

Using simulation and optimization approach to improve outcome through warfarin precision treatment

Chih-Lin Chi¹, Lu He², Kourosh Ravvaz³, John Weissert³, Peter J. Tonellato^{4,5}

¹*School of Nursing & Institute for Health Informatics
University of Minnesota, Minneapolis, MN, USA;*

²*Computer Science and Engineering
University of Minnesota, Minneapolis, MN, USA;*

³*Aurora Health Care, Milwaukee, WI, USA;*

⁴*Department of Biomedical Informatics, Department of Pathology
Harvard Medical School, Boston, MA, USA;*

*Zilber School of Public Health
⁵University of Wisconsin-Milwaukee, Milwaukee, WI, USA*

Email: cchi@umn.edu, Peter_Tonellato@hms.harvard.edu

We apply a treatment simulation and optimization approach to develop decision support guidance for warfarin precision treatment plans. Simulation include the use of ~1,500,000 clinical avatars (simulated patients) generated by an integrated data-driven and domain-knowledge based Bayesian Network Modeling approach. Subsequently, we simulate 30-day individual patient response to warfarin treatment of five clinical and genetic treatment plans followed by both individual and sub-population based optimization. Sub-population optimization (compared to individual optimization) provides a cost effective and realistic means of implementation of a precision-driven treatment plan in practical settings. In this project, we use the property of minimal entropy to minimize overall adverse risks for the largest possible patient sub-populations and we temper the results by considering both transparency and ease of implementation. Finally, we discuss the improved outcome of the precision treatment plan based on the sub-population optimized decision support rules.

Keywords: Precision Medicine; Personalized Treatment; Clinical Trial Simulation; Optimization.

1. Introduction

Precision medicine's fundamental charter is to improve treatment outcomes. Generally, outcome is improved by tailoring treatment personalized to an individual's clinical and genetic characteristics. In this manuscript, we combine both simulation and optimization approaches to personalize treatment using decision support rules, which indicate which particular treatment plan maximizes outcome improvement for patients with a particular set of clinical and genetic characteristics

(throughout this manuscript, patients with a particular set of clinical and genetic characteristics are defined as a ‘sub-population’). We use precision warfarin treatment plans as an example to show our approach, methods and results.

Warfarin is the most widely used anticlotting agent with a highly significant effect on reducing blood clots and preventing strokes and, thus, can dramatically decrease healthcare costs due to clotting-related events. However, administration of this drug is a challenge because of both sides of risks: stroke, if under-dosed, and bleeding, if over-dosed. Over 50 unique warfarin treatment plans have been developed to minimize the complexity of administration and both sides of risks. Warfarin treatments plans predict the initial warfarin dose and suggest a follow-up dose adjustment based on the patient’s lab results, International Normalized Ratio (INR), and the patients individual clinical and/or genetic factors.

The key question to healthcare providers and hospital administrators charged with treating hundreds of thousands (if not millions) of a diverse collection of patients, is which treatment plan is the most “effective”. For warfarin, no single treatment plan exists that provides the “optimal” outcome for every patient in a large diverse patient population. In short, there is no “one-size fits all” treatment plan that optimally reduces adverse events for all individuals. Our approach is based on a fundamental principle, “What are the optimal sub-populations of patients, with corresponding precision-based treatment plan, that minimizes adverse events while simultaneously reduces complexity of implementation across a large diverse patient population.” Our optimization “variable” is the collection of precision treatment plans derived from each individual’s clinical and genetic characteristics^{1,2}. We then vary the treatment plan, and definition of sub-population to create the optimal outcome. The combination of prediction (of treatment outcome as a function of treatment plan) and optimization of grouped outcome (across sub-populations) allows our approach to capture each individuals’ variability; produce precision-driven treatment plans and takes into consideration the complexity of implementing optimized treatment protocols across a diverse patient population²⁻⁴.

Specifically, we selected 5 warfarin treatment plans, each of which consists of 3 different treatment periods and protocols. We demonstrate our approach to find sub-populations and the optimal treatment plan for each sub-population. To avoid confusion, throughout this manuscript, we use the term ‘treatment plan’ for the 30-day course of treatment and ‘protocol’ for the specification of components of the 30 day treatment plan. In addition, ‘personalized treatment plan’ is the treatment plan optimized for an individual whereas ‘precision treatment plan’ is the treatment plan optimized for an entire sub-population.

