Designing Personalized Treatment: An Application to Anticoagulation Therapy Supplementary Material

Abstract from main paper

In this paper, we develop an analytical framework for personalizing the anticoagulation therapy of patients who are taking warfarin. Consistent with medical practice, our treatment design consists of two stages: (i) the initiation stage, modelled using a partially-observable Markov decision process, during which the physician learns through systematic belief updates about the unobservable patient sensitivity to warfarin, and (ii) the maintenance stage, modelled using a Markov decision process, during which the physician relies on his formed belief about patient sensitivity to determine the stable, patient-specific, warfarin dose to prescribe. We develop an expression for belief updates in the POMDP, establish the optimality of the myopic policy for the MDP, and derive conditions for the existence and uniqueness of a myopically optimal dose. We validate our models using a real-life patient data set gathered at the Hematology Clinic of the Jewish General Hospital in Montreal. The proposed analytical framework and case study enable us to develop useful clinical insights, e.g., concerning the length of the initiation period and the importance of correctly assessing patient sensitivity.

Keywords: individualized treatment; stroke prevention; treatment design; warfarin.

S1 Summary of Content

In this supplement, we present supportive material to the main paper, with the same title. In §S2, we describe results of simulation experiments which further justify our two-stage approach. In §S3, we investigate whether there is any evidence that would justify the usage of a three-stage solution approach for our problem. In §S4, we present additional detailed simulation results (Tables S3 and S4) related to the initiation and maintenance stages, as in §8 of the main paper. In Table S1, we present a description of variables in our data set.

S2 Comparing Joint and Separate Solutions

As explained in §3 of the main paper, we solve the two stages of treatment independently by first solving the POMDP (initiation) and then the MDP (maintenance). An alternative is to solve both stages jointly, i.e., to determine dosages throughout so as to minimize the total expected risk, cumulative over both stages. Since solving the two stages jointly leads to a smaller expected risk, we investigate in this section the amount of increase in expected risk which results from going from a joint to a separate solution.

Variable	Description
Demographic information	
Age indicator (AGE IND)	Indicates whether patient is above 75 years of age
Gender	Indicates whether patient is male or female
Past medical history	
Adult onset diabetes mellitus (AODM)	Indicates presence of adult-onset-diabetes mellitus
Rheumatic heart disease (RH)	Indicates presence of rheumatic heart disease
Other comorbidities	Various medical conditions e.g., asthma, cancer, etc.
Hypertension (HYP)	Indicates high blood pressure
Mechanical heart valve (MHV)	Indicates heart valve replacement
Bleeding risk assessment	
Previous cerebrovascular accident (CVA)	Indicates previous occurrence of stroke
Peptic ulcer disease (ULC)	Indicates presence of ulcer of the gastrointestinal tract
Risks of fall	Denotes likelihood of falling (LOW, MEDIUM, HIGH)
Use of other anticoagulant (AC)	Indicates use of other medication

Table S1: Description of all variables in the data set.

We vary the value of p in the initiation stage, ranging from $p = 0.4$ to $p = 0.97$, and let $T = 10$ years be the length of the entire treatment process (since the average age of an AF patient is 75 years old, this choice of length for the entire anticoagulation treatment is reasonable). We compute estimates of the total average risk for the patient for different values of p considered, cumulative over both the initiation and maintenance stages. The risk value corresponding to $p \approx 1$ is roughly equal to the best possible risk value resulting from a joint solution of the two problems (we do not let $p = 1$ since we want to ensure convergence of our belief updates). We compare this "optimal" risk value to alternative risk values, obtained for each p considered, which correspond to solving each stage separately. Our objective is to show that this "optimal" risk value is not much smaller than those obtained when solving the two stages separately.

In the initiation stage, we assume that visits to the clinic are every 3 days; in the maintenance stage, we assume that they are every three weeks. Each simulation point estimate is based on 10,000 independent simulation runs (we deliberately choose a large number of simulation replications to minimize the effect of stochastic noise). In Figure S1, we consider an initial belief probability vector equal to $(0.6, 0.2, 0.2)$ and a medium γ value. In Figure S2, we consider a uniform belief probability vector and a high γ value. As illustrated in Figures S1 and S2, the cumulative risk does not considerably change when jointly or separately solving the problems in both stages; e.g., it is less than 1% in the figures. We ran similar numerical experiments for alternative initial probability vectors and γ values and reached consistent results throughout. These numerical experiments provide further evidence that justifies using a two-stage solution approach in our paper, particularly since a joint solution is not supported by standard POMDP solution

