UNIVERSITÀ DEGLI STUDI DI MILANO

Facoltà di Medicina e Chirurgia

Scuola di Dottorato in Metodologia Clinica

Dipartimento di Scienze della Salute

Venticinquesimo Corso di Dottorato

Settore Scientifico Disciplinare MED 09

PREDICTION OF OPTIMAL WARFARIN MAINTENANCE DOSE USING ADVANCED ARTIFICIAL NEURAL NETWORKS

Dr. GianMarco Podda
Matr. n. R08636

Tutor:
Dott.ssa Elena M. Faioni

Coordinatore del Dottorato:
Chiar.mo Prof. Marco Cattaneo

Anno Accademico 2011-2012

INDEX

SUMMARY

Introduction. The individual response to vitamin K antagonists (VKA) is highly variable, being influenced by clinical factors and genetic variants of enzymes that are involved in the metabolism of VKA (CYP2C)) and vitamin K (VKORC1). Currently, the dose of VKA is adjusted based on measurements of the prothrombin time. In the last years, mathematical algorithms were developed for estimating the appropriate VKA dose, based on different mathematical approaches working on clinical and genetic data. Artificial Neural Networks (ANN) are computerized algorithms resembling interactive processes of the human brain, which allow to study very complex non-linear phenomena like biological systems. Aim. To evaluate the performance of new generation ANN on a large data base of patients on chronic VKA treatment. Methods. Clinical and genetic data from 377 patients (186 m; 191 f) treated with a VKA (warfarin) average weekly maintenance dose (WMD) of 23.7 mg (11.5 SD) were used to create a dose algorithm. Forty-eight variables, including demographic, clinical and genetic data (5 CYP2C9 and 3 VKORC1 genetic variants) were entered into Twist® system, which can select fundamental variables during their evolution in search for the best predictive model. The final model, based on 23 variables expressed a functional approximation of the actual dose within a validation protocol based on a tripartite division of the data set (training, testing, validation). Results. In the validation cohort, the pharmacogenetic algorithm reached high accuracy, with an average absolute error of 5.7 mg WMD. In the subset of patients requiring ≤21 mg (45 % of the cohort) and 21-49 mg (51 % of the cohort) the absolute error was 3.86 mg and 5.45 with a high percentage of subjects being correctly identified (72%, 74% respectively). Conclusion. ANN can be applied successfully for VKA maintenance dose prediction and represent a robust basis for a prospective multicentre clinical trial of the efficacy of genetically informed dose estimation for patients who require VKA.

INTRODUCTION

Vitamin K antagonists (VKA) are widely used in primary and secondary prophylaxis of venous thromboembolism and cardioembolism (Ansell J, 2008). The use of VKA in clinical practice is challenging because VKAs: i) have a narrow therapeutic window ii) have a pronounced inter-individual variability in dose requirement due to demographic and clinical factors, such as age, weight, height, comorbid medical conditions, concurrent medications, diet, ethnicity (Hirsh J, 2003), and genetic factors (Loebstein R, 2001). Among the latter, which are responsible for approximately 30-50% of response variability (Lubitz SA, 2010), polymorphisms of two enzymes are strongly associated with dose requirements of VKAs: isoform CYP2C9 of cytochrome P450 (CYP) and vitamin K epoxide reductase complex subunit 1 (VKORC1). CYP2C9 is the major enzyme involved in the clearance of warfarin, while VKORC1 is essential for the regeneration of reduced vitamin K, which catalyzes the γ-carboxylation of coagulation factors II, VII, IX and X, and other proteins. For the above reasons, treatment with VKA is routinely monitored by the prothrombin time, expressed as International Normalized Ratio (INR), with the aim of constantly maintaining the degree of inhibition of coagulation within the therapeutic range. Due to the many variables that are involved, it is difficult to predict the daily maintenance dose of warfarin, which varies by a factor of 10 among patients (Takahashi H, 2001). The consequences of incorrect dosing can be dramatic, contributing to the development of thromboembolic events, in case of under-dosing, or bleeding complications, in case of over-dosing (Wysowski DK, 2007). It has been reported that warfarin is the most common drug implicated in emergency room visits for adverse drug events in older adults in the United States (U.S. FDA, 2008). Although new oral anticoagulants (NOA) with favourable risk-to-benefit ratios and no need for laboratory moniroring (Tripodi A, 2013), are being gradually introduced in the clinical practice, they will

probably never completely replace VKA. Therefore, there still is a need to improve our ability to predict the effective dose of VKA in each individual patient.

Among Caucasians, the most frequent CYP2C9 variant genotypes are CYP2C9*2 (12-27%) and CYP2C9*3 (8-29%)(Margaglione M, 2000; Scordo MG, 2001; Sanderson S, 2005; Spreafico M, 2008; Sipeky C 2009; Jorgensen AL, 2012), which are associated with reduced clearance of warfarin and explain 17-30% of warfarin dose variability (McClain MR, 2008; Wadelius M, 2009, Jorgensen AL, 2012). VKORC1 variant genotypes are also common among Caucasians and are associated with increased sensitivity to warfarin (Geisen C, 2005), accounting for approximately 23% of warfarin dose variability (Geisen C 2005; McClain MR, 2012,). Warfarin dosing algorithms have been developed to improve our ability to predict the effective warfarin dose in each individual patient. Genotype-based dosing algorithms predict 37-55% of the variation in warfarin dose requirements; neither the addition of race, number of concurrent medications nor the number of concurrent medications interacting with warfarin enhanced algorithm performance (Lubitz SA, 2010).

