Review of Machine Learning Techniques for Warfarin Dosing

Gianluca Truda

TRDGIA001

University of Cape Town

trdgia001@myuct.ac.za

ABSTRACT

Warfarin is a commonly-prescribed anticoagulant for the prevention of blood clots and heart attacks. The narrow therapeutic range and severe side-effects of warfarin make precise administration of the drug essential, but variation in response caused by environmental and genetic factors results in the need for individualised dosing. Machine learning techniques can improve the accuracy of the dosing algorithms, given sufficient datasets. We evaluate fifteen techniques from sixteen papers on warfarin dose prediction, highlighting artificial neural networks (ANNs), support vector regression (SVR), boosted regression trees (BRTs), and random forest regression (RFR) as promising avenues for further research. We also investigate the use of pharmacogenetic factors in model development, finding that they may not be necessary for high accuracy, and suggest four avenues for novel research in this field.

CCS CONCEPTS

• **Applied computing** → **Consumer health**; *Health care information systems*; Genetics; • **Computing methodologies** → *Supervised learning*; *Machine learning approaches*;

1 INTRODUCTION

Many individuals suffer from abnormalities in blood coagulation, which can lead to thrombosis (obstructive blood clots), stroke, or heart attack. One of the most common methods for treating these conditions is the use of anticoagulant drugs, such as warfarin. Whilst oral warfarin treatment is extremely effective, the drug has a narrow therapeutic range and severe side-effects at extreme concentrations. This makes the precise dosing of warfarin an important concern for clinicians. Unfortunately, warfarin metabolism differs across individuals based on age, weight, genetics, diet, drug interactions, and various pre-existing conditions. This makes the task of accurately dosing warfarin a highly individualised endeavour. International standards and dosing protocols have attempted to formalise the procedure, and software tools exist to assist clinicians in making informed dosing decisions, but the high individual variability of warfarin, and the risk of severe bleeding, makes the development of more accurate dosing methods a constant priority. Many studies (see Table 1) have looked at applying statistical models and machine learning techniques to the problem of individualised warfarin dosing, but the datasets are often small and restricted to a specific population group. This review aims to identify promising machine learning techniques that can be trained on a dataset of South African patients. With access to the comprehensive warfarin records provided by the pathology group PathCare (www.pathcare.co.za), it may be possible to use machine learning

principles to construct a model for warfarin dosing that reveals novel insights and offers real benefits to patients in South Africa.

The remainder of this paper is organised as follows: Section 2 provides background information on the drug warfarin, how levels of anticoagulation are measured, and what the current clinical practices for warfarin dosing are. Section 3 examines a number of promising techniques for developing predictive models for warfarin dosing, based on a review of sixteen studies. Section 4 weighs up the evidence for including pharmacogenetic factors in predictive models, despite the drawbacks associated with acquiring genetic data. Section 5 concludes the review and identifies gaps in the current literature that suggest interesting avenues for research.

2 BACKGROUND

2.1 Warfarin as an Anticoagulant

The presence of warfarin causes an anticoagulant effect by inhibiting vitamin K-dependent clotting factors. It has achieved popularity and widespread usage due to its superior bioavailability and relatively predictable onset, but has a very narrow therapeutic range and poses serious risk of bleeding. Warfarin is administered via a racemic mixture of both its stereoisomers, R-warfarin and S-warfarin. S-warfarin is metabolised via the CYP2C9 protein and has higher potency than R-warfarin. Together, both isomers inhibit warfarin's target, VKORC1, limiting its ability to chemically reduce vitamin K, thus decreasing formation of functionally-active clotting factors [17]. This metabolism is illustrated in Figure 1.

There are two different phases of warfarin dosing – the initiation dose, and the maintenance dose. As its name suggests, the initiation dose is the quantity of warfarin administered to a patient beginning anticoagulation therapy, whilst the maintenance dose is used to keep a patient in a therapeutic range once warfarin has already saturated their system. The initiation dose can be optimised to some extent with statistical models, but is typically the easier dose to administer, as patients are monitored regularly and adjustments are made quickly during the initiation of treatment. Simple algorithms guide clinicians in administering an appropriate initiation dose [3]. The maintenance dose, however, can be incredibly unpredictable. At this phase of treatment, the patient is tested very infrequently, so the maintenance dose must be precisely tuned to keep them within therapeutic levels. For the most part, studies in warfarin dosing are focussed on the maintenance dose.

2.2 Individual Variation in Warfarin Response

There is substantial variation in how individuals respond to warfarin. Many clinical and environmental factors, such as age, race, weight, height, and smoking status must be taken into account when determining warfarin dosage. There are also a host of genetic factors worth considering. Around 40% of the individual variation
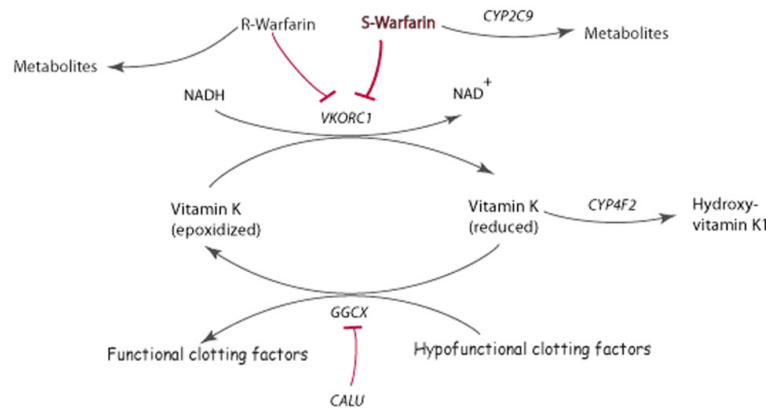
Figure 1: Schematic representation of warfarin metabolism and its mechanism of action. Adapted from Johnson et al. [17]

in dose requirement can be attributed to polymorphisms at only two genes – *CYP2C9* and *VKORC1* [18].

There are also many foods and drugs that interact with warfarin. A 1994 meta-analysis by Wells et al. [45] examined 793 studies to evaluate the evidence for the interactions of drugs and food with warfarin. The analysis revealed high levels of agreement across the studies and 26 drugs and foods were determined to interact with warfarin, including, but not limited to, a number of antibiotics, cardiac drugs, alcohol (with liver disease), three drugs that act on the central nervous system, vitamin K-rich foods, and large quantities of avocado.

2.3   International Normalised Ratio (INR)

The intensity of anticoagulation in a patient is monitored by measuring prothrombin time (PT), which is the time a patient's plasma takes to re-calcify in the presence of thromboplastin [32].

