Chapter 2

DTRs and SMARTs: Definitions, designs, and applications

Kelley M. Kidwell

If it were not for the great variability among individuals, medicine might as well be a science and not an art.
—Sir William Osler, The Principles and Practice of Medicine 1892

2.1 ▪ Introduction to dynamic treatment regimes

Diseases or disorders such as substance abuse, depression, ADHD, autism, HIV, diabetes, and cancer require sequences of treatments over time to address changing characteristics of the disease and the patient. For example, a woman diagnosed with breast cancer may first have surgery, then receive chemotherapy and radiation. If the chemotherapy is deemed too toxic, a break from chemotherapy and/or radiation or lesser dose may be prescribed by her physician. If surgery, chemotherapy, and radiation eradicate the cancer and the woman is hormone receptor positive and post-menopausal, her physician will prescribe an aromatase inhibitor. If there is evidence of disease, additional chemotherapy and a targeted agent may be appropriate to reduce progression. Either way, the physician will continue to follow this woman, prescribing additional medication or modifying prescriptions based on disease and patient characteristics including treatment adherence to elicit the best outcome possible (long survival with little toxicity). To the woman, this seems like standard treatment. To the physician, this is a series of decisions based on information from previous patients with similar treatment history, characteristics, and behaviors. To the statistician, this is an example of a dynamic treatment regime.

Dynamic treatment regimes (DTRs) [245, 307] are also known as adaptive treatment strategies [184, 242], adaptive interventions [72, 249], multistage treatment strategies [390, 392], stepped-care strategies [357], treatment policies [213, 426, 427], and individualized treatment rules [274, 409]. All of these phrases denote sequences of treatments which have been tailored to an individual based on treatment history, patient characteristics and behaviors, and disease history or response to treatment. In plain terms, this is the guideline of treatment that patients hope their physicians have in their minds. A patient often

asks a series of questions which defines their dynamic treatment regime: what will my treatment be, what happens if I do not respond, what happens if I progress, what happens if I cannot tolerate the medication, what happens if I get better? In more mathematical terms, dynamic treatment regimes assign treatments based on a patient's time-dependent covariate history [143]. Thus, treatment assignment is dynamic within a person over time, varying because the person or disease is changing, with the goal of obtaining the best results for that person.

To construct a regime, several pieces of information are needed. First, there must be treatment options. These are not limited to different medications or drugs, but can also include different doses, modes of delivery (e.g., oral or injection), timing schedules, behavioral interventions, or no further treatment. Second, there must be critical decision points at which treatment is assessed and decisions are made to continue, alter, add, or subtract treatment. Regimes do not need to specify the constant addition or intensification of treatment, but rather can allow for less treatment (lower dose or discontinuation). When and how to lessen treatment if there is high burden, side effects, cost, or lasting benefit may be guided by a DTR. Third, tailoring variables allow for these decisions to personalize treatment. Tailoring variables include patient and treatment information which is available up to the time of the critical decision. For instance, these may be previous treatment, response to treatment, adherence to treatment, genetic information, test results, etc. A dynamic treatment regime can then be constructed by combining these three elements to operationalize treatment, provide guidance to physicians, and mimic clinical practice. Most of the methods that exist to construct DTRs are developed for a general k-stage case (k possible treatment decisions), but focus on two-stage DTRs for practical purposes. A general two-stage DTR can be formulated as "Treat with A; if there is response, treat with B; if there is not response, treat with C."

To further illustrate the combination of these elements, consider the treatment of breast cancer. One particular dynamic treatment regime for a 60-year-old, hormone receptor positive woman with stage II breast cancer is, "Following surgery, treat with chemotherapy for six cycles. If there is no evidence of cancer and the lymph nodes are negative, treat with an aromatase inhibitor for five years. If there is evidence of cancer following chemotherapy, continue chemotherapy for another six cycles. If the patient experiences a grade III or higher toxicity on the prescribed chemotherapy, switch to another chemotherapy." Note this dynamic treatment regime includes treatment options (chemotherapies, aromatase inhibitors), critical decision points (what to do following surgery, if cancer remains or not, if there is toxicity), and tailoring variables that personalize treatment (post-menopausal, hormone-receptor positive, stage II, lymph node negative, response to chemotherapy).

This example identifies only one dynamic treatment regime, thus illustrating the difference between a dynamic treatment regime and a simple sequence of treatments. Separating a DTR from a treatment sequence is perhaps the most subtle and often confused aspect of DTRs. A treatment sequence lacks the "dynamic" feature since it does not provide different options depending on the outcome to the first treatment. A treatment sequence would be "Treat with chemotherapy followed by an aromatase inhibitor for post-menopausal women," whereas a DTR includes a treatment specification for each option of the critical decision (usually, treatment if response and treatment if no response), such as that given above. Specifically, in our example, the DTR differs from a treatment sequence because it specifies the treatment sequence dependent on intermediate response (evidence of cancer): after a certain period of time on chemotherapy, for those without evidence of disease, the best treatment option is an aromatase inhibitor, and, for those who have evidence of disease, more chemotherapy elicits the best overall outcome.

Any combination of the three elements mentioned above does not constitute a DTR. A DTR must be constructed with thought and be a treatment regimen either used by physicians in the past or that physicians would consider using. Furthermore, it is exceedingly important that the DTR be "viable" [430], "realistic" [409], or "feasible" [307, 316] to capture those that may experience common contingencies [7] in the course of treatment. Adverse events, such as toxicities or disease progression (frequent occurrences in oncology treatment trials) can and should be considered in the construction of DTRs. To ensure viability, the tailoring variable should be widely accepted in the medical community, and there should be acceptance of a treatment protocol for those who experience a common contingency such as toxicity or other high-burden side effects or treatment dropout.

People enduring chronic diseases are generally treated in a dynamic fashion with a trade-off between short term and long term response. Generally both physicians and patients desire quick remediation as well as long term success. Aggressive initial treatment may set the stage for better overall success or lead to toxicities or other side effects which may hinder success of subsequent treatment. Additionally, treatment must address differences between people and within people over time. There is tremendous heterogeneity among people and within diseases. Because of this, the same treatment may not be the best treatment for everyone or may not even be the best treatment for an extended period of time for that one person.