2. Methods

In this manuscript, we combined all methods to produce rules to decide precision treatment plan with ~1,500,000 clinical avatars, who are simulated Aurora Health Care patients. In this method section, we first discuss the clinical avatars, clinical trial simulation, and optimization. Figure 1 summarizes 4 key steps to produce data for this work, which consists of two major components, Treatment Simulation⁵ and Optimization².

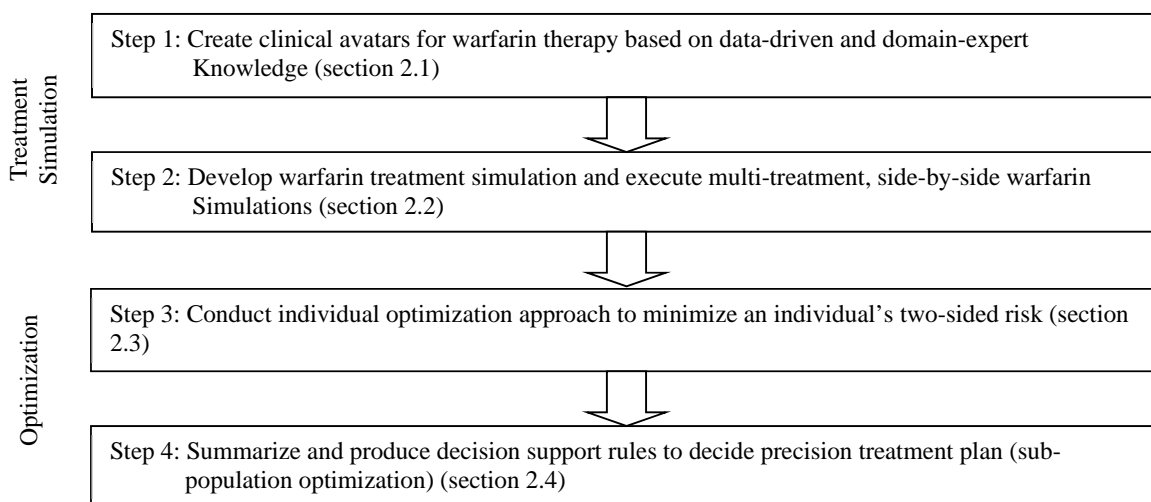


Figure 1. 4-step integrated approach to produce rules to decide precision treatment plan

2.1. Create clinical avatars (simulated patients) for warfarin therapy based on data-driven and domain-expert knowledge

The platform used in this study begins with extracting real patient electronic medical records (EMRs) from Aurora Health Care, a major hospital system in the Midwest. We used only de-identified patient data per IRB approval by an AHC honest broker. We extract from a diverse collection of 14,206 patients with Atrial Fibrillation (AF) over 10 years which were then integrated into that previously validated warfarin clinical trial simulation framework. The simulation framework has reproduced warfarin clinical trial results by using a PK/PD model ⁶ and iterative Bayesian Network Modeling (BNM) implemented in TETRAD ⁷.

We incorporated both data-driven knowledge and domain-expert knowledge to create clinical avatars. Specifically, (1) Distribution and causal relationships among most variables of patient characteristics were discovered from the EMRs, and (2) distribution of genetic tests, CYP2C9 and VKORC1, and causal relationships between other variables were observed from extensive literature review. We used TETRAD to take both types of input parameters to instantiate a Bayesian Network Model of the EMRs coupled with stochastic models to subsequently create clinical avatars. In order to cover variations, we create 100 times more subjects (clinical avatars) than the original subjects (EMR patients). The clinical avatars were then used as input to the following step, the warfarin Pharmacokinetic and Pharmacodynamic (PK/PD) model, which will be described in the following section in details.

2.2. Develop warfarin treatment simulation and execute multi-treatment, side-by-side warfarin simulations

Details of the treatment simulation are described in our preliminary work ⁸. In brief, on the treatment simulation platform, a clinical avatar representing a patient, “takes” the protocol-adjusted warfarin

dose each day of a thirty-day simulation. Since International Normalized Ratio (INR) is not tested daily, the avatar's dose may not be changed until the next scheduled INR test.

In this study, we simulate treatment process for created clinical avatars based on 5 warfarin treatment plans, which are summarized in Table 1. They are AAA, CAA, PGAA, PGPGI, and PGPGA. Each treatment plan includes three protocols for three treatment periods. For example, in the PGPGI treatment plan, modified IWPC PG protocol is used for initial period (days 1 to 3) ⁹, Lenzini PG protocol is used for adjustment period (days 4 and 5), and Intermountain protocol is used for maintenance (days 6 to 30) period. Among all protocols, some protocols (including AHC ¹⁰, IWPC Clinical ⁹, and Intermountain ¹¹) use only INR and others (including IWPC PG ⁹, Modified IWPC PG ¹², and Lenzini PG ¹³) use genotypic information and/or INR to adjust dosage. We note that 'AHC' is the treatment plan or protocols currently used in the Aurora Health Care and our team coded that clinical protocol and put it in our simulation settings. Other treatment plans and protocols can be found in the associated references.