But, thromboplastin varies in reactivity depending on its origin, so the sensitivity of a particular thromboplastin source is tracked with the international sensitivity index (ISI). To simplify and standardise the process of anticoagulant monitoring, the World Health Organisation established the international normalised ratio (INR) as a universal reference value [23]. The INR is determined based on two PT values and an ISI value as follows:

$$INR = \left(\frac{PT_{patient}}{PT_{normal}}\right)^{ISI} \qquad (1)$$

where PT_{normal} is the average of ≥ 20 healthy subjects of both sexes in the same local test system [32]. In essence, the INR is the ratio of measured PT to normal PT calculated in terms of the appropriate ISI for the local test system. INR has become the gold standard measurement for anticoagulation around the world.

According to the American College of Chest Physicians (ACCP), the recommended therapeutic range for oral anticoagulant therapy is an INR between 2.0 and 3.0 for most patients [1]. Often, dosing guidelines specify a specific target INR (e.g. 2.5), instead of a range, but values within 0.5 INR points of this target are considered acceptable, as it has been found that using tighter target ranges for maintenance dosing does not achieve any improvement in anticoagulation control [28]. This means that, for the purposes of treatment, any coagulation measures within this ±0.5 range are

equally good. A 2003 study by Gerald et al. [4] describes how mild, asymptotic elevations in INR can occur and that there existed doubt about whether the warfarin dosage should be dropped accordingly. After performing a randomised, controlled trial of 231 outpatients receiving warfarin, Gerald et al. recommend maintaining the warfarin dose in asymptotic patients with INR ≤ 3.3 and reducing the warfarin dose by < 20% for patients with either an INR > 3.3, or an increased risk of haemorrhage. A final consideration of INR values is the time a patient spends in the therapeutic range. Clinical studies consider time in therapeutic range (TTR) as a superior marker because it indicates a greater stability in the rate of warfarin metabolism. The longer patients spend in the therapeutic range, the more effective their treatment, and the less risk is posed. This makes TTR a valuable tool to assess the quality of the anticoagulation treatment [40].

2.4   Traditional Dosing Protocols

Traditionally, warfarin dosing is determined by a clinician on an individual basis. The dose given is based on medical experience and observations of the patient, and is adjusted based on the INR values measured. In most cases, oral warfarin treatment in outpatients is initiated with a slow-loading regimen that achieves therapeutic anticoagulation within 3-4 weeks. In cases of acute thrombosis, a 5mg loading dose of warfarin is used to reach therapeutic levels sooner [20]. Maintenance doses are determined on a case-by-case basis and adjusted regularly to ensure an INR value in the desired range. The target ranges vary from $2.0 - 3.0$ up to $3.0 - 4.0$ depending on the presence of mechanical heart valves in the patient [20]. The imprecision and possibility of human error involved in manual dosing has resulted in a concerted effort to (at least partially) automate warfarin dosing [20].

2.5   Dosing Equations

There are two widely used equations in clinical practice for warfarin dosing. These are used as a baseline by which to compare many of the machine learning techniques that follow. Gage et al. [10] used an exponential function based on age, body surface area, target INR, whether the patient smokes, whether the patients has deep vein thrombosis (DVT), whether the patient has a pulmonary

embolism, whether the patient is African American, whether the patient takes amiodarone, and whether the patients has polymorphisms at three genes. The IWPC [24] used a least-squares linear regression method to develop their pharmacogenetic equation based on age, height, weight, race, and VKORC1 and CYP2C9 genotypes. The IWPC equation was made available as a web-based software tool (www.warfarindosing.org).

2.6 Software-Assisted Dosing

It is now very common for medical staff to make use of a dosing algorithm to guide their decisions, since they have been found to increase anticoagulant control [21]. Many also make use of software tools to guide this process. Computer-assisted dosing not only increases time spent in the target INR range and reduces the risk of bleeding, but is more cost-effective too [20]. A 2008 study by Poller et al. [33] compared the efficacy of computer-assisted warfarin dosing with dosing by experienced medical staff in a multicenter clinical endpoint study. They compared the performance of two computer-assisted dosing programs (PARMA 5 and DAWN AC) to manual dosing at 32 centres (totalling 13219 patient participants), and concluded that the computer-assisted dosing was both safer and more effective than that of the medical staff. A concurrent study by Jowett et al. [19] compared the cost-effectiveness of computer-assisted dosing and manual dosing at the same 32 centres, finding that the computer-assisted dosing was substantially more cost-effective. Whilst these tools are invaluable to clinicians, there exists much room for improvement. A 2008 study by McDonald et al. [27] used a database of 17396 patient's records to compare the performance of neural networks with a benchmark of DAWN AC predictions. Their findings aligned with those of previous papers – machine learning approaches offer a promise of better results over the current computer-assisted dosing software.

3 PREDICTIVE MODELS FOR WARFARIN DOSING

This section examines sixteen notable studies on warfarin dose prediction (Table 1). The term *prediction* is used to indicate that the models are trained or built on a large fraction of the dataset available, but validated on a small fraction. In a sense, the model tries to *predict* the dose given to each patient in the validation set based on the relationships it has found between the final dose and other factors in the training set. A model that can successfully *predict* warfarin doses given to patients in the past can be used to *suggest* doses for patients in the future.

3.1 Linear Regression (LR)

Linear regression (LR) is utilised heavily in ten of the sixteen studies highlighted in Table 1. LR is used to model the relationship between one continuous dependent variable and two or more independent variables. It is by far the oldest and most widespread technique in the studies sampled, with many using it to establish a baseline for more contemporary approaches to improve upon. An example of this is seen in the 2004 study conducted by Solomon et al. [39], in which a simple multivariate linear regression achieved a reasonable fit ($r = 0.800$), but was outperformed by a back-propagation artificial neural network (ANN) ($r = 0.823$), even at a sample size of
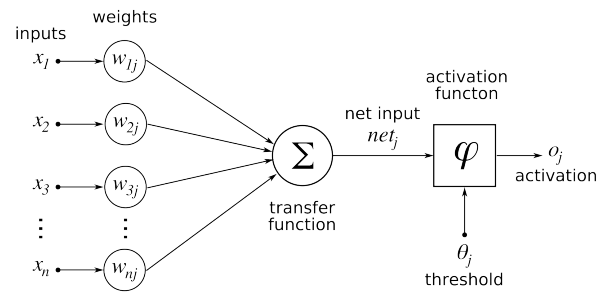
Figure 2: Depiction of a single node/neuron in an artificial neural network (ANN) (Wikimedia Commons).

only 148 patients, which is considered tiny for supervised machine learning. In contrast, a 2013 study by Sharabiani et al. [38] found that multiple linear regression improved performance over previous models [24] and outperformed support vector regression (SVR) and artificial neural network (ANN) methods on their dataset of 326 African American patients. A combination of LR with decision trees, called model-tree (MT) regression, was developed by Quinlan [34] and extended by Hu et al. [15]. The associated M5 algorithm builds an LR model for each leaf node, then "prunes" the children that fail to meet a pre-determined performance threshold. Hu et al. used this technique to great success in their 2012 study [15] and found that MT performed comparably to support vectors (SVR) and other advanced techniques on a dataset of 587 Taiwanese individuals. As found by a number of studies [15, 24, 26, 39, 47], linear regression is seldom on par with other contemporary techniques like ANNs and SVR, but may perform well under certain conditions. For example, Liu et al. [26] note that LR performs as well as SVR, BART, and MARS in the Asian component of their cohort. Whilst linear regression is an invaluable tool in model creation, it appears to be better suited as a baseline against which to compare machine learning techniques (especially in the absence of a reliable clinical dosing control).

