For reprint orders, please contact: reprints@futuremedicine.com

End points and statistical considerations in immuno-oncology trials: impact on multiple myeloma

NCOLOG

Future

Antje Hoering*^{,1}, Brian Durie², Hongwei Wang¹ & John Crowley¹

Unlike conventional cancer treatment, immuno-oncology therapies are commonly associated with delayed clinical benefit and durable responses, as seen with immunooncology therapies for multiple myeloma (MM). Therefore, a longer-term approach to immuno-oncology data assessment is required. Appropriate study designs, end points and statistical methods are essential for evaluating immuno-oncology therapies to assess treatment outcomes, and may better accommodate immuno-oncology clinical trial data. In addition to conventional end points including median progression-free survival (PFS) and overall survival (OS), end points such as hazard ratios for PFS and OS over time, PFS and OS landmark analyses beyond the median, and immune-response end points might provide better indications of the efficacy of immuno-oncology therapies. Long-term data with these agents will allow better prediction of outcomes in MM.

First draft submitted: 17 November 2016; Accepted for publication: 22 February 2017; Published online: 11 April 2017

Unlike conventional therapies that directly kill tumor cells, immunotherapies harness the patient's own immune system to stimulate a response against the cancer [1]. Immunotherapy includes various treatments with different mechanisms of action, which may involve priming or boosting the immune system, T-cell modulation, natural killer cell activation, reducing immunosuppression and enhancing adaptive immunity. Examples of immunotherapies are hematopoietic stem cell transplantation, cell-based approaches (e.g., T-cell infusion, chimeric antigen receptor T-cell therapy), cancer vaccines (e.g., dendritic cell vaccination, DNA vaccination, peptide vaccination, viral vaccine vectors), immunomodulatory drugs and immuno-oncology agents (including immunostimulatory antibodies), which target immune checkpoints and costimulatory pathways. Immuno-oncology agents exhibit kinetics that are characteristically different from conventional cancer therapies, which involve building a cellular immune response followed by tumor regression.

Traditional immunotherapies, such as, IFN and IL-2, have not demonstrated consistent clinical benefit in advanced-stage cancer, possibly due to an incomplete understanding of tumor immunology. Further, the use of conventional trial designs and end points does not capture the novel patterns of response seen with immuno-oncology therapies [2]. Therefore, opportunities to identify effective therapies are potentially being overlooked or not fully appreciated.

The patterns of response to treatment with immuno-oncology agents differ from those seen with conventional therapies [3]. Long-term survival and delayed clinical benefit are common outcomes. Stable disease or responses can occur after conventional progressive disease owing to clinically

1 Cancer Research And Biostatistics (CRAB), Seattle, WA, USA

*Author for correspondence: Tel.: +1 206 839 1789; Fax: +1 206 652 4612; antje@crab.org

Keywords

• immuno-oncology trials • multiple myeloma • study end points

² Samuel Oschin Cancer Center, Cedars-Sinai Medical Center, Los Angeles, CA, USA

insignificant new lesions in the presence of other responsive lesions and to a reduction in total tumor burden. As such, discontinuation of immuno-oncology therapies at the first sign of progressive disease might not always be appropriate [4]. In addition, durable stable disease might represent meaningful antitumor activity in patients who do not meet the criteria for an objective response [3]. Because the clinical benefit of immuno-oncology agents might extend beyond that of traditional cytotoxic agents, alternative statistical methods should be considered to allow treatment efficacy of immuno-oncology therapies to be assessed appropriately [3,5].

