learning applications in medication prescription, especially in pediatrics by systematically comparing multiple studies and approaches regarding their predictive abilities, input variables and output values as well as assessing the quality of the studies by using the standardized IJMEDI checklist.
- This review highlights versatile aspects of different machine learning approaches and the possibilities they offer when it comes to improving the dose prediction for children.

The results of this study can serve as a basis for further research and development in medication prescription using machine learning algorithms and their application.

CRediT authorship contribution statement

Alexa Iancu: Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Ines Leb:** Methodology. **Hans-Ulrich Prokosch:** Supervision. **Wolfgang Rödle:** Conceptualization, Validation, Resources, Writing – original draft, Writing – review & editing.

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 work reported in this paper.

Acknowledgments

PROSPERO-ID: CRD42022336469.

The present work was performed in (partial) fulfilment of the requirements for obtaining the degree “Dr. rer. biol. hum.” from the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) (WR).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2023.105241.

References

[1] W. Rascher, Fehlende Medikamentenzulassungen für Kinder — ein ethisches Dilemma, Pädiatrie 29 (1) (2017) 32–34, https://doi.org/10.1007/s15014-017-1128-1.

[2] C. Lenk, P. Koch, H. Zappel, C. Wiesemann, Off-label, off-limits? Parental awareness and attitudes towards off-label use in paediatrics, Eur. J. Pediatr. 168 (12) (2009) 1473–1478, https://doi.org/10.1007/s00431-009-0956-6.

[3] M. Cohen-Wolkowiez, D.K. Benjamin Jr., Development of therapeutics for children-a tricky balancing act, JAMA Pediatr. 173 (1) (2019) 18–19, https://doi.org/10.1001/jamapediatrics.2018.4026.

[4] T.J. Hwang, L. Orenstein, A.S. Kesselheim, F.T. Bourgeois, Completion Rate and reporting of mandatory pediatric postmarketing studies under the US Pediatric research equity act, JAMA Pediatr. 173 (1) (2019) 68–74, https://doi.org/10.1001/jamapediatrics.2018.3416.

[5] J. Permala, M.A. Hassali, A. Awaisu, A.A. Shafie, Dosing information in a standard drug reference: are pediatrics still therapeutically neglected? Pediatr. Int. 52 (2) (2010) 290–295, https://doi.org/10.1111/j.1442-200X.2009.02958.x.

[6] S. Conroy, I. Choonara, P. Impicciatore, A. Mohn, H. Arnell, A. Rane, C. Knoeppel, H. Seyberth, C. Pandolfini, M.P. Raffaelli, F. Rocchi, M. Bonati, G. Jong, M. de Hoog, J. van den Anker, Survey of unlicensed and off label drug use in paediatric wards in European countries. European Network for Drug Investigation in Children, BMJ 320 (7227) (2000) 79–82, https://doi.org/10.1136/bmj.320.7227.79.

[7] H.A. McPhillips, C.J. Stille, D. Smith, J. Hecht, J. Pearson, J. Stull, K. Debellis, S. Andrade, M. Miller, R. Kaushal, J. Gurwitz, R.L. Davis, Potential medication

