



Published in final edited form as:

Clin Cancer Res. 2018 February 15; 24(4): 730–736. doi:10.1158/1078-0432.CCR-17-1355.

Sequential, Multiple Assignment, Randomized Trial Designs in Immuno-Oncology Research

Kelley M Kidwell, PhD^A, Michael A. Postow, MD^{B,C}, and Katherine S. Panageas, DrPH^D

^ADepartment of Biostatistics, University of Michigan, School of Public Health, Ann Arbor, MI

^BDepartment of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY

^CWeill Cornell Medical College, New York, NY

^DDepartment of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY

Abstract

Clinical trials investigating immune checkpoint inhibitors have led to the approval of anti-CTLA-4 (cytotoxic T-lymphocyte antigen-4), anti-PD-1 (programmed death-1) and anti-PD-L1 (PD-ligand 1) drugs by the United States Food & Drug Administration (FDA) for numerous tumor types. In the treatment of metastatic melanoma, combinations of checkpoint inhibitors are more effective than single agent inhibitors, but combination immunotherapy is associated with increased frequency and severity of toxicity. There are questions about the use of combination immunotherapy or single agent anti-PD-1 as initial therapy and the number of doses of either approach required to sustain a response. In this paper, we describe a novel use of sequential multiple assignment randomized trial (SMART) design to evaluate immune checkpoint inhibitors to find treatment regimens that adapt within individual based on intermediate response and lead to the longest overall survival. We provide a hypothetical example SMART design for BRAF wild-type metastatic melanoma as a framework for investigating immunotherapy treatment regimens. We compare implementing a SMART design to implementing multiple traditional randomized clinical trials. We illustrate the benefits of a SMART over traditional trial designs and acknowledge the complexity of a SMART. SMART designs may be an optimal way to find treatment strategies that yield durable response, longer survival, and lower toxicity.

Introduction

Clinical trials investigating immune checkpoint inhibitors have led to the approval of anti-CTLA-4 (cytotoxic T-lymphocyte antigen-4), anti-PD-1 (programmed death-1) and anti-PD-L1 (PD-ligand 1) drugs by the United States Food & Drug Administration (FDA) for numerous tumor types. Immune checkpoint inhibitors are a novel class of immunotherapy

Corresponding author: Kelley M Kidwell, PhD, Department of Biostatistics, University of Michigan, 1415 Washington Heights, SPH II M2525, Ann Arbor, MI 48104, 734-764-6724, kidwell@umich.edu.

Conflicts of Interest: Dr. Postow declares conflicts of interest with BMS (advisory board, honoraria, research grant), Merck (honoraria, advisory board), Novartis (advisory board), and Array BioPharma (advisory board). Drs. Panageas and Kidwell do not have any conflicts of interest.

agents that block normally negative regulatory proteins on T cells and enable immune system activation. By activating the immune system rather than directly attacking the cancer, immunotherapy drugs differ from cytotoxic chemotherapy and oncogene directed molecularly targeted agents. Cytotoxic chemotherapy or molecularly targeted agents generally provide clinical benefit during treatment and usually not after treatment discontinuation, whereas immunotherapy benefit may persist after treatment discontinuation.

The anti-CTLA-4 drug ipilimumab was approved for the treatment of metastatic melanoma in 2011 and as adjuvant therapy for resected stage III melanoma in 2015. Inhibition of CTLA-4 is also being tested in other malignancies. In melanoma, ipilimumab improves overall survival but is associated with 20% grade 3/4 immune related adverse events (1–6). Agents that inhibit PD-1 and PD-L1 have less immune related adverse events than CTLA-4 blocking agents (7). PD-1 and PD-L1 agents have been approved by the FDA for use in multiple malignancies including, but not limited to, melanoma (nivolumab and pembrolizumab), non-small cell lung cancer (NSCLC) (nivolumab, pembrolizumab, and atezolizumab), renal cell carcinoma (nivolumab), and urothelial carcinoma (atezolizumab) (8–10). Combinations of checkpoint inhibitors that block both CTLA-4 and PD-1 are more effective than CTLA-4 blockade alone (ipilimumab) in patients with melanoma, but combination immunotherapy is associated with increased frequency and severity of toxicity. Although we build our framework on the FDA approved combination of anti-PD-1 therapy and ipilimumab as this reflects the current landscape, one could replace the anti-PD-1 and ipilimumab combination with anti-PD-1 and any drug to reflect novel combination agents that may become available down the pipeline such as inhibitors of indoleamine-2,3-dioxygenase (IDO).

Some individuals may not need combination therapy because they may respond to a single agent and these individuals should not be subjected to increased toxicities associated with combination therapy. Defining this group of individuals, however, is difficult. Many trials are being proposed to evaluate combinations or sequences of immunotherapy drugs alone, in combination with other treatments such as chemotherapy, radiation, and targeted therapies, or with varied doses and schedules (sequential versus concurrent). The goal of these trials is to increase efficacy and decrease toxicity (11).