The introduction of immuno-oncology therapies is likely to have a significant impact on the treatment of multiple myeloma (MM), with improved clinical outcomes and changing treatment paradigms. MM is a malignant disease of monoclonal plasma cells that is associated with a 5-year survival rate below 50% [6]. Owing to the aging population, the incidence of MM is projected to increase by 57% from 2010 to 2030 in the USA [7]. The introduction of autologous stem cell transplantation and new targeted therapies, including proteasome inhibitors (e.g., bortezomib, carfilzomib and ixazomib) and immunomodulatory drugs (e.g., thalidomide, lenalidomide and pomalidomide), has increased response rates and survival in patients with MM [8–10]. Despite these advances, MM is still considered by many to be an incurable disease, primarily because of high intratumoral heterogeneity that increases as the disease progresses [11,12]. However, investigators have recently challenged this belief; indeed, tandem hematopoietic stem cell transplantation and long-term follow-up results from the Total Therapy trial program at the Myeloma Institute of the University of Arkansas for Medical Sciences have shown evidence of a cure [13,14]. Until a definitive cure is developed, there is a need for new agents to improve treatment outcomes for patients with MM. Immunotherapies currently in clinical development for the treatment of MM are presented in **Table 1**.

Immuno-oncology therapies: response assessments & safety considerations ● **Delayed clinical benefit**

Several immuno-oncology therapies have shown a delayed clinical effect [16–18] likely due to their mechanism of action [3]. Because time is required to build an effective immune response, responses to immuno-oncology therapy are expected to take longer than responses to cytotoxic therapy, and may develop after a period of apparent progressive disease (either progression of known lesions or the appearance of new lesions), or may continue after therapy is discontinued. Separation of estimated survival curves might therefore occur several months after the start of treatment with immuno-oncology agents. In addition, patients might experience transient worsening of disease before achieving tumor regression or stable disease [3], including prolonged periods of stable disease that are clinically significant [3,19]. Therefore, caution should be taken when terminating therapy early, as responses can take substantially longer than is typical to become apparent, and immune cell infiltration might be mistaken for tumor progression. Importantly, a delay in the separation of survival curves violates the fundamental study design assumption of proportional hazards, which might reduce the statistical power of a study to differentiate between two treatment arms [2,5,20].

● **Long-term survival**

In some patients, long-term survival might be achievable with immuno-oncology therapies, which is reflected by a plateau in survival curves. Long-term survival has been observed with the use of immuno-oncology therapies in certain cancer types, including melanoma [16,21] and chronic myeloid leukemia [22,23]. Also, durable responses and progression-free survival (PFS) benefit were observed in an open-label Phase III trial of nivolumab versus docetaxel in patients with previously treated metastatic squamous cell non-small-cell lung cancer [24]. In addition, a survival benefit with ipilimumab, an antibody that targets CTLA-4, has been demonstrated in two randomized controlled Phase III trials in melanoma [16,21]. Separation of PFS and overall survival (OS) curves has been observed just prior to 4 months after treatment initiation with ipilimumab plus glycoprotein 100 (gp100) peptide vaccine or ipilimumab monotherapy compared with gp100 monotherapy [16]. In another study, long-term follow-up in patients treated with ipilimumab and dacarbazine demonstrated a 5-year survival rate of 18% [25].

In contrast to the patterns observed in solid tumor studies, in some MM trials, early separation of curves has been observed with immunooncology agents in combination with standard MM therapy. In a Phase III study of elotuzumab plus lenalidomide and dexamethasone in patients with relapsed or refractory MM (RRMM; ELOQUENT-2), the relatively early separation of curves, indicating initial disease control, may be due to the enhancement of the efficacy of elotuzumab by the lenalidomide/dexamethasone backbone. Elotuzumab, with its immune-stimulatory component, allows for long-term disease control. Similarly, in the CASTOR and POLLUX studies of daratumumab, an early separation of curves was observed and treatment was in combination with a similar treatment backbone. Thus, the combination of immuno-oncology agents with standard MM therapy (lenalidomide/dexamethasone) may be a valuable treatment option for patients with MM experiencing early relapse or indolent relapse who would benefit from both early and late disease control. In a study of elotuzumab given as monotherapy in patients with RRMM, elotuzumab did not show significant antimyeloma activity as a single agent after a follow-up period of 8 weeks. This lack of response may be due to elotuzumab's mechanism of action and the short follow-up period. Longer follow-up in this study may have shown benefit in these patients. On the other hand, other MM immunotherapy trials have shown the emergence of a longterm survival plateau or delayed separation of PFS and OS curves.