dosing errors in outpatient pediatrics, J. Pediatr. 147 (6) (2005) 761–767, https://doi.org/10.1016/j.jpeds.2005.07.043.
[8] J. Le, J.S. Bradley, Optimizing antibiotic drug therapy in pediatrics: current state and future needs, J. Clin. Pharmacol. 58 (Suppl 10) (2018) S108–S122, https://doi.org/10.1002/jcph.1128.
[9] B. Horen, J.-L. Montastruc, M. Lapeyre-Mestre, Adverse drug reactions and off-label drug use in paediatric outpatients, Br. J. Clin. Pharmacol. 54 (6) (2002) 665–670, https://doi.org/10.1046/j.1365-2125.2002.t01-3-01689.x.
[10] S. Turner, A.J. Nunn, K. Fielding, I. Choonara, Adverse drug reactions to unlicensed and off-label drugs on paediatric wards: a prospective study, Acta Paediatr. 88 (9) (1999) 965–968, https://doi.org/10.1080/08035259950168469.
[11] J.R. Bellis, J.J. Kirkham, A.J. Nunn, M. Pirmohamed, Adverse drug reactions and off-label and unlicensed medicines in children: a prospective cohort study of unplanned admissions to a paediatric hospital, Br. J. Clin. Pharmacol. 77 (3) (2014) 545–553, https://doi.org/10.1111/bcp.12222.
[12] D. Du BOIS. Clinical calorimetry. Arch Intern Med (Chic) XVII(6_2) (1916) 863. 10.1001/archinte.1916.00080130010002.
[13] R.C. Luten, R.L. Wears, J. Broselow, A. Zaritsky, T.M. Barnett, T. Lee, A. Bailey, R. Vally, R. Brown, B. Rosenthal, Length-based endotracheal tube and emergency equipment in pediatrics, Ann. Emerg. Med. 21 (8) (1992) 900–904, https://doi.org/10.1016/s0196-0644(05)82924-5.
[14] J.A. Lack, M.E. Stuart-Taylor, Calculation of drug dosage and body surface area of children, Br. J. Anaesth. 78 (5) (1997) 601–605, https://doi.org/10.1093/bja/78.5.601.
[15] P.J. Munzenberger, P. McKercher, Pediatric dosing–the pharmacist's dilemma, Contemp. Pharm. Pract. 3 (1) (1980) 11–14.
[16] M. Cella, F. Gorter de Vries, D. Burger, M. Danhof, O. Della Pasqua, A model-based approach to dose selection in early pediatric development, Clin. Pharmacol. Ther. 87 (3) (2010) 294–302, https://doi.org/10.1038/clpt.2009.234.
[17] A.A. Vinks, C. Emoto, T. Fukuda, Modeling and simulation in pediatric drug therapy: application of pharmacometrics to define the right dose for children, Clin. Pharmacol. Ther. 98 (3) (2015) 298–308, https://doi.org/10.1002/cpt.169.
[18] F. Li, P. Nandy, S. Chien, G.J. Noel, C.W. Tornoe, Pharmacometrics-based dose selection of levofloxacin as a treatment for postexposure inhalational anthrax in children, Antimicrob. Agents Chemother. 54 (1) (2010) 375–379, https://doi.org/10.1128/AAC.00667-09.
[19] S. Swaminathan, J.G. Pasipanodya, G. Ramachandran, A.K. Hemanth Kumar, S. Srivastava, D. Deshpande, E. Nuermberger, T. Gumbo, Drug concentration thresholds predictive of therapy failure and death in children with tuberculosis: bread crumb trails in random forests, Clin. Infect. Dis. 63 (suppl 3) (2016) S63–S74, https://doi.org/10.1093/cid/ciw471.
