Published in final edited form as:

Stat Methods Med Res. 2017 August ; 26(4): 1605–1610. doi:10.1177/0962280217708661.

Dynamic treatment regimes, past, present, and future: A conversation with experts

Eric B Laber and Marie Davidian

North Carolina State University, Raleigh, USA

Abstract

We asked three leading researchers in the area of dynamic treatment regimes to share their stories on how they became interested in this topic and their perspectives on the most important opportunities and challenges for the future.

Keywords

Optimal dynamic treatment regime; sequential multiple assignment randomized trial; mHealth; clinical trials; treatment sequence

1 Introduction

We are excited to have been asked by the Editor of *Statistical Methods in Medical Research* to serve as Guest Editors for this special issue on the theme of Optimal Dynamic Treatment Regimes.

In the setting of treatment of a disease or disorder, a dynamic treatment regime is a set of sequential decision rules, each corresponding to a key point in the disease or disorder process at which a decision on the next treatment action for a patient must be made. Each rule takes as input information on the patient, such as demographics, prior medical history, genetic and genomic characteristics, evolving physiological and clinical variables, occurrence, and timing of adverse events related to treatment, and so on, and returns a treatment option from among the available, feasible options. Dynamic treatment regimes have also been referred to as adaptive treatment strategies or adaptive interventions.

Dynamic treatment regimes formalize how clinicians use information on individual patients to make treatment decisions in practice, and thinking in terms of dynamic treatment regimes provides a framework for precision medicine, which is focused on tailoring treatment decisions to a patient's characteristics based on evidence. The goal of clinicians and the broader concept of precision medicine are to make these decisions so as to lead to the most

Reprints and permissions: sagepub.co.uk/journalsPermissions.nav

Corresponding author: Marie Davidian, Department of Statistics, North Carolina State University, Raleigh, North Carolina, USA. davidian@ncsu.edu.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

beneficial expected outcome for an individual patient given his or her characteristics as his or her disease or disorder progresses.

Addressing this goal can be accomplished by deducing an optimal treatment dynamic regime. An optimal regime is one that, if used to select treatment actions for all patients in the target population, would lead to the most beneficial outcome on average. Thus, methodology to estimate and evaluate optimal dynamic treatment regimes based on data is of considerable interest and is the focus of a growing body of statistical and domain science research.

In this special issue, we showcase five articles related to optimal dynamic treatment regimes written by teams of well-known researchers. We also had a conversation with three thought leaders in the field during which we asked each to provide their perspectives on current and future opportunities and challenges presented by the explosion of interest in and fast pace of this area: Michael R. Kosorok, W. R. Kenan, Jr Distinguished Professor in the Department of Biostatistics at the University of North Carolina at Chapel Hill; Susan A. Murphy, H. E. Robbins Distinguished University Professor of Statistics at the University of Michigan, who received a MacArthur Foundation Fellowship in 2013 for her path breaking work in this area; and Peter F. Thall, Anise J. Sorrell Professor in the Department of Biostatistics at The University of Texas M.D. Anderson Cancer Center. This article summarizes our conversation. In each section, the responses of our distinguished panel to a specific question related to this theme are presented. We hope that their comments are an interesting accompaniment to the original research contributions that follow.

2 Involvement in dynamic treatment regimes research

**When and how did you first become involved in research on dynamic treatment regimes?
Can you talk a bit about your first excursion into collaborative research on dynamic treatment regimes?**

Susan Murphy—Prior to the mid-1990s, my research was focused on theory, particularly the validation of large sample approximations in problems with a high dimensional parameter space. However, I realized that statisticians whom I greatly admire, such as Brad Efron and David Cox, closely tied their theoretical work to real-life applications. I decided to strive to follow in their footsteps!

Around this time I met a behavioral scientist, Linda Collins at Penn State University, and she ran occasional brainstorming sessions with other behavioral scientists. I resolved to attend these meetings. It took me quite some time to understand the language and ideas. Around 1996, a behavioral scientist, Karen Bierman, led a very perplexing session in which she explained that in a study, children who receive more reading tutoring sessions appeared to do worse than children who received fewer sessions. After more than a month of going back and forth, I realized that a dynamic treatment regime had been implemented in this study. This dynamic treatment regime specified an increase in the amount of tutoring if a child was not reading well. This realization led me to think about what types of inference might be possible in this setting and how we might use data to construct more effective dynamic treatment regimes.

Peter Thall—In 1999 Randy Millikan, a genitourinary oncologist at M.D. Anderson, asked me to provide the statistical design for a clinical trial of combination chemotherapies for advanced prostate cancer. He wanted to keep track of per-course responses and rerandomize patients if and when their frontline chemo failed. I had never heard of a dynamic treatment regime, but we published a paper on the design, forged ahead with the trial, and enrollment was completed some years later. Randy made me wait for longer follow up before we analyzed the data and published the medical paper, and he insisted that we use the regression model that we had specified in the design paper.

Susan Murphy had read our design paper, and she invited me to attend a summer workshop on dynamic treatment regimes at the Statistical and Applied Mathematical Sciences Institute (SAMSI) in North Carolina, which was very enlightening. I enlisted The Harvard Gang (Lu Wang, Andrea Rotnitzky, and Xihong Lin) to do an inverse probability weighted analysis of the prostate cancer trial data that accounted for dropouts.

In the course of that process, when asking Randy the reason for each dropout, I found out that the actual regimes used by the oncologists were a lot more complicated than the per-protocol rules. Of course, prostate cancer isn't cured, we but learned some important things. First, actual oncology practice is a dynamic treatment regime, and often is very complex. This implies that the conventional clinical trial paradigm of evaluating treatments given at one stage of a dynamic treatment regime is an oversimplification that often is misleading. For example, survival time depends on not only frontline therapy, but also whatever salvage therapies are given at the times of one or more disease progressions. Second, dropout or discontinuation of therapy, possibly due to toxicity, actually is an outcome. Third, if one applies a simple "Try a treatment, and if it fails then switch to something else" rule in a clinical trial design and rerandomizes, i.e. conducts a sequential, multiple assignment, randomized trial (SMART) with a switch-away rule, then the patients in the trial probably will survive a lot longer than if they were given one treatment until death or discontinuation. A large effect of this sort is shown in Figure 1 of our response to a letter¹ regarding our paper.²

Michael Kosorok—I believe it was at the Joint Statistical Meetings in 2006 when Susan Murphy came up to me and told me that I should start working on dynamic treatment regimes because she felt my empirical process background would be helpful to the field. In the summer of 2007, I attended the same SAMSI workshop on the topic that Peter attended and began working on the subject with my colleague Donglin Zeng and some students. Our first work was to figure out how to apply nonparametric Q-learning in cancer clinical trials, especially in nonsmall cell lung cancer in collaboration with an oncologist, Mark Socinski. Mark was interested in how timing of the second line of chemotherapy in late-stage nonsmall cell lung cancer patients might be personalized to maximize survival time. Two papers came from this early period.^{3,4}

3 The most important open problems

What do you perceive as the most important open problems in dynamic treatment regimes research (methodologically, scientifically, and/or otherwise)?

PT—The most important problem is cultural. It comes down to poor communication between biostatisticians and physicians. Many statisticians who understand dynamic treatment regimes find it easier to just publish papers in their own journals rather than fight with physicians about statistical methods and trial design. Thought experiments are very important, and we should keep doing them, but as medical researchers who do biostatistics, our methodological advances don't mean a thing unless we can convince the people who actually make decisions to adopt them.

The medical research community has been doing clinical trials the wrong way for decades. Most physicians involved in clinical trials do not want to hear some statistician telling them that their conventional trial designs, and methods for analyzing data, are fundamentally wrong. In the years since the landmark paper by Hernan, Brumback, and Robins in *Epidemiology*,⁵ where they showed that accounting properly for the actual adaptive regimes being used by physicians treating AIDS patients can reverse one's conclusions, the message simply has not gotten through to the larger medical community. Part of the problem is that dynamic treatment regimes often are counterintuitive, and switching one's mind-set from "I want to evaluate new treatments" to "I want to evaluate multi-stage regimes that include new treatments" is very hard. Still, many physicians understand that a dynamic treatment regime is essentially a formalism of how they practice medicine. Also, keeping track of more things and rerandomizing during a trial really is not so much harder than what usually is done. So, I have reasons to be optimistic about the future.

Methodologically, I believe that, inevitably, Bayesian models and methods will replace most frequentist methods for evaluating dynamic treatment regimes and designing trials. This will be especially true for data analysis, where Bayesian nonparametric models will become standard tools. As this process evolves, there will be high impact methodological advances that will change the whole field.

MK—Inference for dynamic treatment regime estimators, including inference for the value function corresponding to a given or estimated dynamic treatment regime, is still one of the most important open problems. Improved approaches to design and analysis of SMARTs and related designs for estimating dynamic treatment regimes are also important. One of the least developed areas is how to unify statistical control theory, Markov Decision Processes, and dynamic treatment regime estimation to minimize sample size. Better and more flexible methods for utilization of observational data, including partly randomized observational data, are also needed for dynamic treatment regime discovery.

SM—I believe the current biggest challenges in dynamic treatment regimes are (1) devising the best experimental trial strategy so as to develop dynamic treatment regimes with replicable effects, (2) extending dynamic treatment regimes methodology for use in tackling to social problems such as welfare and criminal justice, (3) identifying more dynamical health/psychosocial/biological theories upon which to develop dynamic treatment regimes in

a scientifically grounded way, and (4) developing methodology for real-time dynamic treatment regimes, such as in mobile health (mHealth). All of these present many problems. The development of mobile health interventions requires interdisciplinary teams and requires us, statisticians, to take on leadership roles because the computational constraints, bandwidth constraints, the constraints imposed by intervening in a person's real life need to be taken into account in the development of data analytic methods—this is nontrivial and requires us to really expand out of our comfort zone.

4 Connection to Big Data and precision medicine

Where do you see dynamic treatment regimes fitting with the current surge of interest in Big Data and precision medicine? Is this precision medicine/Big Data “obsession” good for the field dynamic treatment regimes?

MK—It seems that some of the most important structures we could discover from Big Data are optimal decision rules for a variety of application domains. In precision medicine, these decision rules would be dynamic treatment regimes, but there are many other domains, including economics, public policy, and city management, for example, where the general approach used in discovering dynamic treatment regimes is applicable. Moreover, dynamic treatment regimes are special cases of precision medicine outputs. Thus, the dynamic treatment regime framework will probably play a central role in Big Data and precision medicine research and vice versa. So, yes, I think this obsession is good for the field.

SM—Dynamic treatment regimes are definitely part of precision medicine in that the goal is to individualize a sequence of treatments based on the evolving course of the individual's health problem. This is to provide the best treatment to an individual by taking into account both what treatments can be provided in future as well as the current circumstances of the individual (past treatment responsiveness, past side effects, current symptoms). Measures of the current circumstances could be obtained directly from a person or from his/her medical record or high dimensional genetic or wearable sensor data. And definitely the focus on precision medicine and Big Data is good for dynamic treatment regime development. The challenge to us is to use this focus wisely to advance human welfare and not merely just to publish more papers.

PT—“Precision medicine” is a fancy name for what doctors have been doing for thousands of years, but without the benefit of high tech covariates. “Big Data” sounds like something profound is going on, but we actually are at the start of a long journey. The old saying that “All of statistics is regression, or would like to be” has never been truer, and we now have a lot of technological toys that are fun to play with. The practical problem is that, when high dimensional heterogeneity is introduced at the patient level, from genomic, proteomic, cell surface marker, or signaling pathway covariates, dynamic treatment regime modeling, analysis, and SMART design rapidly become intractable. If you look at the dimension of a (covariate vector $\times$ dynamic treatment regime) matrix, even in relatively simple cases, there simply are not enough patients to make reliable inferences. This elaborates the well-known genomics problem of a high dimensional covariate vector with a small sample of patients. The various dimension reduction methods currently being applied are interesting, but 30

patients with 60,000 biomarkers are still 30 patients. When I hear a basic scientist begin a talk by identifying a low dimensional “gene signature” or something similar, I simply don’t believe it. I have been asked to design clinical trials by researchers who begin by showing me a table of targeted agents matched with gene signature subgroups, which they ask me to accept on faith, and then they tell me that they expect to accrue 10 to 20 patients per year.

Of course, if the sampling unit is a cell, rather than a human, then things are completely different. The Elephant in the Living Room is the high profile research at the cellular or molecular level, which receives the most attention and the most funding, but seldom translates into treatment advances. The usual story is that an agent does a great job of killing cancer cells in vitro or tumor xenografts in mice, but then has little or no antidisease effects in patients. The poster child for targeted therapy is Imatinib, which keeps patients with chronic myelogenous leukemia alive longer by hitting one biological target. But other cancers are much harder to defeat. In terms of societal benefit, most basic science research has been oversold. Only when patients actually live longer, or suffer less, should victory be declared.

Years ago, Randy Millikan made the prescient observation that we have entered the age of treatment combinations and sequences. Ideally, statisticians who understand both dynamic treatment regimes and the biology of targeted agents might work to provide the missing link between the laboratory and the clinic, but that will be very hard work. Laboratory-based basic scientists seldom talk to statisticians, and they either don’t know or don’t care about dynamic treatment regimes, but they should. Most early phase trials are run with tiny samples using a foolish “3+3” algorithm based on a binary toxicity, ignoring efficacy and the fact that therapy involves multiple cycles. Ideally, early phase trials are where the dose and possibly schedule of a new agent should be optimized and refined, based on multicycle dynamic treatment regimes and much larger samples than are used in conventional phase I trials. It is impossible to know how many promising agents have been shelved wrongly due to the failure to optimize them sensibly and reliably at the start of clinical evaluation. We have begun to develop phase I–II designs that do this for the two-cycle case.⁶ It is a complex problem, even without introducing patient covariates, and I haven’t applied the method to an actual trial yet.

5 Ten years from now

Where would you like to see dynamic treatment regimes research in ten years? What do you see as the major obstacles to realizing this vision?

SM—I would like to see some effective dynamic treatment regimes developed using high quality experimental strategies. The obstacle to this is that, although health/biological/psychosocial theories should be used to inform the experimentation strategy for developing dynamic treatment regimes, most of these theories are static. We really need dynamical theories, theories about how a health disorder evolves and can be targeted over time. I would also like to see some effective real-time mobile health dynamic treatment regimes in 10 years! Again we need better dynamical theories.

PT—I would like to see future dynamic treatment regime research driven by applied problems, which really is what biostatistics should be about. The behavior modification trials that Susan’s team are doing, where many adaptive decisions and interventions are done in real time, implemented by wearable devices, are quite exciting. They could evolve, at least in concept, into similar but much more complex many-stage dynamic treatment regimes for adaptive drug infusions based on real-time biological data to treat all sorts of diseases. For high blood pressure or diabetes, this is essentially the process control problem that engineers have been working on for decades, but in a human body. A key aspect of this is borrowing strength both between and within subjects in real time, in a way that is safe.

The main obstacles, at present, are the unwillingness of most decision makers in medical research to change their conventional behaviors, limited federal funding for biostatistical research, and the hard reality that actually designing and conducting a clinical trial to evaluate dynamic treatment regimes is a difficult, often painful process that may take years to complete.

MK—In 10 years, we will hopefully have a much larger portfolio of implementations of dynamic treatment regimes in solving practical problems in a wide range of research domains. Hopefully, this will allow more practical advances to be made. We may also see realistic implementations of continual reassessment approaches where for some research areas patients are continually being randomized around nearly optimal treatment regimes to allow for continual improvement of patient outcomes. Hopefully, we will also have a better framework for results from SMARTs and other dynamic treatment regime studies to feed into new SMARTs so that we have a network of learning and continual improvement. Greater willingness of the relevant decision makers to fund such projects would very much help this area move forward.

6 Final thoughts

Any final thoughts or words of advice related to dynamic treatment regimes?

MK—Research work for discovering dynamic treatment regimes is inherently multidisciplinary and requires good team science. We need to find ways to train and facilitate teams who can work together to make progress, especially in practical implementation of this work. It is worth the investment.

PT—Methodological advances in biostatistics will be most likely to translate into substantive medical advances if the dynamic treatment regime/SMART community works on actual problems that come from practicing physicians and basic scientists. In my experience, the best medical collaborators often are physician-scientists who spend time in both the lab and the clinic. A biostatistician who wishes to do this sort of collaborative research must understand both the medicine and the biology, as well as the statistics. My advice to young biostatistical researchers is “Talk to your doctor.”

Acknowledgments

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Cancer Institute of the National Institutes of Health under Award Numbers P01CA142538 (Davidian, Laber) and R01HL118336 (Davidian). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Armstrong AJ, Garrett-Mayer EL, Eisenberger M. Response: Re: Adaptive therapy for androgen independent prostate cancer: a randomized selection trial of four regimens. *JNCI*. 2008; 100:682–683.
2. Thall PF, Logothetis C, Pagliaro LC, et al. Adaptive therapy for androgen-independent prostate cancer: a randomized selection trial of four regimens. *J Natl Cancer Inst*. 2007; 99:1613–1622. [PubMed: 17971530]
3. Zhao YF, Kosorok MR, Zeng D. Reinforcement learning design for cancer clinical trials. *Stat Med*. 2009; 28:3294–3315. [PubMed: 19750510]
4. Zhao YF, Zeng D, Socinski MA, et al. Reinforcement learning strategies for clinical trials in non-small cell lung cancer. *Biometrics*. 2011; 67:1422–1433. [PubMed: 21385164]
5. Robins JM, Hernan MA, Brumback B. Marginal. Structural models and causal inference in epidemiology. *Epidemiology*. 2000; 11:550–560. [PubMed: 10955408]
6. Lee JH, Thall PF, Ji Y, et al. Bayesian dose-finding in two treatment cycles based on the joint utility of efficacy and toxicity. *J Am Statist Assoc*. 2015; 110:711–722.