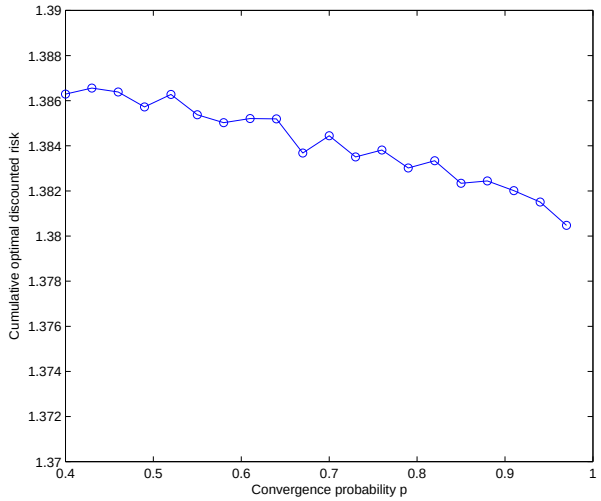

Figure S1: Optimal average cumulative risk, over both stages, for an initial belief vector $(0.6, 0.2, 0.2)$ and medium sensitivity.

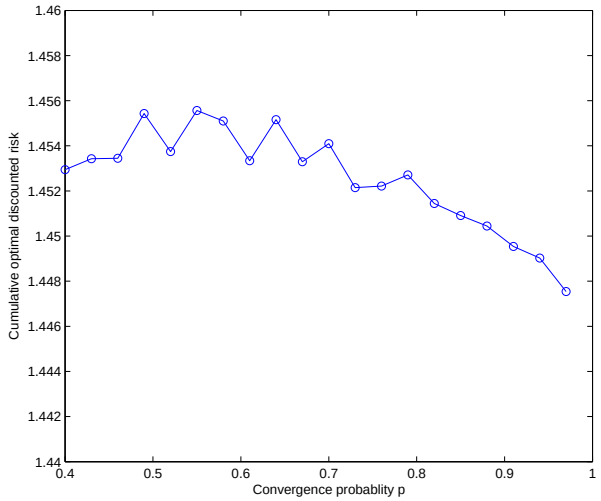

Figure S2: Optimal average cumulative risk, over both stages, for an initial uniform belief vector and high sensitivity.

software, as explained in §3 of the main paper.

S3 Three-Stage Solution Approach

We now investigate whether a three-stage approach, consisting of first minimizing the risk until a certain level of confidence about γ is reached, then maximizing the TTR, and then moving to the maintenance stage, is of any advantage over the two stage approach that we are using in the main paper.

To do so, we numerically solve the POMDP corresponding to the initiation stage under the two objectives: (1) TTR maximization and (2) risk minimization. Then, we simulate patient profiles according to the optimal dosages prescribed by the POMDP, under each criterion. In particular, we varied the convergence probability, p, but focused only on small values ($p = 0.4, 0.5$, and 0.6) to ensure that we restrict attention to the first few visits only. Indeed, our objective is to assess whether the first few stages are very high risk, which would then justify the usage of a three-stage approach.

We considered a uniform initial belief vector ($p_{low} = p_{med} = p_{high} = 1/3$). For each value of p, and each value of γ (low, medium, high), we computed simulation estimates, averaged across 500 independent replications, of the following: average TTR, average number of visits in the initiation stage to reach p, average risk, average number of stroke events during the initiation stage, and average number of bleeding

events during the initiation stage. We report our results in Table S2 below. As explained in §5.3 of the main paper, our risk measure is a *relative risk measure* which cannot be used directly to compute the average numbers of bleeding and stroke events. Instead, in order to compute these numbers, we relied on estimates of probabilities for stroke/ bleeds in the literature, as a function of the underlying INR value; in particular, we used the estimates in Patrick et al. (2009).

Table S2 shows that, overall, clinical performances under the TTR maximization and risk minimization objectives are quite similar. In particular, there does not exist sufficient evidence to justify using a three-stage approach, beginning with risk minimization, then moving to TTR maximization, and finally to the maintenance stage. That is why we do not use such an approach in this paper.

