



Precision medicine: Statistical methods for estimating adaptive treatment strategies

Erica E. M. Moodie¹ · Elizabeth F. Krakow²

Received: 2 March 2020 / Revised: 10 March 2020 / Accepted: 11 March 2020 / Published online: 14 April 2020
© Springer Nature Limited 2020

Series Editors' Note

The beauty of science is that all the important things are unpredictable.
Freeman Dyson

In the typescript which follows, Moodie and Krakow tackle the topical issue of *precision medicine* and statistical methods for estimating adaptive treatment strategies. This may be the most difficult typescript in our series so far for non-statisticians to understand. It even has equations! But please bear with the authors and give it a chance. One needs not to understand the equations to get the thrust of the strategy.

Precision medicine as we discuss elsewhere, is misnamed. In statistics and mathematics precision refers to getting the same answer again and again. It does not mean getting the correct answer, the term for which is accuracy, not precision. However, precision is the current buzz word so there's no point trying to get this straight. When we think about precision we need to consider two elements, reproducibility and replicability. Reproducibility means you give me your data and computer code and I come to the same conclusion you did. Replicability is another matter. I try to replicate your experiment and hopefully reach the same conclusion. In medicine, replicability is obviously more important than reproducibility but things which cannot be reproduced are unlikely to be replicated.

As the authors discuss, one can think about precision medicine as one does a family vacation. A best vacation depends on several co-variables: where you live, your prior travel experiences, advice from family and friends, online reviews, Wikitravel, cost, your travel budget, if you have kids and many other co-variables. Consequently, there is unlikely to be a best vacation for everyone. Yours might be a week at the Ritz Carlton Cancun with dinner at Careyes and ours, a week at the Pfister Hotel in Milwaukee with dinner at Mader's German Restaurant (bring simvastatin). Similarly, it is unlikely there is a best therapy of acute myeloid leukemia, a best donor, a best conditioning regimen, a best posttransplant immune suppressive regimen etc. and certainly no best combination of these co-variables for your patient.

The question Moodie and Krakow tackle is how we can determine the best therapy or combination of therapies for someone receiving a haematopoietic cell transplant. Although the default answer is typically: randomized clinical trials are the gold standard, these inform us of the outcome of a cohort of subjects, not individuals. In many instances, although a new therapy may be shown to be better than an old one in a controlled randomized trial the benefit is not uniformly distributed. Some subjects in the experimental cohort may do worse with the new therapy compared with controls, others better. The question is who are the winners and losers? We cannot do a controlled randomized trial of one person. Moodie and Krakow discuss statistical tools to help us sort this out.

Again, please do not be put off by the equations; forgetaboutit. The overriding message is not so complex, and important. We are always standing by on twitter @BMTStats to help. But don't confuse us with Match.com. And, by the way, Freeman Dyson was a professor at the Institute for Advanced Studies at Princeton but never got his PhD.

Robert Peter Gale, Imperial College London, and Mei-Jie Zhang, Medical College of Wisconsin, Center for International Blood and Marrow Research (CIBMTR).

✉ Elizabeth F. Krakow
efkrakow@fredhutch.org

¹ Department of Epidemiology and Biostatistics, McGill University, 1020 Pine Ave W, Montreal, QC H3A 1A2, Canada

² Fred Hutchinson Cancer Research Center and University of Washington, 1100 Fairview Ave N, Seattle, WA 98109, USA

Introduction

When planning a long-awaited vacation, the decision of where to go to have an “optimal” experience depends on many factors including the home base (Hawaii may be more feasible for someone based on the West Coast of North

America than someone on the East Coast), previous personal travel experiences, and advice from friends or online reviews. Similarly, although the primary and most desirable goal of cancer or graft-versus-host disease (GvHD) treatment is cure, the decision of how to treat someone to achieve an “optimal” outcome may also depend on numerous subject-level covariates analogous to those in our holiday-planning example: where we are presently (general health and performance status, specific organ function, disease burden, mutation landscape, and immune profile), where we have been (response to and adverse effects of previous treatments, perhaps including molecular and immune biomarker responses), and where we could realistically go (anticipated performance of future treatment options, subject to constraints of availability, cost, etc.). The first and second considerations are based on a person’s current and prior data. The third consideration requires considering outcomes of former patients.