3.2 Artificial Neural Networks (ANNs)

Artificial neural networks (ANNs) are the second most common technique employed by the studies in Table 1, with eight out of sixteen studies making use of them. ANNs are a data structure comprised of interconnected nodes that simulate the form of a biological brain using mathematical functions. Simply put, each node (or neuron) is comprised of the following elements: one or more input connections with distinct values, $x_1, x_2, ..., x_n$; a set of relative weights for those input values, $w_1, w_2, ..., w_n$; a single output signal, o_j; and a set of output connections. Each input connection has its own distinct value, x_i, and its own distinct weight, w_i; but every output connection carries the same output signal, o_j, to many other nodes in the network (see Figure 2). The weightings allow each node to alter the relative importance of each input value. In a simple ANN, the node uses a summation function to sum up the dot product of each value-weight pair, $(x_i * w_i)$, and produce a single number. This number is usually passed directly into the transfer function, which compares the value to a threshold and returns a corresponding output. The output of the transfer function

Table 1: Comparison of notable studies on warfarin dose prediction.

Year	Researchers	Factors	Population	Size	Train/Test	Techniques
2004	Solomon et al.[39]	CLN	Israeli	148	70/30	ANN, LR
2007	Miao et al.[29]	CLN + PHCG	Chinese	178	*	LR
2008	McDonald et al.[27]	CLN	Multiethnic (DAWN dataset)	17396	60/40	ANN, ARMAX
2008	Schelleman et al.[35]	CLN + PHCG	Caucasian + African American	259	*	LR
2009	Klein et al. (IWPC)[24]	CLN + PHCG	Multiethnic (IWPC dataset)	5052	80/20	ANN, LARS, LASSO, LR, MARS, MT, RT, SVR
2009	Wadelius et al.[44]	CLN + PHCG	Swedish	181	70/30	LR
2010	Harada et al.[13]	CLN + PHCG	Japanese	97	*	LR
2010	Le Gal et al.[25]	CLN	Canadian	324	*	LR
2011	Cosgun et al.[9]	CLN + PHCG	African American	290	*	RFR, BRT, SVR
2012	Hu et al.[15]	CLN	Taiwanese	587	*	B&V, kNN, MLP, MT, SVR
2013	Sharabiani et al.[38]	CLN	African American	326	80/20	ANN, LR, SVR
2014	Grossi et al.[11]	CLN + PHCG	Caucasian	377	50/50	ANN + TWIST
2014	Isma'eel et al.[16]	CLN + PHCG	Lebanese	174	50/50	ANN, LSM
2014	Zhou et al.[47]	CLN	Chinese (CLIATHVR)	1093	75/25	ANN, LR
2015	Liu et al.[26]	CLN + PHCG	Multiethnic (IWPC dataset)	4798	80/20	ANN, BART, BRT, LASSO, LR, MARS, RFR, RT, SVR
2015	Sharabiani et al.[37]	CLN	Multiethnic (IWPC dataset)	4237	50/50	RVM, SVR

CLN = Clinical, PHCG = Pharmacogenetic. * = undeclared or unavailable. Technique abbreviations are explained throughout section 3.

is often scaled and limited to keep the value within a strict range. The ANNs of concern to warfarin dosing are those with the ability to *learn*. These usually possess a learning function, which modifies the weights on the inputs appropriately, and an error function that calculates the difference between the current output and the desired output. The goal of learning is to minimise these errors. This error value (or some scaled form of it) is often *back-propagated* back to previous layers of the network, so that weighting values of multiple nodes can be updated appropriately before the next learning cycle. Unlike biological neural networks, ANNs group neurons/nodes into layers, with connections running between each. Even the simplest ANNs contain at least three layers – an input layer to receive data from the source files, one or more hidden layers that feed data forward, and an output layer that feeds information back to the outside world [5].

Sharabiani et al. successfully implemented a warfarin dosing algorithm using an ANN in their 2013 study [38]. Their network consisted of two hidden layers and made use of back-propagation to calibrate the weights, but was found not to outperform support vector regression (SVR) and linear regression (LR) on the same dataset of 326 African Americans. This may have been due to the limited size of the dataset. Conversely, Zhou et al. compared an ANN with linear regression in a 2014 paper [47], in which they simply implemented the neural network toolbox in Matlab 7.1. Zhout et al. found that the dose prediction of their ANN was superior to that of an LR model across dose ranges on their dataset of 1093 Chinese patients with heart valve replacements. They attributed this to greater fault-tolerance in ANNs. Liu et al. conducted a 2015 study [26] on 4798 individuals from the multi-ethnic IWPC dataset[46], and evaluated an ANN against eight other techniques for warfarin dose prediction, finding that it performed worse than linear regression across dose levels, but outperformed LR at lower dose levels

and in some ethnicities. Whilst they concluded that ANNs produce a reasonably good model for warfarin dosing, their results suggest that SVR, RFR, RT, BART, LASSO, and MARS are all more promising avenues. Their findings hold more weight given the comprehensiveness of their methodology and the large size of their dataset, but they fail to provide detail on the exact implementation of their ANN, suggesting there may still be room for optimisation. In fact, Hu et al. implemented a multi-layer perceptron (MLP) neural network in a 2012 study [15], which found it to perform almost as well as k-nearest neighbours (kNN) and model tree (MT) approaches. They also found that it benefited more from the extension technique known as *Bagging* (see Section 3.8) than any of the other techniques. Indeed, *Bagged* MLP was one of the highest performing ensemble methods in their entire study, with an average mean absolute error (MAE) of only 0.216.

Despite widespread use of ANNs in many studies, no notable publications have been submitted on the topic of *deep learning* – using ANNs with many more hidden layers – for warfarin dosing. A recent article by Ching et al. [7] describes the mismatch between individuals skilled in deep learning and individuals well-versed in biological and medical fields. They note specifically for the case of warfarin dosing that a lack of standardisation is a challenge for deep learning experts, as many informed data processing steps must be executed before algorithms can be applied. This suggests that deep learning is an avenue worthy of investigation in the future.

