Title: Simulating Sequential Multiple Assignment Randomized Trials to Generate Optimal Personalized Warfarin Dosing Strategies

Running head: Simulating SMARTs to Personalize Dosing

Word count: (manuscript: 3891 words; abstract: 299 words)

Authors (*corresponding):
Benjamin Rich, PhD
McGill University, Department of Epidemiology, Biostatistics, and Occupational Health
1020 Pine Ave W
Montreal, QC H3A 1A2
Canada

*Erica E. M. Moodie, PhD
McGill University, Department of Epidemiology, Biostatistics, and Occupational Health
1020 Pine Ave W
Montreal, QC H3A 1A2
Canada
erica.moodie@mcgill.ca

David A. Stephens, PhD
McGill University, Department of Mathematics and Biostatistics
805 Sherbrooke Street W.
Montreal, QC H3A 2K6
Canada

Financial support provided by Discovery Grants from the Natural Sciences and Engineering Research Council of Canada (EEMM and DAS).
Abstract

Background: Due to the cost and complexity of conducting a Sequential Multiple Assignment Randomized Trial (SMART), it is desirable to pre-define a small number of personalized regimes to study.

Purpose: We proposed a simulation-based approach to studying personalized dosing strategies in contexts were a therapeutic agent’s pharmacokinetic and pharmacodynamics properties are well understood. We take dosing of warfarin as a case study, as its properties are well-understood. We consider a SMART in which there are five intervention points in which dosing may be modified, following a loading phase of treatment.

Methods: Realistic SMARTs are simulated, and two methods of analysis, G-estimation and Q-learning, are used to assess potential personalized dosing strategies.

Results: In settings where outcome modeling may be complex due to the highly non-linear nature of the pharmacokinetic and pharmacodynamic mechanisms of the therapeutic agent, G-estimation provides the more promising method of estimating an optimal dosing strategy. Used in combination with the simulated SMARTs, we were able to improve simulated patient outcomes, and suggest which patient characteristics were needed to best individually tailor dosing. In particular, our simulations suggest that current dosing should be determined by an individual’s current coagulation time as measured by the international
normalized ratio (INR), their last measured INR, and their last dose. Tailoring treatment only based on current INR and last warfarin dose provided inferior control of INR over the course of the trial.

Limitations: The ability of the simulated SMARTs to suggest optimal personalized dosing strategies relies on the pharmacokinetic and pharmacodynamic models used to generate the hypothetical patient profiles. This approach is best suited to therapeutic agents whose effects are well studied.

Conclusions: Prior to investing in a complex randomized trial that involves sequential treatment allocations, simulations should be used where possible in order to guide which dosing strategies to evaluate.

Key words: Adaptive treatment strategies; Adaptive individualized dosing; Anticoagulation therapy; Dynamic treatment regimes; G-estimation; Personalized treatment; Q-learning.
1. Background

In a general treatment setting, there is often interest in adapting treatment to an individual patient if there exists significant heterogeneity in patient response (e.g. if there exist subgroups in whom the treatment may be more or less effective) and if there are serious costs – either health or financial – associated with over- or under-treatment. Under-treatment can reduce patient quality of life, while over-treatment runs the risk of side effects, increased costs, and poor compliance due to treatment fatigue in the context of a clinical trial or real-world care of a chronic condition. It is therefore important to seek effective strategies for when to treat and by how much.

In the context of dosing of a particular drug, balance is typically sought between efficacy and tolerability. The degree to which dosing should be *personalized*, or *individualized*, depends on the variability between individual response to the drug; this individual response to a therapeutic agent can be characterized by its pharmacokinetics (PK), how the agent is dispersed throughout the body, and its pharmacodynamics (PD), the effect of that agent once it is distributed throughout the body. PD encompasses both positive (therapeutic) effects and unintended or adverse effects.
In this work, we will focus on optimal individualized dose finding for warfarin. Warfarin is a highly effective anticoagulant which decreases the risk of thrombosis (clotting). The impact of warfarin can vary considerably, both between people due to genetic and other factors and within individuals over time due in part to dietary choices which can replenish vitamin K and interactions with other medications(1). The appropriate dose required to achieve a measure of blood clotting, the international normalized ratio (INR), within the therapeutic range can vary by five- to twenty-fold(1). Warfarin is widely prescribed in North America and inappropriate dosing is one of the major causes of ER visits resulting from adverse drug effects(2). There is therefore a strong case to be made for finding better strategies for adjusting doses that adapt to an individual’s time-dependent characteristics so as to reduce the rate of adverse events.