Artificial neural network (ANN) are computerized algorithms resembling interactive processes of human brain, which allow studying very complex non-linear phenomena like biological systems. ANNs can learn experiential knowledge expressed through internal connections in a similar way as neurons in the brain and this knowledge can be made available for clinical decision-making (Cross SS, 1995; Dayhoff JE, 2001). Given the complexity of multiple factors interplay behind optimal warfarin response these tool might offer a substantial advantage over other statistical approaches. The aim of this study was to evaluate the performance of new generation ANNs to predict the maintenance dose of warfarin, using a large database of patients on chronic warfarin treatment

METHODS

Patient characteristics

Data from patients who were treated with warfarin between December 2008 and February 2009 were included in the analysis after having obtained written informed consent. The patients were recruited at the Ospedale San Paolo, Università degli Studi di Milano, Milan, Italy. Written informed consent was obtained from all patients and the study was approved by the Ethics Committee of Ospedale San Paolo. The study was conducted in accordance with the Helsenki Declaration.

Patients were enrolled in the study if they were Caucasians, ≥18 years of age, treated with warfarin for any clinical indication, and if their INR values had been in the therapeutic INR range at the last two consecutive laboratory controls. Exclusion criteria were severe liver disease and renal failure. The warfarin maintenance dose in each patient had been identified by specialized medical doctors with the help of a computerized algorithm (Parma 4.1, Instrumentation Laboratory Company, MA, USA) (Manotti C, 2001), which is based on a regression model that considers the weekly variations of INR. Several studies showed that Parma 4.1 significantly improves the quality of treatment, compared to the doctor-prescribed method (Ageno W, 1998; Manotti C, 2001; Poller L, 2004).

Data collection

Demographic data (age, gender, height, weight, body mass index), indication for anticoagulant therapy, concomitant diseases, co-medications, INR target range, warfarin maintenance dose were collected. Concomitant medications were: antihypertensives (ACE-inhibitors, Angiotensin II receptor antagonists, alpha-blockers, vasodilators and dihydropyridines calcium channel blockers), diuretics (loop, thiazide and potassium-sparing diuretics: diuretics), antithrombotic drugs (P2Y12 antagonists and aspirin), proton pump inhibitors,

statins, oral glucose lowering agents, allopurinol, antibiotics, carbamazepine, tapazole, antiarrhythmic drugs (Class I, II, III except amiodarone, IV and V), amiodarone and other drugs (i.e. antidepressant, antipsychotic, anticonvulsant except carbamazepine, L-thyroxin and analgesic).

Laboratory methods

Venous blood was collected in 0.109 M sodium citrate (9:1 vol:vol) and in K_3-EDTA tubes. Plasma and cells were separated after centrifugation at 4000 g for 30 min. DNA was isolated from leukocytes using an automated extraction system (Abbott m2000sp, Abbott Molecular Inc., IL, USA) following the manufacturer's recommendations. The DNA samples were eluted in 100 µl water and stored at –20°C until processed.

The DNA concentration in the samples ranged between 25 and 40 ng/µl, with a purity (measured by the 269/280 ratio) ranging between 1.4 and 1.7, which was considered adequate for further analysis. PT was measured using the human recombinant thromboplastin RecombiPlasTin (HemosILTM, Instrumentation Laboratory, USA), and an Electra 1600 coagulometer (MA, USA). The laboratory performed regular external quality control exercises.

Five common CYP2C9 (*2, *3, *5, *6, *11) and 3 common VKORC1 (3673G→A, 6484C→T, 6853G→C) single nucleotide polymorphisms (SNPs) were assessed through allele specific primer extension PCR. CYP2C9 allele designations refer to those defined by the Cyto-chrome P450 Allele Nomenclature Committee (http://www.cypalleles.ki.se). Genotype analysis was performed using INFINITI™ Analyzer, which is an automated, multiplexing, continuous flow, random access microarray platform (Autogenomics, CA, USA).

Input variables

A total of 48 clinical and genetic variables, known to influence the pharmacological response to warfarin, were available for mathematical modelling (Table 1).

Statistical analysis

Data are shown as means and standard deviations (SD) or medians and interquartile ranges (IQR), according to their distribution, which was evaluated using the Kolmogorov-Smirnov test. The Spearman's correlation test was used to analyse the data. P values <0.05 were considered statistically significant.

The sample selected for the analysis (n = 377) is relatively large, as it is required by the nature of the research question, to allow enough variability to make meaningful inferences as to the predictive capacity of the single variables. Multivariate analysis was carried out with supervised ANN, according to the method already adopted (Penco S, 2008). The choice of a relatively unusual and sophisticated inferential technique such as ANN is motivated by the fact that the underlying relation to be estimated among our independent sample variables and the dependent variable (the optimal dose of warfarin) is extremely complex and there is no reliable a-priori statistical model to refer to. ANNs self-adjust their structure as they learn from their own errors, are able to handle a very high number of variables simultaneously, irrespective of their underlying degree of non-linearity, and lead to structurally robust results even when the underlying statistical process is not well understood, thereby allowing to deal with many sources of inferential trouble such as outliers, collinear interactions among variables and hidden variables (Buscema M, 1998).