Allogeneic transplantation, which represents ongoing immunotherapy in MM wherein donor T cells target host tumor cells, is a good example of the emergence of a long-term survival plateau. Allogeneic transplantation, while potentially curative in a subgroup of patients with MM, is controversial due to a high rate of transplantation-related mortality [26]. As part of a prospective randomized study conducted by SWOG (formerly the Southwest Oncology Group) in patients with MM (Trial S9321), patients with suitable donors were assigned to an allogeneic transplantation arm $(n = 36)$, resulting in high rates of early mortality but a long-term survival plateau in a few patients $(n = 14)$ [27]. In the EBMT-NMAM2000 study conducted by the European Group for Blood and Marrow Transplantation, tandem autologous/reduced intensity conditioning allogeneic transplantation was associated with longer OS but higher mortality than was autologous transplantation at a median follow-up of 96 months [28]. However, follow-up of longer than 5 years is needed to confirm differences in outcomes between the two approaches. In the Total Therapy 2 study at the University of Arkansas for Medical Sciences, the addition of thalidomide to highdose melphalan-based chemotherapy and autologous hematopoietic stem cell transplantation increased the frequency of complete responses and extended event-free survival, but not OS [29]. However, a delayed benefit in OS of thalidomide treatment was observed after 5 years [30]. In a Phase III study of elotuzumab plus lenalidomide and dexamethasone in patients with RRMM (ELOQUENT-2) [31], interim results revealed an early and increasing separation between the PFS curves over time that was maintained at 2 years; follow-up for long-term survival is ongoing. While these studies differ in design and populations studied, and evaluate therapies with differing mechanisms of action, overall they suggest that long-term follow-up, including end points that capture benefit across the entire study population and at milestone time points (e.g., hazard ratios [HRs] over time and 5-year PFS or OS), should be considered when designing immuno-oncology clinical trials involving patients with MM.

● **Immune-related response criteria**

The standard WHO criteria and Response Evaluation Criteria In Solid Tumors (RECIST) [32,33] were designed to assess the effects of cytotoxic and cytostatic anticancer agents, but do not account for the delayed clinical benefit and long-term survival seen with immuno-oncology therapies. The immunerelated response criteria (irRC), a novel set of antitumor assessment criteria for evaluating immune response in patients treated with immuno-oncology therapies, were developed based on RECIST and WHO criteria to account for the unique characteristics of immunooncology agents [34]. The irRC differs from the standard RECIST and WHO response criteria by measuring total tumor burden (the sum of index lesion and new measurable lesions) and comparing it with baseline measurements (preconfirmed immune-related disease progression) at various time points. Use of the irRC might avoid premature discontinuation of therapy in patients who will eventually respond to treatment or have prolonged stable disease. A challenge of using the irRC is uniformly implementing these criteria among treating physicians across various sites and ensuring data interpretability by standardizing assessment methods

All study details were taken from [15].

across immuno-oncology clinical trials. Further clarification of the irRC should be made in an effort to standardize these criteria and minimize discordance between sites and central independent assessments, with the aim of providing more objective and reproducible response assessments in patients treated with immuno-oncology therapies. Nevertheless, since its publication in 2009, the irRC has been shown to be a powerful tool, allowing more comprehensive investigation of immunotherapies in clinical trials [34].

Further study of the utility of the irRC in immuno-oncology trials for different cancer types is warranted. To date, no such studies have been performed in hematologic malignancies, although modification of the irRC to accommodate this is the next step in the evolution of these criteria [34]. The implementation of the irRC is more complicated in the setting of MM compared with solid tumors. Furthermore, RECIST guidelines are not applicable, and the International Myeloma Working Group (IMWG) response criteria are being used instead [35,36]. These IMWG consensus criteria have recently been updated [37]. Notably, according to the IMWG criteria, repeat response assessment for confirmation of progression is critical in the setting of MM, which avoids the problem of assessment errors due to abnormal M-protein spikes that are sometimes observed in patients with MM. The recent update of the IMWG criteria has also proposed new response criteria, including minimal residual disease (MRD) assessment as an end point in MM trials. In addition, using next-generation flow cytometry methods allows for both assessment of MRD and detection of abnormal immune profiles. Ongoing assessment of immune reconstitution for long-term survivors in MM, which may be important for patients receiving immuno-oncology therapy, is addressed as part of these new response criteria, although this is a relatively new area that requires further data. Normalization of the light chain ratios for immunoglobulins appears to be an indicator of normalization of the immune system [38]. Although PFS and OS should remain the primary end points for Phase III immuno-oncology therapy studies, it might be useful to explore secondary immune end points, such as the elimination of antigens or T-cell response. These end points could provide early indications of efficacy, with the caveat that delayed response to treatment is a feature of immuno-oncology therapy.