The long-term effect of immune activation by these drugs is unknown. It is also unknown whether individuals need continued treatment. Oncologists must optimize a balance in clinic, incorporating observed efficacy and toxicity, and informally implement treatment pathways so that treatment may change for an individual depending on the individual's status. Many of these treatment pathways are ad hoc, based on the physician's experience and judgement or information pieced together from several randomized clinical trials. There is a need for formalized, evidence-based treatment pathways to inform decision-making over the course of care. Formal, evidence-based treatment guidelines that adapt treatment based on a patient's outcomes, including efficacy and toxicity, are known as treatment pathways, dynamic treatment regimens (12) or adaptive interventions (13). Specifically, a treatment pathway is a sequence of treatment guidelines or decisions that indicate if, when and how to modify the dosage or duration of interventions at decision stages throughout clinical care (14). For example, in treating individuals with stage III or stage IV Hodgkin Lymphoma, one

treatment pathway is: “Treat with two cycles of doxorubicin, bleomycin, vinblastine and dacarbazine (ABVD). At the end of therapy (6-8 weeks), perform positron emission tomography/computed tomography (PET/CT) imaging. Treat with an additional 4 cycles of ABVD if the scan scores 1-3 on the Deauville scale (considered a negative scan). Otherwise, if the scan scores 4-5 on the Deauville scale (considered a positive scan), switch treatment to escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone (eBEACOPP) for 6 cycles (15).” Note, that one treatment pathway includes an initial treatment followed by subsequent treatment that depends on an intermediate outcome for all possibilities of that intermediate outcome.

Treatment pathways are difficult to develop in traditional randomized clinical trial settings because they specify adapting treatments over time for an individual based on response and/or toxicity. Treatments may have delayed effects such that the best initial treatment is not a part of the best overall treatment regimen. For example, one treatment may initially produce the best response rate, but that treatment may also be so aggressive that for those who did not have a response, they cannot tolerate additional treatment; whereas another treatment may produce a lower proportion of responders initially, but can be followed by an additional treatment to rescue more non-responders and lead to a better overall response rate and longer survival. Thus, treatments in combination or sequence do not necessarily result in overall best outcomes. The sequential, multiple assignment, randomized trial (SMART) (16,17) is a multistage trial that is designed to develop and investigate treatment pathways. SMART designs can investigate delayed effects, treatment synergies and antagonisms, and provide robust evidence about the timing, sequences, and combinations of immunotherapies. Furthermore, treatment pathways may be individualized to find baseline and time-varying clinical and pathologic characteristics associated with optimal response.

In this paper, we describe a novel use of SMART design to evaluate immuno-oncologic agents. We provide a hypothetical example SMART design for metastatic melanoma as a framework for investigating immunotherapy treatment. We compare implementation of a SMART design to implementation of multiple traditional randomized clinical trials. We illustrate the benefits of a SMART over traditional trial designs and acknowledge the complexity of a SMART. SMART designs may be an optimal way to find treatment strategies that yield durable response, longer survival, and lower toxicity.

Sequential, Multiple Assignment, Randomized Trial Design

A SMART is a multi-stage randomized trial in which each stage corresponds to an important treatment decision point. Participants are enrolled in a SMART and followed throughout the trial, but each participant may be randomized more than once. Subsequent randomizations allow for unbiased comparisons of post-initial randomization treatments and comparisons of treatment pathways. The goal of a SMART is to develop and find evidence of effective treatment pathways that mimic clinical practice.

In a generic 2-stage SMART, participants are randomized between several treatments (usually 2-3) (Figure 1). Participants are followed, and an intermediate outcome is assessed over time or at a specific time. Based on the intermediate outcome, participants may be

classified into groups, and they may be re-randomized to subsequent treatment. The intermediate outcome is a measure of early success or failure that allows the identification of those who may benefit from a treatment change. This intermediate outcome, also known as a tailoring variable, should have only a few categories so that it is a low dimensional summary that is well defined, agreed upon, implementable in practice, and gives early information about the overall endpoint. This intermediate outcome does not need to be defined as response/non-response or more specifically, as tumor response, but rather it may be defined differently, such as adherence to treatment, a composite of efficacy measures, or efficacy and toxicity measures. It is imperative that the intermediate outcome is validated and replicable. Although the 2-stage design is most commonly used, SMARTs are not limited to 2 stages, such as a SMART that investigated treatment strategies in prostate cancer (18).

A SMART is similar to other commonly used trial designs, but has unique features that enable the development of robust evidence of effective treatment strategies. The SMART design is a type of sequential factorial trial design in which the second stage treatment is restricted, based on the previous response. A SMART design is similar to a crossover trial in that the same participants are followed throughout the trial and participants may receive multiple treatments. However, in a SMART, subsequent treatment is based on the response to the previous treatment, and a SMART design takes advantage of treatment interactions as opposed to washing out treatment effects (i.e., a SMART does not require a period of time in between treatments to eliminate carryover effects from the initial treatment on the assessment of the second stage treatment).

We focus this overview on SMART designs that are non-adaptive. In a non-adaptive SMART the operating characteristics of the trial, including randomization probabilities and eligibility criteria are pre-determined and fixed throughout the trial. Treatment may adapt within participant based on intermediate response, but randomization probabilities or other trial operating characteristics do not change for future participants based on previous participants' results.