In particular, we work with the family of Supervised ANNs, that is to say, with ANN that tackle problems where an external, objective target output can be fixed, so that they learn by examples (the training set, that is, a suitable sub-sample of the whole database), calculating an error function during the training phase, and adjusting the connection strengths in order to minimize the error function until a satisfactory and stable level of accuracy in the

prediction/classification task is reached. This type of ANNs thus computes a function of the form: y = f(x,w*), where x is the input, y is the output and w* is the set of ANN weights (the function parameters) that encode the ANN's approximate reconstruction of the structure of the function.

In order to cut down of the number of irrelevant variables in the database (i.e., the variables that do not carry any meaningful information for the prediction task), which cause a loss in the power of our inferences, we have employed a special 'artificial organism' called TWIST (Buscema M, 2005), suitably designed for sorting out the most relevant variables for the sake of prediction/classification. It consists of a combination of two already known systems: T&T and IS. The T&T system is a robust data re-sampling technique that is able to arrange the source sample into sub-samples, all of which possessing a similar probability density function. In this way, the database is split into two or more sub-samples in order to train, test and validate the ANN models as effectively as possible on the basis of the available data. The IS system is an evolutionary 'wrapper' system that selects variables in order to minimize their number while preserving the actual amount of task-relevant information contained in the data-set. The combined action of these two systems allows us to increase substantially the inferential power of our ANN system, while circumventing at the same time a few major technical issues. Both systems are based on a Genetic Algorithm, the Genetic Doping Algorithm (GenD) developed at Semeion Research Centre (Rome, Italy) (Buscema M, 2004). The TWIST system is described in detail in the appendix, and Figure 1 below is a snapshot of TWIST at work during the variables selection task.

The TWIST pre-processing singles out the variables that prove to be most significant for the prediction/classification task, while producing at the same time the training set and the testing set, which are extracted from a probability distribution very close to the one that provided the best performance in the task. As to the prediction/classification task, it is carried

out by means of a supervised, Multi Layer Perceptron, with four hidden units (Haykin, 1998). The protocol scheme is reported in figure 1. The patient population was divided into three groups according to weekly maintenance warfarin dose: ≤ 21 mg, 21-49 mg and ≥49 mg.

The warfarin dose that was predicted by ANN was compared with the actual doses by univariate linear regression. The mean absolute error (MAE), the mean of the absolute difference between the predicted and actual dose, and the coefficient of determination (R^2) were used to measure the predictive accuracy of ANN . Analysis were performed using SPSS 18.0 (Inc., IL, USA).

RESULTS

Patients' characteristics

A total of 377 patients met the inclusion criteria and were included in the analysis. The most relevant clinical characteristics of the enrolled patients are described in table 2. The most frequent clinical indication for anticoagulation was atrial fibrillation (69%); other indications included heart valve prosthesis (10%) and pulmonary embolism (8%). The large majority of patients, 325, 86%) were on concurrent drug treatment: on average, they were taking 3 (IQR 1-4) medications potentially interacting with warfarin. The median weekly maintenance dose (WMD) of warfarin was 22.5 mg (IQR 16.3-28.8mg). Thirteen patients whose INR values were not within the therapeutic range were erroneously included in the analysis: their median weekly maintenance dose was 21.4 mg (IQR 12.2-30.0 mg), the INR was higher than 3.0 (INR 3.7 and 4.3) in 2, and lower than 2.0 in 11 (median INR 1.5, IQR 1.5-1.7). We believe that this mistake cannot modify the results of the study: the difference between the warfarin weekly dose determining the target INR and the wrong one was low: 1.7 mg (21.4 mg vs 19.7 mg).
The frequencies of CYP2C9 and VKORC1 genotypes are reported in table 3. Variant VKORC1 genotypes were present in about 70% of patients, while variant CYP2C9 genotypes were present in 38%.

Effect of variables on warfarin maintenance dose

Age and atrial fibrillation were inversely associated with the average therapeutic dose of warfarin, while weight, height, BMI, deep vein thrombosis and other clinical indications for treatment were directly associated with it (Table 4)

The average WMD of warfarin was positively associated with the wilde type genotype of CYP2C9 (*1/*1) and negatively associated with genotypes *1/*3 and *2/*3. Our results are in line with a recent meta-analysis that showed that, compared with subjects with wild

genotype (*1/*1), WMD of warfarin is 22%, 36%, 43%, 53%, and 76% lower among individuals with *1/*2, *1/*3, *2/*2, *2/*3, and *3/*3 genotypes (McClain MR, 2008). Wilde-type VKORC1 genotypes (3673GG, 6484CC, 6853GG) were posively associated with the maintenance dose of warfarin, while genotype variants 3673AA, 6484TT and 6853CC were inversely associated with it. Our results are consistent with previous reports (Geisen C, 2005; McClain AR, 2008; Spreafico M, 2008; Wadelius M, 2009; Jorgensen AL, 2012). No statistically significant associations with warfarin dose and concomitant medications were found.

Artificial neural network analysis

Forty-eight variables, including demographic, clinical and genetic data (5 CYP2C9 and 3 VKORC1 genetic variants) were entered into the Twist® system, which can select fundamental variables during their evolution in search for the best predictive model. The twist system selected 23 variables (table 5) carrying the maximal amount of information to build up a predictive model. The final model, based on these 23 variables, expressed a functional approximation of the actual dose of warfarin within a validation protocol based on a tripartite division of the data set (training, testing, validation). In this procedure the study sample was randomly divided into two main sub-samples: the tuning set sub-sample and the prediction sub-sample, which accounted for about 50% of the tuning sample. The tuning data set was in turn subdivided in two halves, the training sample and the testing sample. This was done five times. During the training phase, the ANN learned a model of data distribution and then, on the basis of such a model, blindly made a functional approximation of dependent variable in the testing set. Training and testing sets were then reversed and consequently 10 analyses for every model employed were conducted. The best performing model was then selected and used to predict the warfarin dose in the prediction set, which had been taken apart during all the procedure. Also in this case the prediction was done blindly.

