

Statistical Challenges in the Design of Late-Stage Cancer Immunotherapy Studies

Rosemarie Mick¹ and Tai-Tsang Chen^{2,3}

Abstract

The past several years have witnessed a revival of interest in cancer immunology and immunotherapy owing to striking immunologic and clinical responses to immune-directed anti-cancer therapies and leading to the selection of "Cancer Immunotherapy" as the 2013 Breakthrough of the Year by *Science*. But statistical challenges exist at all phases of clinical development. In phase III trials of immunotherapies, survival curves have been shown to demonstrate delayed clinical effects, as well as long-term survival. These unique survival

kinetics could lead to loss of statistical power and prolongation of study duration. Statistical assumptions that form the foundations for conventional statistical inference in the design and analysis of phase III trials, such as exponential survival and proportional hazards, require careful considerations. In this article, we describe how the unique characteristics of patient response to cancer immunotherapies will impact our strategies on statistical design and analysis in late-stage drug development. *Cancer Immunol Res*; 3(12); 1292–8. ©2015 AACR.

Introduction

The past several years have witnessed a revival of interest in cancer immunology and immunotherapy owing to striking immunologic and clinical responses to immune-directed anti-cancer therapies (1, 2). Some of the most promising therapies include checkpoint inhibitors, which block signaling pathways, cancer vaccines, and adoptive T-cell strategies that exploit chimeric antigen receptor (CAR)-engineered T cells. The immune checkpoint inhibitors have demonstrated the greatest success to date. Cytotoxic T lymphocyte-associated antigen 4 (CTLA-4), also known as CD152, blockade was shown to increase anti-tumor immunity in mice in selected tumor systems (3, 4), which facilitated the clinical development of anti-CTLA-4 inhibitors. Ipilimumab, a fully human immunoglobulin G1 (IgG1) monoclonal antibody targeting CTLA-4, was approved by the FDA in 2011 for the treatment of patients with previously treated and untreated unresectable metastatic melanoma based on significant improvement in overall survival (OS) in two randomized phase III clinical trials (5, 6). During the same time period, Sipuleucel-T, a type of therapeutic cancer vaccine and an autologous cellular immunotherapy, also demonstrated efficacy and was approved for treatment in metastatic hormone-refractory prostate cancer (7).

The clinical success of anti-CTLA-4 blockade and the subsequent identification of other candidate checkpoints signaled a new era of cancer immunotherapy research (8–10). The hallmark success of anti-CTLA-4 blockade along with promising early

clinical results of checkpoint inhibitors, such as programmed death-1 (PD-1; refs. 11, 12) and the PD-1 ligand, PD-L1 (13), led to the selection of "Cancer Immunotherapy" as the 2013 Breakthrough of the Year by *Science* (14). Nivolumab, a fully human IgG4 (Immunoglobulin G subclass 4) monoclonal antibody to PD-1, showed promising efficacy in multiple tumor types (15–18). The significant survival benefit led to decisions by data monitoring committees to discontinue ahead of schedule four randomized phase III nivolumab clinical trials in metastatic melanoma (19), squamous and nonsquamous non-small cell lung cancer (NSCLC; refs. 20, 21), as well as renal cell carcinoma (22). Pembrolizumab, a humanized IgG4 PD-1-blocking antibody, also has promising efficacy in melanoma (23, 24) and NSCLC (25). Both nivolumab and pembrolizumab were approved for metastatic melanoma and NSCLC, with a specific indication for pembrolizumab for treatment of NSCLC in patients whose tumors show PD-L1 expression. In addition, this class of agents shows evidence of efficacy in other tumor types, such as hematologic malignancies (26) and metastatic bladder cancer (27). A new wave of studies is now under way to evaluate various combinations of checkpoint inhibitors. Complementary activity has been shown in both previously treated and untreated metastatic melanoma patients for the combination of two immune checkpoint inhibitors, ipilimumab and nivolumab (28, 29), which led to the approval of this combination to treat patients with BRAF V600 wild-type unresectable or metastatic melanoma.

The field of adoptive T-cell therapies is also advancing at a rapid pace. In the past few years, durable responses have been observed in early phase trials of adoptive T-cell therapies that exploit chimeric antigen receptor (CAR)-engineered T cells directed to malignant B cells in chronic and acute leukemia (30, 31). The adoptive transfer of tumor-infiltrating lymphocytes (32, 33), as well as several immune costimulatory agents targeting CD137 and CD40, also shows promising activity and will serve as potential candidates in combination with other immunotherapies (34, 35).

The advancement of cancer immunology does not come without challenges. The knowledge we have accumulated in this field

¹Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania. ²Department of Global Biometric Sciences, Bristol-Myers Squibb, Princeton, New Jersey. ³Department of Biostatistics, Columbia University, New York, New York.