		TTR maximization					**Risk minimization**				
p	γ	**TTR**	**Visits**	**Risk**	**# Stroke**	**# Bleeds**	**TTR**	**Visits**	**Risk**	**# Stroke**	**# Bleeds**
0.4	low	0.234	3.85	1.40	0.00397	0.000704	0.243	4.01	1.40	0.00414	0.000742
	med	0.449	8.79	1.42	0.0170	0.00103	0.452	9.47	1.42	0.0179	0.00110
	high	0.322	4.89	1.61	0.0134	0.000442	0.336	6.066	1.57	0.0149	0.000553
0.5	low	0.270	6.59	1.42	0.00681	0.00124	0.277	6.27	1.42	0.00645	0.00117
	med	0.454	19.3	1.40	0.0344	0.00229	0.447	18.4	1.41	0.0342	0.00216
	high	0.402	8.81	1.59	0.0221	0.000804	0.388	8.46	1.60	0.0223	0.000769
0.6	low	0.334	10.1	1.39	0.0106	0.00185	0.337	9.44	1.39	0.00997	0.00176
	med	0.460	28.7	1.39	0.0498	0.00344	0.457	29.2	1.39	0.0506	0.00348
	high	0.409	13.3	1.53	0.0332	0.00121	0.410	15.1	1.53	0.0370	0.00139

Table S2: Comparison of various clinical measures under the TTR maximization and risk minimization criteria in the initiation stage.

S4 Detailed Simulation Results

In this section, we present additional numerical results which support §8, and in particular Figures 4-9, of the main paper.

S4.1 Initiation Stage

We present simulation results quantifying the performances of both the optimal solution for the POMDP and the myopically optimal solution. Our aim is to assess how different these two solutions are. Our results are based on 500 independent simulation replications. In each simulation replication, we proceed as follows. We choose the optimal initial dose and simulate the resulting INR. Since the decision maker does not have knowledge of γ but instead has a prior belief about it, the initial dose depends on this initial belief. For

example, we found that for a uniform initial belief, the optimal initial dose for the POMDP is equal to 5 mgs, whereas it is equal to 2.5 mgs for all other initial beliefs considered. Based on the observed INR, we update our belief about γ; based on this new belief, we follow the decision maker's optimal policy (either for the POMDP, or myopically) to select the new optimal dose. Those steps are repeated until convergence, i.e., until the belief probability that γ is either low, medium, or high exceeds the desired threshold.

In Table S3, we present simulation estimates for the TTR and risk values, and for the required number of visits until convergence, for both the optimal policy of the POMDP and the myopically optimal policy. We do so in order to quantify the performance of each policy. We consider different actual patient sensitivities: low, medium, or high (corresponding to the different rows in the table). We consider different initial belief vectors which are characterized by different initial belief probabilities for those three sensitivity levels (these vectors correspond to the different subparts, in italics, of the table). We also vary the convergence criterion for our belief updating procedure (convergence probability column in the table). This is the final desired belief probability that γ is either low, medium, or high; we consider the following final belief probabilities: 0.995, 0.9. and 0.7.

In the TTR column, we report simulation point estimates for the average TTR values. First, for each simulation replication, we estimate the proportion of visits where the INR falls in the desired therapeutic range. Those resulting estimates are then averaged across the independent simulation replications to yield average TTR values. In the visits column, we report simulation point estimates for the number of visits required to reach the corresponding desired final belief probabilities. First, we determine the number of visits required in each simulation replication; then, we average those numbers across all simulation replications. In the risk column, we report simulation point estimates for the average risk values. At each clinical visit, in each simulation replication, we compute the risk, corresponding to the prescribed dose, based on our functional expression for the risk. We then average those risk values, first across all clinical visits, and then across all simulation replications.

We also computed 95% confidence interval estimates corresponding to all performance measures. We do not include those in the table because of space restrictions; the half-widths of those intervals were typically found to be smaller than 5% of the corresponding point estimates, throughout. For illustrative purposes, we consider a patient who had a previous cerebrovascular event, but does not have any other statistically significant fixed effect from Table 1 of the main paper, i.e., we let $\beta = 0.5753$. We let $\sigma_\epsilon^2 = 0.102$, as estimated from data.

S4.1.1 Main Insights

Table S3 shows that *the performance of the myopically optimal policy is close to that of the POMDP optimal policy*, given our discretization of the problem which is consistent with medical practice. Indeed, for all initial belief vectors and all patient sensitivities considered, the average TTR and risk values are close under those two policies. Thus, our study gives numerical support to solving the single-stage problem instead of the POMDP; this is practically useful given the numerical complexity of the POMDP.