Finally, the pharmacogenetic algorithm obtained by the average of four independent ANN reached an average absolute error of 5.7 mg with a R^2 =478 (table 6). In the subsets of patients requiring ≤21 mg (45 % of the cohort), 21-49 mg (51 % of the cohort) and ≥49 mg warfarin(4 % of the cohort) the absolute error was 3.86 mg, 5.45 mg and 24.2 mg and the percentage of subjects being correctly identified was 72%, 74% and 0% respectively.

DISCUSSION

The number of patients on oral anticoagulant therapy continues to increase worldwide. Warfarin remains the most frequently prescribed oral anticoagulant drug because of its low cost, proven efficacy and the availability of antidotes to treat patients with bleeding complications. Because genetic variation contributes to the observed high inter-individual variability in dose-requirements, which can potentially increase the risk of thrombosis or bleeding, the US Food and Drug Administration (FDA) supported the use of genotyping to guide warfarin dosing (http://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm). In the present study, we tested the efficacy of warfarin dosing during maintenance treatment, based on a large number of genetic and clinical factors that were analyzed by artificial neural networks (ANN), which are effective tools for clinical decision-making. In the validation cohort, our model displayed high accuracy and explained an average absolute error of 5.7 mg/week and from 50% to 71% of weekly warfarin dose variability. The dose prediction was particularly accurate for the subset of patients requiring weekly doses of warfarin ≤21 mg (correct dose identification in 72% of patients) or 22-48 mg (74%). In the subset of patients requiring higher weekly warfarin doses (≥49 mg), the algorithm did not perform successfully, most likely because only 15 patients belonged to this subgroup. Previous studies reported on the peformance of pharmacogenetic algorithms in warfarin dosing (Lenzini P, 2010; IVPC 2009, Gong IY, 2011; Zambon CF, 2011). Table 7 shows that the performance of our algorithm appears to be superior to those of the 2 published studies that mostly enrolled Caucasian patients (Table 7) (IVPC 2009, Zambon CF, 2011). The results of the other published studies can be hardly compared to ours, because they mostly enrolled non-Caucasian patients, who have different prevalences of CYP2C9 and VKORC1 polymorphisms (Jorgensen AL, 2012). The higher performance of our model may be explained

by the high accuracy of ANN to identify complex, non-linear relationships among genetic and clinical variables in a global analysis and by the much higher number of variables (23) that were retained in our final model, compared to other models (Moreno L, 1995).

Considering the promising results of our study, we believe that the performance of the ANN model should be tested a prospective study assessing: i) the time within therapeutic INR range during VKA therapy, ii) the number of thromboembolic and bleeding events and, iii) cost effectiveness, compared to the best care. A randomized trial compared standard dosing regimen with two genotype-guided algorithms, based on the IWPC (IWPC 2009) and Gage algorithms (Gage BF, 2008) that incorporate both clinical and genetic factors. Primary outcomes were percentage of out-of-range INR at 1 and 3 months and percentage of time in therapeutic range. The combined genotype-guided prescription cohort demonstrated superior outcomes with respect to both primary end points at 3 months (30% vs 42% for out-of-range and 71% vs 59% for therapeutic range). Moreover, serious events were significantly less frequent in the genotype-guided cohort (4.5% vs 9.4% of patients; $p < 0.001$). It should be noted that there was no difference in the primary outcome between the two genotype-based algorithms.

In the last years, new oral direct anticoagulants (NOA) have been developed as alternatives to warfarin, with at least equal efficacy and safety, wider therapeutic range, less complex pharmacodynamics, and no need for laboratory monitoring. However, there are patients, like those with severe renal function impairment, who cannot be treated with NOA. Moreover, NOA are very expensive and no specific antidotes to treat patients with bleeding complications have been developed yet. Therefore, it can be easily predicted that replacement of warfarin by NOA will be a slow and partial process. As a consequence, efforts are still needed to improve our ability to predict the correct dose of warfarin.

Our study has several limitations: (i) being a retrospective study, we were only able to compare the predicted dose with the really administered dose; moreover the study methodology do not allow to consider clinical outcomes like adverse events; (ii) it was conducted in a single centre and the small sample size and the homogeneous characteristics of this population limit the external validity; (iii) the induction period of warfarin therapy was not evaluated.

In conclusion, the ANN model that was used in this study appears to be an accurate tolls for the identification of the warfarin maintenance dose, at least for Caucasians patients. Thus our results suggest that ANN can be applied successfully for VKA maintenance dose prediction and represent a robust basis for a prospective multicentre clinical trial of the efficacy of genetically informed dose estimation for patients who require VKA.

REFERENCES

Ageno W, Turpie AG. A randomized comparison of a computer-based dosing program with a manual system to monitor oral anticoagulant therapy. Thromb Res. 1998 Sep 1;91(5):237-40.

Ansell J, Hirsh J, Hylek E, Jacobson A, Crowther M, Palareti G; American College of Chest Physicians. Pharmacology and management of the vitamin K antagonists: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). Chest. 2008 Jun;133(6 Suppl):160S-198S.

Buscema M. Artificial neural networks and complex social systems. I. Theory. Subst Use Misuse. 1998 Jan;33(1):v-xvii, 1-220.