Corresponding Author: T.-T. Chen, Bristol-Myers Squibb, Route 206 & Province Line Road, J44-05, Princeton, NJ, 08540. Phone: 609-252-5781; E-mail: tai-tsang.chen@bms.com

doi: 10.1158/2326-6066.CIR-15-0260

©2015 American Association for Cancer Research.

clearly indicates the need for customized clinical trial designs. Late-stage clinical trials with time-to-event endpoints, such as OS or progression-free survival (PFS), are usually designed with a set of assumptions that reflects the "survival kinetics" of the treatments under investigation (36). The most commonly adopted assumptions include exponential survival and proportional hazards. Exponential distribution is a reasonable statistical model for time-to-event endpoints in the trial design. It implies that the probability of the event of interest decays at a rate (i.e., hazard rate), which remains the same during the observation period, and all individuals will eventually become events. The treatment difference between time-to-event curves is assumed to remain constant and is quantified by the ratio of the hazard rates (i.e., HR). Once the magnitude of the treatment difference that is clinically meaningful and worth detecting is defined, the clinical trial is designed to obtain the number of events required based on these assumptions, along with accrual information and the control of desired false-positive rate (i.e., type I error rate) and false-negative rate (i.e., type II error rate). Due to the size and duration of a randomized phase III clinical trial, interim analyses are usually implemented to ensure that trial participants are not exposed to an unsafe or ineffective treatment. These interim analyses offer the opportunities to either terminate the studies early for unexpected large treatment effects or lack of efficacy that suggests a positive result would be extremely unlikely when the study is complete (i.e., futility; refs. 37–40).

It is now well understood that chemotherapy and immunotherapy do not share the same survival kinetics. Patients treated with chemotherapy show early antitumor effects. When compared with a control arm, improved OS or PFS for the experimental arm is demonstrated by an early separation between the survival curves. In contrast with chemotherapy's immediate and direct cytotoxic effect on a tumor, immunotherapy stimulates the patient's immune system to mount an antitumor response. The time needed to mount this response inserts the delay in clinical effect observed consistently across many trials. Conventional statistical approaches to trial design and statistical analysis are being re-evaluated in the setting of immunotherapy due to the differences in these key attributes (41). The focus of this article is to examine the impact of these unique characteristics of cancer immunotherapies on statistical trial design and the methods used to analyze late-stage randomized phase III clinical trials.

Unique Characteristics Exhibited by Cancer Immunotherapy

Historically, the development of cancer agents has progressed from the early-phase studies assessing safety to later-phase studies evaluating efficacy (42). These cancer agents are usually classified by their mechanisms of action as either cytotoxic or cytostatic. The key attributes exhibited by conventional cytotoxic agents include rapid antitumor activity and undesired side effects caused by the lack of selectivity between normal and cancerous cells. On the other hand, cytostatic agents often show minimal or less severe toxicity while inhibiting or suppressing cellular growth and division, with little or no tumor shrinkage. Although the selection of development strategies and endpoints may vary depending on the mechanisms of action, the statistical framework for clinical development largely stays intact. In particular, for late-stage randomized phase III clinical trials, a constant HR is usually assumed to characterize the treatment effect.

When ipilimumab was in clinical development for metastatic melanoma, an observation made in one of the randomized phase II clinical trials led the clinical trial researchers to challenge the *status quo* of the study designs for cancer immunotherapy. In this phase II dose-ranging study (43), patients with previously treated melanoma were randomized to receive 0.3 mg/kg, 3 mg/kg, or 10 mg/kg of ipilimumab. At the time of final analysis, the study showed that the best overall response rate (BORR) only ranged from 0% to 11% between the three arms. However, the OS demonstrated two important phenomena rarely observed in previous cancer trials, that is, an 8-month delayed clinical effect and dose-response effect on long-term survival. This observation resulted in the change of primary endpoint from BORR to OS in the pivotal metastatic melanoma phase III study (5) before the study had been unblinded. The final analysis of the study confirmed the unique survival kinetics with a 4-month delay and long-term survival of approximately 20%.

The overlapping of the survival curves in the first 4 months indicated a delay of clinical effects for the treatment arms that included ipilimumab as a single agent or in combination with gp100 (Fig. 1). The delayed separation in survival curves clearly implied that the probability of survival did not decay at a constant rate between treatment arms, so the assumption of proportional hazards was likely violated. These key attributes also existed in other time-to-event endpoints. Postow and colleagues (29) reported on a trial that compared ipilimumab as a single agent with ipilimumab combined with nivolumab for newly diagnosed advanced melanoma. A delay in clinical effect of 3 months in PFS was observed for the combination arm. And in a study that compared PFS for ipilimumab versus nivolumab versus ipilimumab plus nivolumab, the PFS curves for all three arms overlapped during the first 3 months (28).

Durable responses and long-term survivors have been noted for ipilimumab (44), the checkpoint inhibitor with the most mature long-term data. A milestone survival analysis of treatment-naïve advanced melanoma demonstrated a 5-year survival rate of 18% (95% confidence interval, 13.6–23.4%) for patients treated with ipilimumab plus dacarbazine (45). A pooled OS analysis comprised of 1,861 melanoma patients from 10 prospective, and two retrospective studies showed that the OS curve began to plateau at 3 years with a 22% survival rate, with follow-up of up to 10 years in some patients (46). Similar phenomena were also observed in recurrence-free survival and led to the approval of ipilimumab for patients with stage III melanoma who were at high risk of recurrence following complete surgical resection (47).

As our experience with checkpoint blockade inhibitors and other immunologic approaches such as adoptive T cells and vaccines matures, it is becoming evident that conventional statistical approaches to trial design and statistical analysis must be questioned and new methods must be tailored to fit the features of cancer immunotherapy. The following sections of this article describe in greater detail how the unusual characteristics of cancer immunotherapies affect our strategies for statistical design and analysis.