By following the same participants over the trial, a SMART enables the development of evidence for treatment pathways that specify an initial treatment, followed by a maintenance treatment for responders and rescue treatment for non-responders. These treatment pathways are embedded within a SMART design, but within the trial participants are randomized to treatments based on the intermediate outcome to enable unbiased comparisons and valid causal inference. The end goal of the trial is to provide definitive evidence for treatment pathways to be used in practice. The SMART design has been used in oncology (19,20), mental health (21), and other areas (22), but this is the first description of using a SMART in immuno-oncology to our knowledge.

Hypothetical Melanoma Sequential, Multiple Assignment, Randomized Trial

Ipilimumab and anti-PD-1 therapy currently are approved to treat metastatic melanoma. However, combinations of these and other immunotherapy drugs may cause toxic events, and it remains unclear whether patients should start with these combinations or start with single agent anti-PD-1 therapy and receive these additional treatments upon disease

progression. There are also questions about the number of doses required to sustain a response for single agent or combination therapy. The best treatment strategy is unknown that may provide enough therapy for sustained response, and limit toxicities. A SMART design may address these questions to provide rigorous evidence for the best immunotherapy treatment pathway in individuals. Our proposed example focuses on patients with BRAF wild-type metastatic melanoma to avoid complexities of additionally considering incorporation of BRAF and MEK inhibitors into the treatment regimen of patients with BRAF mutant melanoma.

In a hypothetical SMART design to investigate treatment strategies including anti-PD-1 therapy and ipilimumab, participants may be randomized in the first stage to receive 4 doses of single agent anti-PD-1 therapy (pembrolizumab 2 mg/kg or nivolumab 240 mg) or combination nivolumab (1 mg/kg) and ipilimumab (3 mg/kg) (Figure 2, note these drugs could be replaced with any novel immunotherapy or approved drug). During follow-up, participants would be evaluated for their tumor response; the intermediate outcome in this SMART would be defined by disease response after 4 doses of immunotherapy (week 12). Although Response Evaluation Criteria in Solid Tumors (RECIST) could be used to define disease response, favorable response could also be defined as any decline in total tumor burden, even in the presence of new lesions, as specified by principles related to Immune Related Response Criteria (23).

In the second stage of the trial, responders to either initial treatment would be re-randomized to continue versus discontinue their initial treatment. Specifically, participants who responded to single agent anti-PD-1 would be re-randomized to continue current treatment for additional doses up to 2 years or to discontinue treatment and participants who responded to the combination of anti-PD-1 + ipilimumab would be re-randomized to continue anti-PD-1 maintenance or discontinue treatment. Participants who did not respond to single agent anti-PD-1 by 12 weeks would be re-randomized to receive ipilimumab or the combination of anti-PD-1 and ipilimumab. Participants who did not respond to the combination therapy would receive the standard of care (e.g., oncogene-directed targeted therapy if appropriate, chemotherapy, or considered for clinical trials; Figure 2). As newer drugs become available and are promising for non-responders to combination therapy, we anticipate that there could be an additional randomization for these non-responders to explore additional treatment pathways. All participants would be followed for at least 28 months. The overall outcome of the trial would be overall survival. Any participant who experienced major toxicity at any time or progressive disease in the second stage would be removed from the study and treated as directed by the treating physician.

Participants belong to one subgroup (Figure 2) in a SMART. Two subgroups make up one treatment pathway since a treatment pathway describes the clinical guidelines for initial treatment and subsequent treatment for both responders and non-responders (Figure 2). While there are 7 subgroups that a participant may belong to, there are 6 embedded treatment pathways in this SMART design. The 6 treatment pathways include:

1. First begin with single agent anti-PD-1 therapy. If no response to single agent anti-PD-1 therapy, then switch to single agent ipilimumab. If response to single agent anti-PD-1, then continue treatment (subgroups 1 and 3);
2. First begin with single agent anti-PD-1 therapy. If no response to single agent anti-PD-1 therapy, then switch to single agent ipilimumab. If response to single agent anti-PD-1, then discontinue treatment (subgroups 1 and 4);
3. First begin with single agent anti-PD-1 therapy. If no response to single agent anti-PD-1 therapy, then add ipilimumab to anti-PD-1 therapy. If response to single agent anti-PD-1 therapy, then continue treatment (subgroups 2 and 3);
4. First begin with single agent anti-PD-1 therapy. If no response to single agent anti-PD-1 therapy, then add ipilimumab to anti-PD-1 therapy. If response to single agent anti-PD-1 therapy, then discontinue treatment (subgroups 2 and 4);
5. First begin with combination anti-PD-1 therapy + ipilimumab. If no response to combination anti-PD-1 therapy + ipilimumab, then receive standard of care. If response to combination anti-PD-1 therapy + ipilimumab, then continue treatment (subgroups 5 and 6);
6. First begin with combination anti-PD-1 therapy + ipilimumab. If no response to combination anti-PD-1 therapy + ipilimumab then receive standard of care. If response to combination anti-PD-1 therapy + ipilimumab, then discontinue treatment (subgroups 5 and 7).