Buscema M, Grossi E, Snowdon D, Antuono P, Intraligi M, Maurelli G, Savarè R. Artificial neural networks and artificial organisms can predict Alzheimer pathology in individual patients only on the basis of cognitive and functional status. Neuroinformatics. 2004;2(4):399-416.

Buscema M, Grossi E, Intraligi M, Garbagna N, Andriulli A, Breda M. An optimized experimental protocol based on neuro-evolutionary algorithms application to the classification of dyspeptic patients and to the prediction of the effectiveness of their treatment. Artif Intell Med. 2005 Jul;34(3):279-305.

Cross SS, Harrison RF, Kennedy RL. Introduction to neural networks. Lancet. 1995;346:1075–1079.

Dayhoff JE, Deleo JM. Artificial neural network: opening the black box. Cancer. 2001;91:1615–1635.

Gage BF, Eby C, Johnson JA, Deych E, Rieder MJ, Ridker PM, Milligan PE, Grice G, Lenzini P, Rettie AE, Aquilante CL, Grosso L, Marsh S, Langaee T, Farnett LE, Voora D, Veenstra DL, Glynn RJ, Barrett A, McLeod HL. Use of pharmacogenetic and clinical factors to predict the therapeutic dose of warfarin. Clin Pharmacol Ther. 2008 Sep;84(3):326-31.

Geisen C, Watzka M, Sittinger K, Steffens M, Daugela L, Seifried E, Muller CR, Wienker TF, Oldenburg J. VKORC1 haplotypes and their impact on the inter-individual and inter-ethnical variability of oral anticoagulation. Thromb Haemost. 2005 Oct;94(4):773-9.

Gong IY, Tirona RG, Schwarz UI, Crown N, Dresser GK, Larue S, Langlois N, Lazo-Langner A, Zou G, Roden DM, Stein CM, Rodger M, Carrier M, Forgie M, Wells PS, Kim RB. Prospective evaluation of a pharmacogenetics-guided warfarin loading and maintenance dose regimen for initiation of therapy. Blood. 2011 Sep 15;118(11):3163-71.

Haykin, S. Neural Networks: A Comprehensive Foundation, second edition 1998 (New York: Prentice Hall).

Hirsh J, Fuster V, Ansell J, Halperin JL; American Heart Association/American College of Cardiology Foundation.Hirsh J, Fuster V, Ansell J, Halperin JL; American Heart Association/American College of Cardiology Foundation guide to warfarin therapy. J Am Coll Cardiol. 2003 May 7;41(9):1633-52.

Jorgensen AL, Fitzgerald RJ, Oyee J, Pirmohamed M, Williamson PR. Influence of CYP2C9 and VKORC1 on Patient Response to Warfarin: A Systematic Review and Meta-Analysis. PLoS One. 2012;7(8):e44064.

Lenzini P, Wadelius M, Kimmel S, Anderson JL, Jorgensen AL, Pirmohamed M, Caldwell MD, Limdi N, Burmester JK, Dowd MB, Angchaisuksiri P, Bass AR, Chen J, Eriksson N, Rane A, Lindh JD, Carlquist JF, Horne BD, Grice G, Milligan PE, Eby C, Shin J, Kim H, Kurnik D, Stein CM, McMillin G, Pendleton RC, Berg RL, Deloukas P, Gage BF. Integration of genetic, clinical, and INR data to refine warfarin dosing. Clin Pharmacol Ther. 2010 May;87(5):572-8.

Loebstein R, Yonath H, Peleg D, Almog S, Rotenberg M, Lubetsky A, Roitelman J, Harats D, Halkin H, Ezra D. Interindividual variability in sensitivity to warfarin--Nature or nurture? Clin Pharmacol Ther. 2001 Aug;70(2):159-64.

Lubitz SA, Scott SA, Rothlauf EB, Agarwal A, Peter I, Doheny D, Van Der Zee S, Jaremko M, Yoo C, Desnick RJ, Halperin JL. Comparative performance of gene-based warfarin dosing algorithms in a multiethnic population. J Thromb Haemost. 2010 May;8(5):1018-26.

Manotti C, Moia M, Palareti G, Pengo V, Ria L, Dettori AG. Effect of computer-aided management on the quality of treatment in anticoagulated patients: a prospective, randomized, multicenter trial of APROAT (Automated PRogram for Oral Anticoagulant Treatment). Haematologica. 2001 Oct;86(10):1060-70.

Margaglione M, Colaizzo D, D'Andrea G, Brancaccio V, Ciampa A, Grandone E, Di Minno G.. Genetic modulation of oral anticoagulation with warfarin. Thromb Haemost 2000;84:775–778.

McClain MR, Palomaki GE, Piper M, Haddow JE. A rapid-ACCE review of CYP2C9 and VKORC1 alleles testing to inform warfarin dosing in adults at elevated risk for thrombotic events to avoid serious bleeding. Genet Med. 2008 Feb;10(2):89-98.

Moreno L, Pineiro JD, Sanchez JL, Manas S, Merino J, Acosta L, Hamilton A. Brain maturation estimation using neural classifier. IEEE Trans Biomed Eng. 1995 Apr;42(4):428-32.

Penco S, Buscema M, Patrosso MC, Marocchi A, Grossi E. New Application of Intelligent Agents in Sporadic Amyotrophic Lateral Sclerosis Identifies Unexpected Specific Genetic Background. BMC Bioinformatics. 2008, 9, 254-266.