Impact of Survival Kinetics on Study Design

The long-term survival and delayed clinical effect observed in multiple cancer immunotherapy phase III clinical trials with time-to-event endpoints (5, 6, 15–25) pose unique statistical challenges in study design. Four basic models summarize how the

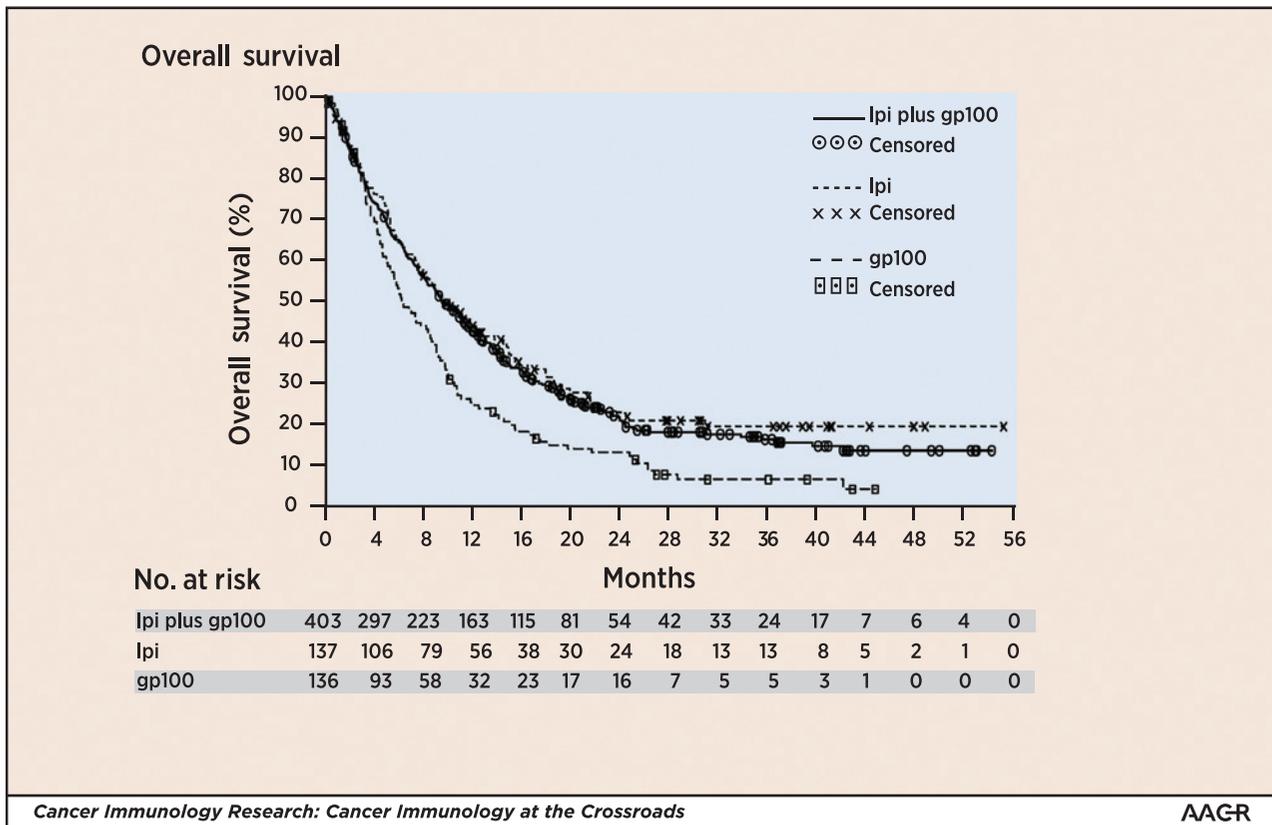


Figure 1. OS in metastatic melanoma. The overlap of the survival curves in the first 4 months indicates a delay of clinical effects for the treatment arms that included ipilimumab. Redrawn from ref. 5 and reprinted with permission of the Massachusetts Medical Society.

long-term survival and delayed clinical effects affect the study design with time-to-event endpoints (41). We briefly summarize these four models below.

Proportional Hazards Model (PHM): Conventionally, studies with time-to-event endpoints are designed with an exponential distribution assumption in which we assume a constant hazard rate for each group. This implies that we assume anything that affects the hazards does so by the same ratio during the observation period (i.e., proportional hazards), and the relative clinical effect of the experimental group over the control group is observed from the beginning. In addition, all patients enrolled in the study are subject to the event of interest; therefore, the survival curves will eventually drop down to zero survival probability (Fig. 2A).

Proportional Hazards Cure Rate Model (PHCRM): When a proportion of patients are expected to remain alive or free of disease even after long follow-ups, the entire population is assumed to consist of patients who are either susceptible or nonsusceptible to the event of interest. PHCRM assumes proportional hazards between two treatment groups among susceptible patients, while the long-term survival rates are also assumed proportional between two groups among nonsusceptible patients, such that the risk ratio of the entire population is consistent within each of these two subpopulations. (Fig. 2B).

Nonproportional Hazards Model (NPHM): Nonproportional hazards imply the HR changes over time during the observation period. The simplest NPHM is under the piecewise exponential

distribution assumption with two intervals, in which the HR within each interval remains constant. A special case of NPHM is the PHM when hazards in both groups remain constant and the HRs are identical in both intervals (Fig. 2C).

Nonproportional Hazards Cure Rate Model (NPHCRM): Both long-term survival and delayed clinical effect are incorporated in the NPHCRM (Fig. 2D).