For example, consider Lisa, who often has headaches. Usually, Lisa takes two aspirin and her headache goes away quickly, but sometimes aspirin is not enough. Aspirin may be right for Lisa when she has a minor headache and has just eaten a meal, but, if her headache is worse and she has an empty stomach, acetaminophen is the better choice. Lisa's brother Mark also suffers from headaches, but he finds acetaminophen works best for him for minor headaches and ibuprofen plus thirty minutes in a dark room work best when headaches are more intense. The best headache treatment differs between Lisa and Mark (even though they are related and close in age) and differs within Lisa and Mark depending on particular characteristics. This simple example with inter- and intra-person heterogeneity is amplified in the setting of complex diseases and disorders. Chronic diseases tend to affect multiple systems and therefore are accompanied with several issues or comorbidities. Dynamic treatment regimes are appropriate in this setting as they can guide what issue to treat when and adapt to responses over time.

Although chronic diseases (and some acute diseases that require continuous treatment) have been treated throughout history in terms of dynamic treatment regimes, medicine has rarely been studied this way. Rather, evidence for optimal treatment generally arises from randomized control trials or large observational cohorts where a particular treatment is compared to the standard of care, placebo, or another treatment, at one point in time for a somewhat homogeneous group of patients. Dynamic treatment regimes were defined and became an area in statistical research to address the issue that "one size does not fit all" and much treatment is ongoing and tailored to the individual. Furthermore, the ultimate outcome of a patient generally does not only depend on the most recent treatment, but rather depends on some amount of the entire treatment and response history. To address these issues with observational data, Robins wrote a series of papers [307, 308, 310, 312] that sparked this area and influenced many methods in both the observational and experimental settings (see Section 2.3).

Dynamic treatment regimes fit into the realm of personalized medicine. Treatment choices may be selected due to genetic/genomic information (the common definition of personalized medicine) or on more standard characteristics such as tumor features, test results, adherence information, side effects, burden, environmental characteristics, etc.

All disease areas have made a push for personalized medicine, often basing treatment on genetic information. High false discovery rates, low allele frequencies, technical variability, and large data issues hinder genetic personalization, but we can use standard clinical data to construct optimal dynamic treatment regimes that lead to better outcomes for patients. This approach to treatment that exploits various resources, assesses intermediate outcomes, and continues to tailor treatment is incorporated in many chronic care treatment models [11, 423, 53]. These models emphasize a repeated and dynamic approach to treatment.

Beyond personalizing or tailoring treatment, dynamic treatment regimes can identify and evaluate delayed effects. It is typically assumed either that the overall treatment effect is additive, so that the overall outcome of a particular treatment sequence is the sum of all its parts, or that only the last treatment is responsible for the immediate outcome. While this may hold for some treatments in some disease settings, delayed effects of treatments limit these assumptions. Delayed effects are those which do not occur immediately after treatment, but may affect a person or his disease later in time or set the stage for subsequent treatment. It is important to assess both intermediate and long term outcomes because what appears to be the best initial treatment may not yield the best overall outcome. There can be very real differences between local and global treatment comparisons. Thus, it is important to avoid myopic treatment decisions (decisions that only look one step ahead [186]) if the ultimate plan of treatment includes a series of therapies. Delayed effects can be thought of as treatment interactions, such that a particular initial treatment may enhance the effect of a particular second treatment (positive synergy) or lead to a higher proportion of toxicities or side effects that reduce the effect of future treatment (negative effect) [242].

A real life example of this phenomenon occurred in the analysis of data from a sequential trial investigating treatment regimens in prostate cancer [393]. In this trial, four treatments were offered in the first stage with the possibility of switching to one of the other three treatments based on no response to the first. One particular treatment, cyclophosphamide, vincristine, and dexamethasone (CVD), had the lowest probability of response after the first stage: 15% compared to the other treatments with response probabilities: 25%, 47%, and 42%. But, when CVD was followed by another specific treatment, ketoconazole plus doxorubicin alternating with vinblastine plus estramustine (KA/VE), the overall probability of success was the highest of all other treatment regimes at 58%. CVD seemed to enhance the salvage effect of the second treatment (positive synergy) [29]. These delayed effects are common in the treatment of chronic diseases and would be missed in standard one-stage trials and analysis [242].

2.2 ▪ Potential outcomes framework

In order to understand the methods in this book and make inference from observational or experimental data concerning dynamic treatment regimes, we present the underlying framework here. The potential outcomes or counterfactual framework quantifies treatment effects of dynamic treatment regimes and constructs estimands of interest. This framework was introduced by Neyman [256] to analyze causal effects of time-independent treatments in randomized trials. Extensions by Rubin [327] to observational studies and Robins [307, 305] to time-dependent treatments in observational and randomized studies apply to the setting of dynamic treatment regimes. Potential outcomes are the set of all possible outcomes for an individual where each possible outcome is associated with a unique treatment (sequence).

In a simpler setting, for example the treatment of a headache, let the outcome be the time it takes for the headache to subside after taking treatment. When one takes aspirin, this may be 15 minutes. When one takes acetaminophen this may take 30 minutes, or if one takes nothing this may take 60 minutes. All of these different times are potential outcomes such that a person will only experience one, depending on which treatment he/she decides to take, but they are all possible if he/she decided to take another treatment.

In the setting of dynamic treatment regimes, consider a regime with two decision points where the treatment decision at stage one is denoted by A_1 and at stage two is denoted by A_2. We are targeting the construction of a dynamic treatment regime such as, "First treat with A_1=1; if response, then treat with A_2=1; if no response, treat with $A_2 = 2$." Baseline information available up to stage one is given by X_1 and additional information up to stage two (which may depend on treatment at stage one) is given by X_2. Let Y be the outcome of interest. Then, our observed data are (X_1, A_1, X_2, A_2, Y). Let $H_1 = X_1$ be the covariate history up to the beginning of the first stage and $H_2 = (X_1, A_1, X_2)$ be the covariate and treatment history up to the start of the second stage. The treatment decision at stage one depends on baseline covariates such that $D_1 = d_1(X_1)$ and the treatment decision at stage two depends on information up to this stage including previous treatment denoted by $D_2 = d_2(X_1, A_1, X_2 A_1)$. Potential outcomes include $X_2^*(a_1)$, the potential covariate at the beginning of the second stage if treatment $A_1 = a_1$ was given, and $Y^*(a_1, a_2)$, the potential outcome at the end of stage two if treatment regime (a_1, a_2) was followed.