Table 1. Five distinct treatment plans included in our treatment simulations.

Treatment Plans \ Treatment Periods	Initial protocol (days)	Adjustment protocol (days)	Maintenance protocol (days)
AAA (Clinical)	AHC (1-2)	AHC (3-7)	AHC (8-30)
CAA (Clinical)	IWPC Clinical (1-2)	AHC (3-7)	AHC (8-30)
PGAA (Pharmacogenetics)	IWPC PG (1-2)	AHC (3-7)	AHC (8-30)
PGPGI (Pharmacogenetics)	Modified IWPC PG (1-3)	Lenzini PG (4-5)	Intermountain (6-30)
PGPGA (Pharmacogenetics)	Modified IWPC PG (1-3)	Lenzini PG (4-5)	AHC (6-30)

In practice, each simulated patient receives warfarin treatment based on the treatment plans and protocols. Their physiological response to the dosing is measured by International Normalized Ratio (INR), i.e., measurement to understand how thin the blood is. Generally, if the INR is larger or smaller than the therapeutic window (often the range $2.0 < \text{INR} < 3.0$), then the protocol calculates the correct change in dose to adjust INR into therapeutic range. In our computational simulation the INRs are predicted using a PK/PD model ⁶ based on clinical and genotypic characteristics of the clinical avatar.

Similar to actual treatment process, on this platform, the clinical avatar ingests the warfarin dose computed by the treatment plan or protocol, then the PK/PD model predict INR, and, then, the warfarin dose is, again, adjusted based on the INR. Typically, warfarin is taken every day, and INR test is measured based on the schedule documented on the protocol.

In this study, we simulated 30-day warfarin treatment process using five treatment plans for each clinical avatar based on adapted PODSS algorithm ^{3,4}, which was originally developed in machine learning settings. Each treatment plan includes three protocols, each of which corresponds to a treatment period, and a different protocol uses a distinct algorithm to adjust dosage. As a result, we generate several INR values in the 30-day treatment for a protocol. Using each clinical avatars

modeled INR values, we calculated Time in Therapeutic Range (TTR; 0-100%) by way of linear interpolation between protocol based check days¹⁴ with therapeutic range defined as 2-3. The TTR is typically used as surrogate outcome in warfarin studies to represent prevention of the both side of risks, stroke and bleeding. In general, the higher TTR, the lower the two-sided risks, and vice versa. Finally, we have 5 TTRs, each of which is associated with a specific treatment plan. These TTRs are personalized prediction results and the values are very different for various patients based on their clinical and genetic characteristics. These predicted TTR outcome can be further used for optimization to identify personalized treatment in the following steps.

2.3. Conduct individual optimization approach to minimize an individual's two-sided risk

In the simulation, when a treatment plan has a larger predicted TTR, the treatment plan has a higher probability and chance to actually minimize two-side of risks. Hence, the purpose is to obtain a treatment plan to maximize predicted TTR (or minimize risks). More details were discussed in our preliminary work¹. Implement the individual optimization approach is straightforward. We first prioritize TTRs (each corresponding to a protocol) using the matrix of data developed from the above warfarin treatment simulations for each patient and then identify the protocol with the largest TTR. The treatment plan or protocol with the greatest TTR indicates the optimal reduction in the two-sided risks for a clinical avatar. As a result, based on each clinical avatars' clinical and genetic characteristics, in this process, we identify a label of treatment plan that minimizes the two-sided risks for each clinical avatar.

2.4. Summarize and produce decision support rules to decide precision treatment plan (sub-population optimization)

The difference between this and the last step is optimization target: an individual or a sub-population.

The advantage of this step is 'practicability', which means we do not need sophisticated program and clinical decision support system (CDSS) to take multiple variables of a patient to find the 'personalized treatment plan'. Instead, one can use simple criteria and easily follow the decision support rules created in this step to find the 'precision treatment plan'. A rule, for example, indicates, PGAA treatment plan is predicted to minimize two-sided risks for the patient sub-population, age < 64.95 yrs and VKORC = A/A or G/A.