Consistent with intuition, Table S3 shows that *the required number of patient visits sharply increases as our desired final belief probability about γ increases.* For example, for a uniform initial belief probability vector, a low γ value, and a desired final belief probability equal to 0.995, the average required number of visits is roughly equal to 52, whereas it is only 13 for a desired belief probability of 0.7.

More interestingly, Table S3 shows that, with a uniform initial belief vector, *it usually takes longer to learn about a medium sensitivity, compared to low or high sensitivities.* For example, for a desired final belief probability of 0.995, the average required number of visits is roughly equal to 52 for high γ, whereas it is roughly 141 for medium γ. The intuition behind this observation is that patients with either low or high sensitivities are easier to identify due to their more extreme reactions to prescribed dosages.

Table S3 also quantifies the advantage of selecting an adequate initial belief vector. For example, we consider the case where the true sensitivity is low, and where we start with an initial belief vector which assigns 0.6 to a low sensitivity, and 0.2 to both medium and high sensitivities. Then, the average number of required visits until convergence is 29, which is considerably smaller than with a uniform initial probability vector. However, we find that the initial probability vector does not have a strong impact on either the average risk or the TTR. *Our recommendation to physician's is to allocate a significant effort in pre-diagnosing the patient, e.g., via preliminary medical tests, so as to form a better initial belief about patient sensitivity; this would lead to substantially reducing the length of the initiation stage.*

S4.2 Maintenance Stage

In Table S4, we consider three different actual values for γ, 0.02, 0.08, and 0.12, which correspond to the different subparts (in italics) of the table. For each actual value of γ, we compute the optimal dose prescription assuming perfect knowledge of this γ; this is d_R^* in Lemma 5.1. We also compute the corresponding optimal expected risk. It is not hard to see why, given d_R^*, this optimal risk is the same across all values of the true underlying γ.

We vary the physician's belief about γ by systematically varying μ_M and σ_M^2, and study corresponding

Table S3: Simulation estimates for the average TTR, cumulative average risk, and average required number of visits across 500 independent replications. Initial beliefs about γ are initial probability vectors for low, medium, and high γ values, respectively. The convergence probability is the desired final belief probability about the value of γ.

		POMDP Solution			**Myopic Solution**		
Convergence probability	γ	**TTR**	**Visits**	**Risk**	**TTR**	**Visits**	**Risk**
	Uniform initial belief						
0.995	low	0.389	52.0	1.37	0.400	32.2	1.37
	medium	0.470	141	1.38	0.466	145	1.38
	high	0.463	127	1.44	0.462	125	1.44
0.9	low	0.358	24.1	1.38	0.369	21.6	1.38
	medium	0.467	74.6	1.38	0.466	73.7	1.38
	high	0.454	57.6	1.47	0.455	50.3	1.46
0.7	low	0.320	13.6	1.40	0.332	13.9	1.39
	medium	0.476	40.3	1.37	0.471	39.9	1.38
	high	0.443	27.1	1.47	0.445	23.8	1.50
	Initial belief: (low: 0.6, medium: 0.2, high: 0.2)						
0.995	low	0.410	27.0	1.36	0.406	22.8	1.37
	medium	0.470	138	1.38	0.469	141	1.38
	high	0.462	123	1.45	0.463	124	1.45
0.9	low	0.346	10.7	1.375	0.358	15.2	1.39
	medium	0.463	74.6	1.39	0.464	76.0	1.38
	high	0.450	52.5	1.48	0.450	58.9	1.48
0.7	low	0.236	4.3	1.40	0.218	4.6	1.40
	medium	0.455	44.0	1.40	0.461	44.6	1.39
	high	0.417	23.9	1.54	0.432	27.6	1.49
	Initial belief: (low: 0.2, medium: 0.6, high: 0.2)						
0.995	low	0.416	43.8	1.361	0.399	43.9	1.37
	medium	0.467	116	1.38	0.473	119	1.37
	high	0.465	152	1.44	0.463	147	1.45
0.9	low	0.400	25.0	1.36	0.374	33.2	1.38
	medium	0.460	44.0	1.40	0.466	45.6	1.38
	high	0.459	78.7	1.45	0.462	79.5	1.44
0.7	low	0.380	15.6	1.37	0.353	21.5	1.38
	medium	0.445	11.4	1.42	0.476	14.7	1.37
	high	0.447	50.2	1.48	0.448	49.9	1.48
	Initial belief: (low: 0.2, medium: 0.2, high: 0.6)						
0.995	low	0.402	39.7	1.37	0.398	33.9	1.37
	medium	0.468	158	1.38	0.473	162	1.38
	high	0.455	97.9	1.46	0.464	102	1.45
0.9	low	0.376	17.4	1.38	0.380	25.0	1.38
	medium	0.470	89.8	1.38	0.466	94.7	1.38
	high	0.427	25.4	1.53	0.431	26.7	1.50
0.7	low	0.369	11.8	1.30	0.340	15.5	1.39
	medium	0.469	55.7	1.39	0.466	59.6	1.38
	high	0.291	5.80	1.61	0.337	6.70	1.54