Three assumptions are necessary to estimate the effects of dynamic treatment regimes in this framework: (1) stable unit treatment value assumption (SUTVA); (2) sequential ignorability or no unmeasured confounding assumption; and (3) positivity. These assumptions are presented in the two-stage setting, but application to greater than two stages follows naturally.

(1) SUTVA [327] corresponds to the assumption that there is no interference between individuals or that each person's potential outcome is not influenced by the treatment of other people. In other words, this is consistency so that the potential outcome under the observed treatment (or dynamic treatment regime) is equal to the observed outcome. This assumption connects the observed data to the potential outcomes such that $X_2^*(a_1) = X_2$ and $Y^*(a_1, a_2) = Y$ when a_1 and a_2 are the treatments actually received. Thus, effects of dynamic treatment regimes can be written as functions of the multivariate distribution of the observable data.

(2) The next assumption of sequential ignorability (also known as no unmeasured confounding or conditional exchangeability) specifies that, conditional on the observed time-dependent covariate and treatment history up to time t_j, the treatment assignment, A_j, at time t_j is made independent of the potential outcomes of the individual. In the two-stage setting, $j = 1, 2$, this implies that for any regime (a_1, a_2), $A_1 \perp\!\!\!\perp [X_2^*(a_1), Y^*(a_1, a_2)] | H_1$ and $A_2 \perp\!\!\!\perp Y^*(a_1, a_2) | H_2$. This assumption always holds under sequential randomization (the experimental setting of a SMART introduced in Section 2.3), but must be evaluated on subject matter grounds in observational studies.

(3) Finally, the last assumption is of positivity, which defines the set of feasible regimes so that for every covariate-treatment history up to time t_j that has a positive probability of being observed, there must be a positive probability that the corresponding treatment dictated by the treatment regime will be observed. If this is not true, treatment regimes may need to be redefined so that the effect of the dynamic treatment regime can be estimated.

The ultimate goal is to treat using the *optimal* dynamic treatment regime which is evidence based. Bellman's principle of optimality explains, "An optimal policy has the

property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision" [27]. Thus, the aim is to find the best (feasible) treatment regime that has the maximal average outcome if everyone in the population followed that regime. The optimal dynamic treatment regime leading to this maximal average outcome could be found using dynamic programming [27] if we knew the distribution of the potential outcomes $\{X_1, X_2^*(a_1), Y^*(a_1,a_2)\}$ for each treatment sequence (a_1,a_2). This optimal decision at each stage may not be equal to the best myopic decision, thus requiring a research strategy which extends beyond a piecemeal conglomeration of myopic decisions [186]. This could include investigation of regimes in the observational or experimental setting. Most often in the observational setting, however, the multivariate distribution of potential outcomes is not known (violation of sequential ignorability), motivating the need for experimental design (SMART, see Section 2.4) to evaluate optimal decisions in a multistage process [186, 288].

2.3 ▪ Modes of constructing, estimating, or comparing dynamic treatment regimes

Using the potential outcomes framework, dynamic treatment regimes can be constructed, estimated, and compared in an observational or experimental setting. The assumptions described in the previous section must hold in both settings to estimate treatment effects and make inference about regimes.

2.3.1 ▪ Observational setting

For some diseases like chronic kidney failure, diabetes, and HIV it may more ethical and/or realistic to use observational studies to shed light on potential dynamic treatment regimes. This is what Lavori and Dawson define as a "data mining or epidemiologic approach" [188] due to the large number of potential dynamic treatment regimes that can be identified. Observational settings include electronic medical records or hospital databases, randomized encouragement trials, and cohort studies [63]. Dynamic treatment regimes have been examined in these settings using methods of G-estimation of structural nested models, other variations of G-estimation, and inverse-probability-of-treatment-weighted estimation of marginal structural models [323, 410, 302, 236, 263, 77]. In all of these methods, consideration of confounders is especially important since there may be many time-varying covariates that affect future treatments.

In order to make causal inference from the observed data, the assumptions are crucial. The no-unmeasured-confounding assumption is extremely difficult to satisfy in observational settings because it requires that treatment can only depend on observed past patient characteristics and treatment history and is therefore independent of future possible observations and the outcome. In a trial, this is given, but in observational data it cannot be verified that all relevant data associated with the treatment decision up to the decision point are available. Furthermore, even if confounders have been identified and collected, standard methods to adjust for confounders do not directly apply to time-varying treatment [188]. Another assumption for unbiased estimates from observational data is that there is no model misspecification. This assumption can be weakened if using doubly robust estimators [316, 21].

Observational data are typically high dimensional due to the multitude of possible treatments and time-varying confounders. Murphy [241] and Robins [316] developed methods to find optimal regimes in this high dimensional setting, but, due to the wealth

of information in covariate history, the optimal regime could be an intricate set of guidelines that are not practical. To circumvent this issue, others have focused on identifying more realistic regimes using a smaller subset of available information [263]. These methods are more practical and generally easier to implement, but they may overly limit the field of regimes since the estimability depends on that regime occurring in the data by more than one individual. Observational data allows inference for "viable" dynamic treatment regimes, but only those which are viable in the specific dataset, limited to the sample size, patient population, environmental setting, etc. There may be other possible dynamic treatment regimes of interest, but the particular dataset may not have such patients following these regimes or enough of these patients for estimation.

It is still beneficial to use observational data to construct dynamic treatment regimes and continue to create methods for inference from observational data. This resource is generally much cheaper than conducting a trial. Moreover, analyzing observational data may be a first step in understanding the treatment regimens being used and which ones tend to have best outcomes. Analyses may lead to choosing promising dynamic treatment regimes to study in an experimental setting. Additionally, it is possible that a randomized trial is not ethical or feasible due to aspects of treatment, the rarity of disease, or costs related to some aspect of the trial.