We used an entropy function implanted in a decision tree algorithm to create the rule set. The reason and details has been discussed in our previous method paper². In brief, we aim to maximize overall outcomes by simultaneously optimize two targets when producing decision support rules: (1) identify a treatment plan to minimizes the two-sided risks for a sub-population and (2) identify the sub-population that maximal proportion of patients can benefit from the identified treatment plan. Due to our specifically designed data structure and entropy/impurity property, we simultaneously optimize these two targets to create decision support rules.

Table 2. Independent variables used to create the decision tree

Variable Name	Variable Type	Variable Range
Weight	Continuous	71-490
Height	Continuous	46.1-85
Age	Continuous	18-102
Race	Discrete(Nominal)	Asian/Black or African American/White
AMI	Binary	0/1
BMI	Continuous	8.7-91.3
CYP2C9	Discrete(Ordinal)	*1/*1 (value=5), *1/*2 (value=4), *1/*3 (value=3), *2/*2 (value=2), *2/*3 (value=1)
Fluvastatin	Binary	0/1
VKORC1G	Discrete(Ordinal)	A/A (value=1), G/A (value=2), G/G (value=3)
Gender	Binary	0/1
Smoker	Binary	0/1

Table 2 summarizes independent variables used for the multi-class decision tree algorithm implemented in Matlab^{15,16}. All independent variables are numeric. The original variable type of CYP2C9 and VKORC1 are categorical. We transform these categorical variables into ordinal variables based on the dosage amount¹⁷ for appropriate therapy. In the Step 3, we have created a label of optimal personalized treatment plan for a clinical avatar. Dependent variable of the decision tree is that label. The dataset consists ~1,500,000 clinical avatars, each of which includes a set of patient characteristics and the optimal treatment plan label. Now the computational problem becomes using a decision tree algorithm to identify patient sub-population with the maximal proportion of individuals that benefit from one or fewer optimal treatment plan label(s).

3. Results

3.1. Observation of the data

Table 3 summarizes comparison of variable distribution between EMR and clinical avatars. There is no statistically significant difference ($P < 0.05$) between the two populations for both continuous and discrete variables.

Table 3. Characteristics of the Original Aurora AF Warfarin and Aurora Clinical Avatar Populations.

Characteristic		Aurora AF Warfarin Population (mean±SD)	Aurora Clinical Avatar Population
Age	year	67.3±14.43	67.2±14.47
Weight	lb	199.24±54.71	199.24±54.6
Height	in	66.78±4.31	66.53±4.32
Gender, %	Female	53.14	53.10
	Male	46.86	46.90
Race, %	White	95.18	95.22
	African-American	4.25	4.19
	Asian	0.39	0.40
	Am. Indian/Alaskan	0.18	0.18
	Pacific Islander	0.0001	0.0001
Tobacco, %	No	90.33	90.67
	Yes	9.66	9.33
Amiodarone, %	No	88.45	88.49
	Yes	11.54	11.51
Fluvastatin, %	No	99.97	99.98
	Yes	0.03	0.02
CYP2C9, %	*1/*1	65.77 ^a	67.39
	*1/*2	14.6 ^a	14.86
	*1/*3	9.11 ^a	9.25
	*2/*2	6.41 ^a	6.51
	*2/*3	1.93 ^a	1.97
	*3/*3	0 ^a	0
VKORC1, %	G/G	38.54 ^a	38.37
	G/A	44.02 ^a	44.18
	A/A	17.33 ^a	17.45

^aAurora patient population's genotypic characteristics were derived from published genotype distributions.

Figure 2 summarizes distribution of personalized treatment plan labels (derived from the Step 3) observed in all clinical avatars. Each clinical avatar has a personalized treatment plan label optimized from one of the five treatment plans. As described in the method section, the optimal label is decided based on the highest TTR across the 5 available treatment-plan options. That optimal-treatment-plan label is a class label used in the subsequent decision tree algorithm. PGAA treatment plan (37.8%) is the majority followed by CAA (23.4%) and AAA (22.7%) treatment plans.

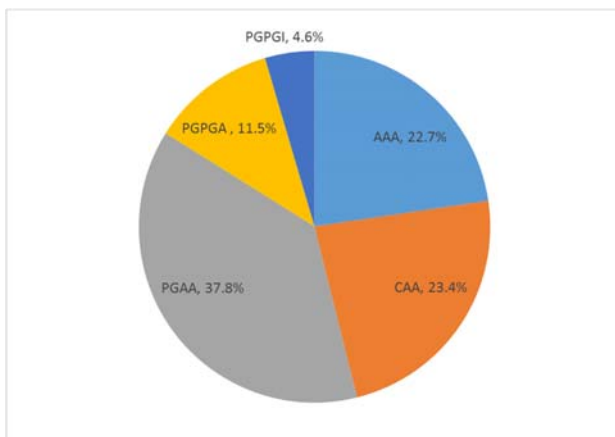


Figure 2. Proportion of personalized protocol derived from individual optimization.