Sequential multiple assignment randomized trial (SMARTs) have been developed to evaluate multi-interval treatment strategies. In these trials, participants are randomly assigned to a treatment option; at the next treatment interval, each trial participant is re-randomized among the set of available treatment options given the participant’s history in that trial. Such trials are often designed to test a relatively small number of treatment strategies (e.g., two treatment intervals, with tailoring on a single variable such as whether
the participants responded to treatment in the first interval). However, SMARTs can also be used to explore and suggest adaptive strategies whose therapeutic value would then need to be verified in a subsequent confirmatory trial. For example, in a continuous dose setting such as warfarin, trial participants might undergo a loading phase in which all participants receive the same dose. At the first randomization, participants will be assessed to see whether their INR is high, acceptable, or low. Participants whose INR is high may be randomized to receive a dose that is 5-25% lower than their current dose (i.e. the randomization would determine the amount by which the dose would be decreased), while participants whose INR was too low may be randomized to receive a dose that is 5-25% higher than their current dose. Participants whose INR was acceptable (i.e. close to the target) would maintain their current dose. At the second randomization, the same procedure would be adopted wherein all participants requiring a dose adjustment would receive one, but the specific amount by which the dose would be adjusted would be determined randomly.

We propose a simulation-based approach to studying personalized dosing strategies in contexts were a therapeutic agent’s pharmacokinetic and pharmacodynamic properties are well understood, using warfarin as a case study. We consider a hypothetical SMART in
which there are five intervention points in which dosing may be modified, following a loading phase of treatment, and evaluate three classes of dosing rules estimated using two different algorithms.

2. Methods

2.1 Setting: Warfarin

Warfarin is used to prevent thromboembolic events in patients at high risk. Indications for warfarin use include atrial fibrillation, deep vein thrombosis and prosthetic heart valve replacement. Warfarin acts on the vitamin K cycle, inhibiting the conversion of inactive vitamin K epoxide to the reduced form required for the activation of clotting factors II, VII, IX and X. Warfarin has been in use since the 1950s, and though it has been extensively studied, its dosing continues to present significant challenges (3, 4). Since the discovery of an association between warfarin dose requirement and genetic polymorphisms in the genes CYP2C9 and VKORC1 involved in warfarin metabolism and mechanism of action respectively (5, 6), genetic testing has been viewed as a key component in warfarin dose determination (7, 8). While it continues to hold great promise, the uptake of genetic testing
into clinical practice has not occurred (9), partly because the clinical benefit of genotype-guided dosing has not been confirmed in randomized clinical trials (10). A reason for this is that the genotype contributes only static information which may be useful in the initiation phase (11), however this information quickly loses relevance as INR measurements are made—methods for dynamic treatment optimization such as are discussed in this paper are needed (12, 13). Indeed, many algorithms for adapting treatment have been proposed, e.g. (14-17). One algorithm based on principles of system and control engineering (14) appeared to do better than a standard algorithm currently used in the Netherlands, but nonetheless the new algorithm’s proposal had to be overruled by the physician over 20% of the time (18). Thus, in practice clinical care continues to rely largely on expert physicians who use their experience and judgement combined with trial and error to achieve stable maintenance doses.

2.2 Data-generating models

We make the simplifying assumption that warfarin dosing is continuous. In reality, possible doses will be limited by the form in which the drug is distributed (e.g. whole milligrams),
and so the simplification makes the implicit assumption that optimal doses may be rounded to achieve near-optimal results.

In pharmaceutical sciences, it is a standard paradigm to assume that the effect of a drug is completely mediated through the true systemic exposure to the drug, so that the dose-effect relationship can be decomposed into two distinct components: (1) the dose-concentration relationship (the PK aspect) and (2) the concentration-effect relationship (the PD aspect).