2.3.2 ▪ Experimental setting

To address some of the shortcomings of analyzing observational data and to provide prospective evidence of dynamic treatment regime effects, there is literature on clinical trial design to examine dynamic treatment regimes [184, 186, 387, 390, 393, 242]. This literature refers to the sequential multiple assignment randomized trial (SMART) where individuals are randomized multiple times and follow specific dynamic treatment regimes. The intention of this type of design is to develop dynamic treatment regimes, estimating the outcomes for each regime in the trial and then selecting the most promising dynamic treatment regime to compare to standard of care in a followup randomized control trial [242]. This objective addresses the dimensionality challenge of not just answering the question of "What treatment when?" but also to address how best to use tailoring variables and information prior to selecting treatment. Since it is unlikely that all of these components have been optimized prior to conducting a SMART, the goal is to conduct a series of trials which build upon one another, developing and refining promising dynamic treatment regimes leading to a confirmatory trial similar to the multiphase experimental approach [40, 73, 71, 62]. While this is ideal, the time and cost of trials may limit this intent and rather treat SMARTs as confirmatory trials.

Just as a randomized control trial is a fixed design that (generally) compares two or more treatments, a SMART is a fixed design that compares or constructs two or more treatment regimes. Thus, a SMART is a trial and a DTR is a treatment guideline carried out by a physician. The aim of a SMART is to construct effective dynamic treatment regimes. The same individuals begin a SMART and are followed throughout multiple randomizations until the end of the trial with fixed randomization probabilities and other trial operational characteristics. Therefore, SMARTs are able to address questions about the best treatment at certain points in time, the best sequences of treatments (or best modes of treatment delivery) depending on intermediate outcomes, the best intermediate outcomes to direct treatment, and how to individualize sequences of treatments based on biological, diagnostic, and/or other patient information. There should be one primary objective of the trial just like in any trial, but SMARTs may lead to more secondary and exploratory aims due to tailoring dynamic treatment regimes. In order for SMARTs to

be feasible, the intermediate outcome must be available for assessment within a relatively short time period, likely not to exceed one year. This is mainly due to the scope and relevance of conducting a trial. Thus, for some diseases with long assessment periods (for example, treatment of some breast or prostate cancers), a SMART is not an appropriate choice. However, SMARTs may still be relevant in treating other comorbidities or mental health issues in these types of diseases.

The most common SMART design includes two stages: an initial stage of randomization to one of two or more treatments followed by a period of followup. At some point in time, or over a period of time, response to the initial treatment and patient characteristics are assessed to then subsequently re-randomize individuals to second-stage treatment. Depending on the intermediate outcome status, one may or may not be re-randomized to a treatment option. Up-front consent of sequential randomizations is recommended so that individuals are randomized to subsequent treatment once eligible. This allows for the usage of data until randomization for balance between treatment assignments of responders and nonresponders. Conceptionally, there is no difference between up-front or sequential randomization and both can be handled accordingly through analysis, but sequential randomization may allow for identification of other potential tailoring variables at each step and more balanced randomization.

2.4 ▪ SMART designs

For more concrete examples, see Figure 2.1 for three of the most commonly used SMART designs. To keep the trial feasible and timely, SMARTs are mostly designed with two-stages and two to three treatment options at each stage. SMART designs, however, can accommodate more than two stages and more than two to three treatment options at each stage. For the three trial designs in Figure 2.1, treatments need not be unique. For example, in **A**, treatments C and D may be the same as G and H, or E and F may be the same as I and J; in **B**, treatments D and E may be the same as G and H; and in **C**, D or E may be the same as G. Additionally, first-stage treatments can also be an option in the second stage so that in design A, E or F may be treatment B, or I or J may be treatment A.

The SMART design in Figure 2.1 **A** is one where all patients are re-randomized, but the treatment assignment depends on response to initial treatment and perhaps the specific initial treatment (if treatments G, H, I, J differ from C, D, E, F). This design has been used for trials investigating treatments for alcohol dependence [267] and drug addiction [155]. In the trial for alcohol dependence, the "extending treatment effectiveness of Naltrexone design" (ExTENd) [267], all patients started with Naltrexone, thus treatments A and B were the same. Treatments C (Naltrexone) and D (Naltrexone and telephone disease management) were the same as treatments G and H, and treatments E (combined behavioral intervention, medical management, and placebo) and F (combined behavioral intervention, medical management, and Naltrexone) matched I and J.

Thus, in ExTENd, the difference between those who began with A versus those that began with B was not the initial medication, but rather the definition of the intermediate outcome of response. Explicitly, patients were randomized to be in a group where the nonresponse status was assessed with a stringent criterion (two or more heavy drinking days during the first eight weeks of Naltrexone treatment) or in a group with a lenient criterion for nonresponse (five or more heavy drinking days during the first eight weeks of Naltrexone). Thus this trial was able to explore if nonresponse to Naltrexone could be identified early (only two drinking days) or was better defined later (five drinking days), and additionally assess the best treatment for nonresponders and for responders to prevent

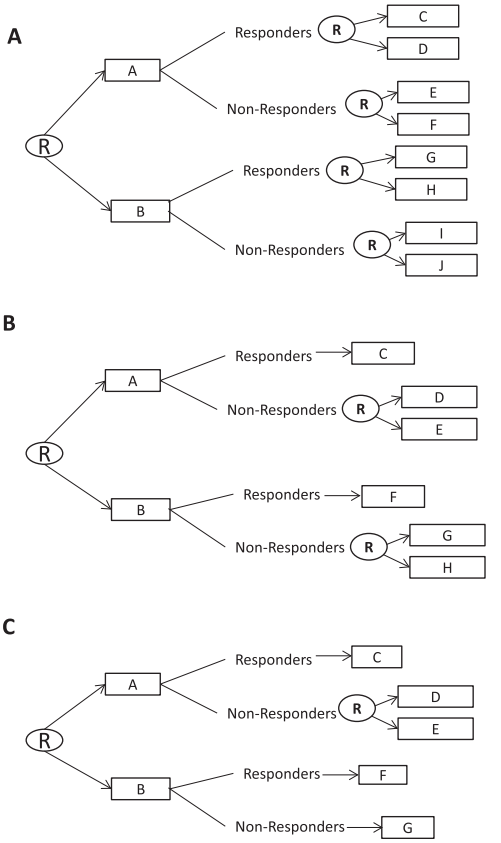

Figure 2.1. *Three of the most common two-stage SMART designs used in practice. All designs include two treatment options at stage one, A or B, and up to two treatment options depending on intermediate outcome (here, responder) status.* ***A*** *represents a SMART where both responders and nonresponders are re-randomized to treatment that depends on responder status.* ***B*** *represents a SMART where only nonresponders are re-randomized.* ***C*** *represents a SMART where re-randomization depends on both responder status and initial treatment.*