Of course, in many people treatment decisions need to be taken several times in a series of treatment choices using updated data on the subject’s current condition and prior experiences. Examples of everyday decisions haematopoietic cell transplant physicians and immunotherapists encounter include:

- (1) Treatment of people with acute myeloid leukemia (AML) posttransplant with measurable residual disease or relapse. Many receive hypomethylating therapy, intensive chemotherapy, donor lymphocyte infusions, and/or second transplants in diverse sequences.
- (2) Immune suppression strategies to prevent and treat acute and chronic GvHD [1–3].
- (3) Sequencing autotransplant vs. chimeric antigen receptor (CAR)-T-cell therapy in persons with advanced lymphomas or sequencing new therapies for chronic lymphocytic leukemia.

To fulfill the goal of precision medicine, adaptive treatment strategies (ATS) use data from observed patient experiences to develop treatment recommendations tailored to a given person. In this tutorial, we give an overview of two specific methods of statistical estimation of ATS, each from one of two broad classes of approaches. We illustrate these approaches using recent analyses of data from the Center for International Blood and Marrow Transplant Research (CIBMTR) [3].

Two approaches to estimating optimal ATS for a single treatment decision

Developing an optimal ATS requires several steps. The first is to clearly define the research question by specifying: (1)

how many treatment decision nodes there are and what treatment options are available at each; (2) what subject-level data relevant to the possible decisions are available; (3) what subject-level data were used to assign treatment (e.g., was it a randomized trial or were hospital-specific or country-specific treatment guidelines used); and (4) which outcome(s) is to be optimized. The latter may be difficult to specify. For example, targeting cure or maximizing disease-free survival (DFS) may overlook important considerations such as adverse effects, quality-of-life (QoL), or cost. A utility combining several competing outcomes could be used. However, constructing such a utility requires substantial input from physicians and patients [4].

There are two broad approaches to estimating ATS: (1) regression-based (indirect) methods; and (2) value-search methods. We outline one example of each approach in a hypothetical simple, one-decision setting for a continuous outcome. We then discuss how these approaches are extended to other outcome types and multiple decision stages. We also discuss a recent application to the two-stage decision-making approach to preventing and, if needed, treating GvHD using immune suppressive therapies in the course of an allotransplant for persons with AML or myelodysplastic syndrome.

We begin with notation for the single decision (or single stage) case. We consider a binary treatment Z , pre-treatment covariates X , and a binary outcome Y . For example, we may wish to decide whether to give someone anti-thymocyte globulin (ATG) prophylaxis in addition to standard tacrolimus/methotrexate (ATG is denoted as Z) to maximize GvHD-free, relapse-free survival (GFRS, our Y) taking into consideration factors such as sex match, age, pretransplant conditioning regimen intensity, donor histocompatibility and relatedness, and type of graft (collectively denoted X). In fact, much of the statistical literature on ATS focuses on a continuous outcome, for instance a biological marker or QoL. However, our focus in this example is on outcomes relevant to transplant physicians.

Regression-based (indirect) estimation of a single-stage ATS

One class of estimating approaches to ATS relies on the familiar method of regression. Specifically, a model is fit for the probability of observing the outcome, Y , as a function of the treatment, the covariates, and interactions between these. In general terms, we write:

$$\text{logit}(\Pr[Y = 1|Z, X; \beta, \varphi]) = g(X; \beta) + \gamma(X, Z; \varphi). \quad (1)$$

It is common to choose both g and γ to be linear functions. As a trivial example, if we take Z to be 1 when treatment is

ATG and 0 otherwise, then we may specify:

$$\text{logit}(\Pr[Y = 1|Z, X; \beta, \varphi]) = (\beta_0 + \beta_1 \text{HLAmismatch}) + Z(\varphi_0 + \varphi_1 \text{HLAmismatch}).$$