A SMART may have several scientific aims, some of which may resemble those of traditional trials and some, pertaining to the treatment pathways, differ. It is important, as in standard trials, to identify and power on a primary aim. Subsequent aims and multiple comparisons may be additionally powered for using any type I error control method (24). In metastatic melanoma, the SMART may be interested in answering one of following 4 questions:

1. Does a treatment strategy that begins with single agent anti-PD-1 or combination anti-PD-1 and ipilimumab therapy lead to the longest overall survival?
2. For responders to initial therapy, does continuing or discontinuing treatment provide the longest overall survival?
3. For non-responders to single agent anti-PD-1 therapy, does ipilimumab or the combination of ipilimumab and anti-PD-1 therapy provide the longest overall survival?
4. Is there a difference in the overall survival between the 6 embedded treatment pathways?

Questions similar to numbers 1, 2 and 3 could be answered in three separate, traditional, parallel arm clinical trials. The traditional paradigm would run a single stage trial (e.g., single agent versus combination therapy) to determine the most effective therapy. A first trial may investigate single agent anti-PD-1 versus the combination of anti-PD-1 and ipilimumab. Another trial with a randomized discontinuation design could identify if continuing or

discontinuing treatment leads to longer overall survival for individuals who received the most effective therapy (e.g., anti-PD-1 alone or in combination with ipilimumab). And a third trial could determine for those refractory to anti-PD-1 therapy, if ipilimumab or the combination of ipilimumab and anti-PD-1 therapy results in longer survival. For each of these three traditional trials, power and analyses are standard in terms of powering for and analyzing a two-group comparison with a survival outcome.

If question number 1, 2 or 3 is the primary aim of a SMART, the sample size and analysis plan is also standard; however, for question numbers 2 and 3, the calculated sample size must be inflated. For question 2, the sample size must be inflated based on the assumed response rates to first stage therapies. Specifically, if 40% respond to single agent therapy and 55% to combination therapy, the calculated two group comparison sample size must be increased by these amounts to ensure that in the SMART there will be sufficient responders in the second stage. For question 3, the sample size must also be inflated for the expected percentage of non-responders to anti-PD-1 therapy. Similarly, in a standard, one stage trial to address question 2 (or 3), more patients would need to be screened to account for the response status, but unlike a SMART, the non-responders (responders) would not be followed. Furthermore, implementing three separate trials may not provide robust evidence for entire treatment pathways and instead provide evidence only for the best treatments at specific time points.

For a SMART powered on question number 1, 2 or 3, the analysis of treatment pathways would be exploratory and hypothesis generating to be confirmed in a follow-up trial. Instead, the SMART may be powered to compare the embedded treatment pathways (question number 4), in contrast with the stage-specific differences. Comparisons of pathways require power calculations and analytic methods specific to SMART designs. Currently, the only sample size calculator available for a SMART design with a survival outcome compares two specific treatment pathways using a weighted log-rank test. This calculator is only applicable for designs similar to the hypothetical melanoma SMART if the non-responders to anti-PD-1 therapy were not re-randomized (i.e., if there were only 4 embedded treatment pathways instead of 6) (25). Any other SMART design (e.g. the design in Figure 1 or our hypothetical design in Figure 2) or any other test (e.g. a global test of equality across all treatment pathways or finding the best set of treatment pathways using multiple comparisons with the best) requires statistical simulation. Other sample size calculations exist for survival outcomes, but do not have an easy to implement calculator (26,27). Methods are available to estimate survival (28,29) and compare (25,26,30–32) treatment pathways with survival outcomes and R packages (33) can aid in the analysis.

In this example, we calculate sample sizes of implementing three single trials versus implementing one trial using a SMART design. For the first single stage trial, we assume a log-rank test, survival rates of 80% and 68%, respectively, at 1 year for combination and single agent anti-PD-1, exponential survival distributions, 1 year for accrual, and an additional 2.5 years of follow-up. The same assumptions were applied for continuing initial treatment versus discontinuing the initial (this is a conservative sample size for this trial since the survival rates at 1 year would likely be closer together and require more patients). In order to have the same assumptions across the single stage trials and SMART design, the

survival rate at 1 year for those who did not respond to single agent anti-PD-1 therapy and received ipilimumab was set at 68% and for those who received the combination anti-PD-1 and ipilimumab was set to 74%. Parameters for the SMART were specified to mimic the single stage settings with the additional assumptions of a response rate to initial therapy being 40% and 1 year survival rates of 69%, 68%, 75%, 74%, 80%, 74% for the treatment pathways 1 through 6, respectively. For the SMART, a weighted log-rank test of any difference in the 6 treatment pathways was used for power via simulation (30,33). With these assumptions, 570 participants are required to observe any difference in the 6 embedded treatment pathways within 1 SMART (Table 1). This sample size is less than the 1,142 participants that are required by summing the sample sizes with the same assumptions using 3 traditional, single stage trials. We note that using a global test in the SMART allows for less participants and that potentially one of the trials in the single stage trial setting may be dropped based on previous trial results. But, a SMART allows us to answer many questions simultaneously and find optimal treatment pathways potentially ignored in the single stage setting.