relapse. The primary objective of this trial was to test the main effect of the treatments for nonresponders (combined behavioral intervention, medical management and Naltrexone versus combined behavioral intervention, medical management and placebo) in terms of the percent of heavy drinking days and percent of drinking days over the last two months of the study. Secondary aims included exploring possible tailoring variables for the treatment options for nonresponders (for example, to assess if the average daily pill count of Naltrexone was associated with better outcome for those who continued Naltrexone or those who did not get Naltrexone).

There are eight embedded dynamic treatment regimes in this type of trial as shown in Figure 2.2 **A**. For the alcohol dependence trial, one such dynamic treatment regime was, "First treat with Naltrexone and medical management. If the person has two or more heavy drinking days during the first eight weeks (stringent criterion of response), then continue Naltrexone and medical management and add cognitive behavioral intervention. If the person has at most one heavy day of drinking in the first eight weeks, then continue the Naltrexone and add telephone disease management."

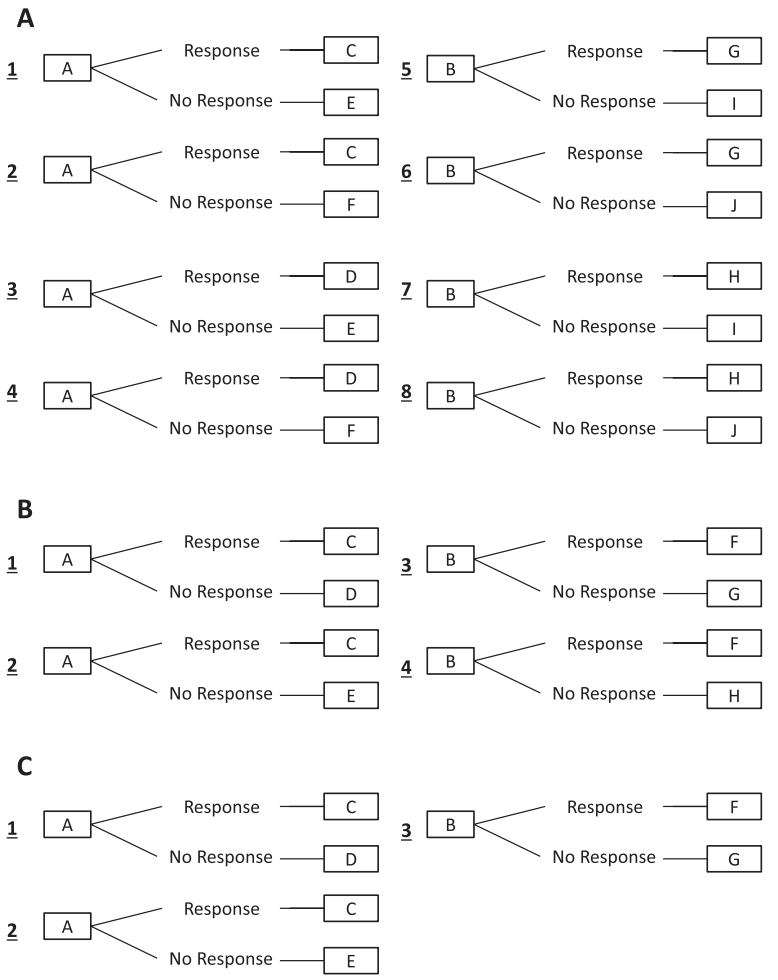

Figure 2.2. *Explicit list of every dynamic treatment regime from the corresponding SMART design in Figure 2.1. Design* ***A*** *includes eight dynamic treatment regimes, design* ***B*** *has four dynamic treatment regimes, and design* ***C*** *includes three dynamic treatment regimes.*

It is easy to see here that a dynamic treatment regime differs from a simple sequence of treatment or one branch of a SMART. Rather, one dynamic treatment regime includes two branches of the SMART and incorporates the treatment sequences for the options at the critical decision point. Thus some individuals are consistent with more than one dynamic treatment regime. In Figure 2.1 **A**, the responders to treatment A who received treatment C are consistent with regimes labeled 1 and 2 in Figure 2.2 **A**. Likewise, the responders to A who received treatment D are consistent with regimes 3 and 4. Thus, we can use the information from all individuals who are consistent with a dynamic treatment regime to estimate the overall outcome for that regime.

It is also important to note that the definition of a dynamic treatment regime does not include randomization. Randomization is a design property of a SMART that is not part of an embedded dynamic treatment regime. Randomization plays the role of ethical and unbiased treatment assignment.

The design in Figure 2.1 **A** is very general, but the most commonly used SMART design is that featured in Figure 2.1 **B**, where randomization only occurs for one group depending on response status. Often, it is of interest to assess the best second-stage therapy to re-engage nonresponders, so that only nonresponders are randomized. On the other hand, for some diseases there may not be any or there is only one second-stage option for nonresponders, but there is a question of the best maintenance treatment or waning of treatment for responders, so that responders are the only group that is re-randomized. In designing this type of SMART, a simple place to start is to specify two treatments for the first stage (for example, A and B). Then, in the second stage, randomize the nonresponders (or responders to the first-stage treatment) to either switch treatment (from A to B or B to A) or augment treatment (intensify A or intensify B). In this case, one dynamic treatment regime would be "First treat with A; if there is response, continue A; if there is no response, switch to B." There are four embedded dynamic treatment regimes in this design when two treatments are considered at each stage as shown in Figure 2.2 **B**.