● **Minimal residual disease**

MRD is a measurement of the depth of a response to treatment in hematologic cancers, using sensitive technology that can detect the presence of cancer when standard tests cannot. MRD is a particularly attractive new concept in hematologic cancer studies, as it may provide a powerful indication of treatment efficacy and remission early on, before standard outcomes like OS can be measured.

Incorporating a novel surrogate end point, such as MRD, into trial designs presents significant challenges, such as the standardization of these measurements and the selection and availability of the sensitive technology platforms needed to detect MRD. The 2016 update to the IMWG criteria defines new response categories of MRD as disease assessment criteria in patients with MM. Based on these new consensus criteria, it is likely that MRD will become more widely used as a surrogate end point in future MM studies. However, obtaining approval from regulatory bodies even for end points such as MRD that are well accepted in the community is an additional challenge. We propose that surrogate end points be incorporated into study designs as interim analyses which would allow for an earlier read on the data; however longer-term end points like OS and PFS should remain primary end points.

● **Immune-related adverse events**

In contrast to conventional cancer treatment, such as chemotherapy, most adverse events (AEs) seen with immuno-oncology therapies are inflammatory in nature, caused by the modulation of immune activity. This spectrum of adverse effects commonly affects the skin, liver, bowel and endocrine system [16]. As conventional safety analyses might not accurately capture the safety profile of immuno-oncology agents, alternative analyses are necessary to characterize AEs of immune system origin. In studies evaluating the efficacy of the CTLA-4 inhibitor ipilimumab, a method was developed to identify immune-related AEs [16]. In studies evaluating the efficacy of immuno-oncology agents, careful monitoring of safety, immune-related AEs and immune-mediated adverse reactions is required during both the trial and long-term follow-up.

Study design issues & solutions ● **Statistical concepts**

To achieve statistical power in immuno-oncology clinical trials, alternative statistical methods should be considered when calculating the required number of events for the final analysis under a delayed-separation assumption [5]. Although statistical survival analysis models are typically nonparametric (e.g., Kaplan–Meier curves and log-rank tests), parametric survival models, such as the exponential or Weibull models, can reveal additional insights.

A minimum notation will help to illustrate the statistical issues raised by the introduction of these new agents. Let the familiar survival curve be denoted by $S(t)$. An important concept in survival analysis is the hazard function, which, simplistically, is the probability of death shortly after time t, given survival to time t. From the calculus, this is given by the derivative of $-S(t)$ divided by S(t). The simplest possible model for the survival curve, containing one parameter, is the exponential model $S(t) = \exp(-\lambda t)$, for which the hazard function is the constant, λ . **Figure 1A** shows two exponential survival curves, with hazards of 0.25 and 1.25 (lower hazard means better survival). Also shown in **Figure 1A** are the HR (which in this case is a constant, 0.2) and the median survival for the two curves (points at which the curves cross 50%). For this example, the medians are 2.75 and 0.55 years, respectively. This illustrates a further property of the exponential model: that the HR (here 0.2) is the inverse of the ratio of the medians $(2.75/0.55 = 5)$.

● **The Weibull model**

The exponential model is extremely useful for calculating sample size and power at the design stage of the clinical trial process, but it is overly simplistic and often does not fit the observed data. A slightly more complicated two-parameter model called the Weibull model is given by $S(t) = \exp(-\lambda t^{\rho})$, where the parameter ρ is often called the 'shape' parameter. For a Weibull survival curve, the hazard is, in general, not a constant but a function of time, given by $\rho \lambda t^{\rho-1}$. For $\rho = 1$, this reduces to the exponential model.