3.2. Derive personalized treatment rules from the data

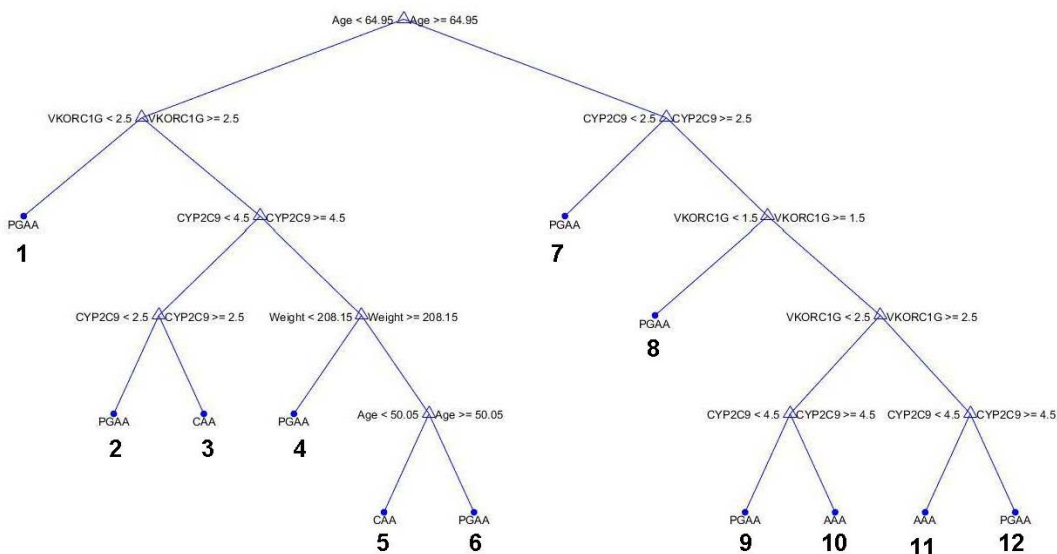


Figure 3. Personalized Treatment Rules extracted from ~1,500,000 clinical avatars by using a multi-class decision tree algorithm. The number of clinical avatars in leaf nodes 1 to 12 ranged from 19,518 to 382,823.

Figure 3 shows decision support rules derived from all those clinical avatars using multi-class decision tree algorithm. When the tree is very large, we have many complicated and a large number of rules, which results in a challenge for practical use. On the other hand, when the tree is very small, the effect of personalized treatment will diminish (extreme example is no tree at all and there is no benefit due to personalized treatment). Therefore, we empirically test and decide 12 decision support rules to maximize visualization and ease of use while maintaining performance of grouping by pruning the tree¹⁸.

In this decision tree, we have 12 leaf nodes, each of which has a number corresponding to the nodes in the following figures. In each leaf node, we also include the label of optimal sub-population treatment plan, which is a predicted class of that node in the decision tree algorithm.

We note that the ‘actual’ prediction has been done in simulation and individual optimization (Steps 2 and 3). Here, the decision tree algorithm is used to summarize and produce decision support rules using prediction data created in the Steps 2 and 3. Similar to Figure 2, the optimal sub-population treatment plan for most nodes is PGAA followed by CAA and AAA.