Examples of this design in practice include trials in the areas of ADHD [272], acute myelogenous leukemia [363, 364], small-cell lung cancer [401], neuroblastoma [226, 225], diffuse large cell lymphoma [137], multiple myeloma [224], and metastatic malignant melanoma [17]. Many of the oncology trials re-randomized the responders to one of two maintenance treatments (often one type being no further treatment), while the nonresponders either continued treatment to see if longer treatment period would be effective, there was no further treatment, or the protocol specified standard of care (for the physician to treat appropriately). The first stage of these trials lasted anywhere between 22 days and 8.5 months. Intermediate outcomes consisted of medical definitions of remission, response, disease progression, having completed a number of cycles, or having obtained a specific test score. Followup spanned from months to years. Two other trials mimic this design, but begin with randomization to three treatment arms [156, 379], thus including six embedded dynamic treatment regimes.

In mental health, there have been two very large trials, STAR*D [333] and CATIE [342, 368], which are versions of this SMART design where the nonresponders are re-randomized in subsequent stages. STAR*D (Sequenced Treatment Alternatives to Relieve Depression) was the largest and longest trial to evaluate the effectiveness of different treatments for people suffering from major depressive disorder. This trial, funded by the National Institute of Mental Health (NIMH), enrolled more than 4,000 individuals over 7 years at 41 clinical sites. Due to the complexity of the design (and therefore analysis), a simplified version is presented. There were four possible levels of treatment where all participants began by receiving 14 weeks of citalopram (level 1). If they did not respond adequately (based on a symptom severity scale), they could choose to be randomized to one of four different treatments (switch group) or to be randomized to add one of three treatments to citalopram (augment group, level 2). Based on response to the switched or augmented treatment, participants could again choose to switch to one of two treatments or augment to one of two options (level 3). If there remained inadequate response, participants were randomized to one of two treatment options (level 4). All of those with adequate response remained on that treatment and were followed for year. Thus this trial is an extended version with three randomization stages, more than two treatment options at the first randomization stage, and groups of randomization based on individual choice (switch or augment).

CATIE (Clinical Antipsychotic Trials of Intervention Effectiveness) was also funded by NIMH to evaluate treatment sequences of antipsychotic drugs. This trial was smaller than STAR*D, with 1,460 participants who were followed for 18 months. Individuals were first randomized to one of five treatments and were assessed for psychotic symp-

toms and side effects. Participants could switch medication within 18 months if the first treatment was not effective or tolerable. If the treatment was effective and tolerable, then the individual remained on that treatment. Therefore, CATIE is a very similar to the SMART design in Figure 2.1 **B**, but with five treatment options at stage 1 and four treatment options at stage 2. See Chapter 11 for more details on analysis of the CATIE trial by imputing missing data. Since these trials were very large and some of the first SMART designs created while the terminology and methods were developing, they illustrate issues with missing data and the need to define viable treatment regimes. Both of these trials have been studied in the SMART/dynamic treatment regime literature, and regimes have been analyzed using machine learning or inverse probability weighting [277, 460, 352, 351].

Many of the oncology trials listed as examples following the SMART design in Figure 2.1 **B** were designed, run, and analyzed prior to the vocabulary of a SMART design and methods to analyze dynamic treatment regimes [168]. Rather, these trials were built out of necessity to answer pertinent treatment questions. Therefore, many of these trials were analyzed stage-specifically such that the outcome (survival) was compared between first-stage treatments ignoring second-stage treatment and between second-stage treatments ignoring initial treatment. These older trials focused on the myopic results and not the construction of optimal dynamic treatment regimes due to the limitation of available statistical methods at the time of the trial and analysis.

For example, the neuroblastoma trial was analyzed in 1999 and in 2009 [226, 225]. In 2009, the article made two stage-specific comparisons and conclusions: first-stage: bone transplantation is not advantageous over chemotherapy, $p = 0.39$; second-stage: 13-cis-retinoic acid is advantageous over no further treatment for those without disease progression regardless of induction therapy, $p = 0.006$. They also compared the groups of individuals who received a specific set of first- and second-stage treatments (not dynamic treatment regimes) and concluded that there was a marginally significant difference in survival between the groups with bone transplantation followed by 13-cis-retinoic acid versus chemotherapy followed by 13-cis-retinoic acid, $p = 0.054$. These conclusions, however, differ from those based on dynamic treatment regimes: all dynamic treatment regime-based analyses have found no significant difference in survival between any dynamic treatment regimes [213, 136, 377, 231, 169]. For more discussion on why stage-specific analysis may differ from dynamic treatment regime-based analysis, see Section 2.4.2.

A more recent SMART of this type merges oncology and mental health, studying neurobehavioral treatment for patients with metastatic malignant melanoma undergoing high dose interferon-alpha therapy. This trial was designed to investigate dynamic treatment regimes, not just main effects [17]. This study proposed the randomization of 70 individuals to one of two treatments (escitalopram or methylphenidate) for six weeks with the goal of treating mood and neurovegetative symptoms. At the end of six weeks, a questionnaire, the Hamilton Depression Scale (HAM-D), will evaluate symptoms, and, based on this score, randomization to subsequent treatment will occur. For those who respond to the initial treatment (HAM-D$\leq$11), the initial treatment will continue as they begin their interferon treatment. For those who do not respond to the initial treatment (HAM-D$>$11), individuals will be randomized between switching to the other initial treatment or add the other treatment to the initial treatment. Therefore, an example of a dynamic treatment regime in this trial is "First treat with escitalopram for six weeks. After six weeks, if the HAM-D score is less or equal to 11, continue this treatment while undergoing interferon treatment. If the HAM-D score is greater than 11, add methylphenidate to the escitalopram." The primary aim of this trial was to select the optimal dynamic treatment regime in terms of maximizing adherence to 12 weeks of interferon therapy (the number of treatments tolerated). This estimation aim requires marginal mean mod-

els [245]. Secondary aims included comparing the main effects of the initial treatment in terms of specific questionnaire measures (usual methods comparing two randomized groups) and comparing the second-stage option of switching or adding more therapy in terms of higher adherence for the nonresponders (usual analysis conditional on no response to initial treatment).