Pairs of Weibull survival curves are shown in **Figure 1B & C**. In **Figure 1B**, the parameter ρ is the same for the two curves. This means that the HR is again a constant (here 0.2), and not a function of time. However, the ratio of median survival times is no longer the inverse of the HR $(2.0/0.65 \neq 5)$. In **Figure 1C**, the shape parameters also differ between the two curves, and the HR is a function of time. The two survival curves in **Figure 1C** are quite similar for a period of time before separating, meaning that the Weibull model might be a good fit to the data from clinical trials in immuno-oncology.

The Weibull model also allows for the inclusion of covariates of survival times, and can describe data from long-tailed distributions, which are common in immuno-oncology trials. As the Weibull model is both a proportional hazards model and an accelerated failure time model, treatment effects can be defined either in terms of HR or relative increase or decrease in survival time [39]. The model offers a compromise between exponential and nonparametric approaches, and can be a useful alternative in the analysis of survival data [40]. This model, along with a mixture model incorporating a cure fraction, has been suggested for use in clinical studies involving patients with long-term survival [41–43]. The SWOG Phase III study investigating the addition of gemtuzumab ozogamicin to induction or maintenance therapy in acute myeloid leukemia, in which the survival curve reached a plateau, is an example of a study in which the Weibull model may have been of use [44]. One drawback of the existing simple parametric models is their lack of flexibility owing to strong assumptions about the shape of the baseline hazard function. Piecewise exponential distributions have also been recognized as a simple and flexible tool in survival analyses, and can perhaps be valuable in immuno-oncology trials [45].

● **Early interim analyses**

Standard group sequential designs and interim analyses have been used in Phase III studies to allow early study termination for success or futility. Standard designs and response criteria should be reassessed when considering immunooncology studies [4]. Early interim analysis should be approached carefully because a drug with delayed separation of the survival curves might confer a misleadingly negative early result. In a Phase III trial of tremelimumab in metastatic melanoma, an early interim analysis showed no survival benefit and the study was terminated; however, extended follow-up showed delayed separation of the survival curves [46].

Data analysis issues & solutions

Traditionally, the most commonly used statistical methods for time-to-event analyses are the log-rank test and Cox regression analysis [47], which have maximal statistical power under the proportional hazards (constant

HR) assumption. However, these methods do not emphasize the treatment effect on longterm survivors. Also, as previously noted, the delayed separation of curves often seen in immuno-oncology studies violates the fundamental study design assumption of proportional hazards, which can lead to a loss of statistical power.

future science group FSS

● **Nonproportional hazards & weighted log-rank test**

For agents, such as immuno-oncology therapies, that show a delayed clinical effect, the possibility of nonproportional hazards should be considered when estimating the size of the study to reduce the likelihood of false-negative conclusions [4]. A sufficient number of patients need to be randomized to ensure that the total number of events is reachable, while the study duration should be long enough to ensure an adequate follow-up duration for all patients to capture any long-term survival benefit. The delayed clinical effect and the long tail of survival curves highlight the importance of using landmark survival analyses, as well as the HR and median OS, to benchmark long-term survival outcomes. Indeed, in a Phase III study of treatment-naive patients with advanced melanoma, a milestone survival analysis with a minimum follow-up of 5 years demonstrated improved outcomes with ipilimumab plus dacarbazine. The median OS was 11.2 months (95% CI: 9.5–13.8 months) in the ipilimumab plus dacarbazine group versus 9.1 months (95% CI: 7.8–10.5 months) in the dacarbazine-only group, with an HR of 0.69 (95% CI: 0.57–0.84) [48]. Milestone survival rates at given time points, such as, 1- or 2-year OS rates and HRs, are warranted in immunooncology therapy trials, provided the minimum follow-up is long enough to allow robust estimation of these rates. Alternative summary measures of treatment differences should also be considered [49–52]. Of these, particularly promising is the estimated probability that a patient on a particular treatment dies earlier than if that same patient received another treatment.