[20] A.N. Edginton, B. Shah, M. Sevestre, J.D. Momper, The integration of allometry and virtual populations to predict clearance and clearance variability in pediatric populations over the age of 6 years, Clin. Pharmacokinet. 52 (8) (2013) 693–703, https://doi.org/10.1007/s40262-013-0065-6.
[21] L.E. Kelly, S. Ito, D. Woods, A.J. Nunn, C. Taketomo, M. de Hoog, M. Offringa, A comprehensive list of items to be included on a pediatric drug monograph, J. Pediatric Pharmacol. Therapeut. 22 (1) (2017) 48–59, https://doi.org/10.5863/1551-6776-22.1.48.
[22] W. Rödle, S. Wimmer, J. Zahn, H.-U. Prokosch, B. Hinkes, A. Neubert, W. Rascher, S. Kraus, D. Toddenroth, B. Sedlmayr, User-centered development of an online platform for drug dosing recommendations in pediatrics, Appl. Clin. Inform. 10 (4) (2019) 570–579, https://doi.org/10.1055/s-0039-1693714.
[23] M.J. Page, J.E. McKenzie, P.M. Bossuyt, I. Boutron, T.C. Hoffmann, C.D. Mulrow, L. Shamseer, J.M. Tetzlaff, E.A. Akl, S.E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M.M. Lalu, T. Li, E.W. Loder, E. Mayo-Wilson, S. McDonald, L.A. McGuinness, L.A. Stewart, J. Thomas, A.C. Tricco, V.A. Welch, P. Whiting, D. Moher, The PRISMA 2020 statement: an updated guideline for reporting systematic reviews, BMJ 372 (2021), n71, https://doi.org/10.1136/bmj.n71.
[24] F. Cabitza, A. Campagner, The need to separate the wheat from the chaff in medical informatics: Introducing a comprehensive checklist for the (self)-assessment of medical AI studies, Int. J. Med. Inf. 153 (2021), 104510, https://doi.org/10.1016/j.ijmedinf.2021.104510.
[25] O. Altay, M. Ulas, M. OZER, E. GENC. An Expert System to Predict Warfarin Dosage in Turkish Patients Depending on Genetic and Non-Genetic Factors. In: 2019 7th International Symposium on Digital Forensics and Security (ISDFS). IEEE; 62019, pp. 1–6.
[26] S. Alzubiedi, M.I. Saleh, Pharmacogenetic-guided Warfarin dosing algorithm in African-Americans, J. Cardiovasc. Pharmacol. 67 (1) (2016) 86–92, https://doi.org/10.1097/FJC.0000000000000317.
[27] N. Asai, C. Doi, K. Iwai, S. Ideno, H. Seki, J. Kato, T. Yamada, H. Morisaki, H. Shigeno. Proposal of Anesthetic Dose Prediction Model to Avoid Post-induction Hypotension Using Electronic Anesthesia Records. In: 2019 Twelfth International Conference on Mobile Computing and Ubiquitous Network (ICMU). IEEE; 112019, pp. 1–4.
[28] S.S. Chen, N.A. Zaborek, A.R. Doubleday, S.C. Schaefer, K.L. Long, S.C. Pitt, R. S. Sippel, D.F. Schneider, Optimizing levothyroxine dose adjustment after thyroidectomy with a decision tree, J. Surg. Res. 244 (2019) 102–106, https://doi.org/10.1016/j.jss.2019.06.025.
[29] E. Cosgun, N.A. Limdi, C.W. Duarte, High-dimensional pharmacogenetic prediction of a continuous trait using machine learning techniques with application to warfarin dose prediction in African Americans, Bioinformatics 27 (10) (2011) 1384–1389, https://doi.org/10.1093/bioinformatics/btr159.
[30] A. Coulet, N.H. Shah, M. Wack, M.B. Chawki, N. Jay, M. Dumontier, Predicting the need for a reduced drug dose, at first prescription, Sci. Rep. 8 (1) (2018) 15558, https://doi.org/10.1038/s41598-018-33980-0.