Another possible SMART design is shown in Figure 2.1 **C** where only nonresponders to a particular first-stage treatment are re-randomized. This may be an appropriate design if there are two or more possible first-stage treatments of interest, but only particular options are feasible or ethical based on the first-stage treatment. For example, this design was used for a trial investigating treatment for nonverbal children 5–8 years old with autism spectrum disorders [164] called the Adaptive CCNIA Developmental and Augmented Intervention Study. Initial treatments included the combination of (1) joint attention/joint engagement intervention (JAE) and enhanced milieu teaching (EMT) and (2) JAE and augmentative and alternative communication (AAC). AAC includes a device, so the investigator did not think it was appropriate to re-randomize the children who received the device in the initial stage to a treatment which did not include the device, especially if the child had responded to the device. Therefore, those who started with the combination of JAE and AAC followed a single dynamic treatment regime defined as "First treat with JAE and AAC and if there is response, continue the same therapy. If there is not response, intensify this therapy." Those that started with JAE and EMT could follow two dynamic treatment regimes since nonresponders were randomized to either an intensified version or a switch to JAE and AAC.

Another interesting aspect of this autism SMART is the definition of the intermediate outcome of response. Investigators assessed the child's improvement in communication via seven assessments and created two variables for each assessment. The variables included taking the difference between the average of the first two intervention sessions and the last two sessions and the difference between the assessment at the start of the trial and at the three-month visit. Response was defined as 25% or more improvement on at least 50% of the 14 measures [193]. Thus, response may not always be a simple definition, but it should be agreed upon by various experts in the field for implementation (see Chapter 15). The primary objective of this trial was to compare the main effects of the first-stage treatment, JAE and EMT versus JAE and AAC, on three different outcomes measuring communication. Additionally, the investigators were interested in interactions between baseline variables and differences in effects of the three dynamic treatment regimes on the outcomes.

2.4.1 ▪ Powering SMARTs

In designing a SMART, like any other clinical trial, the primary and secondary objectives must be well specified. A SMART allows for more options of these aims than a standard one-stage trial. SMARTs **A, B** or **C** in Figure 2.1 lend themselves to many different primary objectives, which are outlined in Table 2.1 [260, 168]: main effects, estimating or comparing dynamic treatment regimes, and further tailoring of the regimes. Standard clinical trials share some of the same objectives as SMARTs, but are not able to detect treatment interactions or delayed effects, investigate dynamic sequences of treatments, or tailor treatment regimens. Like standard trials that go through phases from pilots to large phase III trials, SMARTs may also be conducted similarly. A SMART can be a pilot study to address feasibility [6], a smaller-scale phase II study with increased levels of significance to estimate effects for future trials, or a large phase III trial to test specific hypotheses (of stage-specific treatment or dynamic treatment regimes) or estimate treatment effects.

Table 2.1. *Possible aims of a SMART. Aims (1) and (2) are similar to questions addressed by standard one-stage randomized clinical trials (RCTs), but aims (3) and (4) give further insight into dynamic treatment regimes (DTRs). All are possible through a sequential multiple assignment randomized trial (SMART). A selection of references provided are to methods and applied work for continuous, binary, or survival outcomes.*

Aim or clinical question	RCT	SMART	References
1. What is the best first-stage treatment (this may be a comparison of two or more treatments or of different time schedules, doses or delivery methods; hypothesis test)?	X	X	Usual RCT methods
2. What is the best second-stage treatment for (non)responders to a particular first-stage treatment (hypothesis test)?	X	X	Usual RCT methods scaled by (non)response probability
3. Does the outcome significantly differ between two or more DTRs (hypothesis test or estimation)?		X	[260, 85, 86, 249, 213, 426, 136, 427, 205, 104, 377, 430, 425], Chapters 5, 3
4. For a specific DTR, can we improve individual outcomes by further tailoring treatment by baseline or time-varying characteristics?		X	[352, 250, 458, 128, 451], Chapter 3

The most familiar objective for a SMART is the main effects objective. This aim complements the *ideal* goal of SMARTs to develop DTRs through a series of trials [242] and mimics that of a standard randomized trial, so power and analyses are standard. Main effects objectives addressed in rows 1 and 2 of Table 2.1 include focusing on either the comparison between treatments at stage 1 or between treatments conditional on response at stage 2. With this aim, trials may be sized as usual (for powering on the second-stage treatment comparisons, sample size must be upscaled based on response probability) with typical analyses. Most clinical trials are designed based on the main effects comparison, so its value resides in physicians and statisticians being familiar with it and with the standard tools for power and analysis. Additionally, as a bonus, the effects of the embedded dynamic treatment regimes can be estimated or the best definition of the tailoring variable may be investigated. This aim, however, may be too limited for a SMART given the resources involved and that a single SMART may be the definitive trial investigating optimal dynamic treatment regimes.

Differing from a standard trial, the main objective of a SMART may instead be based on dynamic treatment regimes as shown in Table 2.1 rows 3 and 4. A SMART may be powered to compare two or more specific DTRs using a hypothesis test or to estimate the overall outcomes of the regimes to find the most promising. The hypothesis test aim is more of a confirmatory trial to find the best DTR which may result in an optimal DTR to be implemented in practice after the trial or another trial demanding the comparison of the optimal DTR to the standard of care (likely a nondynamic treatment). As opposed to hypothesis testing and in line with the goal of developing dynamic treatment regimes, the objective of the trial may be to find the most promising DTR to study in future trials or estimating effects to size future confirmatory trials. These goals require special analytic tools based on dynamic treatment regimes and may pose a learning curve for statisticians and physicians. Aims based on DTRs, however, may be closer to answering relevant treatment questions in the field. Many methods exist for this aim (for a subset of references, see Table 2.1), but since this is a relatively new field, there are still outstanding questions and not yet as many resources, or any resources on particular issues, as with standard clinical trials.

Another objective of a SMART may be to provide detailed tailoring or more personalization of treatment. For example, by comparing or estimating effects of DTRs, the best DTR resulting from the trial is likely to be some form of "First treat with A and, if response, continue A; if no response, switch to B." But, further tailoring may provide increased insight into dynamic treatment regimes such that the form is more similar to "First treat with A; for those who respond, continue A; for those who do not respond, but were adherent to A, switch to B; for those who do not respond, but were not adherent to A, switch to C." This is still a relatively simple definition of a DTR, but it has been further tailored by adherence. One can imagine even more personalized regimes by exploring the effects on particular subgroups based on baseline or intermediate outcomes. This is unlikely to be a primary goal of a SMART, but very likely to be a secondary or exploratory aim.