The impact of the survival kinetics introduced by cancer immunotherapy has been simulated in a hypothetical two-arm randomized phase III study with 3-month delay in clinical effect and a 10% long-term survival rate in the control arm (41). If the separation in the Kaplan–Meier curves was not accounted for in the study design or it occurred at a later time, the study would lose power, whereas the length of the trial would increase with more long-term survivors. We adopted the same randomized clinical trial design used in the simulation study (41) to illustrate how different magnitudes of delayed clinical and long-term survival effects would impact the statistical power and study duration. The study was designed based on the conventional exponential distribution assumption (i.e., the PHM), with 1:1 randomization ratio. A total of 512 events were required to detect an HR of 0.75 between two treatment arms in a log-rank test, with an experiment-wise two-sided type I error rate of 5% and power of 90%. The accrual duration for 680 randomized patients was assumed to be 34 months, which translated to a monthly accrual of 20 patients. The study duration was estimated to be approximately 48 months after the first patient was randomized. Note that the

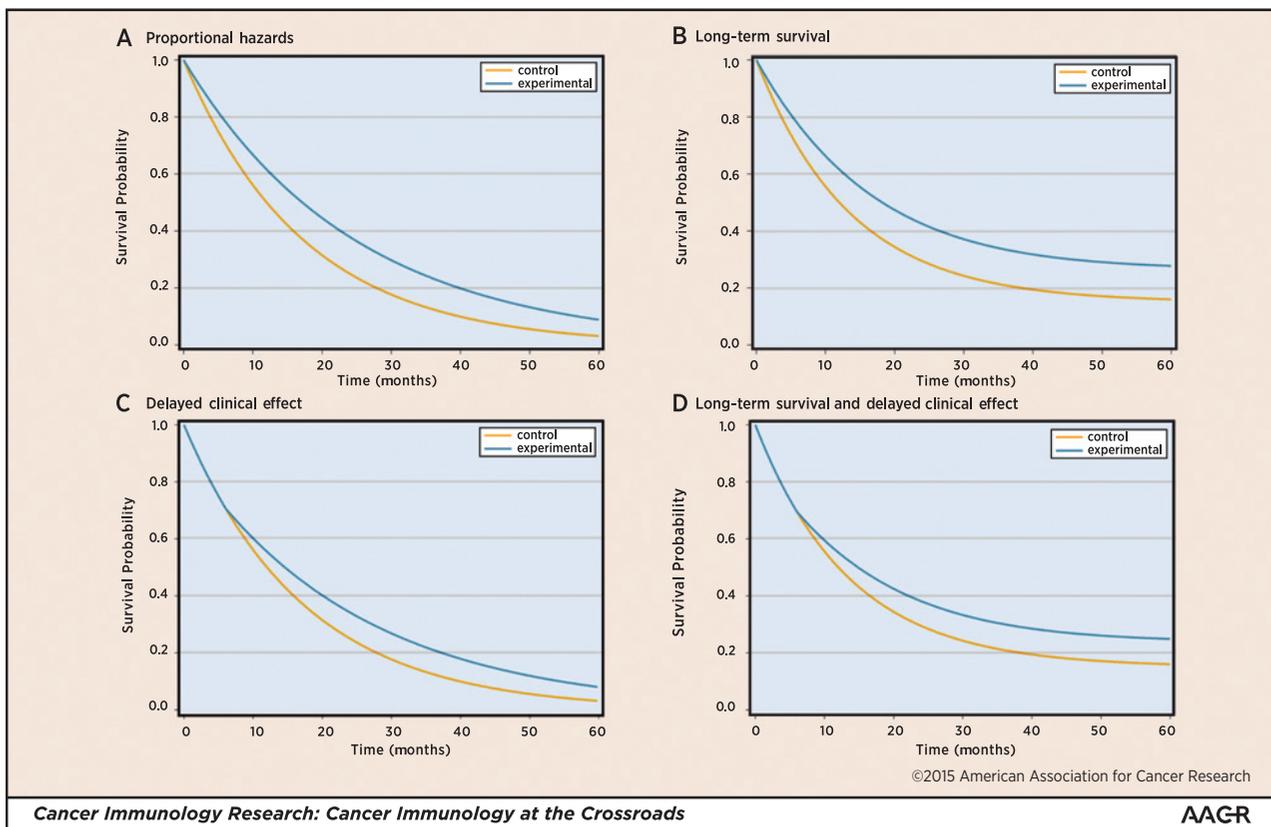


Figure 2.

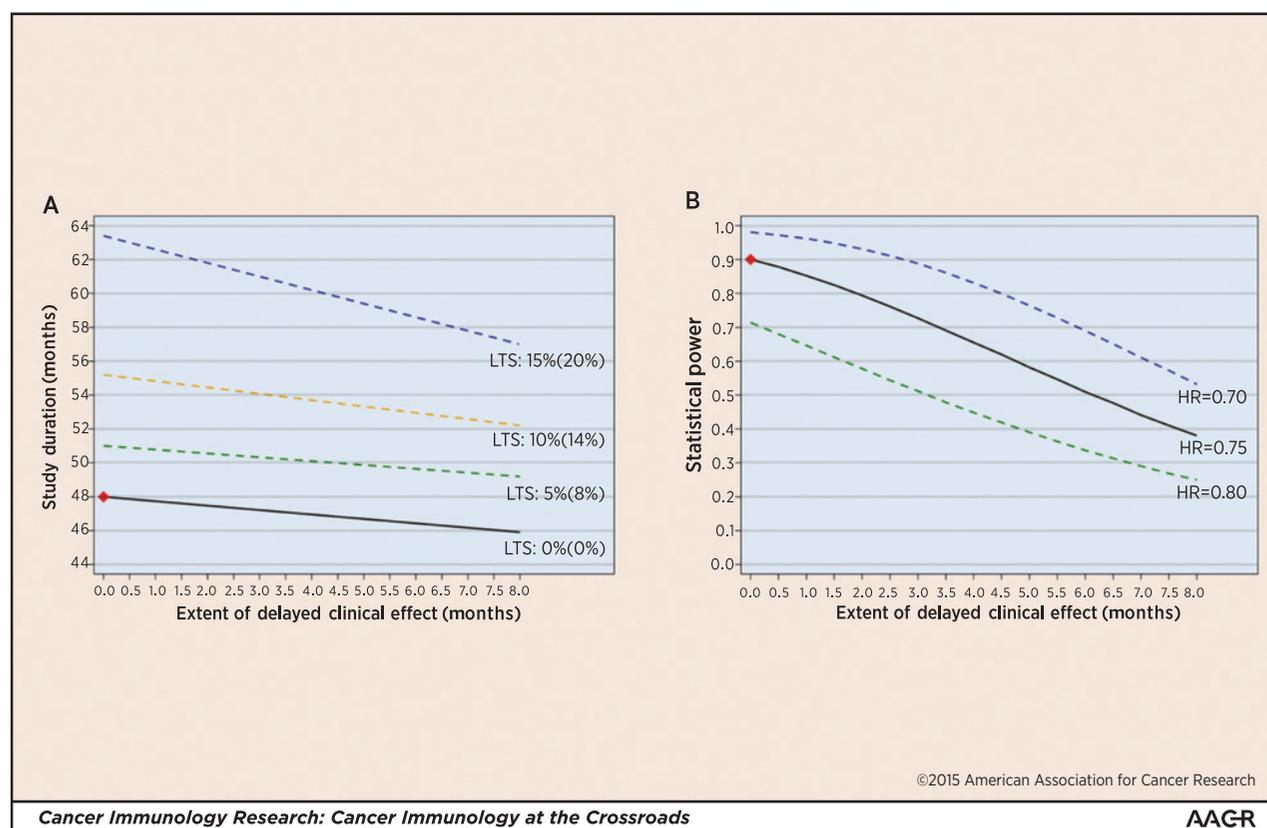
Graphical presentation of Kaplan-Meier survival curves with various combinations of long-term survival and delayed clinical effect. The blue curve represents the cancer immunotherapy-containing regimen, and the orange curve represents a control treatment, respectively. A, PHM with an exponential assumption. B, PHCRM with long-term survival. C, NPHM with delayed clinical effect. D, NPHCRM with both long-term survival and delayed clinical effects.

proportional long-term survival rates were implemented in order to maintain the risk ratio at 0.75 so the impact on the trial duration could be independently assessed without inducing the loss of statistical power simultaneously.

When a study is designed assuming exponential survival, it is assumed that all patients are susceptible to the event of interest during the course of the study. A mixture of patients who are susceptible and nonsusceptible (long-term survivors) to the event of interest would lead to the prolongation of the study duration because the events could only be observed in the subset of the susceptible population. The aforementioned design implied that 75% (512 of 680) of all randomized patients were expected to become events. If the long-term survival rate of the entire study exceeded 25% of the study population, then the study would not be completed with the prespecified number of events within a reasonable monitoring time window. The total trial duration lengthened from 3 months to as long as 16 months under PHCRM when the proportion of long-term survival survivors in the control group increased from 5% to 15% (Fig. 3A). Note that the total study duration decreased slightly with an increasing length of delayed clinical effect for a given control arm long-term survival effect. This was due to the higher number of events observed in the treatment arm prior to the separation of the curves. Consequently, this led to shorter trial duration in studies designed with time-to-event endpoints (i.e., event-driven studies).

On the other hand, the presence of delayed clinical effect would lead to a prolongation of the trial duration. When the treatment effects were assumed to be identical during delay (i.e., $HR = 1$), as shown in Fig. 3B (black curve), the statistical power decreased from 90% (PHM) per study design (red diamonds in Fig. 3) to less than 40%, with the delay of the clinical effect increasing from 0 to 8 months, when the post-delay separation HR was fixed at 0.75. Note that the HR post-delay separation also affected the statistical power. The statistical power would improve or worsen depending on the magnitude of the observed treatment effect post separation relative to what had been defined in the study design. When the HR improved from 0.75 to 0.70, the statistical power increased to the range of 53% to 98% (Fig. 3B, blue curve). With an HR of 0.80 post-delay separation, the statistical power fell into the range of 25% and 71% (Fig. 3B, orange curve). The consequences of discounting these unique survival kinetics in study design were examined (48) for a multicenter, randomized, double-blind two-arm phase III study conducted in patients with treatment-naïve stage III or IV melanoma receiving ipilimumab plus dacarbazine versus placebo plus dacarbazine (6). The trial duration was prolonged for an additional 2 years due to decreased hazards in the second half of the study.

Another important aspect of the study design due to the presence of delayed clinical effect concerns the necessity and the timing of interim analyses. Due to the life-threatening nature of certain diseases, such as cancer, interim analyses have become a

**Figure 3.**

The presence of delayed clinical effect and long-term survival would lead to a loss of statistical power and prolongation of trial duration. The red diamonds in both panels represent the study design with no delayed clinical effect and 90% power when HR = 0.75 and median OS for the control arm = 12 months. The trial duration was estimated to be 48 months. A, under the PHCRM, the prolongation of the trial duration ranged from 3 months to 16 months when the proportion of long-term survivors (LTS) in the control group increased from 5% to 15%, with the numbers in the parentheses representing the proportions of LTS for the entire study. The slight decrease in trial duration was caused by the increasing number of events in the experimental group during the delay. B, when HR after separation remained at 0.75, the statistical power decreased from 90% to 38% with corresponding increase in duration of delayed clinical effect from 0 to 8 months. An observed relative treatment difference of HR = 0.7 would lead to an absolute increase in statistical power of 8% to 15% (blue dashed curve) for various durations of delayed clinical effect. A worse observed treatment effect of HR = 0.8 is shown as a green dashed curve with statistical power ranging from 25% to 71%.

staple in the majority of randomized phase III clinical trials, with the possibility of terminating the trial early due to either unexpectedly large treatment effects, lack of treatment effect, or excessive toxicity. However, the inclusion of interim analyses needs careful considerations, because they may lead to an unnecessary waste of resources and false conclusions. This is crucial in studies with cancer immunotherapy in which the delayed clinical effect is expected. If the interim analyses are placed at times when the true treatment effect has not been demonstrated, they are likely to lead to false-negative conclusions. The study design example provided earlier indicates that a 4-fold decrease in true-positive rate (i.e., statistical power) and an 8-fold increase in false-negative rate (i.e., type II error rate) would be seen at the interim analysis with 50% information fraction if the delayed clinical effect was not considered in the study design (41).

Conclusions and Future Directions

The recent introduction of cancer immunotherapies has heralded an era of significant change in the treatment of cancers. The indirect mechanisms of action of cancer immunotherapy stimu-

late the patient's own immune system to fight against cancers by targeting antigens expressed on cancer cells and kick-starting antitumor effects. Some of the approaches used in cancer immunotherapy, e.g., immune checkpoint inhibitors, hold the promise of long-term survival, but with more delayed clinical effects. The introduction of these new therapies warrants a re-examination of the appropriate methodology for study design.

Statistical models used in study designs typically assume that everybody in the study population is susceptible to the event of interest and will eventually become an event. The emerging evidence and increasing knowledge of new therapies have led clinical trial researchers to challenge the *status quo* of study designs in cancer research. Due to the different mechanisms of action induced by immunotherapy compared with more conventional chemotherapies or targeted therapies, these newly introduced immuno-oncologic agents have delayed clinical benefit. This is reflected in the delayed separation of Kaplan–Meier curves in event-driven studies. In addition, a proportion of patients may sustain remission and do not require therapy within a reasonable monitoring time window. Failure to account for these phenomena in the study design could lead to the erroneous discarding of

new therapies that could potentially benefit patients. It is crucial to tailor study designs to the characteristics exhibited by cancer immunotherapy.

The impact on statistical power and trial duration in the presence of various combinations of long-term survival and delayed clinical effects needs thorough discussion when a study is designed based on the conventional assumption of exponential distribution. In the clinical trial example described here, we have observed an increase of as much as 16 months in the trial duration with 20% long-term survivors in the entire study. Furthermore, the delayed clinical effect led to a 45% to 50% absolute loss of statistical power as the duration of delay lengthened to 8 months. Some of the design scenarios that were not considered here included those with nonproportional long-term survival rates. If a PHCM study was designed and long-term survival was only observed in the experimental arm, this design would yield an overpowered study. On the other hand, if the long-term survival existed in the control group only, or the magnitude of the proportional long-term survival was weaker than that of the proportional hazards in the susceptible population, the study would be underpowered. Although the example assumed there was no treatment difference (i.e., HR = 1) prior to separation, relative treatment effect of other magnitudes could be implemented in the study design as well.

Therefore, the nonproportional hazards assumption needs to be considered in the sizing of the study to reduce the risk of false-negative conclusions when a delayed clinical effect is expected. If a proportion of long-term survivors are expected, sufficient patients need to be randomized to ensure the total number of events is reachable. In the meantime, the sample size also needs to be small enough to ensure sufficient follow-up to capture the long-term survival benefit. The timing of the superiority or futility interim analysis also requires careful consideration in order to control the true-positive rate and false-negative rate, respectively. A delicate balance between the study assumptions, the sample size, the trial duration, and timing of the interim analysis, when applicable, becomes essential in the success of clinical cancer research. Even if these unique survival kinetics are built into the study design, it is recommended that a given trial design be evaluated with a spectrum of duration of delays and long-term survival effects in order to quantify the potential risk of misspecifying the magnitudes of these attributes.

References

- Mellman I, Coukos G, Dranoff G. Cancer immunotherapy comes of age. *Nature* 2011;480:480–9.
- Chen D, Mellman I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* 2013;39:1–10.
- Leach DR, Krummel MA, Allison JP. Enhancement of antitumor immunity by CTLA-4 blockade. *Science* 1996;271:1734–6.
- Yang YF, Zou JP, Mu J, Wijesuriya R, Ono S, Walunas T, et al. Enhanced induction of antitumor T-cell responses by cytotoxic T lymphocyte-associated molecule-4 blockade: the effect is manifested only at the restricted tumor-bearing stages. *Cancer Res* 1997;57:4036–41.
- 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.
- 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.
- Kantoff PW, Higano CS, Shore ND, Berger ER, Small EJ, Penson DF, et al. Sipuleucel-T immunotherapy for castration-resistant prostate cancer. *N Engl J Med* 2010;363:411–22.
- Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nature* 2012;12:252–64.
- Sharma P, Allison JP. The future of immune checkpoint therapy. *Science* 2015;348:56–61.
- Postow MA, Callahan MK, Wolchok JD. Immune checkpoint blockade in cancer therapy. *J Clin Oncol* 2015;33:1974–82.
- Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, activity and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med* 2012;366:2443–54.
- Hamid O, Robert C, Daud A, Hodi FS, Hwu W-J, Kefford R, et al. Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. *N Engl J Med* 2013;369:134–44.
- Brahmer JR, Tykodi SS, Chow L, Hwu W-J, Topalian SL, Hwu P, et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N Engl J Med* 2012;366:2455–65.
- Couzin-Frankel J. Breakthrough of the year 2013. cancer immunotherapy. *Science* 2013;342:1432–3.
- Weber JS, D'Angelo SP, Minor D, Hodi FS, Gutzmer R, Neyns B, et al. Nivolumab versus chemotherapy in patients with advanced melanoma

Alternative clinical endpoints for cancer immunotherapy in late-stage drug development should also be considered in order to mitigate the challenge of accelerating the drug development process when the impact of cancer immunotherapy is derived from long-term follow-ups. Clinical endpoints, such as milestone survival, immune-related response, (immune-related) PFS, disease control rate, tumor growth rate, or quality of life, warrant a closer investigation in the clinical benefit assessment of cancer immunotherapy. Among these potential surrogate endpoints, milestone survival is also a survival endpoint but with cross-sectional assessment at a prespecified time point (48). The long-term survival phenomenon can be directly evaluated by comparing the milestone survival rates at given time points, such as 1-year or 2-year milestone OS rates, provided the minimum follow-up duration is sufficiently long to allow robust estimation of the milestone survival rates. Milestone survival can also serve as an intermediate endpoint in a study designed with OS as the primary endpoint.

When the unique survival kinetics, such as long-term survival or delayed clinical effect are present, statistical methods, such as weighted log-rank test (49) and milestone survival (48), are potentially more powerful than the most commonly used approaches for time-to-event analyses, such as nonparametric log-rank test and Cox regression analysis. In the setting of a mixture of patients, those who are cured and those who are not cured, one would need to consider a model tailored to this setting, such as mixture cure rate models (41, 50). Previous research has also provided a rich body of literature of statistical methodology that can be utilized to account for nonproportional hazards and/or long-term survival in the study design and analysis (51–55).

Recent advances in the understanding of tumor immunology have led to the development of new cancer immunotherapies. As new information emerges from ongoing clinical studies, it should be continuously incorporated into study designs and analyses in order to optimize the drug development process.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Published online December 7, 2015.