Poller L. International Normalized Ratios (INR): the first 20 years. J Thromb Haemost. 2004 Jun;2(6):849-60.

Sanderson S, Emery J, Higgins J. CYP2C9 gene variants, drug dose, and bleeding risk in warfarin-treated patients: a HuGEnet systematic review and meta-analysis. Genet Med. 2005 Feb;7(2):97-104.

Scordo MG, Aklillu E, Yasar U, Dahl ML, Spina E, Ingelman-Sundberg M., , Allele and genotype frequencies of CYP2C9, CYP2C19 and CYP2D6 in an Italian population. 2004 Br J Clin Pharmacol. 2001 Oct;52(4):447-50.

Sipeky C, Lakner L, Szabo M, Takacs I, Tamasi V, Polgar N, Falus A, Melegh B. Interethnic differences of CYP2C9 alleles in healthy Hungarian and Roma population samples: elationship to worldwide allelic frequencies. Blood Cells Mol Dis. 2009 Nov-Dec;43(3):239-42.

Spreafico M, Lodigiani C, van Leeuwen Y, Pizzotti D, Rota LL, Rosendaal F, Mannucci PM, Peyvandi F. Effects of CYP2C9 and VKORC1 on INR variations and dose requirements during initial phase of anticoagulant therapy. Pharmacogenomics. 2008 Sep;9(9):1237-50.

Takahashi H, Echizen H. Pharmacogenetics of warfarin elimination and its clinical implications. Clin Pharmacokinet. 2001;40(8):587-603.

Tripodi A. The Laboratory and the New Oral Anticoagulants. 2013; Clin Chem. 2013 Feb;59(2):353-62.

U.S. FDA. Clinical Pharmacology Subcommittee of the Advisory Committee for Pharmaceutical Science 2005, Warfarin. http://www.fda.gov/ohrms/dockets/ac/cder05.html#PharmScience. Accessed on May 25, 2008.

Wadelius M, Chen LY, Lindh JD, Eriksson N, Ghori MJ, Bumpstead S, Holm L, McGinnis R, Rane A, Deloukas P. The largest prospective warfarin-treated cohort supports genetic forecasting. Blood. 2009 Jan 22;113(4):784-92.

Wysowski DK, Nourjah P, Swartz L. Bleeding complications with warfarin use: a prevalent adverse effect resulting in regulatory action. Arch Intern Med. 2007 Jul 9;167(13):1414-9.

Zambon CF, Pengo V, Padrini R, Basso D, Schiavon S, Fogar P, Nisi A, Frigo AC, Moz S, Pelloso M, Plebani M. VKORC1, CYP2C9 and CYP4F2 genetic-based algorithm for warfarin dosing: an Italian retrospective study. Pharmacogenomics. 2011 Jan;12(1):15-25.

TABLES

Table 1: Variables that have been considered in our analysis

1	Male gender	18	CYP2C9*2*3	35	Antihypertensive drugs
2	Female gender	19	CYP2C9*3*3	36	Antiarrhythmic drugs
3	Age	20	CYP2C9*1*5	37	Diuretics
4	Weight	21	CYP2C9*5*5	38	Amiodarone
5	Height	22	CYP2C9*1*6	39	Antithrombotic drugs
6	Body Mass Index	23	CYP2C9*6*6	40	Proton pump inhibitors
7	Deep vein thrombosis	24	CYP2C9*1*11	41	Statins
8	Pulmunary embolism	25	CYP2C9*11*11	42	Oral glucose lowering agents
9	Atrial fibrillation	26	VKORC1 3673 GG	43	Allopurinol
10	Heart Valve	27	VKORC1 3673 GA	44	Antibiotics
11	Cardiomyopathy	28	VKORC1 3673 AA	45	Carbamazepine
12	Stroke	29	VKORC1 6484 CC	46	Tapazole
13	Other indications	30	VKORC1 6484 CT	47	Other drugs
14	CYP2C9*1*1	31	VKORC1 6484 TT	48	Warfarin weekly maintenance dose
15	CYP2C9*1*2	32	VKORC1 6853 GG		
16	CYP2C9*2*2	33	VKORC1 6853 GC		
17	CYP2C9*1*3	34	VKORC1 6853 CC		

Table 2: Patients' characteristics

Variables	
Age(years), median (IQR)	76 (70-80)
Men, n (%)	186 (49%)
Body mass index, median (IQR)	18.0 (16-20)
Indication for anticoagulation, n (%)	
- Atrial fibrillation	261 (69%)
- Artificial heart valves	39 (10%)
- Pulmonary embolism	30 (8%)
- Deep vein thrombosis	21 (6%)
- Cardiomyopathy	15 (4%)
- Stroke	13 (3%)
- Other indications	14 (4%)
Number of concomitant medications per patient, median (IQR)	3 (1-4)
Patients with no concomitant medications, n (%)	53 (14%)
Numebr (%) of patients on concomitant medication with, :	
- Antihypertensive drugs	209 (55%)
- Antiarrhythmic drugs	192 (51%)
- Diuretics	133 (35%)
- Amiodarone	69 (18%)
- Antithrombotic drugs	68 (18%)
- Proton pump inhibitors	58 (15%)
- Statins	50 (13%)
- Oral glucose lowering agents	32 (9%)
- Allopurinol	18 (5%)
- Antibiotics	13 (3%)
- Carbamazepine	1 (0.3%)
- Tapazole	1 (0.3%)
- Other drugs	123 (32%)
Weekly mean therapeutic warfarin dose, mg per week (±SD)	23.7 mg (±11.5)
Weekly median therapeutic warfarin dose, mg per week (IQR)	22.5 mg (15.6-29.4)

SD= standard deviation; IQR= interquartile range.