We use a two-level hierarchical population PK/PD model, with random effects on the main structural PK and PD parameters. At the individual level, the PK is described by a standard one-compartment model while the PD follows a standard indirect response model (19). The form of the assumed model creates a delay between administration of dose and observed response. For complete details of parameters used as well as the differential equations that determine the time-courses of drug concentration in the central (blood) compartment and the observed INR, see (20). The delay in response is a realistic model for the warfarin setting in which anticoagulant effects are not instantaneous(19, 21), but only manifest themselves as the body’s current supply of vitamin K reduced is depleted. From
the perspective of estimating an optimal adaptive dosing strategy, however, the delayed response represents a further challenge since the optimal strategy is likely to be one which correctly accounts for carry-over effects from one treatment interval to the next.

In our hypothetical trial, warfarin was taken orally once daily for 21 days, with doses potentially modified on days 4, 7, 10, 13, 16 and 19. The first six days are viewed as a loading phase, so that any INR measures made prior to day 7 are considered baseline variables. The trial begins in earnest on day 7, and is made up of five treatment intervals (days 7-9, 10-12, 13-15, 16-18, and 19-21). In the notation that follows, L(0) is the INR measured on day 7 shifted by 2.5, so that L(0)=0 corresponds to an INR of 2.5, lying exactly at the mid-value of the therapeutic range. D(1) is the dose assigned on the same day. Baseline measures of INR (shifted by 2.5) and dosing from the day 4 (during the loading phase) are denoted L(-1) and D(0), respectively. For simplicity, patients are assumed to be fully compliant in our hypothetical trial, always taking the assigned dose of warfarin at the same time each day. On days when dosage is adjusted, INR is assumed to be measured immediately prior to the administration of the dose.
Doses are determined in the hypothetical study as follows. On day 1 (loading phase), all individuals receive the same dose of 15 mg, which we represent with $A(-2) = 0$. On days 4, 7, 10, 13, 16 and 19 (intervals $j = -1, 0, 1, 2, 3, 4, 5$, respectively), set $A(j) = -0.6*\text{L}(j-1) + 0.8*A(j-1) + \varepsilon$ where $\varepsilon$ is a normal random variable with mean 0 and variance 0.04 representing a small degree of physician discretion built into the trial protocol. Doses are determined by a transformation of this treatment variable, $D(j) = 15*\exp(A(j))$. In this simulation, we have taken $A(j) = 0$ to correspond to a dose of 15 mg/day as the reference dose; this is somewhat higher than the typical maintenance doses for warfarin observed in practice, and could be adjusted along with the PK/PD parameters in the model. We generate a sample of size $n=2000$ in our hypothetical trial. Figure 1 plots sample profiles for three randomly selected participants in the hypothetical trial. Finally, we define our outcome $Y$ to be the negative of the sum of the absolute deviations of the INR from 2.5. The outcome, then, is a non-positive score measuring the cumulative deviations of the INR from the mid-value of the therapeutic range over the course of the follow-up and is thus similar to the time in therapeutic range, an accepted endpoint in the evaluation of warfarin dosing efficacy (3, 13).
2.3 Models for adaptive strategies and methods of estimation

Having simulated a trial, that is using the data which were generated on hypothetical trial participants using known PK/PD characteristics of warfarin, we now proceed to statistical estimation and prediction. The goal is to learn about the best strategies for adapting the warfarin dosing, based on individual characteristics that vary over time. We will consider two approaches to the estimation, Q-learning and G-estimation, described below. These methods will allow us to estimate dosing-rule parameters, which we can then evaluate by predicting the response to different adaptive dosing strategies to see which leads to the best outcomes.

Three potential dosing strategies are considered and evaluated:

Strategy I: Select the treatment level $A$ that maximizes the quadratic function

$$[\psi_{j0} + \psi_{j1}L(j-1) + \psi_{j2}A(j-1)]A + \psi_{j3}A^2$$

Strategy II: Select the treatment level $A$ that maximizes the quadratic function

$$[\psi_{j0} + \psi_{j1}L(j-1) + \psi_{j2}A(j-1) + \psi_{j3}L(j-1)A(j-1)]A + \psi_{j4}A^2$$

Strategy III: Select the treatment level $A$ that maximizes the quadratic function

$$[\psi_{j0} + \psi_{j1}L(j-1) + \psi_{j2}A(j-1) + \psi_{j3}L(j-2)]A + \psi_{j4}A^2$$
Parameters $\psi_{j0}$, $\psi_{j1}$, $\psi_{j2}$, $\psi_{j3}$, and possibly $\psi_{j4}$ are estimated using either Q-learning or G-estimation, both described below. The three strategies differ in the degree of tailoring of the dose, as the three dosing strategies differ according to what information is incorporated into the individualization of the treatment level. Strategy I personalizes dosing based on most recent treatment and current INR; strategy II further personalizes by allowing effect modification between current INR and most recent treatment; strategy III tailors current treatment to the individual based on current INR, most recent treatment, and previous INR. Further note that each strategy is based on a quadratic function of treatment; by introducing curvature into the dependence of the outcome on the treatment, we are allowing for the best treatment (i.e. the treatment that maximizes the outcome) to potentially take its value at some point other than the extreme ends of the treatment range.