[31] J.M. Gonzalez-Cava, R. Arnay, J.A. Mendez-Perez, A. León, M. Martín, J.A. Reboso, E. Jove-Perez, J.L. Calvo-Rolle, Machine learning techniques for computer-based decision systems in the operating theatre: application to analgesia delivery, Logic J. IGPL 29 (2) (2021) 236–250, https://doi.org/10.1093/jigpal/jzaa049.
[32] E. Grossi, G.M. Podda, M. Pugliano, S. Gabba, A. Verri, G. Carpani, M. Buscema, G. Casazza, M. Cattaneo, Prediction of optimal warfarin maintenance dose using advanced artificial neural networks, Pharmacogenomics 15 (1) (2014) 29–37, https://doi.org/10.2217/pgs.13.212.
[33] Z.-C. Gu, S.-R. Huang, L. Dong, Q. Zhou, J. Wang, B. Fu, J. Chen, An adapted neural-fuzzy inference system model using preprocessed balance data to improve the predictive accuracy of warfarin maintenance dosing in patients after heart valve replacement, Cardiovasc. Drugs Ther. (2021), https://doi.org/10.1007/s10557-021-07191-1.
[34] P.-J.-H. Hu, T.-H. Cheng, C.-P. Wei, C.-H. Yu, A.L.F. Chan, H.-Y. Wang, Managing clinical use of high-alert drugs: a supervised learning approach to pharmacokinetic data analysis, IEEE Trans Syst., Man, Cybern a. 37 (4) (2007) 481–492, https://doi.org/10.1109/TSMCA.2007.897700.
[35] Y.-H. Hu, F. Wu, C.-L. Lo, C.-T. Tai, Predicting warfarin dosage from clinical data: a supervised learning approach, Artif. Intell. Med. 56 (1) (2012) 27–34, https://doi.org/10.1016/j.artmed.2012.04.001.
[36] Y.-H. Hu, C.-T. Tai, C.-F. Tsai, M.-W. Huang, Improvement of adequate digoxin dosage: an application of machine learning approach, J Healthc Eng. 2018 (2018) 3948245, https://doi.org/10.1155/2018/3948245.
[37] X. Huang, Z. Yu, X. Wei, J. Shi, Y. Wang, Z. Wang, J. Chen, S. Bu, L. Li, F. Gao, J. Zhang, A. Xu, Prediction of vancomycin dose on high-dimensional data using machine learning techniques, Expert. Rev. Clin. Pharmacol. (2021,) 1–11, https://doi.org/10.1080/17512433.2021.1911642.
[38] S. Imai, Y. Takekuma, T. Miyai, M. Sugawara, A new algorithm optimized for initial dose settings of vancomycin using machine learning, Biol. Pharm. Bull. 43 (1) (2020) 188–193, https://doi.org/10.1248/bpb.b19-00729.
[39] T.E. Klein, R.B. Altman, N. Eriksson, B.F. Gage, S.E. Kimmel, M.-T.-M. Lee, N. A. Limdi, D. Page, D.M. Roden, M.J. Wagner, M.D. Caldwell, J.A. Johnson, Estimation of the warfarin dose with clinical and pharmacogenetic data, N. Engl. J. Med. 360 (8) (2009) 753–764, https://doi.org/10.1056/NEJMoa0809329.
[40] A.E. Levy, M. Biswas, R. Weber, K. Tarakji, M. Chung, P.A. Noseworthy, C. Newton-Cheh, M.A. Rosenberg, Applications of machine learning in decision analysis for dose management for dofetilide, PLoS One 14 (12) (2019) e0227324.
[41] X. Li, R. Liu, Z.-Y. Luo, H. Yan, W.-H. Huang, J.-Y. Yin, X.-Y. Mao, X.-P. Chen, Z.-Q. Liu, H.-H. Zhou, W. Zhang, Comparison of the predictive abilities of pharmacogenetics-based warfarin dosing algorithms using seven mathematical models in Chinese patients, Pharmacogenomics 16 (6) (2015) 583–590, https://doi.org/10.2217/pgs.15.26.
[42] D. Li, J. Gao, N. Hong, H. Wang, L. Su, C. Liu, J. He, H. Jiang, Q. Wang, Y. Long, W. Zhu, A clinical prediction model to predict heparin treatment outcomes and provide dosage recommendations: development and validation study, J. Med. Internet Res. 23 (5) (2021) e27118.
[43] K.E. Liu, C.-L. Lo, Y.-H. Hu, Improvement of adequate use of warfarin for the elderly using decision tree-based approaches, Methods Inf. Med. 53 (1) (2014) 47–53, https://doi.org/10.3414/ME13-01-0027.
[44] Z. Ma, P. Wang, Z. Gao, R. Wang, K. Khalighi, Ensemble of machine learning algorithms using the stacked generalization approach to estimate the warfarin dose, PLoS One 13 (10) (2018) e0205872.
[45] R. Liu, X. Li, W. Zhang, H.-H. Zhou, Comparison of nine statistical model based warfarin pharmacogenetic dosing algorithms using the racially diverse international warfarin pharmacogenetic consortium cohort database, PLoS One 10 (8) (2015) e0135784.
[46] S. Min, T. Papaz, A.N. Lambert, U. Allen, P. Birk, T. Blydt-Hansen, B.J. Foster, H. Grasemann, L. Hamiwka, C. Litalien, V. Ng, N. Berka, P. Campbell, D. Claude, C. L. Saw, K. Tinckham, S. Urschel, S.L. van Driest, R. Parekh, S. Mital, An integrated clinical and genetic prediction model for tacrolimus levels in pediatric solid organ transplant recipients, Transplantation (2021), https://doi.org/10.1097/TP.0000000000003700.
[47] A.E. Olesen, D. Grønlund, M. Gram, F. Skorpen, A.M. Drewes, P. Klepstad, Prediction of opioid dose in cancer pain patients using genetic profiling: not yet an option with support vector machine learning, BMC Res. Notes 11 (1) (2018) 78, https://doi.org/10.1186/s13104-018-3194-z.
[48] A. Roche-Lima, A. Roman-Santiago, R. Feliu-Maldonado, J. Rodriguez-Maldonado, B.G. Nieves-Rodriguez, K. Carrasquillo-Carrion, C.M. Ramos, I. Da Luz Sant'Ana, S. E. Massey, J. Duconge, Machine learning algorithm for predicting warfarin dose in caribbean hispanics using pharmacogenetic data, Front. Pharmacol. 10 (2019) 1550, https://doi.org/10.3389/fphar.2019.01550.
[49] H. Schelleman, J. Chen, Z. Chen, J. Christie, C.W. Newcomb, C.M. Brensinger, M. Price, A.S. Whitehead, C. Kealey, C.F. Thorn, F.F. Samaha, S.E. Kimmel, Dosing algorithms to predict warfarin maintenance dose in Caucasians and African Americans, Clin. Pharmacol. Ther. 84 (3) (2008) 332–339, https://doi.org/10.1038/clpt.2008.101.
[50] A. Sharabiani, H. Darabi, A. Bress, L. Cavallari, E. Nutescu, K. Drozda. Machine learning based prediction of warfarin optimal dosing for African American patients. In: 2013 IEEE International Conference on Automation Science and Engineering (CASE). IEEE; 082013, pp. 623–628.
[51] A. Sharabiani, A. Bress, E. Douzali, H. Darabi, Revisiting Warfarin dosing using machine learning techniques, Comput. Math. Methods Med. 2015 (2015), 560108, https://doi.org/10.1155/2015/560108.

[52] L. Su, C. Liu, D. Li, J. He, F. Zheng, H. Jiang, H. Wang, M. Gong, N. Hong, W. Zhu, Y. Long, Toward optimal heparin dosing by comparing multiple machine learning methods: retrospective study, JMIR Med. Inform. 8 (6) (2020) e17648.
[53] J. Tang, R. Liu, Y.-L. Zhang, M.-Z. Liu, Y.-F. Hu, M.-J. Shao, L.-J. Zhu, H.-W. Xin, G.-W. Feng, W.-J. Shang, X.-G. Meng, L.-R. Zhang, Y.-Z. Ming, W. Zhang, Application of machine-learning models to predict tacrolimus stable dose in renal transplant recipients, Sci. Rep. 7 (2017) 42192, https://doi.org/10.1038/srep42192.
[54] Y. Tao, Y. Zhang. "WarfarinSeer": a predictive tool based on SMOTE-random forest to improve warfarin dose prediction in Chinese patients. In: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE; 122018, pp. 1022–1026.
[55] Y. Tao, Y.J. Chen, L. Xue, C. Xie, B. Jiang, Y. Zhang, An ensemble model with clustering assumption for warfarin dose prediction in Chinese patients, IEEE J. Biomed. Health Inform. 23 (6) (2019) 2642–2654, https://doi.org/10.1109/JBHI.2019.2891164.
[56] G. Truda, P. Marais, Evaluating warfarin dosing models on multiple datasets with a novel software framework and evolutionary optimisation, J. Biomed. Inform. 113 (2021), 103634, https://doi.org/10.1016/j.jbi.2020.103634.
[57] N.S. Tyler, C.M. Mosquera-Lopez, L.M. Wilson, R.H. Dodier, D.L. Branigan, V. B. Gabo, F.H. Guillot, W.W. Hilts, J. El Youssef, J.R. Castle, P.G. Jacobs, An artificial intelligence decision support system for the management of type 1 diabetes, Nat Metab. 2 (7) (2020) 612–619, https://doi.org/10.1038/s42255-020-0212-y.
[58] L. van Nguyen, H.D. Nguyen, Y.-S. Cho, H.-S. Kim, I.-Y. Han, D.-K. Kim, S. Ahn, J.-G. Shin, Comparison of multivariate linear regression and a machine learning algorithm developed for prediction of precision warfarin dosing in a Korean population, J. Thromb. Haemost. (2021), https://doi.org/10.1111/jth.15318.
[59] N.A. Zaborek, A. Cheng, J.R. Imbus, K.L. Long, S.C. Pitt, R.S. Sippel, D.F. Schneider, The optimal dosing scheme for levothyroxine after thyroidectomy: a comprehensive comparison and evaluation, Surgery 165 (1) (2019) 92–98, https://doi.org/10.1016/j.surg.2018.04.097.
[60] X. Zhu, W. Huang, H. Lu, Z. Wang, X. Ni, J. Hu, S. Deng, Y. Tan, L. Li, M. Zhang, C. Qiu, Y. Luo, H. Chen, S. Huang, T. Xiao, D. Shang, Y. Wen, A machine learning approach to personalized dose adjustment of lamotrigine using noninvasive clinical parameters, Sci. Rep. 11 (1) (2021) 5568, https://doi.org/10.1038/s41598-021-85157-x.
[61] R.P. Owen, R.B. Altman, T.E. Klein, PharmGKB and the International Warfarin Pharmacogenetics Consortium: the changing role for pharmacogenomic databases and single-drug pharmacogenetics, Hum Mutat. 29 (4) (2008) 456–460, https://doi.org/10.1002/humu.20731.
[62] A.E.W. Johnson, T.J. Pollard, L. Shen, L.-W.-H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L.A. Celi, R.G. Mark, MIMIC-III, a freely accessible critical care database, Sci. Data 3 (2016), 160035, https://doi.org/10.1038/sdata.2016.35.
[63] M.N. Jeffres, The whole price of vancomycin: toxicities, troughs, and time, Drugs 77 (11) (2017) 1143–1154, https://doi.org/10.1007/s40265-017-0764-7.
[64] M.N. Neely, G. Youn, B. Jones, R.W. Jelliffe, G.L. Drusano, K.A. Rodvold, T. P. Lodise, Are vancomycin trough concentrations adequate for optimal dosing? Antimicrob. Agents Chemother. 58 (1) (2014) 309–316, https://doi.org/10.1128/AAC.01653-13.
[65] S.E. Kimmel, B. French, S.E. Kasner, J.A. Johnson, J.L. Anderson, B.F. Gage, Y. D. Rosenberg, C.S. Eby, R.A. Madigan, R.B. McBane, S.Z. Abdel-Rahman, S. M. Stevens, S. Yale, E.R. Mohler, M.C. Fang, V. Shah, R.B. Horenstein, N.A. Limdi, J.A.S. Muldowney, J. Gujral, P. Delafontaine, R.J. Desnick, T.L. Ortel, H.H. Billett, R.C. Pendleton, N.L. Geller, J.L. Halperin, S.Z. Goldhaber, M.D. Caldwell, R. M. Califf, J.H. Ellenberg, A pharmacogenetic versus a clinical algorithm for warfarin dosing, N. Engl. J. Med. 369 (24) (2013) 2283–2293, https://doi.org/10.1056/NEJMoa1310669.
[66] M. Pirmohamed, G. Burnside, N. Eriksson, A.L. Jorgensen, C.H. Toh, T. Nicholson, P. Kesteven, C. Christersson, B. Wahlström, C. Stafberg, J.E. Zhang, J.B. Leathart, H. Kohnke, M.-V. der Zee, H. Anke, P.R. Williamson, A.K. Daly, P. Avery, F. Kamali, M. Wadelius, A randomized trial of genotype-guided dosing of warfarin, N. Engl. J. Med. 369 (24) (2013) 2294–2303, https://doi.org/10.1056/NEJMoa1311386.
[67] T.I. Verhoef, G. Ragia, A. de Boer, R. Barallon, G. Kolovou, V. Kolovou, S. Konstantinides, S. Le Cessie, E. Maltezos, F.J.M. van der Meer, W.K. Redekop, M. Remkes, F.R. Rosendaal, R.M.F. van Schie, A. Tavridou, D. Tziakas, M. Wadelius, V.G. Manolopoulos, M.-V. der Zee, H. Anke, A randomized trial of genotype-guided dosing of acenocoumarol and phenprocoumon, N. Engl. J. Med. 369 (24) (2013) 2304–2312, https://doi.org/10.1056/NEJMoa1311388.
[68] B. Furie, Do pharmacogenetics have a role in the dosing of vitamin K antagonists? N. Engl. J. Med. 369 (24) (2013) 2345–2346, https://doi.org/10.1056/NEJMe1313682.
[69] P.B. Jensen, L.J. Jensen, S. Brunak, Mining electronic health records: towards better research applications and clinical care, Nat. Rev. Genet. 13 (6) (2012) 395–405, https://doi.org/10.1038/nrg3208.
[70] A. Gottlieb, G.Y. Stein, E. Ruppin, R.B. Altman, R. Sharan, A method for inferring medical diagnoses from patient similarities, BMC Med. 11 (2013) 194, https://doi.org/10.1186/1741-7015-11-194.
[71] L. Shahmoradi, A. Darrudi, G. Arji, A. Farzaneh Nejad, Electronic health record implementation: a SWOT analysis, Acta Med. Iran 55 (10) (2017) 642–649.
[72] M.R. Cowie, J.I. Blomster, L.H. Curtis, S. Duclaux, I. Ford, F. Fritz, S. Goldman, S. Janmohamed, J. Kreuzer, M. Leenay, A. Michel, S. Ong, J.P. Pell, M. R. Southworth, W.G. Stough, M. Thoenes, F. Zannad, A. Zalewski, Electronic health records to facilitate clinical research, Clin. Res. Cardiol. 106 (1) (2017) 1–9, https://doi.org/10.1007/s00392-016-1025-6.
[73] L. Cerruti, N. Bleyzac, M. Tod, Semi-mechanistic model for predicting the dosing rate in children and neonates for drugs mainly eliminated by cytochrome metabolism, Clin. Pharmacokinet. 57 (7) (2018) 831–841, https://doi.org/10.1007/s40262-017-0596-3.
[74] B.J. Anderson, P. Larsson, A maturation model for midazolam clearance, Paediatr Anaesth. 21 (3) (2011) 302–308, https://doi.org/10.1111/j.1460-9592.2010.03364.x.
[75] T.N. Johnson, A. Rostami-Hodjegan, G.T. Tucker, Prediction of the clearance of eleven drugs and associated variability in neonates, infants and children, Clin. Pharmacokinet. 45 (9) (2006) 931–956, https://doi.org/10.2165/00003088-200645090-00005.
[76] I.E. Templeton, N.S. Jones, L. Musib, Pediatric dose selection and utility of PBPK in determining dose, AAPS J. 20 (2) (2018) 31, https://doi.org/10.1208/s12248-018-0187-8.
[77] M. Rashid, M. Sarfraz, M. Arafat, A. Hussain, N. Abbas, M.W. Sadiq, M.F. Rasool, N. I. Bukhari, Prediction of lisinopril pediatric dose from the reference adult dose by employing a physiologically based pharmacokinetic model, BMC Pharmacol. Toxicol. 21 (1) (2020) 56, https://doi.org/10.1186/s40360-020-00429-y.
[78] E. Manolis, G. Pons, Proposals for model-based paediatric medicinal development within the current European Union regulatory framework, Br. J. Clin. Pharmacol. 68 (4) (2009) 493–501, https://doi.org/10.1111/j.1365-2125.2009.03484.x.
[79] S. Ekins, J.S. Freundlich, R.C. Reynolds, Are bigger data sets better for machine learning? Fusing single-point and dual-event dose response data for Mycobacterium tuberculosis, J. Chem. Inf. Model. 54 (7) (2014) 2157–2165, https://doi.org/10.1021/ci500264r.
[80] N. Kureshi, S.S.R. Abidi, C. Blouin, A predictive model for personalized therapeutic interventions in non-small cell lung cancer, IEEE J. Biomed. Health Inform. 20 (1) (2016) 424–431, https://doi.org/10.1109/JBHI.2014.2377517.
[81] T. Dietterich, Overfitting and Undercomputing in Machine Learning, Computing Surveys. 27 (1995) 326–327.
[82] I.H. Sarker, A.S.M. Kayes, P. Watters, Effectiveness analysis of machine learning classification models for predicting personalized context-aware smartphone usage, J Big Data. 6 (1) (2019), https://doi.org/10.1186/s40537-019-0219-y.