3.3 Support Vector Regression (SVR)

Support vector regression (SVR) was utilised in six of the sixteen studies highlighted in Table 1. Support vector classification was originally described by Vapnik in 1998 [42]. The support vector machines (SVMs) that underpin this technique simply apply a linear method to the data, allowing unseen examples to be assigned to one
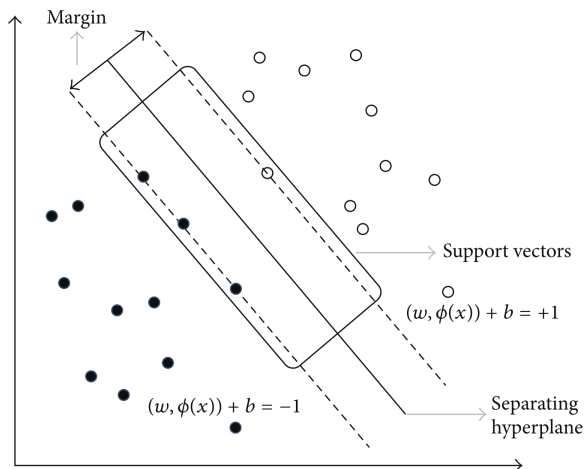
Figure 3: Depiction of a hyperplane and two support vectors classifying data in a two-dimensional feature space, from Sharabiani et al[37].

Figure 4: Illustration of the higher-dimensional mapping of data from input space to feature space in a support vector machine (Wikimedia Commons).

of two classes based on a derived model. The algorithm plots data points (vectors) from two distinct classes in a *feature space*, then separates them with a boundary called a *hyperplane*. The *support vectors* are data points closest to the opposing classes, encapsulating the hyperplane and defining a margin between the classes. The algorithm is primarily concerned not with the hyperplane itself, but with these support vectors (Figure 3).

The true value of support vector machines is that they can be extended into higher-dimensional space through the use of *kernel functions* (Figure 4). One of the simplest kernel functions, for instance, is a linear one: $K(x, y) = x * y$. By increasing the dimensionality of the feature space, support vector machines can be used to accurately classify complex data points with any number of features [12]. Optimisation of classification requires a loss function, and the use of certain loss functions allows support vector machines to be used for support vector regression (SVR) [14]. The performance of an SVM can be optimised by tuning its parameters through a technique known as *K-fold cross validation*, where K is some integer.

Cosgun et al. made use of SVR in their 2011 study [9], in which they used the statistical program R to implement SVR through the e1071 package. Their C and γ parameters were calibrated via *five*-fold internal cross-validation (CV) within the training set, and ϵ was constrained after optimisation of the aforementioned two. Whilst they found that their SVR outperformed models in previous studies (e.g. Klein et al. [24]), this was specifically on an African American cohort of only 290 individuals. Moreover, they found that SVR was not as effective as random forest regression (RFR) on that dataset. Hu et al. found in their 2012 study [15] that SVR outperformed k-nearest neighbour (kNN), Multi-layer perceptron (MLP) neural network, and model-tree regression (MT). The ensemble method of *Bagging* (see Section 3.8) was found to increase the performance of SVR further, with Hu et al. concluding that *Bagged* SVR and *Bagged Voting* with four classifiers were the two best prediction models for their dataset of 587 Taiwanese individuals. In contrast,
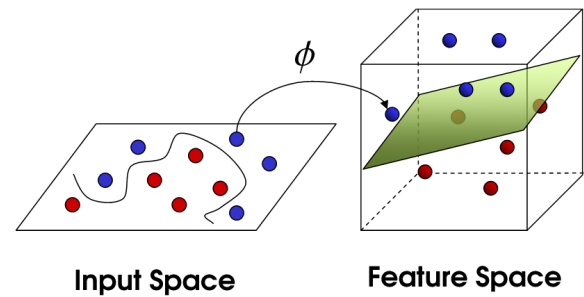
the 2013 Sharabiani et al. study [38] compared both SVR and ANN methods with linear regression on an African American cohort of 326 individuals, finding that SVR was less effective than their linear regression models. Despite this, Liu et al. found in a 2015 study [26] that SVR outperformed linear regression on 4798 individuals from the multi-ethnic IWPC dataset[46]. Given that the cohort in this study was both larger and more diverse than the cohort in the Sharabiani study [38], it would stand to reason that SVR is still a promising method. However, Liu et al. note many other techniques that perform at least as well as SVR on the same dataset. There are some notable limitations of SVR as identified by Sharabiani et al. [37]. Firstly, as the number of data points in the training set grows linearly, the number of support vectors grows linearly with it. Secondly, SVR provides hard boundary decisions, but it would be useful (in the case of warfarin dosing) to have a level of certainty associated with each new data point. Thirdly, cross-validation (CV) is required to estimate the complexity parameter. Relevance vector machines (RVMs) are an attempt at overcoming these concerns.

3.4 Relevance Vector Machines (RVMs)

Relevance vector machines are sparse Bayesian learners, meaning they make use of a probabilistic Bayesian learning framework. This can make them more efficient than support vector machines. RVMs provide the value of *probabilistic* predictions, but have other benefits like automatic estimation of nuisance parameters, and the facility to utilise non-Mercer kernels [41]. They have the identical functional form to support vector machines and can thus be used for both classification and regression.

Sharabiani et al. [37] found that relevance vector machines (RVMs) were effective at *classifying* patients into those requiring > 30mg per week and those requiring < 30mg per week. Thereafter, they used different linear regression models to predict exact doses for each class of patient. Sharabiani et al. found that this biphasic approach was 15% better than the IWPC model[24] on the IWPC's dataset[46] of over 4000 multi-ethnic individuals.

3.5 Regression Trees (RTs)

Regression trees (RTs) were much less common in the literature, with only three of the sixteen studies in Table 1 making use of them.
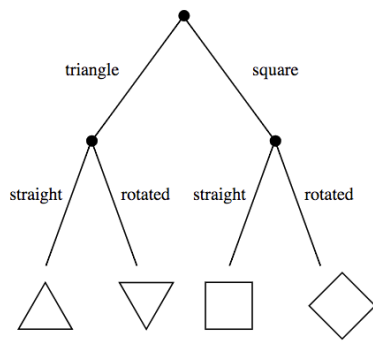
Figure 5: A very simple decision tree, showing classification of shapes (Wikimedia Commons).

RTs are a class of decision tree – a decision support tool that uses a tree-like data graph that models an algorithm containing only conditional ("if") statements (see Figure 5). The primary advantage of RTs is that they accept continuous values as target variables, which makes them useful in warfarin dose prediction. The ensemble method of *Boosting* (see Section 3.8) was used by some studies to increase accuracy.

In 2011, Cosgun et al. [9] found that *boosted* regression trees (BRTs) outperformed algorithms in published reports at the time [24], but did not perform as well as random forest regression (RFR) on their dataset of 290 African Americans. Moreover, Liu et al. supported this finding in their 2015 study [26] on 4798 individuals from the multi-ethnic IWPC dataset, in which the boosted regression tree (BRT) method outperformed linear regression and was one of the most effective of the nine techniques they compared. In that same study, Liu et al. also implemented a Bayesian additive regression tree (BART), which was one of the top-performing approaches. BART uses a non-parametric Bayesian regression approach to sum a collection of weak learners (decision trees) to make a single strong learner. It is distinct from random forest regression (RFR) in that it constrains each tree using a regularisation prior, as per Bayesian reasoning [8].

