**Precision medicine and Dynamic Treatment Regimes**

Personalized medicine, that is the specific tailoring of disease treatment to a patient according to their individual characteristics, is somewhat an obvious notion, and dates back through the history of medicine.[51,52] Indeed, R. Murugan, writing in the Lancet Respiratory Medicine, refers to Hippocrates as a 'proponent of personalized medicine', who 'assessed several factors, such as a patient's constitution, age, and build, and the time of year to help his decision-making when prescribing treatment.'[52] The covariates of age, gender and race are clear contenders for important personalized medicine factors, well-known to be related to treatment outcomes, and are often included in subgroup and regression analyses.[53,54] In many fields, such as oncology, a decision whether to prolong treatment or not may depend on a patient's current quality of life.[55] In short we are used to the idea of personalizing medical decisions.

Similarly the idea that chronic treatment should take into account the on-going patient response is certainly not new. On the one hand, if a disease is cured, treatment should likely cease. On the other, if a patient fails to respond to a first line treatment, often the decision is made to switch to a second or third. Examples include treatment of diseases with antibiotic resistant strains, such as tuberculosis, and treatment of cancers when a line of chemotherapy does not reduce a tumor.[56,57]

What is a fairly recent idea however, is that in the age of big data, high performance computing, artificial intelligence and machine-learning, the opportunity to tailor treatment to the individual algorithmically now exists.[18,19] To distinguish the notion of using precise algorithmic, formulaic and somewhat

mathematical rules of how a patient should be treated, often factoring in advanced measurements and variables such as those involving genomics and proteomics, from the basic notion that each patient should be treated as an individual, the term precision-medicine is often used.[58] And with the advance of computing and technology, the opportunity for precision-medicine is blossoming.[20]

In the statistical literature precision-medicine rules/algorithms to treat each individual patient are often referred to as either individualized treatment rules, emphasizing the dependence on individual patient characteristics, or else either adaptive treatment strategies, or dynamic treatment regimes, emphasizing the dependence on patient response to on-going treatment.[59,60] In this article we just use the designation of dynamic treatment regime (DTR) to denote these concepts.

The precise notion of a dynamic treatment rule that is then optimal for the population, may be formulated as follows:

First consider the problem of deciding on a one-off treatment for a patient, where N treatment options exist. We notate the treatment option a patient is given by T, and then we assume T=1,…,N. For each patient we have an outcome Y, a (one-dimensional) measure of the success of treatment – we assume a convention that higher values of Y correspond to more successful treatment, and thus are preferable. We denote the $i^{th}$ patient's potential outcome were they to be given treatment T as $Y^*_i(T)$. Assume now that the clinicians have collected information X from each patient (X may be a vector including many different covariates), which they believe provide potentially predictive information as to the results of treatment T for the patient. Then to the clinicians' best knowledge, the $i^{th}$ patient with potential outcome $Y^*_i(T)$, may be predicted to have an expected outcome of $E[Y^*(T)|X_i]$, where the random variable $Y^*(T)$ is the potential outcome for a given

treatment, over patients sampled from the appropriate population. That is, conditional on the treatment and patients' particular covariates $X_i$, we cannot distinguish patients' outcomes. The clinicians' task then becomes for each patient, choose T to maximize the expected outcome $E[Y^*(T)|X]$. We will term this T as the optimal treatment choice.

We note that this is of course not the sole definition of the optimal treatment. For example, perhaps the optimal treatment might reasonably be defined to be the one which gives the highest median, not mean, outcome over the population.[61] Or perhaps we may want to consider a multidimensional Y, which contains different (possibly competing) outcome measures.[62,63] Or perhaps we might want to balance the success of treatment with its cost, regarding a small increase in success not worthwhile if it is accompanied by a dramatic rise in treatment cost.[64] Such extensions have been considered. However defining the optimal treatment as simply that which maximizes the mean outcome is perhaps the most interpretable definition, perhaps the definition most amenable to analysis, and certainly the most prevalent definition in the precision-medicine literature.

So the task of assigning an optimal treatment, that is maximizing the expected outcome $E[Y^*(T)|X]$, becomes essentially bound up with the problem of estimating how $E[Y^*(T)|X]$ varies with X for a given T, that is with estimating the function $E[Y^*(T)|X]=f(X,T)$. One obvious way of doing this would be to posit and solve a regression model for Y* in terms of X and T, for which standard and well-practiced techniques and methods may be used. However other methods are certainly available. The fact that we often might require the covariate space X to be high-dimensional could cause problems with the modeling. To ameliorate this problem we might turn to machine-learning methods to assist with the regression.[65,66,67] Alternatively, other techniques may be used. For example, instead of regarding the problem as a

regression where the treatment outcome must be estimated, it might be regarded as a classification problem, where each patient is classified by their personal optimal treatment group. In this way the problem of estimating the outcome $Y^*(T|X)$ is circumvented. Outcome-weighted-learning is a technique of this type, which employs the support vector machinery to give decent performance with a higher dimensional covariate space.[68,69] In short, there are a large number of methods to estimate the optimal treatment, and these topics are a very active area of biostatistical research.[34,29,33,35]

The situation becomes more elaborate when there are a number of treatment decisions to be made. We may model this situation as having a (discrete) number, M, of timepoints $t_1$, $t_2$,…, $t_M$, and at timepoint $t_j$, we have treatment options $T_j=1,\ldots,N_j$ (where $T_j=1$ might represent an entirely different treatment from $T_i=1$ if $i\neq j$). Again we consider maximizing an ultimate hypothetical outcome $Y^*(T_1,\ldots,T_M|X)$, which is now dependent on each of the M treatment choices, and again conditional on patient covariates X. However the situation is different from the single treatment circumstance above, in that it is complicated by the fact that the covariate space X may include not only time varying coefficients that may be changing between timepoints, but also measured response to treatments already given. The problem is often stated as to repeatedly select treatment at each subsequent timepoint to maximize the patients ultimate outcome as a function of treatment at that timepoint $Y^*(T_j|X(j))$, where the patient covariates X(j) at timepoint $t_j$ now include the patient history of all measured covariates, from timepoints $t_1$ to $t_j$, including the previous treatments assigned at timepoints $t_1,\ldots,t_{j-1}$.

As before this may be regarded as a sequence of regression problems, positing models for $Y^*(T_j|X(j))$ in terms of X(j), and then choosing $T_j$ at each timepoint $t_j$ to maximize the patient's expected final outcome, measuring results at timepoint

$t_{j+1}$, and selecting a next hoped for optimal treatment until the patient finishes treatment at timepoint $t_M$. However we now have the problem of low power, that is if there are many treatment options at a number of timepoints, the total number of treatment combinations may be very large. Hence, the number of patients we observe at each timepoint on each treatment path may be very low and not sufficient to estimate the regression parameters accurately.

Q-learning is one work around for this issue of low power when there are many potential treatment sequences.[29,33,34,35] This compensates for a low number of samples of each treatment path by estimating the optimal outcome achievable for each particular patient at a timepoint, and then (theoretically) reassigning them so that the optimal treatment path has a significant number of patients on it. More specifically, Q-learning begins by looking at the last timepoint M and regressing to find the optimal treatment $T^*_M$, and predicted optimal outcome $Y^*(T_M|X(M))$, for each patient dependent on their covariates and history. Then it pretends each patient in fact received the best treatment possible at timepoint $t_M$, replacing the actual outcome of the patient $Y(T_M)$ with the estimated $E[Y^*(T^*_M|X(M)]$. In this way it estimates for each patient at timepoint $t_{M-1}$ the optimal expected outcome *if* the best treatment had been given at the subsequent timepoints (in this case only one final stage). This is termed the Q-function. Then the optimal final expected outcomes of patients currently at timepoint $t_{M-1}$ can be regressed as a function of $T_{M-1}$ and $X(M-1)$, and then the best treatment for each patient at timepoint $t_{M-1}$ may be estimated, along with corresponding outcome. The Q-function is updated to represent the optimal final expected outcome of the patients as a function of $X(M-2)$ and $T_{M-2}$ assuming the optimal treatments are then given subsequently. Again the optimal $T_{M-2}$ is chosen and this recursive choice of best onwards treatment and updating of the optimal final expected outcome is repeated, moving back through all the stages, until the best treatment $T_1$ at timepoint $t_1$ is estimated. At this point the optimal treatment sequence for each patient, and the optimal expected outcome

if this best treatment regime is adhered to, has been estimated, and each of the solved regressions provides a decision rule for how to assign optimal treatment to a new patient at each stage. In effect we have found the optimal dynamic treatment rule. [29,33,34,35]