Note that in this equation, $g(X; \beta) = (\beta_0 + \beta_1 \text{HLAmismatch})$ describes the impact of an HLA-mismatched donor (relative to that of an HLA-matched donor) on GRFS under standard immune suppression (tacrolimus/methotrexate), i.e., when $Z=0$. GRFS (or, more accurately, the logit of the probability of experience GRFS) is altered by $\gamma(X, Z = 1; \varphi) = (\varphi_0 + \varphi_1 \text{HLAmismatch})$ when treatment is, instead, ATG. Note that if $(\varphi_0 + \varphi_1 \text{HLAmismatch}) > 0$, then the probability of experiencing GRFS will be higher when ATG therapy is used. Thus, if we can estimate the parameters φ_0 and φ_1 , then we could estimate the optimal rule as “treat with ATG when $(\varphi_0 + \varphi_1 \text{HLAmismatch}) > 0$ ”, otherwise, treat with tacrolimus/methotrexate alone. Returning to the general regression formulation given in Eq. (1), this suggests that the optimal rule can be deduced by: (1) estimating regression parameters of Eq. (1); and (2) specifying the optimal rule to be that which assigns treatment $Z = 1$ whenever $\gamma(X, Z = 1; \varphi) > 0$, assuming that the larger value of Y (in this case, 1) is preferable.

This regression-based approach has the advantage of relying on a familiar statistical tool accompanied by standard approaches to covariate selection, model diagnostics, and so on. This straightforward approach, whether in the single- or multiple-stage setting, is known as Q-learning. However, the approach relies on assuming the model designated to describe Eq. (1) is correctly specified. This implies that all confounders [5] have been measured. Regression-based methods can be made more robust by specifying a very flexible functional form for the regression equation, for example by using splines [6], nonlinear models, or even nonparametric machine learning approaches [7], though the latter may lose some of the interpretability of a more traditional regression model. Regression-based methods can also be made more robust by specifying a propensity score model—i.e., a model for the treatment assignment mechanism—and using this to adjust for confounding so that even if the $g(X; \beta)$ is not correctly specified, the estimated treatment rule will be consistent for the truth in large samples—provided the propensity score model and the component of the model that specifies the treatment rule, $\gamma(X, Z; \varphi)$, are correctly specified. These so-called doubly robust methods include g-estimation [8] and dynamic weighted ordinary least squares [9] although neither of these are well-developed for binary outcomes.

Value-search (direct) estimation of a single-stage ATS

In the ATS literature the expected outcome is sometimes known as the value function. We may consider the

value function under a treatment strategy. E.g., $V^1 = E[Y(Z = 1)]$ represents the value under the treatment strategy “give everyone ATG”, whereas $V^d = E[Y(Z = d(X))]$ where $d(X) = I((\varphi_0 + \varphi_1 \text{HLAmismatch}) > 0)$ is the probability GRFS if all patients were treated according to the rule “treat with ATG when $(\varphi_0 + \varphi_1 \text{HLAmismatch}) > 0$ ”. Value-search methods aim to estimate the value function directly under a series of candidate treatment strategies, d . These strategies could be linear decision rules such as “treat with ATG when $(\varphi_0 + \varphi_1 \text{HLAmismatch}) > 0$ ”, or could involve nonlinear treatment rules such as “treat with ATG when $(\vartheta_0 + 1.2^{\text{Age} - \vartheta_1}) > 0$ ” where the latter cannot easily be estimated with traditional regression-based methods.

A classic value-search method relies on inverse probability of treatment weighting (IPTW). In this approach, the propensity score is used to construct weights that are used to remove confounding, which may exist when treatment is not randomly assigned. The analyst first estimates the propensity score model, constructs inverse probability of treatment weights and then computes a weighted average of the outcomes Y for those individuals who followed a given treatment rule, d . Using the same propensity score model and weights, weighted averages are computed for each candidate treatment rule and the resulting estimates are compared with see which candidate rule returns the greatest expected outcome.

Value-search approaches have the advantage of more easily accommodating rules of any form, not simply linear decision rules. However, more sophisticated approaches than IPTW are generally recommended as IPTW estimators of the value (expected outcome) often have large standard errors making it difficult to distinguish the relative benefit of the candidate rules. These approaches include augmented IPTW [10], residual weighted learning [11], and others (e.g., [12]). Value-search methods do not require the true best adaptive decision rule to be among the candidate rules. The approach will, in large samples, simply select the treatment strategy, which results in the best outcome *among* the candidate strategies being considered.