A SMART would most likely require less time from start to finish than the single stage trials since it is unlikely the single stage trials would be run simultaneously (because the trials based on response to initial treatment would require an actionable result from the first trial) (34). Furthermore, because participants are followed throughout the trial and offered follow-up treatment, individuals may be more likely to enroll in the SMART (i.e. the sample of participants in a SMART may be more generalizable) and adhere to treatment (34).

Beyond the sequences of treatments in a SMART design that are tailored to an individual based on intermediate outcome, additional analyses (like subgroup analyses in traditional trials) may evaluate more individualized treatment pathways. Information including demographic, clinical, and pathologic data collected at baseline and between baseline and the measurement of the intermediate outcome may be used to further individualize treatment sequences for better overall survival. To further personalize treatment pathways, the analysis requires methods specific for SMART data such as Q-learning or other similar methods (35,36). Briefly, Q-learning, borrowed from computer science, is an extension of regression to sequential treatments (37). Q-learning is a series of regressions used to construct a sequence of treatment guidelines that maximize the outcome (e.g., find more detailed treatment pathways that include baseline and time-varying variables associated with the longest survival). It may be as beneficial for some individuals to receive single agent as combination therapy even when combination therapy is better when averaged across all individuals. In addition, a subgroup of individuals may benefit more from single agent therapy because of savings in cost and toxicity compared to combination therapy. These questions are unlikely to be powered for in the SMART, but a priori hypotheses can direct analysis and lead to the identification of more personalized treatment pathways that can be validated in subsequent trials.

Discussion

This paper has focused on an example SMART in BRAF wild-type metastatic melanoma to answer questions about the best treatment pathways including ipilimumab and anti-PD-1

therapy. As new immunotherapies are available for trials, ipilimumab may ultimately be replaced in this type of design by one of the more novel drugs (e.g., inhibitors of the immunosuppressive enzyme IDO or other checkpoint inhibitors such as drugs targeting lymphocyte-activation gene 3 “LAG-3”). Our proposed SMART design could be considered as a template for testing any number of these potential future possible combinations.

A SMART design may be a more efficient trial design to understand which immunotherapy treatment pathways in BRAF wild-type metastatic melanoma lead to the longest overall survival. SMARTs can definitively evaluate the treatment pathways that many physicians use in practice leading to the recommendation of treatments over time based on individual response. One single SMART can enroll and continue to follow participants throughout the course of care to provide evidence for beginning treatment with single agent anti-PD-1 or combination therapy and the optimal number of doses needed to sustain a response while limiting toxicity.

Of course, a SMART design is not limited to providing robust evidence for treatment pathways in BRAF wild-type metastatic melanoma, but can help develop and test treatment pathways that lead to optimal outcomes in other melanomas, cancers, and diseases. We acknowledge our SMART proposal is inherently limited by heterogeneity in some of the treatment pathways, such as in the “Standard of Care” box in subgroup 5. In our melanoma example, this box could include diverse treatments such as chemotherapy, inhibitors of other molecular drivers such as imatinib for patients with KIT mutations, and other potentially effective immunotherapy agents. How the various treatments within this pathway affect overall outcomes remains unknown in our proposed design.

A SMART requires less overall participants and can be implemented and analyzed in a shorter period of time than executing several single stage, standard two-arm trial designs (34). However, a commitment to more participants at the initiation of the trial for a SMART is needed than for individual standard trials and logistics may be more complex in a SMART by re-randomizing participants at an intermediate time point (34). With current technology that can handle multi-site interim randomizations or the ability to randomize participants upfront to follow particular treatment pathways, the increased logistics should not outweigh the benefits of finding optimal immunotherapy treatment pathways from SMART designs.

The SMART design, even when powered on questions regarding the best initial treatment in a pathway or best strategy for responders or non-responders (i.e., question 1, 2 or 3 from the previous section), may be more beneficial than multiple traditional single stage designs. A SMART can conclusively answer one question with additional analyses to address questions concerning treatment pathways that may be relevant to clinical practice, such as how long to remain on immunotherapy. Furthermore, SMART designs can identify treatment interactions when treatments differ in the first and second stages (i.e. a SMART design that differs from that in Figure 2 by re-randomizing to different treatments in the second stage as opposed to continuing or discontinuing initial treatment), and there may be delayed effects of initial treatments that modify the effects of follow-up treatments. Single stage trials cannot evaluate these interactions between first and second stage treatments dependent on intermediate outcomes.

More novel trial designs, including the SMART, may be needed to answer pertinent treatment questions and provide robust evidence for effective treatment regimens, especially in immuno-oncology research where novel combinations are frequently being proposed. A SMART can examine treatment sequences and combinations of immunotherapies and other drugs that lead to the longest overall survival with decreased toxicities. SMART designs may be able to verify potential optimal treatment pathways identified from dynamic mathematical modeling (38). SMARTs may require a paradigm shift for practicing physicians, pharmaceutical companies, and guidance agencies to begin to test and approve treatment regimens that may adapt within an individual along the course of care as opposed to testing and approving agents at particular snapshots in time and piecing these snapshots together trusting that these pieces tell the full story.