Other machine-learning methods instead of Q-learning may be applied. Outcome-weighted-learning has extensions for example that also attempt to combat the low power due to the possible large number of treatment options, such as backwards-outcome-weighted-learning and simultaneous-outcome-weighted-learning.[68,69] Again the various possible methods for finding dynamic treatment regimes and their pros and cons are a subject of very active biostatistical research.[29,33,34,35]

One major statistical issue with the estimation of dynamic treatment regimes is that few of the methods have satisfactory inference methods. That is, while the optimal dynamic treatment regime may be estimated, the results of following this and, importantly, how close this estimate is (in some sense) to the true optimal treatment regime are very non-trivial to estimate.[70,71] Progress continues to be made on this subject.[72,73]

**SMART Studies and Dynamic Treatment Regimes**

To calculate dynamic treatment regimes, whether by a regression based method of estimating $E[Y^*(T_j)|X(j)]$ at a given timepoint j, as essentially required in Q-learning, or whether by classifying patients into groups corresponding to their optimal treatment, as required in outcome-weighted-learning, will require high

quality data, detailing measurements of outcomes and relevant covariates at each timepoint.

Observational data may certainly be used to calculate dynamic treatment regimes, however certain assumptions are necessary.[33,74] We must make the consistency assumption, which states the actual outcome Y for each patient (and the covariates X(j)) observed at each timepoint $t_j$ is indeed the potential outcome for that treatment sequence (that is $Y_i=Y_i^*(T)$ when the $i^{th}$ patient does receive treatment T).[75] We must make the stable unit treatment value assumption, which states that a patient's outcomes and covariates are not affected by how treatment is assigned to them (or assigned to the other patients).[76] And we must make the assumption of no unmeasured confounders, essentially an assumption that treatment assignment is independent of any unmeasured patient covariates.[75]

The important assumption of no unmeasured confounders is unverifiable from observational data.[77] However if treatment is assigned at random (conditional on measured covariates X(j) at timepoint $t_j$) it is automatically met.[34] Hence one proposed method for generating data amenable to calculate dynamic treatment regimes is the Sequential Multiple Assignment Randomized Trial (SMART).[25,28]

In a SMART, patients may be re-randomized to different treatments at subsequent timepoints, with the randomization options and possibilities based on patient characteristics, including previous treatments received and response to them.[28,29] In this way, the investigator hopes that enough relevant treatment combinations are explored over time so that the statistical analysis may pick out salient features influencing outcome, and from this predict how future patients should be treated to maximize each patient’s expected individual outcome.

It may be argued that a SMART study may be more realistic than a classical clinical trial (where patients are randomized independent of response), as it permits non-responders to be switched to a different treatment, as often rightly happens in real-life clinical decisions.[59,78] Nevertheless, there are perhaps inherent difficulties in the implementation (and more so in the associated analyses as discussed above) of a SMART beyond those for classical clinical trials. Increased efforts may need to be made to collect, and process, large enough amounts of data to hope to data-mine dynamic treatment regimes with enough precision, and the possible large number of treatment paths may give requirements for extra and unfamiliar administration and logistics, particularly if treatment re-randomization depends on real time patient responses.

One major issue with the design and implementation of a SMART is that the lack of inference methods regarding the estimation of optimal treatment regimes means powering a SMART for the sole purpose of estimating these is difficult.[73] While theoretical progress on the problem continues to be made, one way around this is to explore the performance through simulation studies. Another simple solution is to power the SMART to perform a classical analysis for a question of interest, such as treatment of embedded fixed non-dynamic treatment regime (or even a pre-chosen dynamic treatment regime).[28,48,17] In the LIBERTI trial we show how this may be done, where the ultimate question of interest is establishing whether there are treatment effects.

Thus SMART studies and their designs, protocols, and analyses may tend to be more complicated than classic randomized clinical trials, and this means extra care may need to be taken in detailing, presenting and explaining these to the appropriate

stakeholders. These issues notwithstanding, SMART studies have already been very usefully employed in many areas, with researchers in psychiatry, oncology and behavioral medicine in particular exploring their advantages and benefits.[79,80,81,82,83,84,34]

**Short-Term Effect Statistical Modeling/Power Calculations**

We study the short-term effect using the change in Vancouver Scar Score at four months and testing three hypotheses regarding the short-term effect. Denoting the change in patient Vancouver Scar Score at four months after treatment with $CO_2$ laser, pulsed-dye laser and just medical therapy as $V_{CO2}$, $V_{PDL}$, $V_{MED}$ respectively, we may rewrite the short-term null hypotheses as:

Short-Term Null Hypothesis 1: $E[V_{PDL}]=E[V_{MED}]$.

Short-Term Null Hypothesis 2: $E[V_{CO2}]=E[V_{MED}]$.

Short-Term Null Hypothesis 3: $E[V_{CO2}]=E[V_{PDL}]$.

We will test the short-term hypotheses, using 2-sided t-tests and (a Hochberg modification of) the Bonferroni multiple comparison procedure, with an overall Type I error rate of 5%.[45,85] We will reject any of the short-term hypotheses with 80% power, at overall size 5%, adjusting for multiple comparisons, if there is an effect size of more than 0.63 standardized units (using standard t-test theory). Pilot studies suggest that we then have an 80% power of observing a minimal clinically significant difference of 0.9 Vancouver Scar Score units in treatment

effect.

**Long-Term Effect Statistical Modeling/Power Calculations**

Our proposed comparison of the long-term (two-year) change in Vancouver Scar Score between sequences of three treatment blocks is less standard than the simple short-term (four-month) change examination. We study the long-term effect using the change in Vancouver Scar Score at twenty-four months and testing two hypotheses regarding the long-term effect. Due to the large number of possible treatment combinations possible over the twenty-four month period, and the need to isolate the effects of differing laser treatments from each other, we must set-up a non-trivial regression model and structure tests carefully.

Specifically, we propose a regression model applied to the treatment effects in each of the blocks. We denote a treatment sequence by a triplet $(i,j,k)$ with each of $i,j,k$ corresponding to (randomized) treatment in the first, second and third block, respectively, and being set to 0, 1, or 2, to represent solely non-surgical medical therapy, $CO_2$ laser, and pulsed-dye laser, respectively. We let the change in Vancouver Scar Score over the study for any given participant on treatment sequence $(i,j,k)$ be denoted by $z_{i,j,k}$. We model the change in scar score by a normal variable, with a global constant variance, as:

$$z_{i,j,k} \sim N(\mu_{i,j,k}, \sigma^2), \quad E[z_{i,j,k}] = \mu_{i,j,k} = \mu + \alpha_{1,i} + \alpha_{2,j} + \alpha_{3,k} + \beta_{i,j} + \gamma_{i,k},$$

subject to the constraints:

$\sum_i \alpha_{x,i}=0$, $\sum_x \alpha_{x,i}=0$, $\sum_j \beta_{p,j}=0$, $\sum_i \beta_{i,q}=0$, $\sum_a \beta_{a,a}=0$, $\sum_k \gamma_{q,k}=0$, $\sum_j \gamma_{j,r}=0$, $\sum_a \gamma_{a,a}=0$.

This model we propose seems clinically reasonable; for example it contains the 'usual' regression model of non-temporal effects and two-way interactions between first and second treatments and also between second and third treatments. The imposed constraints should force the model to identify 'better' and 'worse' treatment options at a given timepoint, as well as 'better' and 'worse' treatment interactions between treatments in the first and second blocks and also between treatments in the second and third blocks. This identification will simply correspond to the signs of the $\alpha_{x,p}$, $\beta_{i,j}$, $\gamma_{j,k}$ respectively. Focusing on the identification of 'better' and 'worse' treatments rather than exact estimation of average treatment group outcomes and effects will potentially give us increased power to determine whether or not there is a laser effect, and subsequently whether or not there is a difference in effect between lasers.