2.3.1 Q-learning

Q-learning is one of the most popular methods of estimating optimal adaptive treatment strategies due to the ease with which it can be explained and implemented. For a continuous valued outcome, Q-learning consists of a sequence of linear regressions (or, equivalently, ordinary least squares fits of linear models). Effectively, the data analyst begins at the final
interval of treatment, and estimates the impact of treatment in the last interval on the outcome. That information is then used to estimate that outcome that would have been observed, for each individual in the study, had the optimal dosing strategy been used rather than the trial protocol. This best-possible outcome, or pseudo-outcome, is then used as the response in a regression in which the impact of treatment in the second-last interval is modelled, and the process repeats until the impact of treatment at each interval is modelled, assuming that in all future intervals, participants received optimal dosing.

In the context of our hypothetical trial, the Q-learning procedure is as follows:

1. Fit a linear regression of the outcome Y on treatment at day 19 (interval j = 5).
   The regression model should contain a quadratic term in the case of a continuous dose, as well as effect modification of the treatment effect by any variables with which dosing might be tailored. For example, to estimate strategy I, Y should be fit as a function of L(j-1), A(j-1), A(j), L(j-1)*A(j), A(j-1)*A(j), and A(j)^2.

2. For j = 4, 3, 2, 1 (corresponding to days 16, 13, 10, and 7),
   a. Use the coefficient estimates of the model(s) in the previous step(s) to generate the best-possible outcome, Y*(j), for each individual in the
study. This corresponds to the outcome that would have been observed had each individual followed the study protocol for treatments up to the current interval, then following a dosing strategy that was optimal (as estimated by the Q-learning procedure thus far).

b. Fit a linear regression of the best-possible outcome $Y^*(j)$ on treatment in interval $j$. As in step (1), the regression model should contain a quadratic term in the case of a continuous dose, as well as effect modification of the treatment effect by any variables with which dosing might be tailored. Again taking the example of strategy I, $Y^*(j)$ should be fit as a function of $L(j-1)$, $A(j-1)$, $A(j)$, $L(j-1)*A(j)$, $A(j-1)*A(j)$, and $A(j)^2$.

3. The optimal dosing strategy is then taken to be the sequence of rules, for $j = 1, 2, ..., 5$ that maximizes the regression function at each stage. For strategy I, this corresponds to assigning treatment $A$ such that $A$ maximizes $[\psi_{j0} + \psi_{j1}L(j-1) + \psi_{j2}A(j-1)]A + \psi_{j3}A^2$, where in place of parameters $\psi_{j0}$, $\psi_{j1}$, $\psi_{j2}$, and $\psi_{j3}$, estimated values are used.

See (22) for a slightly more mathematical introduction to Q-learning and a worked example with two stages of intervention and a binary-valued treatment.
For Q-learning to produce unbiased estimators of the optimal dosing strategy, the regression models assumed in each stage of estimation must be correctly specified. Unfortunately, it may not be easy to specify these models, particularly in the context of our hypothetical trial where the data arise from complex (highly non-linear) PK/PD models.

2.3.2 G-estimation

An alternative method of estimating optimal dosing strategies is G-estimation. This approach has not gained the same popularity as Q-learning in part because it is less simple to implement. Conceptually, however, it is very similar to Q-learning and in some instances, the procedures are actually identical(23). G-estimation requires additional modelling relative to Q-learning, but this cost is offset by a gain in robustness: even if the outcome regression models are not correctly specified, G-estimation can produce unbiased estimators of the optimal dosing strategy provided the treatment models can be correctly specified. This is particularly appealing in the context of a randomized trial – hypothetical or otherwise – where the treatment assignment mechanism is known by design.
Conceptually, G-estimation can be viewed as a series of weighted regressions. However the weighting is rather non-standard: rather than using weights that are constrained to be non-negative and sum to 1, here the weights are constrained to have mean 0, and represent deviations from the expected treatment, given covariates. Estimation proceeds as follows:

1. At the last interval, $j = 5$:
   
   a. Fit a likelihood model to the density of the continuous doses, and from this, obtain fitted estimates of $\pi^{(1)}(j) = E[A(j)|A(j-1),L(j),L(j-1),L(j-2);\alpha]$ and $\pi^{(2)}(j) = E[A^2(j)|A(j-1),L(j),L(j-1),L(j-2);\alpha]$. These models are, in effect, generalized propensity score models (3) for the treatment dose, $A(j)$, and for the squared dose, $A^2(j)$.
   
   b. Assume a linear regression for the outcome $Y$ on treatment at day 19 (interval $j = 5$), say $E[Y|A(j),A(j-1),L(j),L(j-1),L(j-2);\psi,\beta] = m_j(a,l;\psi,\beta)$. As in Q-learning, the regression model should contain a quadratic term in the case of a continuous dose, as well as effect modification of the treatment effect by any variables with which dosing might be tailored.
   
   c. Using the estimated generalized propensity scores found in (1a) as a vector of weights, solve the following equation to find estimates of the model parameters in (1b):
\[ \sum_{l=1}^{n} \left( x_{ij} \left( A_{ij} - \pi^{(1)}(j) \right) \right) \{ Y - m_{ij}(a,l;\psi,\beta) \} = 0 \]

where \( X_{ij} \) denotes the vector of covariates (including a constant) proposed as potential tailoring variables in (1b), i.e. those variables thought to modify the effect of the treatment.

2. For \( j = 4, 3, 2, 1 \) (corresponding to days 16, 13, 10, and 7),
   a. Use the coefficient estimates of the model(s) in the previous step(s) to generate the best-possible outcome, \( Y^*(j) \), for each individual in the study. This corresponds to the outcome that would have been observed had each individual followed the study protocol for treatments up to the current interval, then following a dosing strategy that was optimal (as estimated by the G-estimation procedure thus far).
   b. Fit a weighted regression of the best-possible outcome \( Y^*(j) \) on treatment in interval \( j \). As in step (1a), fit a likelihood model to the density of the continuous doses to obtain fitted estimates of \( \pi^{(1)}(j) \) and \( \pi^{(2)}(j) \). As in (1b), propose a regression model, \( m^*_j(a,l;\psi,\beta) \), for the expected value of \( Y^*(j) \) that includes a quadratic term in the case of a continuous dose, as well as effect modification of the treatment effect by
any variables, $X_{ij}$, with which dosing might be tailored. Solve the equation

$$\sum_{i=1}^{n} \left( \frac{X_{ij}}{A_{ij} - \pi^{(1)}(j)} \right) \left\{ Y^{*}(j) - m^{*}(a,l;\psi,\beta) \right\} = 0.$$ 

3. The optimal dosing strategy is then taken to be the sequence of rules, for $j = 1, 2, ..., 5$ that maximizes the regression model $m^{*}(a,l;\psi,\beta)$ at each stage, or simply $m_j(a,l;\psi,\beta)$ in the final interval where $j = 5$. As in Q-learning, this corresponds to assigning treatment $A$ such that $A$ maximizes $[\psi_{j0} + \psi_{j1}L(j-1) + \psi_{j2}A(j-1)]A + \psi_{j3}A^2$ for strategy I, where in place of parameters $\psi_{j0}$, $\psi_{j1}$, $\psi_{j2}$, and $\psi_{j3}$, estimated values are used.

See (24) for details.

2.4 Evaluation of the individualized dosing strategies

We compare the results of our study in an idealized fashion: because our trial participants are simulated, we have access to an “unlimited supply” of individuals. Hence, following our trial in 2000 individuals, we can generate outcomes for a further 1000 individuals under a variety of different dosing strategies. We then compare the distribution of outcomes under each of seven different dosing strategies: (1) the trial protocol, (2)-(4) using strategies I-III
with parameters estimated by Q-learning, and (5)-(7) using strategies I-III with parameters estimated using G-estimation. Additionally, we can show, for each individual in our new test group, which of the seven strategies leads to the most favourable outcome.