Currently, there are a few easy to use sample size calculators for designing SMARTs to compare DTRs. These easy to use applications are generally conservative and based on test statistics using marginal means model variances, thus methods in Chapter 3 provide more efficient calculations based on optimal semiparametric variances. With this caveat, sample size calculators for continuous, binary and survival outcomes with the goal of estimation to find the most promising DTR or comparison of DTRs may be found at `https://methodology.psu.edu/downloads`.

2.4.2 ▪ Comparisons to SMARTs

There are many benefits to using a SMART design, but to continue to clarify them, it is helpful to contrast the SMART design to other designs.

Separate one-stage trials The closest analogue to a SMART with the same goal of constructing dynamic treatment regimes is to combine information from several separate one-stage trials [186, 193, 60]. This approach pieces together the best myopic decisions, which may not be the same as the optimal regime. A piecemeal strategy starts at the beginning of a treatment regime and defines the best treatment independently at each stage based on the best outcome at that stage alone, whereas an optimal regime is found using a recursive algorithm [27] considering the initial state and decisions. A recursive approach is not valid in combining results from multiple trials with different study samples requiring the need for a SMART to follow the same sample across treatment regimes, collect confounding information, and to capitalize on the possibility of delayed, diagnostic, or prescriptive effects of the initial treatment (discussed in Section 8.1).

By following the same individuals throughout the trial, a SMART is generally more efficient than combining data or conclusions from several trials and has the ability to analyze information that could be missed. For example, if treatment A showed the best response in the first trial, only treatment A would be used in the second trial which then tests subsequent treatment for responders and nonresponders. But, if there is a delayed effect such that treatment B followed by C for responders and E for nonresponders leads to the best DTR, separate trials would never find this result (the best DTR from the trial used as an example at the end of Section 8.1 would never be found). Moreover, information up to randomization of the second-stage treatment can be considered in randomization (for stratification) in a SMART and can additionally be used in analysis to assess whether these characteristics or intermediate outcomes provide information for individualizing the second treatment. This advantage of SMARTs, assessing prescriptive or diagnostic effects, may be lost in separate trials due to recall bias or not all having uniform initial treatment. Last, SMARTs can generally recruit from a wider pool of individuals due to taking ad-

vantage of heterogeneity between people and in disease characteristics, whereas standard one-stage trials attempt to sample from a somewhat homogeneous population. Since these heterogeneous individuals are followed throughout the stages of treatment, cohort effects are prevented and results may be generalizable to a wider group of individuals with the disease of interest. Furthermore, there may be less dropout because individuals who do not respond have the option of obtaining further treatment. This is likely to be a selling point for participation in a SMART.

Crossover and factorial designs SMARTs are similar to crossover trials and factorial trials, but differ from both. Crossover trials involve giving two or more treatments to the same group of individuals. The individual receives the first treatment and after a washout period receives the second treatment. Crossover trials reduce confounding factors associated with between-subjects designs and therefore have lower sample size than their two-arm standard trial counterparts since an individual serves as his own control. A SMART does not require a washout period because it is specifically interested in the sequence of treatments. Also, in a SMART design, the second treatment is given based on intermediate response, whereas all individuals in a crossover trial receive the same second treatment. SMARTs use the between-subjects confounding to their advantage to find tailored treatment sequences.

Factorial designs can address treatment interactions, unlike crossover trials, and test differences in treatment effects. The simplest factorial design is a 2×2 design where treatment assignments are set up in a 2×2 table and individuals fall into one cell by following that combination of treatments. For example, consider treatments A and B. We can set up a factorial design where we have treatment A or placebo for A versus treatment B or placebo for B. Here, the four cells include treatment A and placebo for B, treatment A and treatment B, treatment B and placebo for A, placebo for A and placebo for B. This trial maximizes those receiving some type of active treatment and can evaluate treatments A and B independently, as well as in combination. Treatment allocation does not depend on intermediate outcomes, and is not sequential but simultaneous, thus differing from a SMART. A SMART may be thought of a version of a sequential factorial design tailored to intermediate response.

Adaptive designs Adaptive designs have become more popular and many confuse SMARTs with these designs due to some of the vocabulary (adaptive treatment strategies). Adaptive designs include a wide variety of designs where some trial feature changes for future individuals based on past individuals. These designs are very flexible and therefore there are numerous types of trial designs that fall under the umbrella of adaptive designs. Designs can be adaptive both in the learning phase and confirmatory phase of trials [160]. For example, there are adaptive dose finding designs (phase 1 trials) where dose assignment may depend on accumulating information based on toxicity or efficacy. Combined or seamless phase designs may use data collected in phase I or II to modify aspects of the trial leading to the next phase (II or III). In more confirmatory studies, randomization probabilities may change due to promising treatment effects, or a trial may stop early due to interim analysis in group sequential design, sample size may be re-estimated, or some combination of these changes may be made to trial characteristics as information is accumulated. These trials all modify an operational characteristic of the trial while it is ongoing based on collected data. This design is adaptive between individuals such that changes occur for future individuals. In a SMART, the adaptation occurs within individuals such that the same individual may change treatment based on his or her own information, not on all of the information from everyone in the trial up to the

point of re-randomization. In a SMART, randomization probabilities and all aspects of the design are fixed at the beginning of the trial. There has been some work to integrate these two type of designs into adaptive SMARTs [387, 393, 433, 66, 191], but generally these are distinct types of trial design.

2.5 ▪ Conclusion

Dynamic treatment regimes naturally express the treatment guideline of many acute and chronic diseases which require sequences of treatments tailored to the individual. These regimes can be constructed using observational or experimental data, and treatment effects can be estimated and compared using a SMART design. SMARTs have the potential to address treatment questions unlike those answered with a standard randomized clinical trial pertaining to the timing, sequencing, and tailored of treatments and other interventions. The SMART differs from adaptive designs due to following the same patients throughout with fixed design parameters and therefore can investigate delayed effects and take advantage of treatment interactions. There are a number of SMARTs in practice, particularly in the areas of mental health and oncology, but the number of SMARTs will likely grow as the terminology and methods disseminate. SMARTs and their resulting product of dynamic treatment regimes exemplify personalized medicine.