- who progressed after anti-CTLA-4 treatment (CheckMate 037): a randomised controlled open-label, phase 3 trial. *Lancet Oncol* 2015;16:375–84.
16. Rizvi NA, Mazières J, Planchard D, Stinchcombe TE, Dy GK, Antonia SJ, et al. Activity and safety of nivolumab, an anti-PD-1 immune checkpoint inhibitor, for patients with advanced, refractory squamous non-small-cell lung cancer (CheckMate 063): a phase 2, single-arm trial. *Lancet Oncol* 2015;16:257–65.
 17. Gettinger SN, Horn L, Gandhi L, Spigel DR, Antonia SJ, Rizvi NA, et al. Overall survival and long-term safety of nivolumab (anti-programmed death 1 antibody BMS-936558, ONO-4538) in patients with previously treated advanced non-small-cell lung cancer. *J Clin Oncol* 2015;33:2004–12.
 18. McDermott DF, Drake CG, Sznol M, Choueiri TK, Powderly JD, Smith DC, et al. Survival, durable response, and long-term safety in patients with previously treated advanced renal cell carcinoma receiving nivolumab. *J Clin Oncol* 2015;33:2013–20.
 19. Robert C, Long GV, Brady B, Dutriaux C, Maio M, Mortier L, et al. Nivolumab in previously untreated melanoma without BRAF mutation. *N Engl J Med* 2015;372:320–30.
 20. Brahmer J, Reckamp KL, Bass P, Crinò L, Eberhardt WEE, Poddubskaya E, et al. Nivolumab versus docetaxel in advanced squamous-cell non-small cell lung cancer. *N Engl J Med* 2015;373:123–35.
 21. Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small cell lung cancer. *N Engl J Med* 2015;373:1627–39.
 22. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med* 2015;373:1803–13.
 23. Robert C, Ribas A, Wolchok JD, Hodi FS, Hamid O, Kefford R, et al. Anti-programmed-death-receptor-1 treatment with pembrolizumab in ipilimumab-refractory advanced melanoma: a randomised dose-comparison cohort of a phase 1 trial. *Lancet* 2014;384:1109–17.
 24. Robert C, Schachter J, Long GV. Pembrolizumab versus ipilimumab in advanced melanoma. *N Engl J Med* 2015;372:2521–32.
 25. Garon EB, Rizvi NA, Hui R, Leighl N, Balmanoukian AS, Eder JP, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med* 2015;372:2018–28.
 26. Ansell SM, Lesokhin AM, Borrello I, Halwani A, Scott EC, Gutierrez M, et al. PD-1 blockade with nivolumab in relapsed and refractory Hodgkin's lymphoma. *N Engl J Med* 2015;372:311–9.
 27. Powles T, Eder JP, Fine GD, Braiteh FS, Loriot Y, Cruz C, et al. MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature* 2014;515:558–62.
 28. Larkin J, Chiarion-Sileni V, Gonzalez R, Grob JJ, Cowey CL, Lao CD, et al. Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. *N Engl J Med* 2015;373:23–34.
 29. Postow MA, Chesney J, Pavlick AC, Robert C, Grossmann K, McDermott D, et al. Nivolumab and ipilimumab versus ipilimumab in untreated melanoma. *N Engl J Med* 2015;372:2006–17.
 30. Porter DL, Levine BL, Kalos M, Bagg A, June CH. Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. *N Engl J Med* 2011;365:725–33.
 31. Grupp SA, Kalos M, Barrett D, Aplenc R, Porter DL, Rheingold SR, et al. Chimeric antigen receptor-modified T cells for acute lymphoid leukemia. *N Engl J Med* 2013;368:1509–18.
 32. Tran E, Turcotte S, Gros A, Robbins PF, Lu Y-C, Dudley ME, et al. Cancer immunotherapy based on mutation-specific CD4+ T cells in a patient with epithelial cancer. *Science* 2014;344:641–4.
 33. Feldman SA, Assadipour Y, Kriley I, Goff SL, Rosenberg SA. Adoptive cell therapy—tumor-infiltrating lymphocytes, T-cell receptors, and chimeric antigen receptors. *Semin Oncol* 2015;42:626–39.
 34. Yonezawa A, Dutt S, Chester C, Kim J, Kohrt HE. Boosting cancer immunotherapy with anti-CD137 antibody therapy. *Clin Cancer Res* 2015;21:313–20.
 35. Vonderheide RH, Glennie MJ. Agonistic CD40 antibodies and cancer therapy. *Clin Cancer Res* 2013;19:1035–43.
 36. Marubini E, Valsecchi MG. Analysing survival data from clinical trials and observational studies. New York, NY: John Wiley & Sons, Inc.; 1995.
 37. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1979;35:549–56.
 38. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. *Biometrika* 1977;64:191–99.
 39. Pampallona S, Tsiatis A, Kim KM. Interim monitoring of group sequential trials using spending functions for the type I and II error probabilities. *Drug Inf J* 2001;35:1113–21.
 40. Lachin JM. A review of methods for futility stopping based on conditional power. *Stat Med* 2005;24:2747–64.
 41. Chen TT. Statistical issues and challenges in immuno-oncology. *J Immunother Cancer* 2013;1:1–9.
 42. Friedman LM, Furberg CD, DeMets D. Fundamentals of clinical trials. 4th ed. New York, NY: Springer-Verlag; 2010.
 43. Wolchok JD, Neyns B, Linette G, Negrrier S, Lutzky J, Thomas L, et al. Ipilimumab monotherapy in patients with pretreated advanced melanoma: a randomised, double-blind, multicentre, phase 2, dose-ranging study. *Lancet Oncol* 2010;11:155–64.
 44. McDermott D, Lebbe C, Hodi FS, Maio M, Weber JS, Wolchok JD, et al. Durable benefit and the potential for long-term survival with immunotherapy in advanced melanoma. *Cancer Treat Rev* 2014;40:1056–64.
 45. Maio M, Grob J-J, Aamdal S, Bondarenko I, Robert C, Thomas L, et al. Five-year survival rates for treatment-naïve patients with advanced melanoma who received ipilimumab plus dacarbazine in a Phase III trial. *J Clin Oncol* 2015;33.
 46. Schadendorf D, Hodi FS, Robert C, Weber JS, Margolin K, Hamid O, et al. Pooled analysis of long-term survival data from phase II and phase III trials of ipilimumab in unresectable or metastatic melanoma. *J Clin Oncol* 2015;33:1889–94.
 47. Eggermont AMM, Chiarion-Sileni V, Grob JJ, Dummer R, Wolchok JD, Schmidt H, et al. Adjuvant ipilimumab versus placebo after complete resection of high-risk stage III melanoma (EORTC 18071): a randomised, double-blind, phase 3 trial. *Lancet Oncol* 2015;16:522–30.
 48. Chen TT. Milestone survival: a potential intermediate endpoint for immune checkpoint inhibitors. *J Nat Cancer Inst* 2015;107:djv156.
 49. Harrington DP, Fleming TR. A class of rank test procedures for censored survival data. *Biometrika* 1982;69:553–66.
 50. Othus M, Barlogie B, LeBlanc ML, Crowley JJ. Cure models as a useful statistical tool for analyzing survival. *Clin Cancer Res* 2012;18:3731–6.
 51. Lakatos E. Sample sizes based on the log-rank statistic in complex clinical trials. *Biometrics* 1988;44:229–41.
 52. Yang S, Prentice R. Semiparametric analysis of short-term and long-term hazard ratios with two-sample survival data. *Biometrika* 2005;92:1–17.
 53. Fine GD. Consequences of delayed treatment effects on analysis of time-to-event endpoints. *Drug Inf J* 2007;41:535–39.
 54. Hasegawa T. Sample size determination for the weighted log-rank test with the Fleming-Harrington class of weights in cancer vaccine studies. *Pharm Stat* 2014;13:128–35.
 55. Zhang D, Quan H. Power and sample size calculation for log-rank test with a time lag in treatment effect. *Stat Med* 2009;28:864–79.

Cancer Immunology Research

Statistical Challenges in the Design of Late-Stage Cancer Immunotherapy Studies

Rosemarie Mick and Tai-Tsang Chen

Cancer Immunol Res 2015;3:1292-1298.

Updated version Access the most recent version of this article at:
<http://cancerimmunolres.aacrjournals.org/content/3/12/1292>

Cited articles This article cites 52 articles, 11 of which you can access for free at:
<http://cancerimmunolres.aacrjournals.org/content/3/12/1292.full#ref-list-1>

Citing articles This article has been cited by 6 HighWire-hosted articles. Access the articles at:
<http://cancerimmunolres.aacrjournals.org/content/3/12/1292.full#related-urls>

E-mail alerts [Sign up to receive free email-alerts](#) related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, use this link <http://cancerimmunolres.aacrjournals.org/content/3/12/1292>. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.