Our mean model has in total 28 parameters, allowing for all values $i$, $j$, and $k$ (including those corresponding to unobserved treatment groups). However, the constraints ensure we have only eleven free parameters. As we have twelve overall treatment groups, only one more treatment group than free parameters, we cannot significantly relax the model further, and the requirement to have our given constraints is somewhat unavoidable (in particular a three-way interaction term must be omitted).

We may interpret these parameters as follows: $\mu$ is the grand mean, $\alpha_{x,p}$ represents the main order effect of having treatment $p$ at timepoint $x$, $\beta_{p,q}$ represents the interactive effect of treatment $p$ at the first timepoint and treatment $q$ at the

second, and $\gamma_{p,q}$ represents the interactive effect of treatment $p$ at the second timepoint and treatment $q$ at the third.

We note $\mu + \alpha_{x,w}$ corresponds to the expected outcome for participants assigned treatment $w$ at timepoint $x$ and randomized at other timepoints equally between treatments. Other more elaborate interpretations may similarly be given. Thus if, for example, $\alpha_{x,u} - \alpha_{x,0} \geq q$, we have an interpretation that there is a treatment, $u$, with an average outcome different from the current standard by more than $q$.

Regarding this SMART trial as a subset of experimental runs of a full $3^3$ factorial trial,[17,31,32] the $\alpha_{x,w}$ have a nice interpretation as the main treatment effects and indeed $\mu + \alpha_{x,w}$ are then marginal effects in the sense that they average the treatment effects of specifying treatment in one block, or factor, over the levels of the remaining two factors.

Under the null hypothesis of no laser effect, the model is certainly true, and all parameters except $\mu$ are 0, in particular all the $\alpha_{x,i}$ must be 0. Further if there is no difference between the effect of $CO_2$ and pulsed-dye laser, the standard regression estimators of the coefficients must yield $\alpha_{x,1} = \alpha_{x,2}$ for each timepoint $x$. Testing $\alpha_{x,i}$ to determine if they obey these relationships is then a valid test of the long-term null hypotheses. Focusing on determining whether there is a laser effect using just the $\alpha_{x,w}$, not all the coefficients in the model, should hopefully increase the power by reducing the degrees of freedom in the tests. We may write (relaxations of) our long-term hypotheses then as:

Long-Term Null Hypothesis 1: (No laser effect) $\alpha_{1,0} = \alpha_{2,0} = \alpha_{3,0} = 0$.

Long-Term Null Hypothesis 2: (No difference in lasers) $\alpha_{1,1}-\alpha_{1,2}=\alpha_{2,1}-\alpha_{2,2}=\alpha_{3,1}-\alpha_{3,2}=0$.

For testing both these long-term null hypotheses, we may then use this regression model and F(2,157)-tests with Type I errors of 5%. We regard the more important determination in the investigation of long-term effects to simply be whether there is any laser effect (that is testing Long-Term Null Hypothesis 1), as opposed to whether there is a difference between effects corresponding to different types of laser (that is testing Long-Term Null Hypothesis 2). Therefore to maximize power when we examine the more important aspect, we use step-down testing, first testing Long-Term Null Hypothesis 1, and only if this is rejected testing Long-Term Null Hypothesis 2. Step-down testing will preserve an overall total Type I error rate of 5%.[46,47,86]

The sample size will give good power to reject the long-term null hypotheses (according to standard linear model theory). The precise conditions on the coefficients which match with a particular test power are slightly difficult to interpret. However we may make the following observations:

1: the power of rejecting Long-Term Null Hypotheses 1 is over 80%, if choosing solely non-surgical medical therapy at a particular timepoint, and assigning random treatment at the other timepoints, results in an outcome difference of more than 0.35 standardized units in comparison to simply randomly assigning treatment at all timepoints.

2: the power of rejecting Long-Term Null Hypotheses 2 (if Long-Term Null Hypotheses 1 is rejected) is over 80%, if choosing $CO_2$ laser therapy at a particular

timepoint, and assigning random treatment at the other timepoints, results in an outcome difference of more than 0.37 standardized units in comparison to assigning pulsed-dye laser therapy at the given timepoint then assigning random treatment at the remaining two timepoints.

More involved power estimates and their corresponding clinical interpretations convince us that we have more than sufficient power. We therefore believe this is a powerful and effective model to compare the overall population effects of the twelve sequences, with the focus being more to identify better and worse options than to precisely extrapolate other unobserved treatment sequences or effects. The model also seems statistically reasonable and it appears to have very good power for rejecting our hypotheses, despite having comparatively large degrees of freedom (as is needed to realistically model the effects of twelve treatment sequences).

**Precision Medicine Dynamic Treatment Statistical Modeling/Power Calculations**

We intend to analyze optimal patient individualized dynamic treatment regimes using machine-learning techniques involving Q-learning[23,87,34] and Outcome-Weighted-Learning.[68,69] We will then contrast this with the estimated effects of a population-level treatment regime, hence obtaining insights into possible benefits of personalizing treatment, or otherwise, in this field.

Accurate estimation of individualized dynamic treatment regimes is the Holy Grail of the research (providing answers automatically to the other hypotheses) but is also expected to be the most challenging, due to the complicated estimations that

will need to be performed.

We will not specify the precise analysis we will use, for we expect significant advances in the precision-medicine methodology whilst the study is underway, and we intend to take advantage of these. We do intend our first exploration to be with the Q-learning framework,[34] and we will also explore the use of the simultaneous-outcome-weighted-learning framework.[68,69]

Explicit sample size calculations are not possible in any straightforward manner for applying Q-learning or outcome-weighted-learning to the data from LIBERTI. However the sample size of 180 (including 10% inflation to allow for dropout) is predicted to be sufficient to provide good insights into an individualized dynamic treatment regime. More specifically, we expect to find a superior personalized treatment (with results within approximately 90% of the true optimal treatment) with probability at least 80%, based on published simulation results for similar studies.[50]

**Simulation Studies**

The modeling and power calculations we detailed for testing the short-term hypotheses are standard and straightforward. However those for testing the long-term hypotheses are less so, therefore for further reassurance we should explore somewhat the performance of our procedure through simulation studies. As explained, modeling and power calculations for a precision-medicine dynamic treatment regime is extremely difficult, and simulations should definitely be used to evaluate the

feasibility of estimating these from data generated by the LIBERTI trial.

Simulations suggest that our models and analysis will on the one hand have very good power to detect a laser effect, and on the other be likely to construct a beneficial dynamic treatment regime for patients. We present two of our simulations below for illustration.

To demonstrate possible performance of the regression model to detect laser effect on a two-year outcome, we take a somewhat Bayesian stance. We simply model the two-year mean expected change in Vancouver Scar Score for each of the twelve treatment sequences as random independent samples from a normal distribution with mean 0 and variance hyperparameter $\sigma_\mu^2$. That is we model $\mu_{i,j,k} \sim N(0,\sigma_\mu^2)$ for each of the twelve sequences $(i,j,k)$ to be observed. We then model 168 patients, equally distributed between these twelve treatment sequences, who have a two-year change in Vancouver Score Scar given by independent samples from $z_{i,j,k} \sim N(\mu_{i,j,k},\sigma^2)$, where $z_{i,j,k}$ is the outcome for a given participant on treatment path $(i,j,k)$. As no relation is assumed between similar treatment paths (eg. knowing the results of even an apparently close treatment path imparts nothing at all about the results of another) this is a worst case scenario for our regression model, which benefits when there are predictable main effects and treatment interactions.

We apply the regression model to our simulated data and determine the power as a function of $\sigma_\mu/\sigma$. The corresponding power curves for rejecting Long-Term Null Hypothesis 1, (no treatment effect), and for rejecting both firstly Long-Term Null Hypothesis 1, and then subsequently Long-Term Null Hypothesis 2 (no difference between types of lasers) are shown in Figure 2. It appears we have good power in these circumstances, with the power becoming acceptable when the variance between

the expected mean outcomes of the treatment paths means, $\sigma_\mu^2$, approaches that of the variance of actual patients on each path, $\sigma^2$. This can be regarded as an appropriate minimal clinically significant difference we would want to detect in this scenario.