3.6   Random Forest Regression (RFR)

The random forest algorithm is used to improve the robustness of decision trees. Instead of a single tree modelling the data, a collection (or *forest*) of different trees is generated and the average of their outputs is used to give an accurate prediction. The random forest approach is popular for its ability to maintain accuracy despite missing data, its tendency to not overfit the model to the dataset, and its propensity for large, multi-dimensional datasets. Random forest *regression* (RFR), however, is less effective than the random forest classifier. RFR also suffers from the "black box" problem – there is little control over the workings of the model. Often, only a single tuning parameter exists [9].

Cosgun et al. evaluated random forest regression (RFR) in their 2011 study [9] on 290 African Americans. Their algorithm used a bootstrap-based cross-validation (CV) approach (see Section 3.8), both as a performance booster and to prevent overfitting to the dataset. They found that RFR outperformed boosted regression

trees and support vector regression, and achieved an average R^2 value of 0.664, compared to the IWPC model[24] at $R^2 \approx 0.275$. In the comparison of nine methods on the large, diverse IWPC dataset[46] by Liu et al. [26], RFR outperformed linear regression at all dose ranges, and performed comparably to boosted regression trees (RTs).

3.7   Multiple Adaptive Regression Splines (MARS)

The multiple adaptive regression splines (MARS) technique was utilised in only two of the studies from Table 1. MARS is a statistical modelling technique and is largely considered an extension of linear models that allows non-parametric regression. It is often able to outperform simpler regression techniques as it automatically models non-linearities and interactions present in the dataset, making it more flexible.

The IWPC study[24] that included MARS did not provide direct comparisons to other algorithms, but in the 2015 study [26] by Liu et al., MARS demonstrated the best performance of all nine methods investigated. This is notable, both because Liu et al. compared MARS to many highly-effective techniques, and because they made use of a large (N=4798) and highly-diverse cohort. However, due to the limited evidence for its effectiveness, it still requires more investigation as a technique.

3.8   Ensemble Methods

Boosting is a meta-algorithm that reduces bias and variance in supervised learning [6]. It can be used, as in the case of regression trees (RTs), to convert weak learners to strong ones. Boosting was used successfully by Cosgun et al. [9] and Liu et al. [26] to improve the performance of their regression trees.

Bagging (or bootstrap aggregating) is a homogeneous ensemble technique that reduces variance and the likelihood of overfitting [15]. Bagging generates multiple sample datasets through repeated sampling (with replacements) from the training dataset. Each of the derived datasets can then be used to build a distinct model. The average of each model's prediction is accurate more often than any of the individual predictions [30]. Bagging can be applied to virtually any prediction model.

Voting is a technique that combines the predictions of multiple models. Each model is trained on the dataset independently and assigned a weight based on its prediction accuracy. A combination rule uses the combined predictions and their weightings to produce an overall prediction [15].

Notably, Hu et al. combined both Bagging and Voting (B&V in Table 1) by using voting as the base classifier bagging extension, which they called "Bagged Voting" [15]. Bagging alone was able to improve the performance of their multi-Layer Perceptron (MLP) and Support Vector Regression (SVR) techniques, and "Bagged Voting" produced extremely low MAE and $\delta(E)$.

While Boosting and Bagging are both successful, Opitz et al. [30] note that while Bagging is nearly always more accurate than using a single model, it is occasionally much less accurate than Boosting.

4 PHARMACOGENETIC AND CLINICAL APPROACHES

As illustrated in Table 1, all notable studies on warfarin dose prediction use one of two factor classes to develop models – either only clinical factors (CLN), or both clinical and pharmacogenetic (CLN + PHCG) factors. Common *clinical* factors include age, body mass, height, other medications the patient is taking, whether or not the patient smokes, and what other diseases the patient suffers from. *Pharmacogenetic* factors include all the genotypes for each of the SNPs (single nucleotide polymorphisms) associated with warfarin metabolism and interaction. The two most notable of these are found in *CYP2C9* and *VKORC1*, which explain approximately 40% of the individual variation in dose requirement [18]. Whilst many studies have reported improved model performance when using both clinical and pharmacogenetic factors [24, 29, 31, 36, 43], there is doubt as to whether genotype-guided (pharmacogenetic) dosing is actually clinically beneficial [2, 3, 15, 22]. Genetic testing is still very expensive, takes additional time and resources, and is unavailable in many parts of the world [15]. It is therefore important to consider whether the advantages (if they exist) of including pharmacogenetic data in models outweigh the potential disadvantages.

A 2005 study by Sconce et al. [36] discussed how previous dosing algorithms did not account for genetic and environmental factors. They studied age, body size, and the *CYP2C9* and *VKORC1* genotypes in 297 patients with stable anticoagulation and INR in the range $2.0 - 3.0$. They found that age, height, and *CYP2C9* genotype contributed to S-warfarin clearance, but only age and body size contributed to R-warfarin clearance. Similarly, a 2007 study by Miao et al. [29] assessed the contribution of VKORC1, CYP2C9, age, body size, and weight on warfarin dose required by a Chinese cohort of 178 patients, finding that all but body size had a significant impact on required warfarin dose. This was further supported when Wadelius et al. [43] found that, when accounting for multiple testing and linkage disequilibrium, the genes CYP2C9 and VKORC1 had statistically significant association with warfarin metabolism in 201 patients. However, they also noted the importance of non-genetic predictors like age, bodyweight, and drug interaction. Notably, a 2007 randomised trial by Anderson et al. [3] compared genotype-guided and standard warfarin dosing in 206 patients *initiating* anticoagulation treatment. Whilst they did find that the pharmacogenetic algorithms improved the accuracy and efficiency of warfarin dose initiation, there was no resulting reduction in out of range INRs observed in the patients. This suggests that the benefits of pharmacogenetic factors may be limited to the initiation phase, without extending a measurable benefit in the more important maintenance phase. Most significantly, a 2009 publication by the International Warfarin Pharmacogenetics Consortium (IWPC) [24] used clinical and genetic data from 4043 patients in the IWPC dataset[46] to compare dosing algorithms that used only clinical data with those that used both clinical and genetic data. Averaging the results across SVR, RT, MT, MARS, LARS, and Lasso, in addition to linear regression, they found that adding pharmacogenetic factors increased the predictive effectiveness. More recently, Pirmohamed et al. [31] detailed a randomised trial of pharamacogenetic dosing of warfarin. They compared standard dosing protocols to genotype-guided protocols for initiation of warfarin therapy in 455

patients split into case- and control- groups, finding that pharmacogenetic dosing lead to a higher percentage of time spent in the therapeutic range. Despite this, a 2009 meta-analysis by Jonas et al. [18], which evaluated the current evidence for pharmacogenetic dosing algorithms by reviewing over 20 papers, concluded that the potential clinical benefits were still unclear. Finally, a 2013 paper by Kimmel et al. [22] compared pharmacogenetic algorithms and clinical algorithms for warfarin dosing, noting that the clinical utility of genotype-guided dosing of warfarin has only been tested in small clinical trials or observational studies, with "equivocal results." Their study recruited 1015 patients and compared the time the INR was in the therapeutic range from days 4 to 28 of warfarin therapy. They concluded that genotype-guided dosing of warfarin did not improve anticoagulation control during the first 4 weeks of therapy.

In summary, it is apparent that there is still no consensus about whether or not pharmacogenetic factors provide notable benefit to the model. More advanced machine learning approaches have produced significant gains in performance over previous models, despite many of them [15, 27, 37–39, 47] not incorporating any pharmacogenetic factors whatsoever. There is also doubt about whether or not the benefits of pharmacogenetic factors observed in models of initiation dosing translate to maintenance dosing [3]. It therefore seems that a better approach to improving warfarin prediction is found in making superior use of the data available.

5 CONCLUSIONS

Linear regression (LR) is a useful tool in analysing any dataset, but is not as fault-tolerant as other techniques. It has been used across many studies as a baseline against which to compare new methods, but is not a promising avenue in and of itself. Artificial neural networks (ANNs) have a good deal of promise, but often fail to outperform linear regression. Multi-layered perceptrons (MLPs) are one path to improving neural network performance, and deep learning may be another. Support vector regression (SVR) also exhibited notable performance across a number of studies, but has some limitations. Ensemble techniques improve SVR performance considerably, and adaptation to relevance vector machines (RVMs) overcomes many of the limits SVR encounters. Specifically, relevance vector machines (RVMs) were found by one study to provide an effective means of classifying patients into two classes before more specialised models were applied. This approach could be used to further improve performance on almost any technique that responds poorly to highly-varied data. Boosted regression trees (BRTs) and Bayesian additive regression trees (BARTs) performed exceptionally well across various studies and offer a viable path to improving warfarin dose prediction. Random forest regression (RFR) was used in two studies with high levels of success, and is another promising avenue for optimisation. Finally, multiple adaptive regression splines (MARS) saw superb performance in one study, but requires more research.

At present the literature is comprehensive and diverse, but there are still many potential gaps for interesting future research. Many suspect that the use of pharmacogenetic factors in models improves accuracy, but with the high expense and low availability of genetic

tests in developing regions, there is room for more work in producing models of high precision using only clinical factors. Moreover, the datasets used to train dosing algorithms are exclusive to regions in the developed world, like the United States, Europe, and China, but are used throughout the world. This means that regions of differing ethnicity, such as South Africa, may experience better results from a model trained on a local cohort. In addition to this, training models on the local PathCare dataset may result in entirely novel results being obtained. Finally, one of the most promising gaps in the literature is the topic of deep learning, which has of yet seen no application in warfarin dosing. Given the tremendous effect deep learning has had on other fields, it seems quite likely that it may still offer accuracy gains in the context of warfarin dosing.

REFERENCES

[1] G. W. Albers, J. E. Dalen, A. Laupacis, W. J. Manning, P. Petersen, and D. E. Singer. 2001. Antithrombotic therapy in atrial fibrillation. *Chest* 119, 1 SUPPL. (2001). https://doi.org/10.1378/chest.119.1

[2] Douglas C Anderson. 2009. Pharmacogenomics and warfarin. *American journal of health-system pharmacy : AJHP : official journal of the American Society of Health-System Pharmacists* 66, 2 (1 2009), 121. https://doi.org/10.2146/ajhp080635

[3] Jeffrey L. Anderson, Benjamin D. Horne, Scott M. Stevens, Amanda S. Grove, Stephanie Barton, Zachery P. Nicholas, Samera F.S. Kahn, Heidi T. May, Kent M. Samuelson, Joseph B. Muhlestein, and John F. Carlquist. 2007. Randomized trial of genotype-guided versus standard warfarin dosing in patients initiating oral anticoagulation. *Circulation* 116, 22 (2007), 2563–2570. https://doi.org/10.1161/CIRCULATIONAHA.107.737312

[4] Gerald A. Banet, Amy D. Waterman, Paul E. Milligan, Susan K. Gatchel, and Brian F. Gage. 2003. Warfarin dose reduction vs watchful waiting for mild elevations in the international normalized ratio. *Chest* 123, 2 (2003), 499–503. https://doi.org/10.1378/chest.123.2.499

[5] Balkrishna Charudatta Bangal. 2010. *Automatic generation control of interconnected power systems using artificial neural network techniques*. Ph.D. Dissertation.

[6] Leo Breiman. 1996. Bias, variance, and arcing classifiers. (1996).

[7] Travers Ching, Daniel S Himmelstein, Brett K Beaulieu-jones, Alexandr A Kalinin, Brian T Do, Gregory P Way, Enrico Ferrero, Paul-michael Agapow, Michael Zietz, Michael M Hoffman, Wei Xie, Gail L Rosen, Benjamin J Lengerich, Johnny Israeli, Jack Lanchantin, Stephen Woloszynek, Anne E Carpenter, Avanti Shrikumar, Jinbo Xu, Evan M Cofer, Christopher A Lavender, Srinivas C Turaga, Amr M Alexandari, Zhiyong Lu, David J Harris, Dave Decaprio, Yanjun Qi, Anshul Kundaje, Yifan Peng, Laura K Wiley, Austin Huang, Anthony Gitter, and Casey S Greene. 2018. Opportunities and obstacles for deep learning in biology and medicine. *The Royal Society* (2018). https://doi.org/10.1098/rsif.2017.0387

[8] Hugh A. Chipman, Edward I. George, and Robert E. McCulloch. 2012. BART: Bayesian additive regression trees. *Annals of Applied Statistics* 6, 1 (2012), 266–298. https://doi.org/10.1214/09-AOAS285

[9] Erdal Cosgun, Nita A. Limdi, and Christine W. Duarte. 2011. 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

[10] B. F. Gage, C. Eby, J. A. Johnson, E. Deych, M. J. Rieder, P. M. Ridker, P. E. Milligan, G. Grice, P. Lenzini, A. E. Rettie, C. L. Aquilante, L. Grosso, S. Marsh, T. Langaee, L. E. Farnett, D. Voora, D. L. Veenstra, R. J. Glynn, A. Barrett, and H. L. McLeod. 2008. Use of pharmacogenetic and clinical factors to predict the therapeutic dose of warfarin. *Clinical Pharmacology and Therapeutics* 84, 3 (2008), 326–331. https://doi.org/10.1038/clpt.2008.10

[11] Enzo Grossi, Gian Marco Podda, Mariateresa Pugliano, Silvia Gabba, Annalisa Verri, Giovanni Carpani, Massimo Buscema, Giovanni Casazza, and Marco Cattaneo. 2014. 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

[12] Lutz H Hamel. 2011. *Knowledge discovery with support vector machines*. Vol. 3. John Wiley & Sons.

[13] Takumi Harada, Noritaka Ariyoshi, Hitoshi Shimura, Yasunori Sato, Iichiro Yokoyama, Kaori Takahashi, Shin Ichi Yamagata, Mizuho Imamaki, Yoshio Kobayashi, Itsuko Ishii, Masaru Miyazaki, and Mitsukazu Kitada. 2010. Application of Akaike information criterion to evaluate warfarin dosing algorithm. *Thrombosis Research* 126, 3 (2010), 183–190. https://doi.org/10.1016/j.thromres.2010.05.016

[14] Kurt Hornik, David Meyer, and Alexandros Karatzoglou. 2006. Support vector machines in R. *Journal of statistical software* 15, 9 (2006), 1–28.

[15] Ya Han Hu, Fan Wu, Chia Lun Lo, and Chun Tien Tai. 2012. Predicting warfarin dosage from clinical data: A supervised learning approach. *Artificial Intelligence in Medicine* 56, 1 (2012), 27–34. https://doi.org/10.1016/j.artmed.2012.04.001

[16] Hussain A. Isma'Eel, George E. Sakr, Robert H. Habib, Mohamad Musbah Almedawar, Nathalie K. Zgheib, and Imad H. Elhajj. 2014. Improved accuracy of anticoagulant dose prediction using a pharmacogenetic and artificial neural network-based method. *European Journal of Clinical Pharmacology* 70, 3 (3 2014), 265–273. https://doi.org/10.1007/s00228-013-1617-2

[17] J. A. Johnson, K. E. Caudle, L. Gong, M. Whirl-Carrillo, C. M. Stein, S. A. Scott, M. T. Lee, B. F. Gage, S. E. Kimmel, M. A. Perera, J. L. Anderson, M. Pirmohamed, T. E. Klein, N. A. Limdi, L. H. Cavallari, and M. Wadelius. 2017. Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Pharmacogenetics-Guided Warfarin Dosing: 2017 Update. *Clinical Pharmacology and Therapeutics* 102, 3 (2017), 397–404. https://doi.org/10.1002/cpt.668

[18] Daniel E. Jonas and Howard L. McLeod. 2009. Genetic and clinical factors relating to warfarin dosing. *Trends in Pharmacological Sciences* 30, 7 (2009), 375–386. https://doi.org/10.1016/j.tips.2009.05.001

[19] Sue Jowett, S. Bryan, L. Poller, A. M.H.P. Van Den Besselaar, F. J.M. Van Der Meer, G. Palareti, C. Shiach, A. Tripodi, M. Keown, S. Ibrahim, G. Lowe, M. Moia, A. G. Turpie, and J. Jespersen. 2009. The cost-effectiveness of computer-assisted anticoagulant dosage: Results from the European Action on Anticoagulation (EAA) multicentre study. *Journal of Thrombosis and Haemostasis* 7, 9 (9 2009), 1482–1490. https://doi.org/10.1111/j.1538-7836.2009.03508.x

[20] David M. Keeling, Trevor Baglin, Campbell Tait, Henry Watson, David Perry, Caroline Baglin, Steve Kitchen, and Michael Makris. 2011. Guidelines on oral anticoagulation with warfarin - fourth edition. *British Journal of Haematology* 154, 3 (2011), 311–324. https://doi.org/10.1111/j.1365-2141.2011.08753.x

[21] Y. K. Kim, R. Nieuwlaat, S. J. Connolly, S. Schulman, K. Meijer, N. Raju, S. Kaatz, and J. W. Eikelboom. 2010. Effect of a simple two-step warfarin dosing algorithm on anticoagulant control as measured by time in therapeutic range: A pilot study. *Journal of Thrombosis and Haemostasis* 8, 1 (2010), 101–106. https://doi.org/10.1111/j.1538-7836.2009.03652.x

[22] Stephen E. Kimmel, Benjamin French, Scott E. Kasner, Julie A. Johnson, Jeffrey L. Anderson, Brian F. Gage, Yves D. Rosenberg, Charles S. Eby, Rosemary A. Madigan, Robert B. McBane, Sherif Z. Abdel-Rahman, Scott M. Stevens, Steven Yale, Emile R. Mohler, Margaret C. Fang, Vinay Shah, Richard B. Horenstein, Nita A. Limdi, James A.S. Muldowney, Jaspal Gujral, Patrice Delafontaine, Robert J. Desnick, Thomas L. Ortel, Henny H. Billett, Robert C. Pendleton, Nancy L. Geller, Jonathan L. Halperin, Samuel Z. Goldhaber, Michael D. Caldwell, Robert M. Califf, and Jonas H. Ellenberg. 2013. A Pharmacogenetic versus a Clinical Algorithm for Warfarin Dosing. *New England Journal of Medicine* 369, 24 (2013), 2283–2293. https://doi.org/10.1056/NEJMoa1310669

[23] T B L Kirkwood and S M Lewis. 1983. Requirements for thromboplastins and plasma used to control oral anticoagulant therapy. *WHO Tech Rep Ser* 687 (1983), 81–99.

[24] Altman R.B. Eriksson B.F. Klein, T.E. 2009. Estimation of the Warfarin Dose with Clinical and Pharmacogenetic Data. *Archives of Internal Medicine* 360, 8 (2009), 753–764. https://doi.org/10.1056/NEJMoa0809329.Estimation

[25] G. Le Gal, M. Carrier, S. Tierney, H. Majeed, M. Rodger, and P. S. Wells. 2010. Prediction of the warfarin maintenance dose after completion of the 10 mg initiation nomogram: Do we really need genotyping? *Journal of Thrombosis and Haemostasis* 8, 1 (2010), 90–94. https://doi.org/10.1111/j.1538-7836.2009.03676.x

[26] Rong Liu, Xi Li, Wei Zhang, and Hong Hao Zhou. 2015. 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). https://doi.org/10.1371/journal.pone.0135784

[27] S. McDonald, C. Xydeas, and P. Angelov. 2008. A retrospective comparative study of three data modelling techniques in anticoagulation therapy. *BioMedical Engineering and Informatics: New Development and the Future - Proceedings of the 1st International Conference on BioMedical Engineering and Informatics, BMEI 2008* 1 (2008), 219–225. https://doi.org/10.1109/BMEI.2008.298

[28] D. J. Meier, S. Seva, and William P. Fay. 2007. A comparison of anticoagulation results of patients managed with narrow vs. standard international normalized ratio target ranges [6]. *Journal of Thrombosis and Haemostasis* 5, 6 (2007), 1332–1334. https://doi.org/10.1111/j.1538-7836.2007.02561.x

[29] Liyan Miao, Jian Yang, Chenrong Huang, and Zhenya Shen. 2007. Contribution of age, body weight, and CYP2C9 and VKORC1 genotype to the anticoagulant response to warfarin: Proposal for a new dosing regimen in Chinese patients. *European Journal of Clinical Pharmacology* 63, 12 (11 2007), 1135–1141. https://doi.org/10.1007/s00228-007-0381-6

[30] David Opitz and Richard Maclin. 1999. Popular Ensemble Methods: An Empirical Study. *Journal of Artificial Intelligent Research* 11 (1999), 169–198. https://doi.org/10.1613/jair.614

[31] Munir Pirmohamed, Girvan Burnside, Niclas Eriksson, Andrea L. Jorgensen, Cheng Hock Toh, Toby Nicholson, Patrick Kesteven, Christina Christersson, Bengt Wahlström, Christina Stafberg, J. Eunice Zhang, Julian B. Leathart, Hugo Kohnke, Anke H. Maitland-van der Zee, Paula R. Williamson, Ann K. Daly, Peter

Avery, Farhad Kamali, and Mia Wadelius. 2013. A Randomized Trial of Genotype-Guided Dosing of Warfarin. *New England Journal of Medicine* 369, 24 (2013), 2294–2303. https://doi.org/10.1056/NEJMoa1311386

[32] L. Poller. 2004. International Normalized Ratios (INR): The first 20 years. *Journal of Thrombosis and Haemostasis* 2, 6 (2004), 849–860. https://doi.org/10.1111/j.1538-7836.2004.00775.x

[33] Leon Poller, M. Keown, S. Ibrahim, G. Lowe, M. Moia, A. G. Turpie, C. Roberts, A. M.H.P. Van Den Besselaar, F. J.M. Van Der Meer, A. Tripodi, G. Palareti, C. Shiach, S. Bryan, M. Samama, M. Burgess-Wilson, A. Heagerty, P. Maccallum, D. Wright, and J. Jespersen. 2008. An international multicenter randomized study of computer-assisted oral anticoagulant dosage vs. medical staff dosage. *Journal of Thrombosis and Haemostasis* 6, 6 (6 2008), 935–943. https://doi.org/10.1111/j.1538-7836.2008.02959.x

[34] J R Quinlan. 1992. Learning with continuous classes. *Machine Learning* 92 (1992), 343–348. https://doi.org/10.1.1.34.885

[35] 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, and S. E. Kimmel. 2008. Dosing algorithms to predict warfarin maintenance dose in Caucasians and African Americans. *Clinical Pharmacology and Therapeutics* 84, 3 (2008), 332–339. https://doi.org/10.1038/clpt.2008.101

[36] Elizabeth A. Sconce, Tayyaba I. Khan, Hilary A. Wynne, Peter Avery, Louise Monkhouse, Barry P. King, Peter Wood, Patrick Kesteven, Ann K. Daly, and Farhad Kamali. 2005. The impact of CYP2C9 and VKORC1 genetic polymorphism and patient characteristics upon warfarin dose requirements: Proposal for a new dosing regimen. *Blood* 106, 7 (2005), 2329–2333. https://doi.org/10.1182/blood-2005-03-1108

[37] Ashkan Sharabiani, Adam Bress, Elnaz Douzali, and Houshang Darabi. 2015. Revisiting warfarin dosing using machine learning techniques. *Computational and Mathematical Methods in Medicine* 2015 (2015). https://doi.org/10.1155/2015/560108

[38] Ashkan Sharabiani, Houshang Darabi, Adam Bress, Larisa Cavallari, Edith Nutescu, and Katarzyna Drozda. 2013. Machine learning based prediction of warfarin optimal dosing for African American patients. In *Automation Science and Engineering (CASE), 2013 IEEE International Conference on.* 623–628. https://doi.org/10.1109/CoASE.2013.6653999

[39] Idit Solomon, Nitsan Maharshak, Gal Chechik, Leonard Leibovici, Aharon Lubetsky, Hillel Halkin, David Ezra, and Nachman Ash. 2004. Applying an artificial neural network to warfarin maintenance dose prediction. *Israel Medical Association Journal* 6, 12 (2004), 732–735.

[40] A I Tekkesin, Y Cakilli, C Turkkan, K Demir, E B Borklu, O Yildirimturk, G Cinier, and A T Alper. 2015. Assessment of warfarin TTR (time in therapeutic range) results in a tertiary reference center. *American Journal of Cardiology* 115 (3 2015), S2–S3. https://doi.org/10.1016/J.AMJCARD.2015.01.051

[41] Michael E Tipping. 2001. Sparse Bayesian learning and the relevance vector machine. *Journal of machine learning research* 1, Jun (2001), 211–244.

[42] Vladimir Vapnik. 1998. *Statistical learning theory. 1998.* Wiley, New York.

[43] Mia Wadelius, Leslie Y. Chen, Niclas Eriksson, Suzannah Bumpstead, Jilur Ghori, Claes Wadelius, David Bentley, Ralph McGinnis, and Panos Deloukas. 2007. Association of warfarin dose with genes involved in its action and metabolism. *Human Genetics* 121, 1 (2007), 23–34. https://doi.org/10.1007/s00439-006-0260-8

[44] Mia Wadelius, Leslie Y. Chen, Jonatan D. Lindh, Niclas Eriksson, Mohammed J.R. Ghori, Suzannah Bumpstead, Lennart Holm, Ralph McGinnis, Anders Rane, and Panos Deloukas. 2009. The largest prospective warfarin-treated cohort supports genetic forecasting. *Blood* 113, 4 (2009), 784–792. https://doi.org/10.1182/blood-2008-04-149070

[45] P S Wells, A M Holbrook, N R Crowther, and J Hirsh. 1994. Interactions of warfarin with drugs and food. , 676–683 pages. https://doi.org/10.7326/0003-4819-121-9-199411010-00009

[46] M Whirl-Carrillo, EM McDonogh, J Herbet, L Gong, K Sangkuhl, C Thotn, R Altman, and E Klein. 2012. Pharmacogenomics Knowledge for Personlized Medicine. *Clinical Pharmacology and Therpeutics* 92, 4 (2012), 414–417. https://doi.org/10.1038/clpt2012.96.Pharmacogenomics

[47] Qin Zhou, Joey Kwong, Jie Chen, Wenzhe Qin, Jin Chen, and Li Dong. 2014. Use of artificial neural network to predict warfarin individualized dosage regime in Chinese patients receiving low-intensity anticoagulation after heart valve replacement. *International Journal of Cardiology* 176, 3 (10 2014), 1462–1464. https://doi.org/10.1016/j.ijcard.2014.08.062