Acknowledgments

Memorial Sloan Kettering Cancer Center Core Grant (P30 CA008748; K. S. Panageas and M. A. Postow; P. I.: C. Thompson) and PCORI Award (ME-1507-31108; K.M. Kidwell).

References

- Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med*. 2010; 363:711–23. [PubMed: 20525992]
- Robert C, Thomas L, Bondarenko I, O'Day S, Weber J, Garbe C, et al. Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. *N Engl J Med*. 2011; 364:2517–26. [PubMed: 21639810]
- Lipson EJ, Drake CG. Ipilimumab: an anti-CTLA-4 antibody for metastatic melanoma. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2011; 17:6958–62.
- Weber J. Ipilimumab: controversies in its development, utility and autoimmune adverse events. *Cancer Immunol Immunother CII*. 2009; 58:823–30. [PubMed: 19198837]
- Blansfield JA, Beck KE, Tran K, Yang JC, Hughes MS, Kammula US, et al. Cytotoxic T-lymphocyte-associated antigen-4 blockage can induce autoimmune hypophysitis in patients with metastatic melanoma and renal cancer. *J Immunother Hagerstown Md* 1997. 2005; 28:593–8.
- Attia P, Phan GQ, Maker AV, Robinson MR, Quezado MM, Yang JC, et al. Autoimmunity correlates with tumor regression in patients with metastatic melanoma treated with anticytotoxic T-lymphocyte antigen-4. *J Clin Oncol Off J Am Soc Clin Oncol*. 2005; 23:6043–53.
- Robert C, Schachter J, Long GV, Arance A, Grob JJ, Mortier L, et al. Pembrolizumab versus Ipilimumab in Advanced Melanoma. *N Engl J Med*. 2015; 372:2521–32. [PubMed: 25891173]
- Topalian SL, Sznol M, McDermott DF, Kluger HM, Carvajal RD, Sharfman WH, et al. Survival, Durable Tumor Remission, and Long-Term Safety in Patients With Advanced Melanoma Receiving Nivolumab. *J Clin Oncol*. 2014; 32:1020–30. [PubMed: 24590637]
- Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Cs szí T, Fülöp A, et al. Pembrolizumab versus Chemotherapy for PD-L1–Positive Non-Small-Cell Lung Cancer. *N Engl J Med*. 2016; 375:1823–33. [PubMed: 27718847]
- Rosenberg JE, Hoffman-Censits J, Powles T, van der Heijden MS, Balar AV, Necchi A, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *The Lancet*. 2016; 387:1909–20.
- Drake CG, Lipson EJ, Brahmer JR. Breathing new life into immunotherapy: review of melanoma, lung and kidney cancer. *Nat Rev Clin Oncol*. 2013; 11:24–37. [PubMed: 24247168]
- Murphy SA, van der Laan MJ, Robins JM. CPPRG. Marginal Mean Models for Dynamic Regimes. *J Am Stat Assoc*. 2001; 96:1410–23. [PubMed: 20019887]