μ_M / *SCV for* γ	*0.02*	*0.04*	*0.08*	*0.1*	*0.12*
$\gamma = 0.02$*; assuming perfect knowledge of* γ*, optimal risk is 1.36 and optimal dose is 9.4 mgs*					
0.01	1.36 (9.3)	1.37 (4.6)	1.39 (2.3)	1.39 (1.9)	1.40 (1.5)
0.5	1.37 (6.1)	1.39 (3.0)	1.40 (1.5)	1.40 (1.2)	1.40 (1.0)
5	1.40 (1.8)	1.40 (0.9)	1.41 (0.5)	1.41 (0.4)	1.41 (0.3)
$\gamma = 0.08$*; assuming perfect knowledge of* γ*, optimal risk is 1.36 and optimal dose is 2.35 mgs*					
0.01	2.25 (9.3)	1.43 (4.6)	1.37 (2.3)	1.37 (1.9)	1.37 (1.5)
0.5	1.55 (6.0)	1.37 (3.0)	1.38 (1.5)	1.38 (1.2)	1.38 (1.0)
5	1.37 (1.8)	1.38 (0.9)	1.40 (0.5)	1.40 (0.4)	1.40 (0.3)
$\gamma = 0.12$*; assuming perfect knowledge of* γ*, optimal risk is 1.36 and optimal dose is 1.6 mgs*					
0.01	5.13 (9.3)	1.69 (4.6)	1.38 (2.3)	1.36 (1.9)	1.36 (1.5)
0.5	2.19 (6.1)	1.42 (3.0)	1.36 (1.5)	1.37 (1.2)	1.37 (1.0)
5	1.36 (1.8)	1.37 (0.9)	1.39 (0.5)	1.39 (0.4)	1.40 (0.3)

Table S4: Expected risk in the maintenance stage as a function of σ_M^2 and μ_M, which are the mean and variance of γ at the beginning of the maintenance stage. In parentheses, we include the optimal dosages required to minimize the risk criterion given that initial belief about γ.

changes in the optimal stable doses and the associated risks. The optimal stable dose, for each combination of μ_M and σ_M^2, is the unique solution to the problem in (5.3) of the main paper. The associated risk is computed as the expected value of our risk function, where we use the true (yet unobservable) value of γ and the (possibly incorrect) dose which is optimal conditional on the physician's belief. Proceeding as such enables us to compute the true risk associated with a belief-based dose prescription.

In order to systematically vary μ_M and σ_M^2, we vary the squared coefficient of variation (SCV) of our distributional belief about γ. The SCV is defined as the variance divided by the square of the mean. We vary the SCV from 0.01 to 5 (this is the SCV column in the table). We consider several values of μ_M, ranging from 0.02 to 0.12. In Table S4, we only report the assumed values for the SCV and μ_M, since the corresponding values of σ_M^2 can be readily deduced. For each value of μ_M and the SCV, we present the optimal expected risk and, in parentheses, the corresponding belief-based optimal dose. Given the physician's belief, the optimal belief-based dose is the same irrespective of the true underlying γ.

Table S4 shows that *the belief formed about* γ *during the initiation stage has a strong impact on both the stable dose in the maintenance stage, and the resulting risk.* This constitutes a strong motivation to learn effectively about γ, as we do in this paper. For example, for $\gamma = 0.08$, $\mu_M = 0.02$, and an SCV for γ equal to 0.01, the risk is 2.25 and the optimal stable dose equal to 9.3 mgs. In contrast, the optimal stable dose for $\gamma = \mu_M = 0.08$, and an SCV for γ equal to 0.01 is 2.3 mgs and the resulting risk is 1.37. Table S4 clearly shows that a wrong initial belief about γ leads to considerably wrong stable dose recommendations.

References

Patrick, A. R., J. Avorn, and N. K. Choudhry. 2009. Cost-effectiveness of genotype-guided warfarin dosing for patients with atrial fibrillation. *Circulation: Cardiovascular Quality and Outcomes,* 2.5: 429-436.