3. Results

Figure 2 depicts the optimal dose suggested by Strategy III with parameters estimated by Q-learning (top panel) and G-estimation (bottom panel); current recommended treatment is shown as a function of most recent treatment, A(j-1), and most recent shifted INR, L(j-1), with the second most recent shifted INR, L(j-2), set to 0. Separate figures can be plotted for each level of L(j-2). Purple regions in the plots indicate that warfarin dose should be reduced, while green regions indicate that warfarin dose should be increased; the intensity of the colour indicates the magnitude of the dose adjustment. Our results indicate that the dosing strategies found by G-estimation lead to a smaller range of possible doses, and thus less aggressive changes in dose than suggested by the dosing strategy found by Q-learning in the first interval, and more aggressive changes in dose for intervals 2-5. See, for example, the first column: for the lowest value of A(j-1) and highest value of L(j-1), the dose change suggested by Q-learning is a darker purple than that given by G-estimation,
indicating that the optimal dose found by Q-learning suggests a greater reduction in warfarin dose than the optimal dose found by G-estimation. Conversely, in intervals 2-5, the lower right quadrant of the G-estimation plots are more purple than the Q-learning plots, indicating that G-estimation is suggesting more aggressive changes in dosing. In applying the six estimated dosing strategies, along with the simulated SMART protocol for dose allocation, to a population of new patients, we observe that the dosing strategies estimated by Q-learning outperform the study protocol, but the strategies found by G-estimation provide even greater gains in patient outcomes; see Figure 3. Finally, we consider which of the seven potential strategies (i.e., the trial protocol, Strategies I-III estimated by Q-learning, and Strategies I-III estimated by G-estimation) provides the best outcome for each of the 1000 new individuals. While Figure 3 clearly demonstrates that dosing Strategy III, when estimated by G-estimation, provides the best outcome on average, Figure 4 further shows that it also provides the best strategy for the greatest number of individuals.
4. Limitations

The ability of the simulated SMARTs to suggest optimal personalized dosing strategies relies on the pharmacokinetic and pharmacodynamic models used to generate the hypothetical patient profiles. This approach is best suited to therapeutic agents whose effects are well studied. The implementation of the PK/PD models is not trivial, and requires the use of differential equation solvers. Hence, expertise in computer programming is required, as is considerable computation time if large data-sets or complex scenarios are considered.

In this paper, we have focused on an idealized setting, in which all study participants are fully compliant and INR is measured without error. Because of the simulated nature of the investigations, it would be straightforward to explore deviations from these simplifying assumptions.
5. Conclusions

Prior to investing in a complex randomized trial, simulations should be used where possible in order to guide which dosing strategies to evaluate. Following such an investigation by simulation, a confirmatory study would be required in which trial participants are randomized to specific dosing strategies such as standard care versus the best strategy found in the simulation study. In the work presented here, we have succeeded in implementing a realistic data-generating model for treatment with warfarin, whose PK and PD characteristics are well-studied, and have compared the use of G-estimation and Q-learning as an optimal dose finding strategy. This approach could easily be extended to more complex models (e.g. two-compartment PK models) and other covariates could be included in either the data-generating model, or explored in the analysis models. More importantly, we have demonstrated that it can be very useful to try to implement realistic simulation settings when the biology of a problem is well understood so as to determine which adaptive dosing strategies should be pursued in clinical trials. Where funds are to be invested in a trial, particularly a sequential randomization trial, it is important to do so in the “smartest” way possible.
References


Figure captions.

Figure 1. Dose (column 1), drug concentration (column 3), and INR (column 3) profiles of three randomly selected participants in the hypothetical trial with data generated according to the warfarin PK/PD model. Data are displayed in continuous time, however INR values are recorded only every three days, as indicated by the open circles.

Figure 2. Optimal strategy for treatment adjustments at each interval as a function of previous treatment and current INR, estimated by Q-learning (top panel) and G-estimation (bottom panel) for different combinations of A(j-1) and L(j-1), assuming L(j-2)=0. Dose is then given by D(j) = 15*exp(A(j)).

Figure 3. The distribution of outcomes under receipt of seven different personalized dosing strategies applied to 1000 new individuals: the trial protocol, three strategies suggested by Q-learning, and three suggested by G-estimation.

Figure 4. Bar plot depicting the proportion of the 1000 new individuals in the test set for which each dosing strategy produced the most favourable outcome.