13. Nahum-Shani I, Qian M, Almirall D, Pelham WE, Gnagy B, Fabiano GA, et al. Experimental design and primary data analysis methods for comparing adaptive interventions. *Psychol Methods*. 2012; 17:457–77. [PubMed: 23025433]
14. Almirall D, Nahum-Shani I, Sherwood NE, Murphy SA. Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research. *Transl Behav Med*. 2014; 4:260–74. [PubMed: 25264466]
15. Press OW, Li H, Schöder H, Straus DJ, Moskowitz CH, LeBlanc M, et al. US Intergroup Trial of Response-Adapted Therapy for Stage III to IV Hodgkin Lymphoma Using Early Interim Fluorodeoxyglucose-Positron Emission Tomography Imaging: Southwest Oncology Group S0816. *J Clin Oncol*. 2016; 34:2020–7. [PubMed: 27069074]
16. Lavori PW, Dawson R. A design for testing clinical strategies: biased adaptive within-subject randomization. *J R Stat Soc Ser A Stat Soc*. 2000; 163:29–38.
17. Murphy SA. An experimental design for the development of adaptive treatment strategies. *Stat Med*. 2005; 24:1455–81. [PubMed: 15586395]
18. Thall PF, Wooten LH, Logothetis CJ, Millikan RE, Tannir NM. Bayesian and frequentist two-stage treatment strategies based on sequential failure times subject to interval censoring. *Stat Med*. 2007; 26:4687–702. [PubMed: 17427204]
19. Kidwell KM. SMART designs in cancer research: Past, present, and future. *Clin Trials Lond Engl*. 2014; 11:445–56.
20. Kelleher SA, Dorfman CS, Plumb Vilardaga JC, Majestic C, Winger J, Gandhi V, et al. Optimizing delivery of a behavioral pain intervention in cancer patients using a sequential multiple assignment randomized trial SMART. *Contemp Clin Trials*. 2017; 57:51–7. [PubMed: 28408335]
21. Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy SA. A “SMART” Design for Building Individualized Treatment Sequences. *Annu Rev Clin Psychol*. 2012; 8:21–48. [PubMed: 22224838]
22. Penn State Methodology Center. [cited 2016 Aug 12] Projects Using SMARTS [Internet]. Available from: <https://methodology.psu.edu/ra/adap-inter/projects>
23. Seymour L, Bogaerts J, Perrone A, Ford R, Schwartz LH, Mandrekar S, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol*. 2017; 18:e143–52. [PubMed: 28271869]
24. Dmitrienko, A., editor. *Mult Test Probl Pharm Stat*. Boca Raton, Fla: CRC Press; 2010. Multiplicity Problems in Clinical Trials: A Regulatory Perspective.
25. Li Z, Murphy SA. Sample size formulae for two-stage random trials with survival outcomes. *Biometrika*. 2011; 98:503–18. [PubMed: 22363091]
26. Feng W, Wahed AS. Supremum weighted log-rank test and sample size for comparing two-stage adaptive treatment strategies. *Biometrika*. 2008; 95:695–707.
27. Feng W, Wahed AS. Sample size for two-stage studies with maintenance therapy. *Stat Med*. 2009; 28:2028–41. [PubMed: 19382105]
28. Lunceford JK, Davidian M, Tsiatis AA. Estimation of survival distributions of treatment policies in two-stage randomization designs in clinical trials. *Biometrics*. 2002; 58:48–57. [PubMed: 11890326]
29. Guo X, Tsiatis AA. A weighted risk set estimator for survival distributions in two-stage randomization designs with censored survival data. *Int J Biostat*. 2005; 1:1–15.
30. Kidwell KM, Wahed AS. Weighted Log-Rank Statistic to Compare Shared-Path Adaptive Treatment Strategies. *Biostatistics*. 2013; 14:299–312. [PubMed: 23178734]
31. Tang X, Wahed AS. Comparison of treatment regimes with adjustment for auxiliary variables. *J Appl Stat*. 2011; 38:2925–38.
32. Lokhnygina Y, Helterbrand JD. Cox regression methods for two-stage randomization designs. *Biometrics*. 2007; 63:422–8. [PubMed: 17425633]
33. Tang X, Melguizo M. DTR: An R Package for Estimation and Comparison of Survival Outcomes of Dynamic Treatment Regimes. *J Stat Softw* [Internet]. 2015; 65 [cited 2017 Apr 3] Available from: <http://www.jstatsoft.org/v65/i07/>.
34. Wolbers M, Helterbrand JD. Two-stage Randomization Designs in Drug Development. *Stat Med*. 2008; 27:4161–74. [PubMed: 18570274]

35. Goldberg Y, Kosorok MR. Q-learning with censored data. *Ann Stat.* 2012; 40:529–60. [PubMed: 22754029]
36. Nahum-Shani I, Qian M, Almirall D, Pelham WE, Gnagy B, Fabiano GA, et al. Q-learning: A data analysis method for constructing adaptive interventions. *Psychol Methods.* 2012; 17:478–94. [PubMed: 23025434]
37. Watkins CJCH, Dayan P. Q-learning. *Mach Learn.* 1992; 8:279–92.
38. Agur Z, Halevi-Tobias K, Kogan Y, Shlagman O. Employing dynamical computational models for personalizing cancer immunotherapy. *Expert Opin Biol Ther.* 2016; 16:1373–85. [PubMed: 27564141]

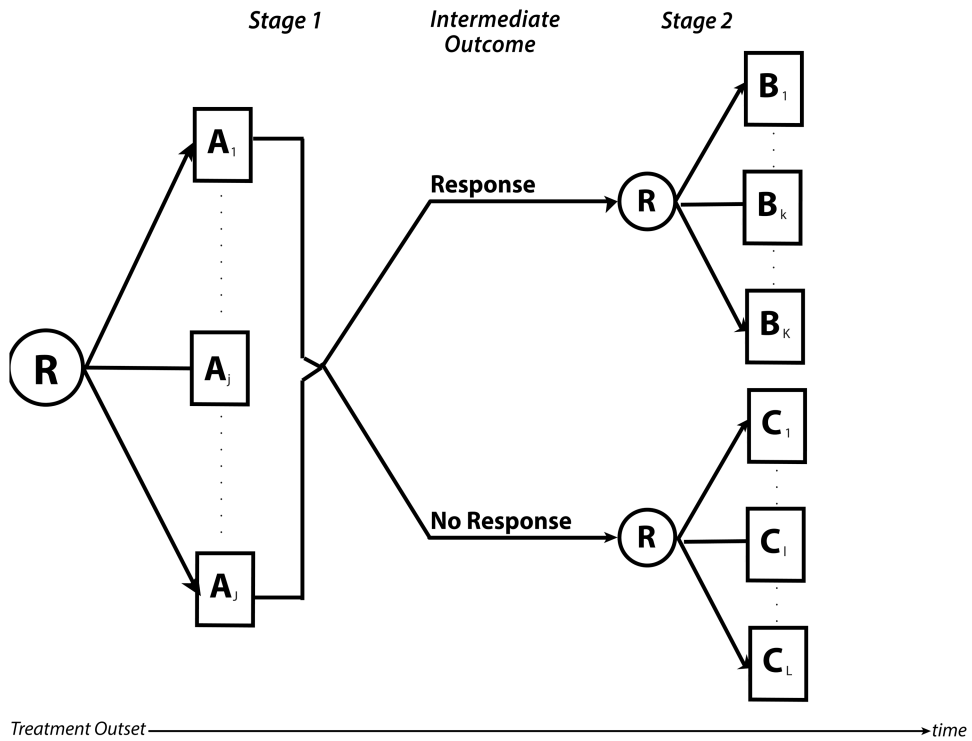


Figure 1.