Two approaches to estimating optimal ATS for a sequence of several treatment decisions

Consider now, a setting where we have to make multiple treatment decisions. Recall, for example, setting (2) above where interest lies in devising the best strategy to prevent and/or treat acute and chronic GvHD. Interest may lie in maximizing the binary outcome of 2-year DFS [3] or maximizing DFS time without restriction to 2 years and allowing censoring [2].

We must now extend notation to accommodate two stages of decision-making, wishing to individualize treatment decisions, considering a binary treatment Z_1 indicating use of ATG given pretransplant to prevent GvHD. Pre-treatment covariates X_1 are measured and may include recipient-, donor- and disease-related variables. In subjects developing GvHD, further interventions of different intensities will be offered. This second stage of treatment is denoted Z_2 . Pre-treatment covariates, denoted X_2 , are again measured before giving GvHD therapy, which may include all or some subset of X_1 , as well as post- Z_1 variables such as time to develop GvHD, current health, investigational biomarkers of GvHD severity and/or functional measures such as the Karnofsky Performance Status score. Again, we consider a binary outcome Y such as 2-year GRFS.

A key concern with multistage interventions is that a treatment may have delayed effects. For example, an intensive therapy may elicit a good short-term response but may compromise subsequent therapy(ies) resulting in a lower long-term success rate [13, 14]. Estimation must therefore proceed either sequentially backwards for regression-based estimators or by considering the ATS for value-search methods.

Regression-based (indirect) estimation of a multistage ATS

Implementation of Q-learning in a multistage decision analysis proceeds by following a general sequential algorithm for a two-stage case:

1. Propose a model for the final outcome as a function of the second stage of treatment, Z_2 , and any elements of $X_2^* = (X_1, Z_1, X_2)$ that (i) may be potential tailoring variables for stage 2 treatment or (ii) may be important predictors of the outcome Y . To be more concise, we can let denote (X_1, Z_1, X_2) :

$$\begin{aligned} \text{logit}(\Pr[Y = 1|Z_2, X_2^*; \beta_2, \varphi_2]) \\ = g_2(X_2^*; \beta_2) + \gamma_2(Z_2, X_2^*; \varphi_2) \end{aligned} \tag{2}$$

where the functions $g_2()$ and $\gamma_2()$ are analogous to those defined in Eq. (1). More specifically, the contrast function $\gamma_2()$ is used to define a (possibly linear) decision rule at the second decision-stage.

2. Estimate the parameters in Eq. (2) and use these to define the optimal (estimated) stage 2 decision rule as “treat with $Z_2 = 1$ whenever $\gamma_2(Z_2 = 1, X_2^*; \hat{\varphi}_2) > \gamma_2(Z_2 = 0, X_2^*; \hat{\varphi}_2)$ ”, or equivalently “treat with $Z_2 = 1$ whenever $\gamma_2(Z_2 = 1, X_2^*; \hat{\varphi}_2) > 0$ ”. Through

$X_2^* = (X_1, Z_1, X_2)$ this rule may account for previous (stage 1) treatment, Z_1 and responses to the treatment, contained in X_2 .

3. We now wish to estimate the optimal first-stage decision. However, using principles much like those in traditional randomized clinical trials, we do not wish to condition on or adjust for any post-(stage 1) treatment variables. We do not wish to condition on the second stage treatment and yet we must ensure that comparisons between the two stage 1 treatment options are “fair” and not simply a reflection of later, downstream treatments. We accomplish this by creating a new, pseudo-outcome we denote \tilde{Y}_1 which we generate for each individual in the sample according to:

$$\begin{aligned} \tilde{Y}_1 = \max\{ & \Pr[Y = 1|Z_2 = 0, X_2^*; \hat{\beta}_2, \hat{\varphi}_2], \\ & \Pr[Y = 1|Z_2 = 1, X_2^*; \hat{\beta}_2, \hat{\varphi}_2] \}. \end{aligned}$$

That is, the pseudo-outcome is the estimated “best possible” probability of the outcome an individual could have based on the estimates for the outcome model specified in Eq. (2). Using this pseudo-outcome is equivalent to performing a stage 1 analysis in a world where all individuals in the sample were treated optimally at the second stage. Note that in this “optimal treatment world”, not everyone would receive the same *treatment* (ATG or standard), but all would be treated according to the same *rule* (“treat with $Z_2 = 1$ whenever $\gamma_2(Z_2 = 1, X_2^*; \varphi_2) > 0$ ”).

4. Propose a model for the pseudo-outcome as a function of first stage of treatment, Z_1 , and X_1 :

$$\begin{aligned} \text{logit}(\Pr[\tilde{Y}_1 = 1|Z_1, X_1; \beta_1, \varphi_1]) \\ = g_1(X_1; \beta_1) + \gamma_1(Z_1, X_1; \varphi_1) \end{aligned} \tag{3}$$

where again $g_1()$ and $\gamma_1()$ are analogous to the functions in Eq. (1).¹

5. Estimate the parameters in Eq. (3), and use these to define the optimal (estimated) stage 1 decision rule as “treat with $Z_2 = 1$ whenever $\gamma_1(Z_1 = 1, X_1; \hat{\varphi}_1) > 0$ ”.

The above algorithm can be adapted to more than two stages simply by computing a new pseudo-outcome for all stages other than the final stage. The final sequence of treatment rules is made up of a sequence with

¹ Note that many regression packages will yield a warning in attempting to fit Eq. (3) as \tilde{Y}_1 is not binary (although it does lie in the interval [0,1]). This warning may safely be disregarded.

components consisting of “treat with $Z_j = 1$ whenever $\gamma_j(Z_j = 1, X_j^*; \hat{\phi}_j) > 0$ ” for each treatment stage j . Other regression-based forms of estimation vary in how the pseudo-outcome is constructed. However, the basic principles of the backwards inductive approach remain.

Value-search estimation of a multistage ATS

The extension of the simple, inverse probability of treatment-weighted estimator to multiple stages is straightforward. As in the single-stage setting, a set of candidate treatment strategies must be posited by the analyst. Propensity score models are fit—now at each decision stage—and inverse probability of treatment weights are constructed at each interval and then multiplied together. As in the single-stage setting, for each candidate strategy of interest, those individuals in the sample who were observed to follow the treatment strategy under investigation are used to compute a weighted average of the outcomes Y , thus yielding an estimate of the value function for that strategy. The resulting estimates of the value functions for each candidate strategy are compared to see which returns the greatest expected outcome.

As in the single-stage setting, there are numerous alternatives to the simple IPTW approach, many of which include some form of outcome modeling and consequently offer greater precision in the estimated value function and thus the choice of preferred treatment strategy.

Further extensions

The methods we describe along with related methods have been extended in several ways. Beyond two treatment options, one can consider multiple distinct treatments or even continuous treatment [15, 16]. Censored continuous outcomes—e.g., “survival time”—can also be accommodated with censoring handled by assuming independent censoring or by using inverse probability of censoring

weights [17–20]. Within regression-based approaches this is done by assuming appropriate models in Eqs. (2) and (3). For example, a Cox model could be assumed for a survival outcome. Except for a few instances (e.g., [3, 21]), other outcome types such as counts or binary outcomes have rarely been considered. We are unaware of any analyses or methods that have aimed to optimize a binary outcome over more than two stages of intervention.

The form of the outcome model can be very general, particularly for Q-learning. For example, we considered a parametric model for DFS time that allowed for a fraction of individuals to be cured [2]. This model allowed for treatment to interact with covariates differently in the cure and the survival components of the model. This flexible approach revealed that although *on average* ATG therapy is not a preferred treatment choice for either GvHD prevention or treatment (see Fig. 1) when the outcome being considered is DFS, significant numbers of people may benefit from pre-transplant ATG and a much smaller fraction of patients with GvHD may benefit from ATG treatment (Fig. 2).

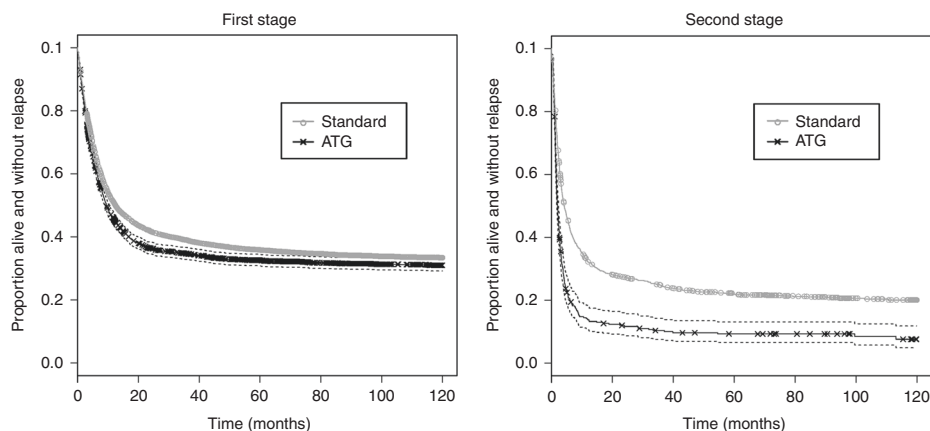
Algorithm validation

Oncologists are already familiar with trials that aim to evaluate two treatments, using either factorial designs [22] or using sequential randomizations where nonresponders are randomized to different salvage therapies or responders are randomized to different consolidation or maintenance treatments or to maintenance vs. no intervention (e.g., [23, 24]). Unfortunately, results of first and second randomizations are often published in separate articles (for example [25–28]). Consequently, insights which might emerge by considering the trajectories as a whole are lost. The implication is that the infrastructure for conducting trials with multiple treatment assignments and sequential randomizations already exists.

The Sequential Multiple-Assignment Randomized Trial (SMART) is the preferred trial design to develop ATS that could serve as decision support tools and practice guidelines.

Fig. 1 Survival comparison for ATG vs. standard GvHD prophylaxis and treatment.

Kaplan–Meier curves show clear overall benefit to standard therapy over ATG at both the first treatment stage (left panel, GvHD prophylaxis) and second treatment stage (right panel, treatment of established GvHD).



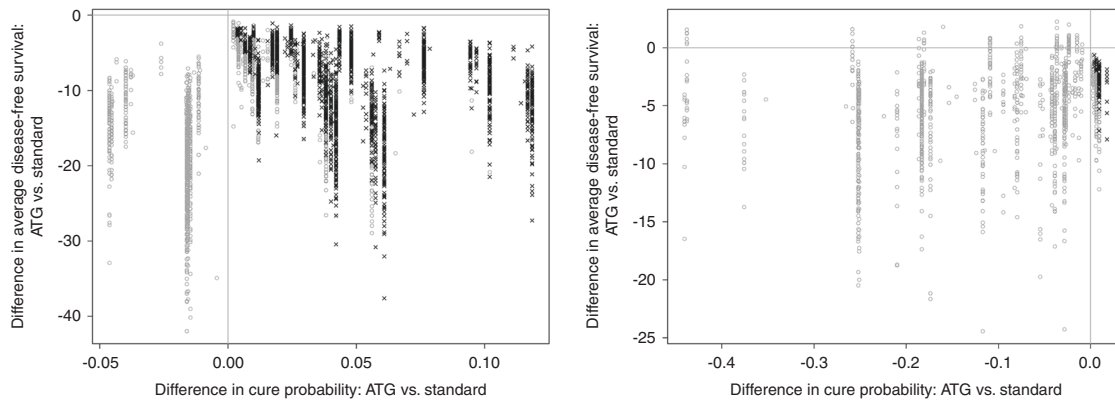


Fig. 2 Difference in expected disease-free survival if not cured (y-axis) vs. difference in probability of cure (x-axis) if given ATG instead of standard treatment. For a small number of people, ATG may be preferred as it confers a higher probability of being cured (no death or relapse) than standard treatment, particularly at the first stage

Compared with developing ATS through retrospective analysis of medical records and registry data, prospective development of ATS through SMART clinical trials has the advantage that randomization reduces selection bias and confounding-by-indication along with reducing the risk of unmeasured confounders. In a SMART, participants are randomized to one of a pre-defined list of treatment options at each critical decision node. This approach allows discovery and testing of tailoring variables while also assessing comparative efficacy of different treatments. SMARTs are used to identify which subset of X_1 and X_2 predicts a good or poor response to a given treatment at the respective decision stage.

To validate an ATS, whether it was developed in a SMART or through retrospective analyses, would require a subsequent conventional randomized trial. For example, subjects could be randomized between the ATS and a different “standard-of-care” treatment sequence. Alternatively, the ATS could be tested in “ecological studies” where it is implemented in some hospitals or over some period, and the outcome of subjects treated under the ATS compared with the outcome of contemporaneous subjects treated in different hospitals that did not use the ATS or to a historical cohort.

Conclusion

In this brief review, we introduced two simple forms of analysis for estimating optimal ATS. These methods can be applied to nonexperimental data such as those arising from clinical practice or registries or can be applied to randomized trials. In particular, the SMART design is specifically targeted at designing treatment algorithms for tailored interventions with multiple decision points [29–31].

An important point to keep in mind is that, in general, analyses aimed at uncovering ATS are exploratory rather

(left panel, prophylaxis), with much smaller numbers expected to benefit at the second GvHD stage (right panel, GvHD treatment). Black crosses indicate patients for whom ATG is the preferred prophylaxis or GvHD treatment.

than confirmatory in nature. SMARTs may be confirmatory in nature but are typically powered for tailoring only on a very small number of covariates such as response to first-stage treatment. Nevertheless, these methods can identify candidate strategies and tailoring variables that appear promising and discard other clearly suboptimal strategies. As big data, expanded access to anonymized electronic medical records and incorporation of novel biomarkers into clinical decision making become the norm, ATS approaches to developing decision support tools are becoming increasingly feasible and potentially useful, both for ‘simple’ decisions and complex ones.

Acknowledgements EEMM acknowledges salary support from a chercheur boursier senior award from the Fonds de recherche du Québec—Santé and research support from the Canadian Institutes of Health Research (Grant #FDN-167267). The findings presented here are the responsibility of the authors and are not the opinion of the CIBMTR. This research has been supported in part by the National Institutes of Health under Award Number R01 HL113548. Data analyzed in the paper were supplied by the Center for International Blood and Marrow Transplant Research (CIBMTR) which was supported in part by National Institutes of Health (NIH/NCI) grant U24-CA076518-20.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Liu Y, Logan B, Liu N, Xu Z, Tang J, Wang Y. Deep reinforcement learning for dynamic treatment regimes on medical registry data. *Health Inf.* 2017;2017:380–5.

2. Moodie EEM, Stephens DA, Alam S, Zhang MJ, Logan B, Arora M, et al. A cure-rate model for Q-learning: estimating an adaptive immunosuppressant treatment strategy for allogeneic hematopoietic cell transplant patients. *Biom J*. 2019;61:442–53.
3. Krakow EF, Hemmer M, Wang T, Logan B, Arora M, Spellman S, et al. Tools for the precision medicine era: how to develop highly personalized treatment recommendations from cohort and registry data using Q-Learning. *Am J Epidemiol*. 2017;186:160–72.
4. Murray TA, Thall PF, Yuan Y. Utility-based designs for randomized comparative trials with categorical outcomes. *Stat Med*. 2016;35:4285–305.
5. Hu ZH, Peter Gale R, Zhang MJ. Direct adjusted survival and cumulative incidence curves for observational studies. *Bone Marrow Transpl*. 2019;55:538–43.
6. Gauthier J, Wu QV, Gooley TA. Cubic splines to model relationships between continuous variables and outcomes: a guide for clinicians. *Bone Marrow Transpl*. 2020;55:675–80.
7. Logan BR, Sparapani R, McCulloch RE, Laud PW. Decision making and uncertainty quantification for individualized treatments using Bayesian Additive Regression Trees. *Stat Methods Med Res*. 2019;28:1079–93.
8. Robins JM. Optimal structural nested models for optimal sequential decisions. *Lect Notes Stat*. 2004;179:189–326.
9. Wallace MP, Moodie EEM. Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics*. 2015;71:636–44.
10. Zhang B, Tsiatis AA, Laber EB, Davidian M. A robust method for estimating optimal treatment regimes. *Biometrics*. 2012;68:1010–8.
11. Zhou X, Mayer-Hamblett N, Khan U, Kosorok MR. Residual weighted learning for estimating individualized treatment rules. *J Am Stat Assoc*. 2017;112:169–87.
12. Zhao Y, Zeng D, Rush AJ, Kosorok MR. Estimating individualized treatment rules using outcome weighted learning. *J Am Stat Assoc*. 2012;107:1106–18.
13. Chakraborty B, Moodie EEM. *Statistical methods for dynamic treatment regimes: reinforcement learning, causal inference, and personalized medicine*. New York, NY: Springer (Statistics for Biology and Health series); 2013.
14. Murphy SA. An experimental design for the development of adaptive treatment strategies. *Stat Med*. 2005;24:1455–81.
15. Schulz J, Moodie EEM. Doubly robust estimation of optimal dosing strategies. *J Am Stat Assoc* (under invited review). 2020.
16. Rich B, Moodie EE, Stephens DA. Optimal individualized dosing strategies: a pharmacologic approach to developing dynamic treatment regimens for continuous-valued treatments. *Biom J*. 2016;58:502–17.
17. Goldberg Y, Kosorok MR. Q-Learning with censored data. *Ann Stat*. 2012;40:529–60.
18. Huang X, Ning J, Wahed AS. Optimization of individualized dynamic treatment regimes for recurrent diseases. *Stat Med*. 2014;33:2363–78.
19. Hager R, Tsiatis AA, Davidian M. Optimal two-stage dynamic treatment regimes from a classification perspective with censored survival data. *Biometrics*. 2018;74:1180–92.
20. Simoneau G, Moodie EEM, Azoulay L, Platt RW. Adaptive treatment strategies with survival outcomes: an application to the treatment of type 2 diabetes using a large observational database. *Am J Epidemiol*. 2020 [Epub ahead of print].
21. Wallace MP, Moodie EEM, Stephens DA. Model selection for G-estimation of dynamic treatment regimes. *Biometrics*. 2019;75:1205–15.
22. Freidlin B, Korn EL. Two-by-two factorial cancer treatment trials: is sufficient attention being paid to possible interactions? *J Natl Cancer Inst*. 2017;109:dxj146.
23. Stone RM, Berg DT, George SL, Dodge RK, Paciucci PA, Schulman PP, et al. Postremission therapy in older patients with de novo acute myeloid leukemia: a randomized trial comparing mitoxantrone and intermediate-dose cytarabine with standard-dose cytarabine. *Blood*. 2001;98:548–53.
24. Habermann TM, Weller EA, Morrison VA, Gascoyne RD, Casileth PA, Cohn JB, et al. Rituximab-CHOP versus CHOP alone or with maintenance rituximab in older patients with diffuse large B-cell lymphoma. *J Clin Oncol*. 2006;24:3121–7.
25. Catovsky D, Richards S, Matutes E, Oscier D, Dyer M, Bezars RF, et al. Assessment of fludarabine plus cyclophosphamide for patients with chronic lymphocytic leukaemia (the LRF CLL4 Trial): a randomised controlled trial. *Lancet*. 2007;370:230–9.
26. Matutes E, Bosanquet AG, Wade R, Richards SM, Else M, Catovsky D. The use of individualized tumor response testing in treatment selection: second randomization results from the LRF CLL4 trial and the predictive value of the test at trial entry. *Leukemia*. 2013;27:507–10.
27. Crump M, Kuruvilla J, Couban S, MacDonald DA, Kukreti V, Kouroukis CT, et al. Randomized comparison of gemcitabine, dexamethasone, and cisplatin versus dexamethasone, cytarabine, and cisplatin chemotherapy before autologous stem-cell transplantation for relapsed and refractory aggressive lymphomas: NCIC-CTG LY.12. *J Clin Oncol*. 2014;32:3490–6.
28. Kuruvilla J, Kouroukis CT, Bengier A, Cheung M, Berinstein N, Couban S, et al. A randomized trial of rituximab vs observation following Autologous Stem Cell Transplantation (ASCT) for relapsed or refractory CD20-positive b cell lymphoma: final results of NCIC CTG LY.12. *Blood*. 2013;122:155.
29. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. *Am J Prev Med*. 2007;32:S112–8.
30. Kosorok MR, Moodie EEM, editors. *Adaptive treatment strategies in practice: planning trials and analyzing data for personalized medicine*. Philadelphia, PA: ASA-SIAM (American Statistical Association-Society for Industrial Mathematics); 2016.
31. Wallace MP, Moodie EE, Stephens DA. SMART thinking: a review of recent developments in sequential multiple assignment randomized trials. *Curr Epidemiol Rep*. 2016;3:225–32.