Table 3

Prevalences of genetic CYP2C9 and VKORC1 genotypes, and allele frequency of CYP2C9

Gene	Genotype	n (%)	Gene	Genotype	n (%)
CYP2C9	*1/*1	235 (62.3%)	VKORC1	3673GG	111 (29%)
	*1/*2	85 (22.5%)		3673GA	189 (50%)
	*1/*3	44 (11.7%)		3673AA	77 (20%)
	*2/*2	4 (1.1%)		6484CC	107 (28%)
	*2/*3	7 (1.8%)		6484CT	195 (52%)
	*3/*3	1 (0.3%)		6484TT	75 (20%)
	*11/*11	1 (0.3%)		6853GG	101 (27%)
				6853GC	195 (51%)
				6853CC	81 (22%)
CYP2C9	Allele	n (%)			
	*1	599 (79.4%)			
	*2	100 (13.3%)			
	*3	53 (7%)			
	*11	2 (0.3%)			

Table 4
Correlation between the weekly therapeutic warfarin dose with clinical and genetic variables.

Variables	r	p
VKORC1 6484TT	-0.31	<0.01
VKORC1 6853 CC	-0.31	<0.01
VKORC1 3673 AA	-029	<0.01
Age	-0.21	<0.01
CYP2C9*1/*3	-0.14	<0.01
CYP2C9*2/*3	-0.14	<0.01
Atrial fibrillation	-0.14	0.006
VKORC1 6484 CC	0.33	<0.01
VKORC1 3673 GG	0.33	<0.01
VKORC1 6853 GG	0.30	<0.01
Body Mass Index	0.21	<0.01
Weight	0.21	<0.01
Height	0.15	<0.05
Deep vein thrombosis	0.15	<0.05
CYP2C9*1/1	0.13	<0.05
Other indications for Warfarin	0.12	<0.05

Spearman test

Table 5. Variables that were selected by the TWIST system

1	Male gender	13	C9*3/*3
2	Female gender	14	C9*2/*3
3	Height	15	VKR3673 AA
4	BMI	16	VKR6484 CT
5	Deep venous thrombosis	17	VKR6853 GC
6	Pulmunary embolism	18	VKR6853 CC
7	Cardiomyopathy	19	Amiodarone
8	Stroke	20	Diuretics
9	C9*1/*1	21	Statins
10	C9*1/*2	22	Tapazole
11	C9*1/*6	23	Other drugs
12	C9*1/*11		

Table 6
Results obtained by ANN on the validation cohort

ANN	n (%)	MAE (mg)*	R^2	% correctly classified
≤ 21 mg/week	170 (45)	3.86 (0.92)	0.67	72
22-48 mg/week	193 (51)	5.45 (1.02)	0.51	74
≥49 mg/week	14 (4)	24.2 (6.3)	0.15	0
Total	377 (100)	5.72 (0.94)	0.48	70

MAE: mean absolute error; R^2 is the coefficient of determination; %= percentage of patients predicted dose within 20% of the actual dose; ANN: artificial network analysis; *mean (SD)

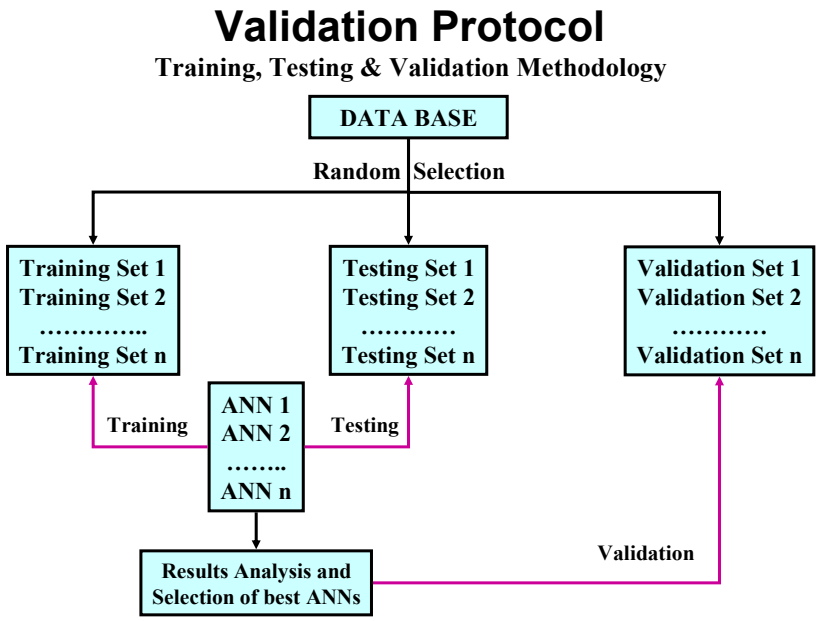
Table 7. Comparison of the performances of different pharmacogenetic algorithms

Study	n	Caucasians (%)	R^2	MAE (mg)	% of patients who were correctly classified
This study	377	100	0.48	5.72	70
IWPC (IWPC, 2009)	1009	56	0.43	8.5	≈45
Zambon (Zambon, 2011)	97	100	0.56	7.0	52

n: number of subjects included; MAE: mean absolute error; R^2 is the coefficient of determination; %= percentage of patients for whom the predicted dose within 20% of the actual dose

FIGURE

Figure 1: validation protocol employed for neural network analysis of the 377 patients



APPENDIX

The TWIST methodology

TWIST (Buscema M, 2005) is an ensemble of two distinct algorithms: T&T and I.S.

A.1 T&T

The "Training and Testing" algorithm (T&T) is based on a population of n ANNs, managed by an evolutionary system. In its simplest form, this algorithm reproduces several distribution models of the complete dataset DΓ (one for every ANN of the population) in two subsets ($d_{\Gamma}^{[tr]}$, the Training Set and $d_{\Gamma}^{[ts]}$, the Testing Set). During the learning process each ANN, according to its own data distribution model, is trained on the subsample $d_{\Gamma}^{[tr]}$ and blind-validated on the subsample $d_{\Gamma}^{[ts]}$.

The performance score reached by each ANN in the testing phase represents its "fitness" value (i.e., the individual probability of evolution). The genome of each "ANN-individual" thus codifies a data distribution model with an associated validation strategy. The n data distribution models are combined according to their fitness criteria using an evolutionary algorithm. The fitness-based selection of "ANN-individuals" determines the evolution of the population; that is, the progressive improvement of performance of each network until the optimal performance is reached, which is equivalent to the optimal splitting of the global dataset into subsets. The evolutionary algorithm ruling this process, named "Genetic Doping Algorithm" (GenD) (Buscema M, 2004), is similar to a genetic algorithm (i.e. it works by crossover and mutation genetic operators), but maintains a constitutional instability across

the evolutionary process, thereby sustaining a natural proliferation of biodiversity and a continuous meta-evolution of the population.

The working of T&T is organized into two phases:

1) Preliminary phase: in this phase an evaluation of the parameters of the fitness function that will be used on the global dataset is performed. During this phase, an inductor $\Omega_{D_\Gamma^{[tr]},A,F,Z}(\times)$ is set up, which consists of an Artificial Neural Network equipped with a standard Back Propagation algorithm. For this inductor, the optimal configuration is determined at the end of different training trials on the global dataset D_Γ. In this way, the configuration that most "suits" the available dataset is determined: The number of layers and hidden units, and some possible generalizations of the standard learning law. The parameters thus determined define the configuration and the initialization of all the ANN-individuals of the population, and will subsequently stay fixed in the following computational phase. Basically, during this preliminary phase there is a fine-tuning of the inductor that defines the fitness values of the population's individuals during evolution.

The accuracy of the ANN performance upon the testing set will be the fitness of that individual (that is, of the trial-specific tentative distribution into two halves of the whole dataset).

2) Computational phase: The system extracts from the global dataset the best training and testing sets. During this phase, the ANN-individuals carry out their computational task, based upon the established configuration and the initialization parameters. From the evolution of the population, managed by the GenD algorithm, the best distribution of the global dataset DΓ into two subsets is generated, starting from the initial population of possible solutions $x = \left(D_\Gamma^{[tr]}, D_\Gamma^{[ts]}\right)$. Preliminary experimental sessions are performed using several

different ANN initializations and configurations, in order to achieve the best partition of the global dataset.

A2. I.S.

Parallel to T&T, TWIST runs I.S. (Input Selection), an adaptive system which is also based on the evolutionary algorithm GenD, and that is able to evaluate the relevance of the different variables of the dataset in a sophisticated way. Therefore, it can be considered on the same level as a feature selection technique.

From a formal point of view, I.S. is an artificial organism based on the GenD algorithm and consists of a population of ANNs, in which each one carries out a selection of the independent variables on the available database. The elaboration of I.S., as for T&T, is developed in two phases:

1) Preliminary phase: An inductor $\Omega_{D_{\Gamma}^{[tr]},A,F,Z}(\times)$ is configured to evaluate the parameters of the fitness function. This inductor is a standard Back-Propagation ANN. The parameters configuration and the initialization of the ANNs are carried out with particular care to avoid possible over-fitting problems that can be present when the database is characterized by a large number of variables that describe a small quantity of data. The number of epochs E_0 necessary to train the inductor is determined through preliminary experimental tests.

2) Computational phase: The inductor carries out its computational task, with the configuration determined in the previous phase and the fixed initialization parameters, to

extract the most relevant variables of the training and testing subsets. Each ANN-individual of the population is trained on the training set $D'^{[tr]}_{\Gamma}$ and tested on the testing set $D'^{[ts]}_{\Gamma}$.

The evolution of ANN-individuals in the population is again based on GenD. In the I.S. approach, the GenD genome consists of n binary values, where n is the cardinality of the original input space. Every gene indicates whether the corresponding input variable is active or not in that particular selection of variables. For each genome, the relevant fitness value is computed as usual. Through the evolutionary algorithm, the different "hypotheses" of variable selection, generated by each ANNs within the population, change over time, at each generation: This leads to the selection of the best combination of input variables. As in T&T, the crossover and mutation genetic operators are applied on the ANNs population; the rates of occurrence for both operators are adaptively self-determined by the system at each generation.

When the evolutionary algorithm no longer improves its performance, the process stops, and the best selection of the input variables is employed on the testing subset. In order to improve the speed and the quality of the solutions that have to be optimized with respect to standard evolutionary algorithms, GenD does not breed best-performing ANN-individuals, but rather most representative ones. The selection criterion is therefore not that of picking up momentarily brilliant but possibly unreliable outliers, but rather reinforcing those characteristics that are stably well performing.