A generic two-stage sequential multiple assignment randomized trial (SMART) design where participants are randomized between any number of treatments A_1 to A_J . Response is measured at some intermediate time point or over time such that responders are re-randomized in the second stage between any number of treatments B_1 to B_K and non-responders are re-randomized between any number of treatments C_1 to C_L . The same participants are followed throughout the trial.

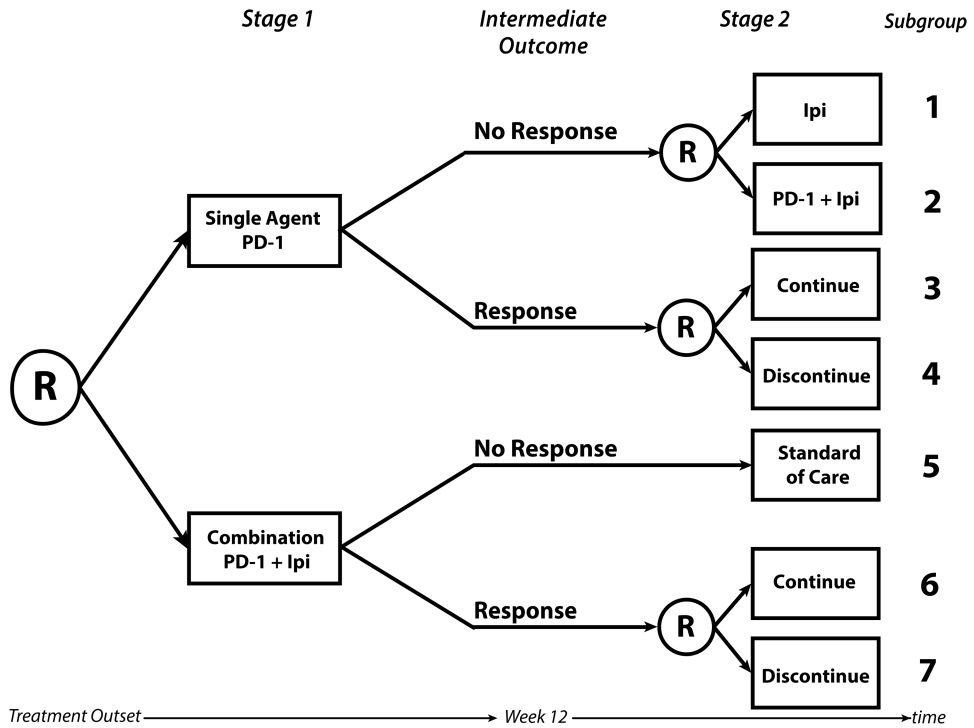


Figure 2.

A hypothetical two-stage sequential multiple assignment randomized trial (SMART) design in the setting of BRAF wild-type metastatic melanoma. Participants are initially randomized to either single agent anti-programmed cell death protein 1 (PD-1) therapy or to a combination of anti-PD-1 therapy plus ipilimumab (Ipi). Note that ipilimumab may be replaced by any novel combination agent. After four doses or approximately 12 weeks, response is measured. Those who did not respond to the single agent are re-randomized to receive ipilimumab or the combination. Those who did respond to single agent anti-PD-1 are re-randomized to continue the single agent or discontinue therapy. Those who did not respond initially to the combination receive standard of care and those who did respond are re-randomized to continue the combination or discontinue therapy. Subgroups 1-7 denote the subgroups that any one participant may fall into. There are 6 embedded treatment pathways in this SMART, each one made up of 2 subgroups: {1,3}, {1,4}, {2,3}, {2,4}, {5,6}, and {5,7}.

Table 1

Comparison of the sample sizes needed for three single trials versus one SMART design. the trials in approach 1 would require a total of 1142 participants versus 570 total participants from one SMART.

Approach	Trial	Hypothesis	Sample Size
1	Single agent anti-PD-1 versus combination	Log-rank test assuming 68% versus 80% survival at 1 year	178
	For responders: Continue initial treatment versus discontinue	Log-rank test assuming 80% versus 68% survival at 1 year	178
	For non-responders to anti-PD-1: Single agent versus combination	Log-rank test assuming 68% versus 74% survival at 1 year	786
2	SMART Design in Figure 2	Log-rank test assuming largest difference in 6 treatment pathways is 80% versus 68% survival at 1 year, 40% response to initial therapy	570

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript