

Article

# Machine Learning in High-Alert Medication Treatment: A Study on the Cardiovascular Drug

Chun-Tien Tai <sup>1,2,3</sup>, Kuen-Liang Sue <sup>4</sup> and Ya-Han Hu <sup>4,5,\*</sup> 

<sup>1</sup> Department of Information Management and Institute of Healthcare Information Management, National Chung Cheng University, Chiayi 621, Taiwan; flamquit@cgmh.org.tw

<sup>2</sup> Chiayi Chang Gung Memorial Hospital, Chiayi 613, Taiwan

<sup>3</sup> Department of Information Management, Shu Zen College of Medicine and Management, Kaohsiung 821, Taiwan

<sup>4</sup> Department of Information Management, National Central University, Taoyuan 320, Taiwan; klsue@mgt.ncu.edu.tw

<sup>5</sup> MOST AI Biomedical Research Center, National Cheng Kung University, Tainan 701, Taiwan

\* Correspondence: yhhu@mgt.ncu.edu.tw

Received: 6 July 2020; Accepted: 14 August 2020; Published: 21 August 2020



**Abstract:** The safety of high-alert medication treatment is still a challenge all over the world. Approximately one-half of adverse drug events (ADEs) are related to high-alert medications, which motivates us to improve the predicament faced in clinical practice. The purpose of this study is to use machine-learning techniques to predict the risk of high-alert medication treatment. Taking the cardiovascular drug digoxin as an example, we collected the records of 513 patients who received the pertinent therapy during hospitalization at a tertiary medical center in Taiwan. Considering serum digoxin concentration (SDC) is the primary indicator for assessing the risk of digoxin therapy, patients with SDC being controlled at the recommended range before their discharge were defined as a low-risk population; otherwise, patients were defined as the high-risk population. Weka 3.9.4—an open source machine learning software—was adopted to develop binary classification models to predict the risk of digoxin therapy by a number of machine-learning techniques, including k-nearest neighbors (kNN), decision tree (C4.5), support vector machine (SVM), random forest (RF), artificial neural network (ANN) and logistic regression (LGR). The results showed that the performance of RF was the best, followed by C4.5 and ANN; the remaining classifiers performed poorly. This study confirmed that machine-learning techniques can yield favorable prediction effectiveness for high-alert medication treatment, thereby decreasing the risk of ADEs and improving medication safety.

**Keywords:** machine learning; high-alert medication; cardiovascular drug; adverse drug events; decision tree

## 1. Introduction

Since *To Err is Human* published by the Institute of Medicine (IOM) in 1999, patient safety has become a global concern [1]. Although governments, nonprofit organizations and medical institutions have proposed various measures and invested much resources to improve patient safety in the past 20 years, incidents of patients being injured due to improper medical care continue to occur every day around the world [2]. Therefore, the avoidance of medical injuries to maintain patient safety is still a difficult problem to solve [3].

The World Health Organization (WHO) and the World Alliance for Patient Safety have jointly proposed two Global Patient Safety Challenges, which provide improvement measures and methods for patient safety issues to reduce the possible harm to patients during medical care. In 2017, the WHO

released the third global patient safety challenge which indicates that if medication errors occur in the medical system, it may cause serious harm to patients [4].

In the United States, medication errors cause at least one death each day and injure nearly 1.3 million people each year [4]. It is estimated that the additional medical costs worldwide caused by medication errors amount to US \$42 billion per year, accounting for 0.7% of global healthcare expenditure [5].

Compared with ordinary drugs, high-alert medications carry greater risks of adverse drug events (ADEs) and have received widespread attention [6]. The definition of high-alert medications is drugs that bear a heightened risk of causing significant patient harm when they are used in error [7]. Therefore, although high-alert medications account for only a small portion of all therapeutic drugs, they are the main cause of adverse drug events (ADEs). Recent studies have shown that up to 50% of ADEs are related to high-alert medications [8].

Medication errors will not only lead to an increase in medical costs, but also cause varying degrees of harm to patients, especially high-alert medications. To date, clinical decision support systems (CDSS) have been widely and maturely applied to most drugs. However, the current application of CDSS in high-alert medications is mainly based on the recommendations of the Clinical Practice Guideline (CPG) to establish prompts in the system, rather than generating predictions of treatment results. Therefore, some scholars have begun to use machine-learning techniques to conduct research on specific high-alert medications [9,10]. The purpose is to provide better clinical decision support to enhance the safety of high-alert medication treatment.

Cardiovascular disease (CVD) is currently the main cause of death in the world [11]. According to WHO statistics, approximately 17.65 million people worldwide died of CVD in 2016, accounting for 31% of all deaths, indicating that approximately one in three people died of CVD. The number of people who have died of CVD is not only twice as high as that of cancer, but more than the sum of all deaths from infectious diseases [12]. Similarly, CVD is also the most common cause of death in Europe. Although the mortality rate of CVD has gradually decreased, more than four million Europeans die each year.

Because of the high mortality rate of CVD, the treatment of this disease has received considerable attention. In terms of medication therapy, cardiovascular drugs can effectively control the symptoms of CVD, but such drugs are also prone to severe ADEs. Levy et al. pointed out that nearly half of ADEs are related to the use of cardiovascular drugs [13].

Among all cardiovascular drugs, digoxin is classified as a high-alert medication. Because of the complex pharmacological properties, the use of digoxin is more likely to cause serious injury to patients. In the treatment process of CVD with digoxin, inappropriate decisions may affect treatment outcome, cause drug toxic reactions and severely lead to death. When physicians decide to prescribe high-alert medications, more careful decision-making is needed. However, physicians need to interpret a large amount of information within a limited time, including the patient's physiology, disease status and laboratory data to determine the treatment regimen. The risk of medication therapy needs to be assessed during decision-making process to reduce the incidence of ADEs. For a busy medical environment, these goals are not easy to achieve.

In the past, the applications of machine-learning techniques in medical domain covered a wide range of ordinary drugs and only a few studies focused on the high-alert medications [14–28]. On the other hand, those related studies used relatively few variables or adopted techniques lacking interpretability. We intend to apply machine-learning techniques, including  $k$ -nearest neighbors (kNN), decision tree (C4.5), support vector machine (SVM), random forest (RF), artificial neural network (ANN) and logistic regression (LGR), to build binary classification models and enhance the safety of high-alert medication treatment.

In order to improve the predicament faced in practice, this study attempts to focus on high-alert medication treatment and investigate cardiovascular drugs that have attracted much attention worldwide. We will take digoxin as an example, retrieve the data needed for research from the medical

record database of a tertiary medical center in Taiwan and then adopt the method of machine learning to construct a risk prediction model. It is hoped that the results will help physicians to make better clinical decisions and thus improve patient medication safety.

## 2. Materials and Methods

### 2.1. Data Source

The source of research data were received from a patient records database of a tertiary medical center in Taiwan. Patients who had accepted digoxin therapy during hospitalization between January 2004 and December 2013 were included in this study. Exclusion criteria were that the patient had taken digoxin within 3 months before hospitalization. Because the half-life of the drug is 36 to 48 h with normal renal function, a 3 month clearance period was applied to ensure the accuracy of the prediction models [29].

The experimental data for this study were collected from 513 clinical cases. According to the variables mentioned in the previous literature, each clinical record included initial dose of digoxin, demographic data and laboratory data related to liver function, renal function and serum K<sup>+</sup>. In addition, a specific disease congestive heart failure (CHF) was also included.

Finally, we considered another important variable known as drug-to-drug interaction (DDI). The variable cannot only affect treatment outcome, but also increase the incidence of ADEs. For this reason, we included DDI in our experiment before constructing the prediction models. According to Thomson Micromedex Database [30], the clinical significance of DDI can be divided into three degrees (major, moderate and minor), the major degree being the most severe DDI. In addition to drug interaction facts, we compared DDI information published by the Ministry of Health and Welfare in Taiwan with the drugs used at the case hospital. Table S1 shows the drugs that induce the major degree DDI when combined with digoxin. During the treatment of digoxin, patients who used any of the drugs listed in Table S1 will be considered to be at risk of DDI.

### 2.2. Variable

The dependent variable (DV) of this study was the risk of digoxin treatment, which can be determined by serum digoxin concentration (SDC). In clinical practice, physicians judge the therapeutic effect of digoxin mainly based on SDC. To eliminate the ADEs rates of digoxin, SDC value is recommended at the range of 0.5–0.9 ng/mL [28,31]. Therefore, patients with SDC being controlled at recommended range before their discharge were defined as a low-risk population; otherwise, patients were defined as the high-risk population. Moreover, a total of 12 independent variables (IVs) were considered in our study (Table 1).

**Table 1.** Definition of variables.

Category	Name	Definition	Type	
IV	Demographic	Gender	Male/female	Categorical
		Age	Actual age on the day of hospitalization	Numeric
		Weight	Body weight measured on the day of hospitalization	Numeric
	Laboratory	ALT *	The record measured before digoxin therapy	Numeric
		AST *	The record measured before digoxin therapy	Numeric
		SCr *	The record measured before digoxin therapy	Numeric
		BUN *	The record measured before digoxin therapy	Numeric
		ALB *	The record measured before digoxin therapy	Numeric
		K <sup>+</sup> *	The record measured before digoxin therapy	Numeric
	History	CHF	The diagnostic records of CHF noted in the past year	Categorical
	Medication	Dose	Initial dose of digoxin during hospitalization	Numeric
		DDI	Patients who used any of the drugs listed in Table S1 during hospitalization.	Categorical
DV	Treatment outcome	Risk	0.5 ≤ SDC ≤ 0.9 ng/mL: low risk SDC < 0.5 or SDC > 0.9 ng/mL: high risk	Categorical

\* ALB—albumin; ALT/AST—alanine aminotransferase/aspartate aminotransferase; BUN—blood urea nitrogen; K<sup>+</sup>—serum potassium; SCr—serum creatinine.

### 2.3. Investigated Classification Techniques

Six well-known classification techniques were considered, including  $k$ NN, C4.5, SVM, RF, ANN and LGR. Table 2 lists the primary applications of these techniques in previous studies in medical domain [14–28].

**Table 2.** Applications of classification techniques in medical domain.

Techniques	Applications
$k$ NN	Modeling liver-related adverse effects of drugs [14]. Predicting drug side effects [15]. Personalized drug concentration predictions [16].
C4.5	Detecting adverse drug events [17]. The analysis of adverse drug reactions [18]. Expert system for adverse drug reaction detection [19]. Antidiabetic drug failure prediction [20].
SVM	Prediction models for drug-induced hepatotoxicity [21]. Prediction of drug-induced eosinophilia adverse effect [22]. Prediction of cardiovascular outcomes [23].
RF	Adverse drug reaction detection [19]. Prediction of drug-induced nephrotoxicity [24].
ANN	Predicting the serum digoxin concentrations of infants [25]. Prediction of the serious adverse drug reactions [26].
LGR	Clinical decision support system to toxicity in patients treated with digoxin [27]. Analysis of patients with digoxin serum concentration [28].

$k$ -nearest neighbors ( $k$ NN) is a well-known case learning method that classifies new cases based on distance functions. For a test case,  $k$ NN first finds its top  $k$ -nearest neighbors in the vector space; the class label of the test case is assigned to the majority class among its neighbors [32]. The Euclidean distance and the Hamming distance are usually used for numeric and categorical variables, respectively.

Decision tree (C4.5) is a commonly used classification technique. The growth of the decision tree is a recursive process. First, an attribute with the best gain ratio is selected as the split node of the tree, and the corresponding branches are generated according to its attribute values. Then, the dataset can be partitioned into subsets, and each subsets can be used for next round of tree partition. The decision tree grows until the termination condition is met. The rules generated by the decision tree are easy to understand and can be applied in practice [33]. Because gain ratio is proposed to measure the improvement of data purity in C4.5, it is also used to measure the importance of IVs [34].

Support vector machine (SVM) is a supervised learning method that maps IVs and a DV to a high-dimensional vector space by using structural risk minimization. Then, it attempts to find a hyperplane to separate the data into two categories. Therefore, a new test case can be mapped into one subspace based on the determined hyperplane [35].

Random forest (RF) is a tree-based ensemble learning algorithm, which adopts bagging and random feature selection techniques [36]. Each decision tree is constructed from bootstrap samples of training data, and the final predicted result is determined by a majority vote. When building multiple trees, a random sample with replacement and the random selection of feature of the training set are used, which reduce the possibility of overfitting in tree growing process.

Artificial neural network (ANN) consists of an input- and an output- layer with a number of hidden layers. Artificial neurons between two adjacent layers are connected; the neurons receive the inputs from the outputs of the front adjacent neurons and convert them into an output value through the transfer functions. The key component in the learning procedure is the connection weights between two neurons. In training phase, it iteratively tunes the connection weights between neurons for minimizing the difference between the estimated output value and the actual output. The learning process repeatedly executes the feedforward and backward phases until any stop criteria is reached [37].

Logistic regression (LGR) is a statistical learning method that uses the dataset to develop a mathematical model for predicting the DV. LGR is a nonlinear regression model that can be applied when the DV is binomial category data; it can find the relationships between the binomial DV and a group of IVs [38]. Specifically, it puts the values of IVs into the logistic function to get a probability value.

#### 2.4. Experimental Setup and Performance Measure

To construct the risk prediction models for digoxin treatment, this study adopted Weka 3.9.4 ([www.cs.waikato.ac.nz/ml/weka](http://www.cs.waikato.ac.nz/ml/weka)), an open source machine learning software to investigate the performance of the selected techniques, including IBk (kNN in Weka), J48 (C4.5 in Weka), SMO (SVM in Weka), random forest (RF in Weka), multilayer perceptron (ANN in Weka), simple logistic (LGR in Weka). To optimize the prediction performance of the models, we used the CV parameter selection meta-learner module implemented in Weka to search the best parameter settings. Table 3 lists the tests of parameter tuning for each classification technique and the final best parameter settings found in this study.

**Table 3.** Parameter settings in Weka.

Technique	WEKA Module	Parameters	Test Range	Best Parameter Setting
kNN	IBk	Number of neighbors	1–10	1
C4.5	J48	Confidence factor	0.1–0.3	0.3
		Minimum number of instances per leaf	2–10	2
SVM	SMO	Kernel	Polykernel, RBF Kernel	Polykernel
RF	Random forest	Number of trees	10–100	60
		Number of attributes to be used in random selection	2–10	10
		Number of hidden nodes	3–10	9
ANN	Multilayer perceptron	Learning rate	0.1–0.5	0.3
		Momentum factor	0–0.5	0
		Maximum number of epochs	300–1000	1000
LGR	Simple logistic	None		N/A

The classification performance of the prediction models was compared to find out the optimal classifier. We split the data into training and testing datasets in a proportion of 70:30 and applied 10-fold cross-validation method in the training dataset to search the best parameter settings. In addition, previous research indicated that the class imbalance problem can significantly affect learning performance [39]. To improve the classification performance, the resample module (spread subsample module in Weka) is used in this study to modify the distribution of instances of two classes in the training set. Finally, we used the best parameter setting to construct the binary classification models and reported the classification results of testing set.

By using a confusion matrix (Table 4), the precision, recall and F-measure were adopted to analysis the performance of each prediction models. These three metrics were defined by the following Formulas:

$$\text{Precision} = TP/(TP + FP), \quad (1)$$

$$\text{Recall} = TP/(TP + FN), \quad (2)$$

$$F\text{-measure} = 2 \cdot \text{Precision} \cdot \text{Recall}/(\text{precision} + \text{recall}) \quad (3)$$

**Table 4.** Confusion matrix.

Classes		Predicted Class	
		High Risk	Low Risk
Actual Class	High risk	TP	FN
	Low risk	FP	TN

Precision is the ratio of correctly predicted high risk patients to the total predicted high risk patients. Recall is the ratio of correctly predicted high risk patients to all of the high risk patients

in the dataset. F-measure is the harmonic mean of precision and recall. In addition, our study also included the area under the curve (AUC) as a performance indicator. To evaluate the performance of the models, we referred to the guidelines proposed by Hosmer and Lemeshow [38]. The performance is defined as “outstanding” if  $AUC \geq 0.9$ , “excellent” if  $0.9 > AUC \geq 0.8$ , “acceptable” if  $0.8 > AUC \geq 0.7$ , “fair to poor” if  $0.7 > AUC \geq 0.5$ .

### 3. Results

#### 3.1. Descriptive Statistics

According to the aforementioned definitions of variables, our research collected and preprocessed the medical records of all patients. Finally, 513 cases were enrolled in this study and the statistical information is shown in Table 5.

**Table 5.** Statistical information for the collected cases.

Category	Name	Range	Unit	Summary Statistics	
IV	Demographic	Gender	Male/female	-	Male: 278 (54.19%) Female: 235 (45.81%)
		Age	19 to 101	years	$\mu = 74.03, \sigma = 12.08$
		Weight	32 to 105	kg	$\mu = 56.26, \sigma = 12.01$
	Laboratory	ALT	5 to 1754	U/L	$\mu = 78.33, \sigma = 161.58$
		AST	11 to 2615	U/L	$\mu = 88.25, \sigma = 215.26$
		SCr	0.27 to 12.1	mg/dL	$\mu = 1.466, \sigma = 1.387$
		BUN	2 to 237	mg/dL	$\mu = 34.33, \sigma = 28.53$
		ALB	1 to 4.3	g/dL	$\mu = 2.625, \sigma = 0.639$
		K+	1.9 to 6.7	meq/L	$\mu = 3.987, \sigma = 0.726$
		History	CHF	Yes/no	-
Medication	Dose	0.125 to 2	mg	$\mu = 0.504, \sigma = 0.232$	
	DDI	Yes/no	-	Yes: 364 (70.96%) No: 149 (29.04%)	
DV	Treatment outcome	Risk	Low/high	-	Low: 169 (32.94%) High: 344 (67.06%)

#### 3.2. Experimental Results

In experimental evaluations, six binary classification models were developed to identify high risk patients. The experimental results are reported in Table 6. AUC, F-measure, precision, and recall were used to assess the performance of the prediction models.

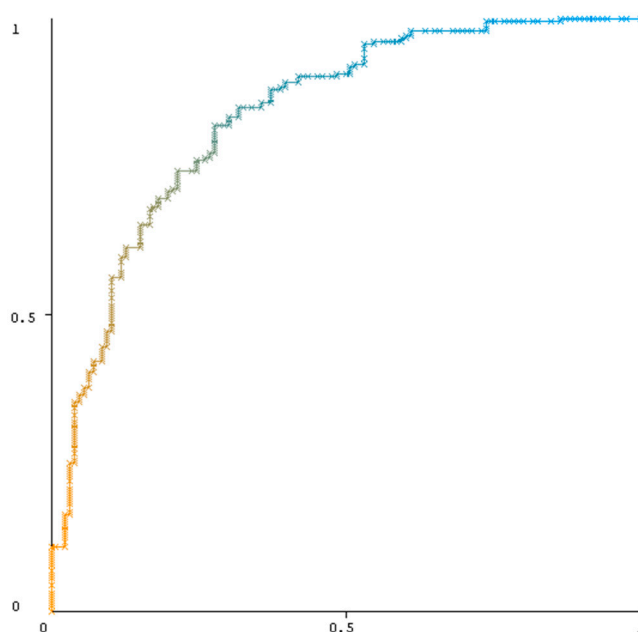
**Table 6.** Results of classification performance.

Classifier	AUC	Precision	Recall	F-Measure
kNN	0.569	0.722	0.691	0.706
C4.5	0.719	0.803	0.734	0.767
SVM	0.554	0.698	0.536	0.607
RF	0.836	0.852	0.742	0.794
ANN	0.688	0.795	0.682	0.734
LGR	0.551	0.689	0.532	0.600

Regarding the AUC values ranged from 0.551–0.836. The results showed that the performance of RF classifier was the best (0.836; excellent discrimination), followed by C4.5 (0.719) and ANN (0.688); the remaining classifiers performed poorly. Figure 1 illustrates the receiver operating characteristic (ROC) curve of RF classifier. In addition, F-measure is the harmonic mean of precision and recall, so it is commonly adopted to compare the whole performance. Regarding the F-measure of the prediction models, the RF result was still the best (0.794), followed by C4.5 (0.767) and ANN (0.734); the other



classifiers performed poorly. Overall, the decision tree-based methods including RF and C4.5 exhibited markedly superior prediction performance in our study.



**Figure 1.** The receiver operating characteristic (ROC) curve of RF classifier.

To understand the performance of the model in more specific realistic scenarios, we further adopted the best classifier RF to investigate the changes when controlling for specific variables, including gender, CHF and DDI. The experimental results are shown in Table 7. The highest AUC value was 0.848 in the female and patients with DDI datasets; the model for patients without DDI was the only one that yielded the result below 0.8.

**Table 7.** Results of classification performance in specific scenarios.

Dataset	AUC	Precision	Recall	F-measure
Male	0.834	0.875	0.824	0.848
Female	0.848	0.901	0.753	0.820
With CHF *	0.804	0.845	0.803	0.824
Without CHF *	0.809	0.866	0.640	0.736
With DDI *	0.848	0.840	0.794	0.816
Without DDI *	0.750	0.793	0.676	0.730

\* CHF—congestive heart failure; DDI—drug-to-drug interaction.

#### 4. Discussion

In recent years, the patient safety of high-alert medication treatment is an urgent issue, especially cardiovascular drugs [1,3]. Previous studies have employed various methods, such as statistical models and pharmacokinetics, to enhance the safety of using high-alert medications. In addition, machine-learning techniques (e.g., ANN) have lately been used to enhance the predictive power of models. However, because the decision-making process of ANN cannot be known and lack of explanatory power, its use in clinical applications is still limited.

Our study investigated various well-known machine-learning techniques. The results indicated that, compared with other techniques (kNN, LGR and SVM), the average performance of decision tree-based methods was superior to that of other techniques. In addition, decision tree-based methods yielded the information of decision rules that can support physicians in making treatment decision of high-alert medications.

In addition to comparing the prediction capabilities of different classification techniques, we further analyzed the influence of each variable on the occurrence of ADEs to provide a reference for clinicians. By computing gain ratios, we disclosed that the most important variables affecting the outcome of the treatment are shown in Table 8. After discussion with the medical specialists, including 2 cardiologists, 1 clinical nurse specialists and 3 clinical pharmacists, the validity of the results is discussed as follows:

**Table 8.** Ranking of the selected variables.

Rank	Variables	Gain Ratios
1	BUN *	0.0182
2	SCr *	0.0151
3	DDI *	0.0103
4	Dose	0.0085

\* BUN—blood urea nitrogen; SCr—serum creatinine; DDI—drug-to-drug interaction.

First, BUN and SCr are commonly used indicators of kidney function. A raised BUN/SCr ratio will increase the hospitalization risk and mortality of CHF patients [40]. Doherty et al. found that the higher the BUN value, the more digoxin is accumulated in the body. Because digoxin is primarily excreted by kidneys, low kidney function may cause the rise of SDC and lead to ADEs such as digoxin toxicity [41]. Our study confirmed that kidney function is highly correlated with the risk of ADEs. This finding is similar with the outcomes of previous studies about digoxin treatment [42,43]. Therefore, physicians need to adjust the dose when decided to prescribe digoxin to patients with renal insufficiency.

Furthermore, we also found that DDI is a significant factor of treatment outcome in high-alert medications. DDI has long received considerable attention in the field of medicine, especially involving high-alert medications, which has a complex pharmacological mechanism [16,17,44–49]. The digoxin related DDIs can easily cause toxicity reaction and ADEs [50]. Although our study only considered drugs that will produce major degree DDI with digoxin, these results are still adequate to confirm the significance of DDI.

Finally, the adequacy of the initial dose is important in high-alert medication therapy. Inappropriate doses resulting in high plasma drug concentrations can lead to severe ADEs that influence the treatment outcome, particularly digoxin [51]. Therefore, previous studies have focused on novel approaches to predict the adequate dose of digoxin [44–49,51]. In practice, the initial dose is considerably different from the maintenance dose, which may increase the risk of ADEs before reaching a stable maintenance dose. In this study, we investigated the initial dose instead of the maintenance dose and demonstrated the influence of initial dose on the treatment outcome.

To further clarify the dependencies between the top four important IVs and the predicted binary outcome, four feature sets were formulated in this study: BUN only, BUN + SCr, BUN + SCr + DDI, BUN + SCr + DDI + Dose. The classification performance results for each feature set are shown in Table 9. Among the four feature sets, the highest AUC value was 0.796 in BUN + SCr + DDI + Dose. The experiment results reveal that the combination of two or more IVs would make better predictions than just one IV alone.

**Table 9.** Classification performance results for each feature set.

Selected IVs	AUC	Precision	Recall	F-measure
BUN	0.702	0.789	0.579	0.668
BUN, SCr	0.766	0.825	0.708	0.762
BUN, SCr, DDI	0.788	0.833	0.747	0.787
BUN, SCr, DDI, Dose	0.796	0.851	0.708	0.773



## 5. Conclusions

The issue of medication safety has received considerable attention and discussion over the past few decades [1]. It is especially critical for high-alert medications that have complex pharmacological properties and easily cause ADEs. Digoxin is not only a high-alert medication, but also a widely used cardiovascular drug in clinical practice. When making an inappropriate treatment decision may lead to severe ADEs and even death. Therefore, the safety of high-alert medications is still a challenge all over the world.

The objective of this study was to predict the high risk patients of high-alert medication treatment. Six machine-learning techniques, including kNN, C4.5, SVM, RF, ANN and LGR, were adopted to construct the prediction models. In this study, the medical records of 513 inpatients were employed to obtain the experimental results and confirm the prediction efficiency of all adopted techniques. Finally, among all classification techniques, the RF exhibited the best performance. Overall, the decision tree-based methods exhibited an average performance superior to that of other techniques and can support clinicians as aids for making treatment decisions.

Although high-alert medications have a narrow therapeutic range and complex drug properties, this research verified that the properly use of medical records and consideration of drug-related variables can achieve better prediction outcomes. In conjunction with clinical experience, the prediction models can facilitate improving medication safety for patients, decreasing the risk of ADEs and economizing the use of medical resources.

Several limitations of this study should be addressed for further research. First, the data used in this study were collected from a tertiary medical center only. Further evaluation of clinical cases in other medical institutions is essential to confirm the validity of the model. Second, we only considered the concomitant medications with major degree DDIs. To construct a more powerful predictive model, including all drugs having DDI when used in combination with high-alert drugs is indispensable. Finally, the main limitation of the machine learning methods is lack of interpretability. How to build trust between physicians and the prediction models is another critical issue.

Improper use of high-alert medication would increase ADEs and may induce serious morbidity and mortality. This study confirmed that machine-learning techniques can yield favorable prediction effectiveness for high-alert medication treatment, thereby decreasing the risk of ADEs and improving medication safety. Future applications in the research of other medications are expected to enhance the patient safety.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-3417/10/17/5798/s1>, Table S1: The drugs inducing major drug-to-drug interaction when combined with digoxin.

**Author Contributions:** Conceptualization, C.-T.T. and Y.-H.H.; methodology, K.-L.S. and Y.-H.H.; software, K.-L.S.; validation, K.-L.S. and Y.-H.H.; formal analysis, C.-T.T., K.-L.S., and Y.-H.H.; resources, Y.-H.H.; data curation, Y.-H.H.; writing—original draft preparation, C.-T.T., K.-L.S. and Y.-H.H.; writing—review and editing, C.-T.T., K.-L.S. and Y.-H.H.; visualization, K.-L.S.; supervision, Y.-H.H.; project administration, C.-T.T.; funding acquisition, C.-T.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** The work was supported by the Ministry of Science and Technology of Taiwan under Grant MOST 107-2410-H-194 -054-MY2.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Mitchell, I.; Schuster, A.; Smith, K.; Pronovost, P.; Wu, A. Patient safety incident reporting: A qualitative study of thoughts and perceptions of experts 15 years after 'To Err is Human'. *BMJ Qual. Saf.* **2016**, *25*, 92–99. [[CrossRef](#)] [[PubMed](#)]
2. Landrigan, C.P.; Parry, G.J.; Bones, C.B.; Hackbarth, A.D.; Goldmann, D.A.; Sharek, P.J. Temporal trends in rates of patient harm resulting from medical care. *N. Engl. J. Med.* **2010**, *363*, 2124–2134. [[CrossRef](#)]

3. Dixon-Woods, M.; Pronovost, P.J. Patient safety and the problem of many hands. *BMJ Qual. Saf.* **2016**, *25*, 485–488. [CrossRef]
4. World Health Organization. WHO Launches Global Effort to Halve Medication-Related Errors in 5 Years. Available online: <http://www.who.int/mediacentre/news/releases/2017/medication-related-errors/en/> (accessed on 20 August 2017).
5. Donaldson, L.J.; Kelley, E.T.; Dhingra-Kumar, N.; Kieny, M.P.; Sheikh, A. Medication without harm: WHO's third global patient safety challenge. *Lancet* **2017**, *389*, 1680–1681. [CrossRef]
6. Salman, M.; Mustafa, Z.U.; Rao, A.Z.; Khan, Q.U.A.; Asif, N.; Hussain, K.; Rashid, A. Serious Inadequacies in High Alert Medication-Related Knowledge Among Pakistani Nurses: Findings of a Large, Multicenter, Cross-sectional Survey. *Front. Pharmacol.* **2020**, *11*, 1026. [CrossRef]
7. Institute for Safe Medication Practices. ISMP List of High-Alert Medications in Acute Care Settings. Available online: <https://www.ismp.org/recommendations/high-alert-medications-acute-list> (accessed on 23 August 2018).
8. Lee, J.; Han, H.; Ock, M.; Lee, S.I.; Lee, S.; Jo, M.W. Impact of a clinical decision support system for high-alert medications on the prevention of prescription errors. *Int. J. Med. Inform.* **2014**, *83*, 929–940. [CrossRef]
9. Sharabiani, A.; Bress, A.; Douzali, E.; Darabi, H. Revisiting warfarin dosing using machine learning techniques. *Comput. Math. Methods Med.* **2015**. [CrossRef]
10. Ma, Z.; Wang, P.; Gao, Z.; Wang, R.; Khalighi, K. Ensemble of machine learning algorithms using the stacked generalization approach to estimate the warfarin dose. *PLoS ONE* **2018**, *13*, e0205872. [CrossRef]
11. Mc Namara, K.; Alzubaidi, H.; Jackson, J.K. Cardiovascular disease as a leading cause of death: How are pharmacists getting involved? *Integr. Pharm. Res. Pract.* **2019**, *8*, 1–11. [CrossRef] [PubMed]
12. Naghavi, M.; Wang, H.; Lozano, R.; Davis, A.; Liang, X.; Zhou, M. GBD 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **2015**, *385*, 117–171.
13. Levy, M.; Kewitz, H.; Altwein, W.; Hillebrand, J.; Eliakim, M. Hospital admissions due to adverse drug reactions: A comparative study from Jerusalem and Berlin. *Eur. J. Clin. Pharmacol.* **1980**, *17*, 25–31. [CrossRef] [PubMed]
14. Rodgers, A.D.; Zhu, H.; Fourches, D.; Rusyn, I.; Tropsha, A. Modeling liver-related adverse effects of drugs using k nearest neighbor quantitative structure-activity relationship method. *Chem. Res. Toxicol.* **2010**, *23*, 724–732. [CrossRef] [PubMed]
15. Zhang, W.; Liu, F.; Luo, L.; Zhang, J. Predicting drug side effects by multi-label learning and ensemble learning. *BMC Bioinform.* **2015**, *16*, 365. [CrossRef] [PubMed]
16. Shakeel, D.; Mir, S.A. Personalized drug concentration predictions with machine learning: An exploratory study. *Int. J. Basic Clin. Pharmacol.* **2020**, *9*, 980. [CrossRef]
17. Glasgow, J.M.; Kaboli, P.J. Detecting adverse drug events through data mining. *Am. J. Health-Syst. Pharm.* **2010**, *67*, 317–320. [CrossRef] [PubMed]
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. Evaluat. Clin. Pract.* **2017**, *23*, 1240–1246. [CrossRef] [PubMed]
19. Mansour, A.M. Decision tree-based expert system for adverse drug reaction detection using fuzzy logic and genetic algorithm. *Int. J. Adv. Comp. Res.* **2018**, *8*, 110–128. [CrossRef]
20. Kang, S.; Kang, P.; Ko, T.; Cho, S.; Rhee, S.J.; Yu, K.S. An efficient and effective ensemble of support vector machines for anti-diabetic drug failure prediction. *Exp. Syst. Appl.* **2015**, *42*, 4265–4273. [CrossRef]
21. Kim, E.; Nam, H. Prediction models for drug-induced hepatotoxicity by using weighted molecular fingerprints. *BMC Bioinform.* **2017**, *18*, 227. [CrossRef]
22. Zhang, H.; Yu, P.; Xiang, M.L.; Li, X.B.; Kong, W.B.; Ma, J.Y.; Zhang, J. Prediction of drug-induced eosinophilia adverse effect by using SVM and naïve Bayesian approaches. *Med. Biol. Eng. Comp.* **2016**, *54*, 361–369. [CrossRef]
23. Chen, T.; Brewster, P.; Tuttle, K.R.; Dworkin, L.D.; Henrich, W.; Greco, B.A.; Massaro, J.M. Prediction of cardiovascular outcomes with machine learning techniques: Application to the Cardiovascular Outcomes in Renal Atherosclerotic Lesions (CORAL) study. *Int. J. Nephrol. Renovasc. Dis.* **2019**, *12*, 49. [CrossRef] [PubMed]

24. Su, R.; Li, Y.; Zink, D.; Loo, L.H. Supervised prediction of drug-induced nephrotoxicity based on interleukin-6 and-8 expression levels. *BMC Bioinform.* **2014**, *15*, S16. [[CrossRef](#)] [[PubMed](#)]
25. Yao, S.H.; Tsai, H.T.; Lin, W.L.; Chen, Y.C.; Chou, C.; Lin, H.W. Predicting the serum digoxin concentrations of infants in the neonatal intensive care unit through an artificial neural network. *BMC Pediatr.* **2019**, *19*, 517. [[CrossRef](#)] [[PubMed](#)]
26. Yen, P.F.; Mital, D.P.; Srinivasan, S. Prediction of the serious adverse drug reactions using an artificial neural network model. *Int. J. Med. Eng. Inform.* **2011**, *3*, 53–59. [[CrossRef](#)]
27. Albert, A.; Serrano, A.J.; Soria, E.; Jiménez, N.V. Clinical decision support system to prevent toxicity in patients treated with digoxin. In *Intelligent Medical Technologies and Biomedical Engineering: Tools and Applications*; IGI Global: Hershey, PA, USA, 2010; pp. 1–21.
28. Kockova, R.; Skvaril, J.; Cernohous, M.; Maly, M.; Kocka, V. Five year two center retrospective analysis of patients with toxic digoxin serum concentration. *Int. J. Cardiol.* **2011**, *146*, 447–448. [[CrossRef](#)]
29. Gheorghade, M.; Adams, K.F., Jr.; Colucci, W.S. Digoxin in the management of cardiovascular disorders. *Circulation* **2004**, *109*, 2959–2964. [[CrossRef](#)]
30. Thomson Reuters. *Thomson Micromedex Database*; Thomson Reuters: New York, NY, USA, 2013.
31. Bauman, J.L.; DiDomenico, R.J.; Viana, M.; Fitch, M. A method of determining the dose of digoxin for heart failure in the modern era. *Arch. Intern. Med.* **2006**, *166*, 2539–2545. [[CrossRef](#)]
32. Aha, D.W.; Kibler, D.; Albert, M.K. Instance-based learning algorithms. *Mach. Learn.* **1991**, *6*, 37–66. [[CrossRef](#)]
33. Quinlan, J.R. *C4. 5: Programs for Machine Learning*; Elsevier: Amsterdam, The Netherlands, 2014.
34. Karegowda, A.G.; Manjunath, A.S.; Jayaram, M.A. Comparative study of attribute selection using gain ratio and correlation based feature selection. *Int. J. Inf. Tech. Knowl. Manag.* **2010**, *2*, 271–277.
35. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berlin, Germany, 2013.
36. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
37. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [[CrossRef](#)]
38. Hosmer, D.W., Jr.; Lemeshow, S.; Sturdivant, R.X. *Applied Logistic Regression*; John Wiley & Sons: Hoboken, NJ, USA, 2013; Volume 398.
39. Tan, P.N.; Steinbach, M.; Kumar, V. *Introduction to Data Mining*; Pearson Education India: Bengaluru, India, 2016.
40. Lin, H.J.; Chao, C.L.; Chien, K.L.; Ho, Y.L.; Lee, C.M.; Lin, Y.H.; Chen, C.Y. Elevated blood urea nitrogen-to-creatinine ratio increased the risk of hospitalization and all-cause death in patients with chronic heart failure. *Clin. Res. Cardiol.* **2009**, *98*, 487. [[CrossRef](#)] [[PubMed](#)]
41. Doherty, J.E.; Flanagan, W.J.; Perkins, W.H.; Ackerman, G.L.; Gammill, J.; Sherwood, J. Studies with tritiated digoxin in anephric human subjects. *Circulation* **1967**, *35*, 298–303. [[CrossRef](#)] [[PubMed](#)]
42. Pawlosky, N.; MacDonald, E.; Patel, R.; Kanji, S. Evaluation of digoxin concentration after loading dose in patients with renal dysfunction. *Can. J. Hosp. Pharm.* **2013**, *66*, 104. [[CrossRef](#)]
43. Bajraktarević, A. Safety and efficacy of digoxin therapy—Where are we now? *J. Pharm. Pharmacol.* **2016**, *4*, 128–134. [[CrossRef](#)]
44. Chen, R.; Zou, S.L.; Wang, M.L.; Jiang, Y.; Xue, H.; Qian, C.Y.; Xia, Z.L. Population pharmacokinetics of digoxin in elderly patients. *Eur. J. Drug Metab. Pharmacokinet.* **2013**, *38*, 115–121. [[CrossRef](#)]
45. Kroese, W.L.G.; Avery, A.J.; Savelyich, B.S.P.; Brown, N.S.; Schers, H.; Howard, R.; Horsfield, P. Assessing the accuracy of a computerized decision support system for digoxin dosing in primary care: An observational study. *J. Clin. Pharm. Ther.* **2005**, *30*, 279–283. [[CrossRef](#)]
46. Muzzarelli, S.; Stricker, H.; Pfister, O.; Foglia, P.; Moschovitis, G.; Mombelli, G.; Brunner-La Rocca, H. Individual dosage of digoxin in patients with heart failure. *QJM-Int. J. Med.* **2011**, *104*, 309–317. [[CrossRef](#)]
47. Suematsu, F.; Yukawa, E.; Yukawa, M.; Minemoto, M.; Ohdo, S.; Higuchi, S.; Goto, Y. Population-based investigation of relative clearance of digoxin in Japanese neonates and infants by multiple-trough screen analysis. *Eur. J. Clin. Pharmacol.* **2001**, *57*, 19–24. [[CrossRef](#)]
48. Yukawa, M.; Yukawa, E.; Suematsu, F.; Takiguchi, T.; Ikeda, H.; Aki, H.; Mimemoto, M. Determination of Digoxin Clearance in Japanese Elderly Patients for Optimization of Drug Therapy. *Drugs Aging* **2011**, *28*, 831–841. [[CrossRef](#)]

49. Zhou, X.D.; Gao, Y.; Guan, Z.; Li, Z.D.; Li, J. Population pharmacokinetic model of digoxin in older Chinese patients and its application in clinical practice. *Acta Pharmacol. Sin.* **2010**, *31*, 753–758. [[CrossRef](#)] [[PubMed](#)]
50. Vamos, M.; Erath, J.W.; Hohnloser, S.H. Digoxin-associated mortality: A systematic review and meta-analysis of the literature. *Eur. Heart J.* **2015**, *36*, 1831–1838. [[CrossRef](#)] [[PubMed](#)]
51. Hu, Y.H.; Tai, C.T.; Tsai, C.F.; Huang, M.W. Improvement of Adequate Digoxin Dosage: An Application of Machine Learning Approach. *J. Healthcare Eng.* 2018. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).