3.3. Analysis of treatment plans in each leaf node

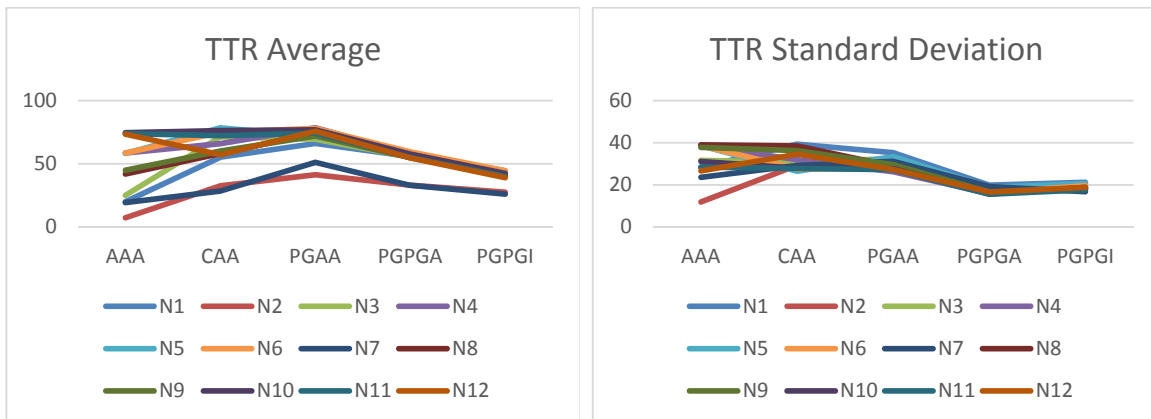


Figure 4. TTR average and standard deviation of the five protocols in different sub-populations (nodes 1 to 12). In general genetic-based treatment plans have lower variations (such as PGAA, PGPGA, and PGPGI).

Before comparing personalized with one-fit-all treatment plans, in this section, we observe TTR average and standard deviation of each one-fit-all treatment plan for the 12 leaf nodes (Figure 4). Specifically, we first follow decision rules in the Figure 3 to identify the 12 patient sub-populations. In each sub-population, we take the average and standard deviation of the 5 treatment plans from each clinical avatar belong to that sub-population. As a result, we have 5 TTR averages and standard deviations observed from the 12 patient sub-populations.

In the Figure 4, we can observe that variations for AAA and CAA are very high. Although the average TTRs for some sub-populations are high, the average TTRs for others are low. Therefore, the overall TTR is not high and, therefore, the chance to become optimal sub-population protocol is in the middle. On the other hand, although the variations for PGPGA and PGPGI treatment plans are consistently low, the TTR averages for most nodes are also consistently low. Thus, the overall TTRs are low across all patient sub-populations. Finally, PGAA treatment plan has both advantages: higher TTR average and lower variations across all patient sub-populations. Therefore, the chance to become sub-population optimal protocol is the highest. Therefore, we can expect PGAA should be the dominated sub-population protocol, followed by AAA and CAA. This expectation corresponds to the rule to decide personalized treatment plan in Figure 3, in which PGAA is the majority followed by AAA and CAA treatment plans.

3.4. Compare personalized with one-fit-all treatment plans

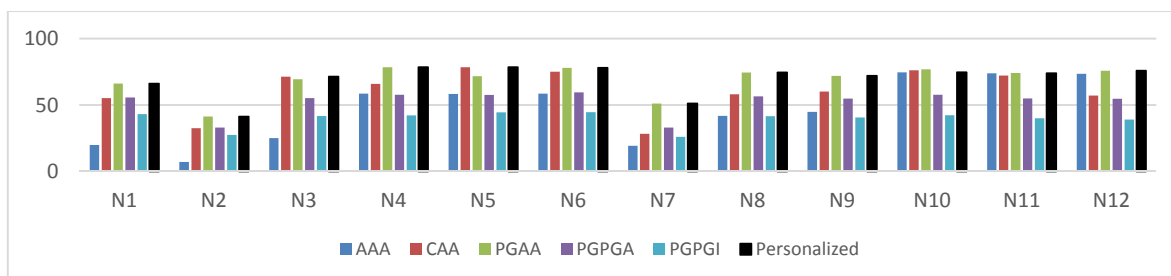


Figure 5. Comparison of TTR average between optimal sub-population and one-fit-all treatment plans.

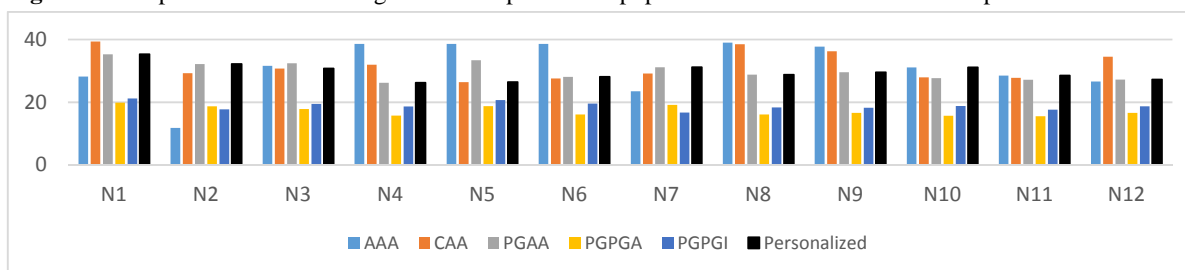


Figure 6. Comparison of TTR standard deviations between optimal sub-population and one-fit-all treatment plans.

Finally, we compare TTR average and standard deviations between the optimal sub-population and one-fit-all treatment plans in Figures 4 and 5, respectively. There are 12 sub-populations, each of which correspond a leaf node (N1~N12). The optimal sub-population treatment plan is the treatment plan (black bar) recommended in the Figure 3 for a specific type of patients. On the other hand, the one-fit-all treatment plans are the treatment plans (color bars) uniformly used on every clinical avatar belong to that sub-population. We note that only computational settings, like simulation, allow a clinical avatar receive warfarin treatment based on multiple treatment plans. From Figure 5, we can observe that TTR of the optimal sub-population treatment plan is almost the highest across all patient sub-populations. On the other hand, variations are, in general, around the middle (Figure 6). In sum, compared to one-fit-all treatment protocol, personalized treatment plan improve TTR outcome, ranged from 15% to 31% (average and median across those 12 leaf nodes are both 24%). Such outcome improvement is the result of optimizing and reusing existing treatment plan using precision-medicine approach.

4. Conclusion

Combing clinical avatars simulation, with optimization can yield clinical decision support rules with respect to dosing strategy for patients with AF. Clinical avatars are developed through BNM model based on the distribution of patient characteristics in Aurora Healthcare EMRs and genotypic characteristics from published literature (Step 1 in the Figure 1). We then simulated individual dose response and outcomes for 30-day warfarin therapy. All ~1.5-million clinical avatars underwent on five simulated treatment regimes (Step 2). Simulated outcomes were then used as input to apply individual optimization to create the label of a treatment plan that maximize TTR outcome (Step 3). Finally, a decision tree algorithm was used to identify decision support rules that minimizes overall

adverse risks (or maximizes TTR) for the largest possible patient sub-populations from these clinical avatars (Step 4). The decision support rules identify treatment plans that maximizing TTR outcome for patients with a specific type of clinical and genetic characteristics.

Leaf nodes of decision support rules demonstrate precision treatment plans for a specific type of individuals. Figures 4 and 5 further show computational evidence of average TTR outcome and variation when comparing between personalized and one-fit-all treatment plans. When comparing TTR average (Figure 5), the precision treatment plan almost has the highest TTR treatment outcome across all sub-populations from node 1 to 12. Compared with one-size-fits-all treatment plan, personalized treatment plan improved TTR outcomes ranged from 15% to 31% (average and median across 12 leaf nodes are both 24%). On the other hand, the standard deviations are about the middle compared to all other one-fit-all treatment plans. Therefore, we argue that without developing a new treatment plan and protocol, the efficacy can be improved by applying precision-medicine approach to re-use existing treatment options.

Many clinical trial simulations (CTS) were used to assess cost-effectiveness¹⁹, health economic²⁰, and discrete-event^{19,21}. Some studies apply CTS to help understand how different genotypes influence dosing decisions²². In this study, we used CTS coupled with sub-population optimization to produce a ‘practical’ precision-medicine approach. Instead of using many patient’s clinical and genetic characteristics, a care-giver can use a few variables at the point of care to optimize treatment plan. By using personalized treatment rules in the Figure 3, small, rural clinics can provide optimized treatment regimes without significant increased burden or expensive software and EMR implementation.

In general, limitations can be classified in two categories: the PK/PD model and clustering approaches. (1) Unlike machine learning, which can integrate many variables to predict INR, PK/PD models use only variables (such as partial variables in the Table 2) that were significantly relevant for prediction. (2) PK/PD model was developed with patient populations distinct from the patient population used to construct the clinical avatar model. (3) Clustering approaches, such as the decision tree algorithm, may not create rules of interest for physicians. This limitation will be address by our ongoing work, which will focus on producing various versions of rule-creation approaches. In this study, bleeding and thrombosis risks are equally important in our study. In specialized scenarios, one may be more important than the other risk. We are also working on incorporating constraints and weightings to produce these specialized rules.

Acknowledgement

We thank NIH-1R01LM011566-01 to support this work.

References

1. Chi C-L, Fusaro VA, Patil P, Crawford MA, Contant CF, Tonellato PJ. An approach to optimal individualized warfarin treatment through clinical trial simulations. In: *2010 5th Cairo International Biomedical Engineering Conference*. IEEE; 2010:116-120.
2. Chi C-L, Ravvaz K, Weissert J, Tonellato PJ. Optimal decision support rules improve personalize warfarin treatment outcomes. In: *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE; 2016:2594-2597.

3. Chi C-L, Street WN, Ward MM. Building a hospital referral expert system with a prediction and optimization-based decision support system algorithm. *J Biomed Inform.* 2008;41(2):371-386.
4. Chi C-L, Street WN, Robinson JG, Crawford MA. Individualized patient-centered lifestyle recommendations: An expert system for communicating patient specific cardiovascular risk information and prioritizing lifestyle options. *J Biomed Inform.* 2012;45(6):1164-1174.
5. Ravvaz K, Weissert J, Ruff C, Chi C, Tonellato P. Personalized Anticoagulation: Optimizing Warfarin Management Using Genetics and Simulated Clinical Trials. *Circ Cardiovasc Genet.* In press.
6. Hamberg A-K, Dahl M-L, Barban M, et al. A PK-PD model for predicting the impact of age, CYP2C9, and VKORC1 genotype on individualization of warfarin therapy. *Clin Pharmacol Ther.* 2007;81(4):529-538. doi:10.1038/sj.clpt.6100084.
7. TETRAD project. <http://www.phil.cmu.edu/tetrad/>. Accessed February 26, 2017.
8. Fusaro VA, Patil P, Chi C-L, Contant CF, Tonellato PJ. A systems approach to designing effective clinical trials using simulations. *Circulation.* 2013;127(4):517-526. doi:10.1161/CIRCULATIONAHA.112.123034.
9. International Warfarin Pharmacogenetics Consortium, Klein TE, Altman RB, et al. Estimation of the warfarin dose with clinical and pharmacogenetic data. *N Engl J Med.* 2009;360(8):753-764. doi:10.1056/NEJMoa0809329.
10. Aurora health care warfarin proposal. *Unpubl Stand Protoc Aurora Heal Care.*
11. Anderson JL, Horne BD, Stevens SM, et al. A randomized and clinical effectiveness trial comparing two pharmacogenetic algorithms and standard care for individualizing warfarin dosing (CoumaGen-II). *Circulation.* 2012;125(16):1997-2005.
12. Kimmel SE, French B, Kasner SE, et al. A pharmacogenetic versus a clinical Algorithm for warfarin dosing. *N Engl J Med.* 2013;369(24):2283-2293. doi:10.1056/NEJMoa1310669.
13. Lenzini P, Wadelius M, Kimmel S, et al. Integration of genetic, clinical, and INR data to refine warfarin dosing. *Clin Pharmacol Ther.* 2010;87(5):572-578. doi:10.1038/clpt.2010.13.
14. Rosendaal FR, Cannegieter SC, van der Meer FJ, Briët E. A method to determine the optimal intensity of oral anticoagulant therapy. *Thromb Haemost.* 1993;69(3):236-239.
15. Coppersmith D, Hong SJ, Hosking JRM. Partitioning nominal attributes in decision trees. *Data Min Knowl Discov.* 1999;3(2):197-217. doi:10.1023/A:1009869804967.
16. Breiman L, Friedman J, Stone CJ, Olshen RA. *Classification and Regression Trees.* Chapman & Hall; 1984.
17. Dean L. *Warfarin Therapy and the Genotypes CYP2C9 and VKORC1.*; 2016. doi:10.1038/clpt.2010.13.
18. Hastie T, Tibshirani R, Friedman JH (Jerome H). *The Elements of Statistical Learning : Data Mining, Inference, and Prediction.* Springer; 2009.
19. Pink J, Pirmohamed M, Lane S, Hughes DA. Cost-Effectiveness of Pharmacogenetics-Guided Warfarin Therapy vs. Alternative Anticoagulation in Atrial Fibrillation. *Clin Pharmacol Ther.* 2014;95(2):199-207. doi:10.1038/clpt.2013.190.
20. Hughes DA, Walley T. Economic Evaluations During Early (Phase II) Drug Development. *Pharmacoeconomics.* 2001;19(11):1069-1077. doi:10.2165/00019053-200119110-00001.
21. Pink J, Lane S, Pirmohamed M, Hughes DA. Dabigatran etexilate versus warfarin in management of non-valvular atrial fibrillation in UK context: quantitative benefit-harm and economic analyses. *BMJ.* 2011;343:d6333. doi:10.1136/BMJ.D6333.
22. Salinger DH, Shen DD, Thummel K, Wittkowsky AK, Vicini P, Veenstra DL. Pharmacogenomic trial design: use of a PK/PD model to explore warfarin dosing interventions through clinical trial simulation. *Pharmacogenet Genomics.* 2009;19(12):965-971. doi:10.1097/FPC.0b013e32833333b80.