Next we also simulate two tailoring variables for each patient (which in this set-up have no effect whatsoever on patient outcome, but just add challenging noise to the estimation of dynamic treatment regimes). We use basic Q-learning (approximating the Q-functions by linear regressions in tailoring variables, current treatment and interactions of current treatment and tailoring variable) to determine a dynamic treatment regime. The precise process is detailed further below in the details for the second main simulation. A further 1680 patients are simulated and treated according to this dynamic treatment regime. Figure 3 shows the results of the experiment. We might expect Q-learning to do very little in this scenario, as there is no structure to exploit between the treatment outcomes. We observe that while the dynamic treatment regime does not always produce a particularly good outcome (for example not infrequently the expected patient benefit under the dynamic treatment regime is above 0), it is reassuringly almost always superior to randomly assigned treatment (even though the Q-functions are extremely misspecified).

To explore the effects of Q-learning in a more realistic set-up, we explore various models for the laser effect. For example we may model percentage change in scar score from one stage to the next as a log-normal variable with mean dependent on current scar score, and a binary time-independent variable representing race, with the added mechanism that the scar score may not increase by more than some percentage randomly assigned to each patient. We may choose coefficients to represent pilot data somewhat, and we may build in certain rules such as: a time

varying effect with treatment in block 2 leading to greatest change, a race effect, such as one laser has no effect on one race, a scar score interaction effect such that one laser is likely beneficial for higher scar scores, but detrimental for lower scar scores, and complex interaction effects such as the effect of laser two being multiplied by the number of treatments of laser one received.

Precisely, we present a simulation where Vancouver Scar Score (VSS) changes throughout treatment according to the following rules:

Patients have tailoring variables of: (Race), which we simulate as binary, and VSS before each treatment is received. We denoted VSS after the $i^{th}$ treatment block is received as $(VSS)_i$ for $0 \leq i \leq 3$. That is $(VSS)_0$ is patient initial VSS before treatment, $(VSS)_1$, that after the first treatment block, $(VSS)_2$ that after the second, and $(VSS)_3$ that after the third and final treatment. We assume no further improvement from the results of the final treatment to the longer-term Vancouver Scar Score at the two-year mark.

Our population is initially simulated by generating 168 patients, with (Race) distributed binomially with p=0.5, and $(VSS)_0 \sim \min(N(10.2,1),13)$ a normal with mean 10.2, variance 1, and truncated above at 13.

For each patient, we simulate:

$$(VSS)_1 \sim \min((VSS)_0 x \exp(N(a, 0.3^2)),\ 0.800x(VSS)_0 + U[0,0.4],\ 13),$$

where N is a normal variable with the value of $a$ determined by:

$$a = -0.500(PDL1)(Race) - 0.500((VSS)_0 - 6)(CDL1)(1 - 5(Race)/6)$$

with PDL1=1 if pulsed-dye laser is assigned in treatment block 1, 0 otherwise, and CDL1=1 if $CO_2$ laser is assigned in treatment block 1, 0 otherwise, and U is a uniform random variable.

Next, we simulate:

$$(VSS)_2 \sim \min((VSS)_1 \times \exp(N(b, 0.15^2)),\ 0.800 \times (VSS)_0 + U[0,0.3],\ 13),$$

where $b$ is determined by:

$$b = -0.625(PDL2)(Race) - 0.625((VSS)_1 - 6)(CDL2)(1+PDL1)(1-5(Race)/6)$$

and now PDL2=1 if pulsed-dye laser is assigned in treatment block 2, 0 otherwise, and CDL2=1 if $CO_2$ laser is assigned in treatment block 2, 0 otherwise.

Finally, we simulate:

$$(VSS)_3 \sim \min((VSS)_2 \times \exp(N(c, 0.15^2)),\ 0.900 \times (VSS)_0 + U[0,0.2],\ 13),$$

where $c$ is determined by:

$$c = -0.375(PDL3)(Race) - 0.400((VSS)_2 - 6)(CDL3)(1+PDL1+PDL2)(1-5(Race)/6)$$

and now PDL3=1 if pulsed-dye laser is assigned in treatment block 3, 0 otherwise, and CDL3=1 if $CO_2$ laser is assigned in treatment block 3, 0 otherwise.

We repeatedly model the scar score trajectories of 168 patients under such rules, then apply Q-learning to predict an optimal dynamic treatment regime. We use a very

simple model for the Q-functions, that is we approximate them through linear regressions (which then necessarily must be misspecified). More precisely, we choose the outcome for the Q-function at each stage to be $(VSS)_3-(VSS)_0$, model each Q-function as a normal variable of unspecified variance, and model population means as follows:

$$E[Q_j]=(a_j+b_j(Race)+c_j(VSS)_{j-1})+(d_j+e_j(Race)+f_j(VSS)_{j-1})(PDL(j))$$
$$+ (g_j+h_j(Race)+i_j(VSS)_{j-1})(CDL(j))$$

for each j=1,2,3. PDL(j) is 1 if the $j^{th}$ treatment is pulsed-dye laser, 0 otherwise, and CDL(j) is 1 if the $j^{th}$ treatment is $CO_2$ laser, 0 otherwise.

The Q-functions are hence seen to be very misspecified, for example due to modeling $(VSS)_3-(VSS)_0$ using a normal distribution instead of $\log((VSS)_3/(VSS)_0)$, due to not allowing for a truncation effect, and due to not explicitly allowing for interactive treatment effects. Nevertheless we solve the corresponding regressions using the data from our 168 patients (assuming that of the original 180 patients 10% dropout), and theoretically reassigning patients to their most beneficial treatment at each stage, then estimate the dynamic treatment regime generated by these Q-functions.

Next we then model the results of applying this dynamic treatment regime to 1680 new patients. Figure 4 shows a plot of simulated final scar score against initial scar score for one simulation run of trial patients, and then new patients following the estimated dynamic treatment regime. It is observable that the dynamic treatment regime appears to weed out less effective or detrimental treatments, and favors better personalized treatments.

Over 100 simulations the average decrease in scar score for the trial patients is 5.6 units, but for the new patients is 7.4 units, roughly a 40% improvement. This, and other simulations, both for the LIBERTI study, and results exploring Q-learning in general, suggest the Q-learning framework may be most beneficial.

We note that the not insignificant performance in these simulations was obtained using even a very crude form of Q-learning, (modeling the Q-functions through linear regressions, with misspecified models for the Q-functions).

We finally also remark how our regression model for examining the long-term null hypotheses would have performed in this second simulation. A laser effect and then subsequently a difference in effect between lasers would be detected, that is both Long-Term Null Hypothesis 1 and Long-Term Null Hypothesis 2 would be rejected in over 80% of these simulations.

**Patient Variables/Evaluation**

From each patient, data will be gathered in several broad areas including: Demographic Data, Injury Factors, Treatment Factors, and Study Assessment Factors.

Demographic Data will include routine information such as:
Age, Gender, Ethnicity, Fitzpatrick Type (a measure of skin pigmentation and response to ultra-violet light),[88] Height, Weight, Body Mass Index (BMI), Medical

Comorbidities, Participant Zip Code, and Alcohol, Tobacco, and/or Recreational Drug Use.

Injury and Treatment Factors related to burn injury will include: Mechanism of Injury, Total Burn Surface Area (TBSA),[89] Utilization of Medications Prescribed for Burn Scars (including narcotics, narcotic analogs, non-steroidal anti-inflammatory drugs (NSAIDs), antihistamines, GABA analogs, antidepressants), Time From Injury to Presentation, and Length of Hospital Stay.

Study Assessment Factors will include: Skin Measurements, Validated Scar Assessment Scales, Quality of Life Metrics, and Photography/Videography Records.

The study assessment factor category denotes the variables that may be used to determine impact of lasers on hypertrophic burn scars. These variables, measured at baseline and each treatment, may also be important tailoring variables. It is therefore most important that these variables are precisely quantified and measured.

For skin measurements, a representative $1cm^2$ area that best reflects the overall scar pattern will be chosen by the investigator and participant for skin measurement at the participant's first treatment visit. This anchor spot will be photographed and identified on a representative diagram with appropriate measurements to allow for continued tracking throughout the study. A similar area of healthy skin will also be identified at the participant's first treatment visit and measured for purposes of color and elasticity comparison. Two nearby spots that are also representative of the overall scar pattern will also be selected. These

will be 1cm$^2$ areas within 5cm of the anchor spot and at least 1cm apart from each other. At each visit a series of measurements will be recorded for the anchor spot and the two nearby spots. Skin measurements will consist of the following:

Elasticity (assessed using a Cutometer® dual MPA 580 (CK Electronics, Cologne, Germany) with an 8 mm spot size),

Color/Hyperemia (assessed using a CR-400 Chromo Meter (Konica Minolta Sensing Americas, Ramsey, NJ) with an 8 mm spot size),

Scar Density/Thickness: (assessed utilizing 2D imaging with a Sonosite M-Turbo (FUJIFILM Sonosite, Inc., Bothell, WA)). □

We will use two validated scar scoring systems to assess the scar, the Vancouver Scar Scale and the Patient and Observer Scar Assessment Scale.[90,42] We will also record the UNC 5p Scar Scale.[3] The VSS is a 14-point scale based on standardized descriptions of pigmentation (0-3), vascularity (0-3), pliability (0-5), and height (0-3). The POSAS is a combined patient and observer scoring system with each part containing six items scored on 10-point scales with standardized descriptions. The clinical observer rates the scar on its vascularity, pigmentation, thickness, relief, and pliability. The patient rates the pain, itching, color, stiffness, thickness, and irregularity of the scar. The UNC 5p consists of a 15-point scale for pruritus (0-3), pain (0-3), paresthesia (0-3), pliability (0-3), and perception (0-3).

To assess overall quality of life, a validated, standardized questionnaire will be used. For adults, this will be the SF-12 Health Survey, which is a modified version of the original SF-36.[91] For children under the age of 18, this will be the

Pediatric Quality of Life Inventory.[92]

Standardized digital photography will be performed to document the locations of the selected study area and anchor spot. Changes to scar appearance over time will also be documented with photographs. Video will be obtained to document changes in scar characteristics with mobility that cannot be captured with still photography. Photo and video images are for documentation purposes, and for controlling the quality of measurements. They will not be used in the scar assessment scales.

**References**

1: American Burn Association. Burn Incidence Fact Sheet [Internet]. Chicago, IL: American Burn Association; 2017 [cited 2017 Nov 1]. Available from: http://ameriburn.org/who-we-are/media/burn-incidence-fact-sheet/

2. Finnerty CC, Jeschke MG, Branski LK, Barret JP, Dziewulski P, Herndon DN. Hypertrophic scarring: the greatest unmet challenge after burn injury. The Lancet. 2016 Oct 7;388(10052):1427-36.

3. Hultman CS, Edkins RE, Wu C, Calvert CT, Cairns BA. Prospective, before-after cohort study to assess the efficacy of laser therapy on hypertrophic burn scars. Annals of plastic surgery. 2013 May 1;70(5):521-6.

4. Arno AI, Gauglitz GG, Barret JP, Jeschke MG. Up-to-date approach to manage keloids and hypertrophic scars: a useful guide. Burns. 2014 Nov 30;40(7):1255-66.

5. Bayat A, McGrouther DA, Ferguson MW. Skin scarring. BMJ: British Medical Journal. 2003 Jan 11;326(7380):88.

6. Brown BC, McKenna SP, Siddhi K, McGrouther DA, Bayat A. The hidden cost of skin scars: quality of life after skin scarring. Journal of Plastic, Reconstructive & Aesthetic Surgery. 2008 Sep 30;61(9):1049-58.

7. Friedstat JS, Hultman CS. Hypertrophic burn scar management: what does the evidence show? A systematic review of randomized controlled trials. Annals of plastic surgery. 2014 Jun 1;72(6):S198-201.

8. Porter C, Tompkins RG, Finnerty CC, Sidossis LS, Suman OE, Herndon DN. The metabolic stress response to burn trauma: current understanding and therapies. The Lancet. 2016 Oct 7;388(10052):1417-26.

9. Hultman CS, Edkins RE, Lee CN, Calvert CT, Cairns BA. Shine on: review of laser- and light-based therapies for the treatment of burn scars. Dermatology research and practice. 2012 Jun 20;2012.

10. Bloemen MC, van der Veer WM, Ulrich MM, van Zuijlen PP, Niessen FB, Middelkoop E. Prevention and curative management of hypertrophic scar formation. Burns. 2009 Jun 30;35(4):463-75.

11. Kerwin, L.Y., Tal, E., Kader, A., Stiff, M.A. and Fakhouri, T.M., 2014. Scar prevention and remodeling: a review of the medical, surgical, topical and light treatment approaches. International journal of dermatology, 53(8), pp.922-936.

12. Fontana CR, Bonini D, Bagnato VS. A 12-month follow-up of hypopigmentation after laser hair removal. Journal of Cosmetic and Laser Therapy. 2013 Apr 1;15(2):80-4.

13. Kidd M, Hultman CS, Van Aalst J, Calvert C, Peck MD, Cairns BA. The contemporary management of electrical injuries: resuscitation, reconstruction, rehabilitation. Annals of plastic surgery. 2007 Mar 1;58(3):273-8.

14. Duke J, Wood F, Semmens J, Edgar DW, Spilsbury K, Rea S. An assessment of burn injury hospitalisations of adolescents and young adults in Western Australia, 1983-2008. Burns. 2012 Feb 29;38(1):128-35.

15. Barber RC, Chang LY, Purdue GF, Hunt JL, Arnoldo BD, Aragaki CC, Horton JW. Detecting genetic predisposition for complicated clinical outcomes after burn injury. Burns. 2006 Nov 30;32(7):821-7.

16. Schwacha MG, Holland LT, Chaudry IH, Messina JL. Genetic variability in the immune-inflammatory response after major burn injury. Shock. 2005 Feb 1;23(2):123-8.

17: Collins LM, Nahum-Shani I, Almirall D. Optimization of behavioral dynamic treatment regimens based on the sequential, multiple assignment, randomized trial (SMART). Clinical Trials. 2014 Aug;11(4):426-34.

18. Hamburg MA, Collins FS. The path to personalized medicine. N Engl J Med. 2010 Jul 22;2010(363):301-4.

19. Mirnezami R, Nicholson J, Darzi A. Preparing for precision medicine. New England Journal of Medicine. 2012 Feb 9;366(6):489-91.

20. Murdoch TB, Detsky AS. The inevitable application of big data to health care. Jama. 2013 Apr 3;309(13):1351-2.

21. Alyass A, Turcotte M, Meyre D. From big data analysis to personalized medicine for all: challenges and opportunities. BMC medical genomics. 2015 Jun 27;8(1):33.

22. Moodie EE, Richardson TS, Stephens DA. Demystifying optimal dynamic treatment regimes. Biometrics. 2007 Jun 1;63(2):447-55.

23. Laber EB, Lizotte DJ, Qian M, Pelham WE, Murphy SA. Dynamic treatment regimes: Technical challenges and applications. Electronic journal of statistics. 2014 Jan 1;8(1):1225.

24. Murphy SA. Optimal dynamic treatment regimes. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2003 May 1;65(2):331-55.

25. Lavori PW, Dawson R. A design for testing clinical strategies: biased adaptive within□subject randomization. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2000 Jan 1;163(1):29-38.

26. Lavori PW, Dawson R. Dynamic treatment regimes: practical design considerations. Clinical trials. 2004 Feb;1(1):9-20.

27. Dawson R, Lavori PW. Placebo free designs for evaluating new mental health treatments: the use of adaptive treatment strategies. Statistics in medicine. 2004 Nov 15;23(21):3249-62.

28. Murphy SA. An experimental design for the development of adaptive treatment strategies. Statistics in medicine. 2005 May 30;24(10):1455-81.

29. Zhao YQ, Laber EB. Estimation of optimal dynamic treatment regimes. Clinical Trials. 2014 Aug;11(4):400-7.

30. Nahum-Shani I, Qian M, Almirall D, Pelham WE, Gnagy B, Fabiano GA, Waxmonsky JG, Yu J, Murphy SA. Experimental design and primary data analysis methods for comparing adaptive interventions. Psychological methods. 2012 Dec;17(4):457.

31: Murphy SA, Bingham D. Screening experiments for developing dynamic treatment regimes. Journal of the American Statistical Association. 2009 Mar 1;104(485):391-408.

32: Box GE, Hunter WG, Hunter JS. Statistics for experimenters: an introduction to design, data analysis, and model building. New York: Wiley; 1978 Jun.

33. Kosorok MR, Moodie EE, editors. Adaptive treatment strategies in practice: planning trials and analyzing data for personalized medicine. Society for Industrial and Applied Mathematics; 2015 Dec 8.

34. Schulte PJ, Tsiatis AA, Laber EB, Davidian M. Q-and A-learning methods for estimating optimal dynamic treatment regimes. Statistical science: a review journal of the Institute of Mathematical Statistics. 2014 Nov;29(4):640.

35. Chakraborty B, Moodie EE. Statistical methods for dynamic treatment regimes. New York: Springer; 2013.

36. Fearmonti R, Bond J, Erdmann D, Levinson H. A review of scar scales and scar measuring devices. Eplasty. 2010;10:e43.

37. Hultman CS, Saou MA, Roach ST, Hultman SC, Cairns BA, Massey S, Koenig HG. To heal and restore broken bodies: a retrospective, descriptive study of the role and impact of pastoral care in the treatment of patients with burn injury. Annals of plastic surgery. 2014 Mar 1;72(3):289-94.

38. Liu A, Moy RL, Ozog DM. Current methods employed in the prevention and minimization of surgical scars. Dermatologic Surgery. 2011 Dec 1;37(12):1740-6.

39. Chan HH, Wong DS, Ho WS, Lam LK, Wei W. The use of pulsed dye laser for the prevention and treatment of hypertrophic scars in Chinese persons. Dermatologic surgery. 2004 Jul 1;30(7):987-94.

40. Batta K, Goodyear HM, Moss C, Williams HC, Hiller L, Waters R. Randomised controlled study of early pulsed dye laser treatment of uncomplicated childhood haemangiomas: results of a 1-year analysis. The Lancet. 2002 Aug 17;360(9332):521-7.

41. Edwards RR, Smith MT, Klick B, Magyar-Russell G, Haythornthwaite JA, Holavanahalli R, Patterson DR, Blakeney P, Lezotte D, McKibben J, Fauerbach JA. Symptoms of depression and anxiety as unique predictors of pain-related outcomes following burn injury. Annals of Behavioral Medicine. 2007 Oct 1;34(3):313-22.

42. Draaijers LJ, Tempelman FR, Botman YA, Tuinebreijer WE, Middelkoop E, Kreis RW, van Zuijlen PP. The patient and observer scar assessment scale: a reliable and feasible tool for scar evaluation. Plastic and reconstructive Surgery. 2004 Jun 1;113(7):1960-5.

43. Anisimov VV, Yeung WY, Coad DS. Imbalance properties of centre□stratified permuted-block and complete randomisation for several treatments in a clinical trial. Statistics in medicine. 2017 Apr 15;36(8):1302-18.

44. Little RJ, Rubin DB. Statistical analysis with missing data. John Wiley & Sons; 2014 Aug 25.

45. Hsu H, Lachenbruch PA. Paired t test. Wiley Encyclopedia of Clinical Trials.

2008.

46. Morrison DF. Multivariate analysis, overview. John Wiley & Sons, Ltd; 1998.

47. Selvin S. F Distributions. Encyclopedia of Biostatistics. 1998.

48: Li Z, Murphy SA. Sample size formulae for two-stage randomized trials with survival outcomes. Biometrika. 2011 Jul 13;98(3):503-18.

49: Wu CJ, Hamada MS. Experiments: planning, analysis, and optimization. John Wiley & Sons; 2011 Sep 20.

50. Zhao Y, Zeng D, Socinski MA, Kosorok MR. Reinforcement learning strategies for clinical trials in nonsmall cell lung cancer. Biometrics. 2011 Dec 1;67(4):1422-33.

51. Abrahams E, Silver M. The history of personalized medicine. Integrative neuroscience and personalized medicine. 2010 Dec 21:3-16.

52. Murugan, Raghavan. "Movement towards personalised medicine in the ICU." The Lancet Respiratory Medicine 3.1 (2015): 10-12.

53. Alosh M, Huque MF, Bretz F, D'Agostino RB. Tutorial on statistical considerations on subgroup analysis in confirmatory clinical trials. Statistics in medicine. 2017 Apr 15;36(8):1334-60.

54. Gabler NB, Duan N, Liao D, Elmore JG, Ganiats TG, Kravitz RL. Dealing with heterogeneity of treatment effects: is the literature up to the challenge?. Trials. 2009 Jun 19;10(1):43.

55. Harrington SE, Smith TJ. The role of chemotherapy at the end of life:”when is

enough, enough?". Jama. 2008 Jun 11;299(22):2667–78.

56. Espinal MA, Kim SJ, Suarez PG, Kam KM, Khomenko AG, Migliori GB, Baéz J, Kochi A, Dye C, Raviglione MC. Standard short-course chemotherapy for drug-resistant tuberculosis: treatment outcomes in 6 countries. Jama. 2000 May 17;283(19):2537–45.

57. Gerber DE, Rasco DW, Le P, Yan J, Dowell JE, Xie Y. Predictors and impact of second-line chemotherapy for advanced non-small cell lung cancer in the United States: real-world considerations for maintenance therapy. Journal of Thoracic Oncology. 2011 Feb 28;6(2):365–71.

58. Katsnelson A. Momentum grows to make 'personalized' medicine more 'precise'. Nat Med 2013 03;19(3):249.

59. Lavori PW, Dawson R. Introduction to dynamic treatment strategies and sequential multiple assignment randomization. Clinical Trials. 2014 Aug;11(4):393–9.

60. Petersen ML, Deeks SG, van der Laan MJ. Individualized treatment rules: Generating candidate clinical trials. Statistics in medicine. 2007 Nov 10;26(25):4578–601.

61. Wang L, Zhou Y, Song R, Sherwood B. Quantile-Optimal Treatment Regimes. Journal of the American Statistical Association. 2017 May 20 (in press).

62. Butler EL, Laber EB, Davis SM, Kosorok MR. Incorporating Patient Preferences into Estimation of Optimal Individualized Treatment Rules. Biometrics. 2017 Jul 25.

63. Laber EB, Lizotte DJ, Ferguson B. Set-valued dynamic treatment regimes for competing outcomes. Biometrics. 2014 Mar 1;70(1):53–61.

64. Huang Y, Laber E. Personalized evaluation of biomarker value: a cost-benefit perspective. Statistics in biosciences. 2016 Jun 1;8(1):43-65.

65. Kneip A, Sarda P. Factor models and variable selection in high-dimensional regression analysis. The Annals of Statistics. 2011;39(5):2410-47.

66. Meinshausen N, Meier L, Bühlmann P. P-values for high-dimensional regression. Journal of the American Statistical Association. 2009 Dec 1;104(488):1671-81.

67. Javanmard A, Montanari A. Confidence intervals and hypothesis testing for high-dimensional regression. Journal of Machine Learning Research. 2014 Oct 1;15(1):2869-909.

68. Zhao Y, Zeng D, Rush AJ, Kosorok MR. Estimating individualized treatment rules using outcome weighted learning. Journal of the American Statistical Association. 2012 Sep 1;107(499):1106-18.

69. Zhao YQ, Zeng D, Laber EB, Kosorok MR. New statistical learning methods for estimating optimal dynamic treatment regimes. Journal of the American Statistical Association. 2015 Apr 3;110(510):583-98.

70. Laber EB, Qian M, Lizotte DJ, Pelham WE, Murphy SA. Statistical inference in dynamic treatment regimes. arXiv preprint arXiv:1006.5831. 2010 Jun 30.

71. Chakraborty B, Murphy S, Strecher V. Inference for non-regular parameters in optimal dynamic treatment regimes. Statistical methods in medical research. 2010 Jun;19(3):317-43.

72. van der Laan MJ, Luedtke AR. Targeted learning of the mean outcome under an

optimal dynamic treatment rule. Journal of causal inference. 2015 Mar 1;3(1):61-95.

73. Luedtke AR, Van Der Laan MJ. Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy. The Annals of Statistics. 2016;44(2):713-42.

74. Cain LE, Saag MS, Petersen M, May MT, Ingle SM, Logan R, Robins JM, Abgrall S, Shepherd BE, Deeks SG, John Gill M. Using observational data to emulate a randomized trial of dynamic treatment-switching strategies: an application to antiretroviral therapy. International journal of epidemiology. 2016 Dec 1;45(6):2038-49.

75. Robins JM. Correcting for non-compliance in randomized trials using structural nested mean models. Communications in Statistics-Theory and methods. 1994 Jan 1;23(8):2379-412.

76. Rubin DB. Bayesian inference for causal effects: The role of randomization. The Annals of statistics. 1978 Jan 1:34-58.

77. Byrd JB, Ho PM. The possibility of unmeasured confounding variables in observational studies: a forgotten fact? Heart 2011 11;97(22):1815.

78. Almirall D, Compton SN, Rynn MA, Walkup JT, Murphy SA. SMARTer discontinuation trial designs for developing an adaptive treatment strategy. Journal of child and adolescent psychopharmacology. 2012 Oct 1;22(5):364-74.

79. Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, Niederehe G, Thase ME, Lavori PW, Lebowitz BD, McGrath PJ. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR* D report. American Journal of Psychiatry. 2006 Nov;163(11):1905-17.

80. Wang L, Rotnitzky A, Lin X, Millikan RE, Thall PF. Evaluation of viable dynamic treatment regimes in a sequentially randomized trial of advanced prostate cancer. Journal of the American Statistical Association. 2012 Jun 1;107(498):493-508.

81. Naar-King S, Ellis DA, Idalski Carcone A, Templin T, Jacques-Tiura AJ, Brogan Hartlieb K, Cunningham P, Jen KL. Sequential multiple assignment randomized trial (SMART) to construct weight loss interventions for African American adolescents. Journal of Clinical Child & Adolescent Psychology. 2016 Jul 3;45(4):428-41.

82. Almirall D, Nahum-Shani I, Sherwood NE, Murphy SA. Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research. Translational behavioral medicine. 2014 Sep 1;4(3):260-74.

83. Liu Y, Zeng D, Wang Y. Use of personalized dynamic treatment regimes (DTRs) and Sequential Multiple Assignment Randomized Trials (SMARTs) in mental health studies. Shanghai archives of psychiatry. 2014 Dec 1;26(6):376.

84. Kidwell KM. SMART designs in cancer research: Past, present, and future. Clinical Trials. 2014 Apr 14:1740774514525691.

85. Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. Biometrika. 1988 Dec 1;75(4):800-2.

86. Lai TL. Sequential analysis. John Wiley & Sons, Ltd; 2001 Apr 1.

87. Song R, Wang W, Zeng D, Kosorok MR. Penalized Q-Learning for Dynamic Treatment Regimens. Statistica Sinica. 2015 Jul;25(3):901-20.

88. Roberts WE. Skin type classification systems old and new. Dermatologic clinics.

2009 Oct 31;27(4):529–33.

89. Giretzlehner M, Dirnberger J, Owen R, Haller HL, Lumenta DB, Kamolz LP. The determination of total burn surface area: how much difference?. Burns. 2013 Sep 30;39(6):1107–13.

90. Baryza MJ, Baryza GA. The Vancouver Scar Scale: an administration tool and its interrater reliability. Journal of Burn Care & Research. 1995 Sep 1;16(5):535–8.

91. Brazier JE, Harper R, Jones NM, O'cathain A, Thomas KJ, Usherwood T, Westlake L. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. Bmj. 1992 Jul 18;305(6846):160–4.

92. Varni JW, Burwinkle TM, Seid M, Skarr D. The PedsQL™* 4.0 as a pediatric population health measure: feasibility, reliability, and validity. Ambulatory Pediatrics. 2003 Dec 31;3(6):329–41.

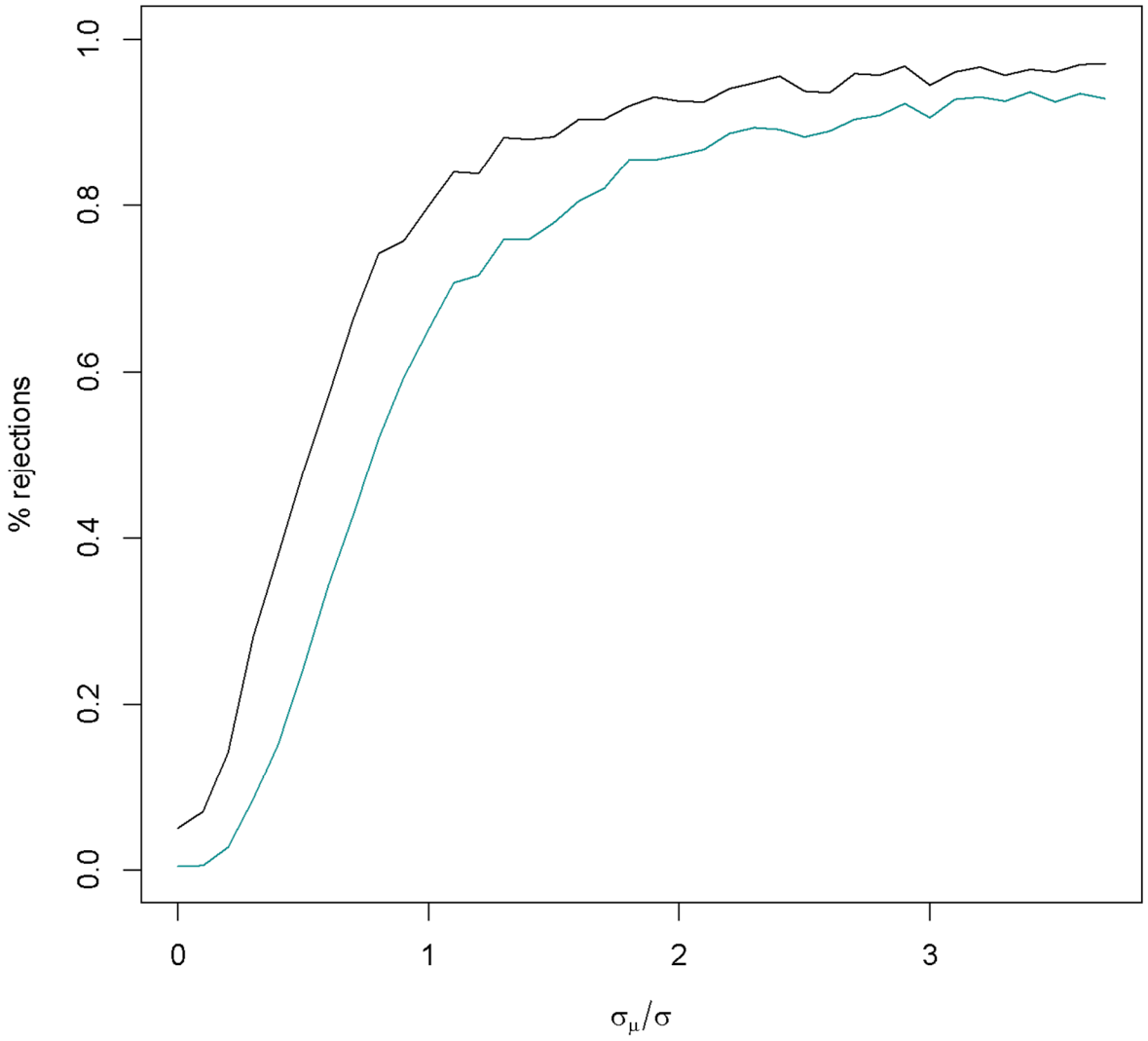


Figure 2: Power curves showing, in black, percentage of times $H_{01}$ was rejected in simulations, and, in blue, percentage of times both firstly $H_{01}$ and then $H_{02}$ was rejected in simulations, as functions of $\sigma_\mu/\sigma$.

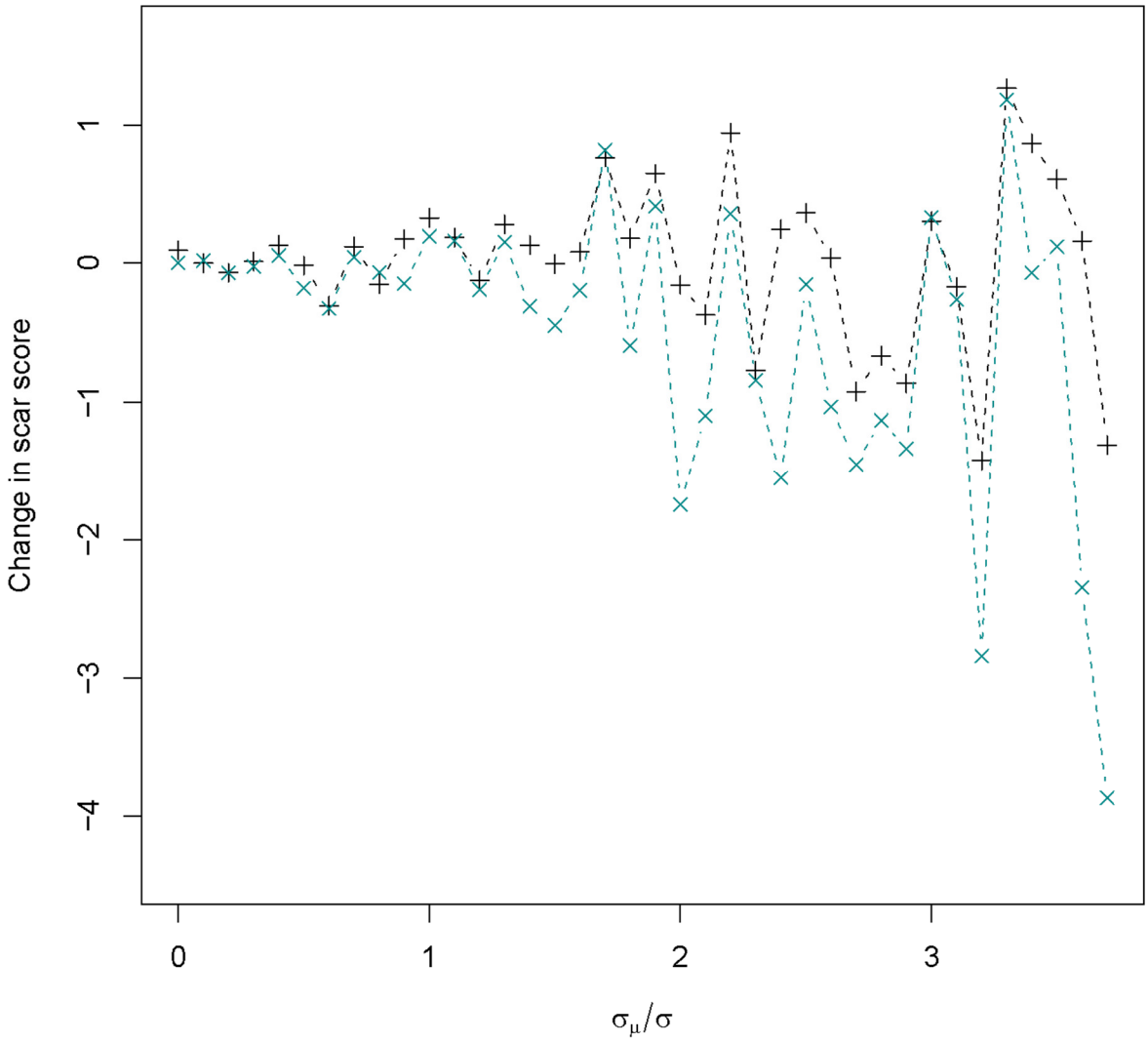


Figure 3: Simulated change of patient scar scores for patients in trial, shown as black '+', and for patients subsequently following the discovered DTR, shown as blue 'x', plotted against $\sigma_\mu/\sigma$.

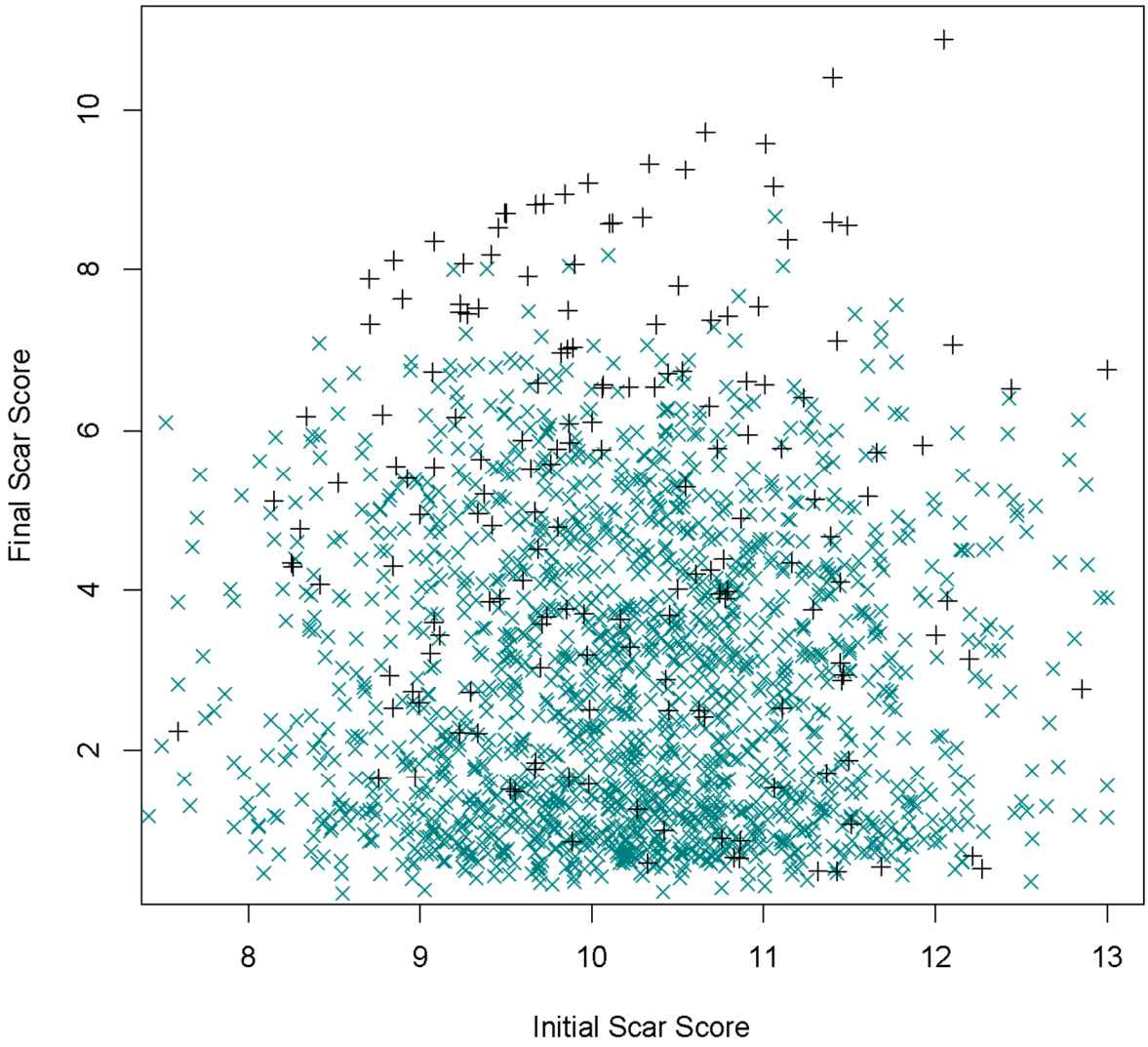


Figure 4: Plot of simulated patient final scar scores against corresponding initial scar scores, for patients in trial, shown as black '+', and for patients subsequently following the discovered DTR, shown as blue 'x'.