The weighted log-rank test is more powerful when nonproportional hazards, such as delayed clinical effect and long-term survival, are present, and can be considered an appropriate alternative to traditional statistical methods [20], along with the Weibull model. To avoid a substantial loss of statistical power, Fine proposed using a weighted log-rank test with the Fleming–Harrington [53] class of weights as the primary analysis in confirmatory studies of immunotherapies focusing on a survival end point [20]. However, there is no consensus on an *a priori* set of weights that should be used for a given trial, as the point at which the survival curves might diverge cannot be predicted at the start. A weighted logrank test, which downweighs the early time period during which survival curves might be similar, requires more events to detect a treatment effect than does the standard log-rank test. In the absence of knowledge of that time period, and thus of the appropriate weights to use, the sample size calculated using a proportional hazards assumption should be increased. In our experience, this inflation factor should be in the order of 10%; however, further investigation is needed.

Perspectives on the changing treatment paradigm in MM predictive markers of anticancer immunity

Immune-response end points might provide early indications of the efficacy of immuno-oncology therapies. In addition, predictive biomarkers of immune response, such as, elimination of antigens and T-cell or antibody response, might better identify subsets of patients who would benefit from treatment. The use of such immuneresponse end points could improve the design and implementation of immuno-oncology trials by determining whether an immuno-oncology therapy has achieved its biologic effect, thus predicting clinical outcomes. However, T-cell immuneresponse assays are highly variable and should be standardized to minimize data variability [5]. In melanoma immuno-oncology studies, markers of immune response correlating with outcomes have been reported [54,55]; however, further development of reliable and reproducible assays for novel markers that correlate with improved survival in immuno-oncology trials for different types of cancers, including MM, is needed.

● **Surrogate end points**

As patients with hematologic cancers often present with slow-growing indolent disease, alternate or surrogate end points are particularly important when considering trial designs, with the goal of expediting drug development and delivering new treatments to these patients faster. In MM, implementation of appropriate surrogate end points in future studies will be important to better establish early signals for efficacy. Surrogate end points evaluating immune responses, biomarkers and MRD in earlier phase studies may be beneficial.

Delivering new treatment options to patients as quickly as possible by expediting the regulatory process of developing drugs is not uncommon for serious and life-threatening diseases, such as cancer. Breakthrough or fast-track designations have become more common for

new cancer therapies including immuno-oncology agents such as daratumumab, which was approved by health agencies based on Phase II data. As such, appropriate surrogate end points in Phase II studies of immuno-oncology agents should be established to increase the confidence of earlier phase data and retain the scientific rigor that is required for regulatory approval.

● **Maintenance therapy**

A longer PFS has been observed in patients with MM undergoing continuous therapy with several drugs, including lenalidomide [56–59]. However, next-generation treatment for MM should include alternative maintenance strategies, as some of these drugs have been associated with a higher frequency of second primary malignancies [60]. The specificity, therapeutic efficacy and low toxicity profile of monoclonal antibodies make them attractive candidates for maintenance therapy. In the solid tumor setting, in contrast to traditional chemotherapies used for the treatment of melanoma, nivolumab and pembrolizumab are currently given until disease progression or unacceptable toxicity [61,62]. In MM, the treatment of residual disease with continuous immuno-oncology therapy after induction therapy might yield better outcomes than induction alone. Although there is some evidence of a beneficial immunomodulatory effect of lenalidomide in the maintenance setting [63], further studies evaluating the efficacy of long-term therapy in patients with MM are needed. However, longterm improved OS does not necessarily mean that a patient is cured. Indeed, a meta-analysis of maintenance therapy in MM showed that a subset of 'cured' patients with long-term survival have some residual disease versus those who are MRD negative [64]. Therefore, long-term survival can be improved with maintenance therapy, as some patients require ongoing immune modulation of residual disease despite improved OS.

A challenge in immuno-oncology is how to optimize the sequence of treatment with immunooncology therapies, as increasing evidence suggests that clinical benefit might be optimized by administering immuno-oncology therapies as early as possible in the treatment paradigm. In a retrospective analysis evaluating the efficacy of ipilimumab therapy before and after BRAF inhibitor treatment in patients with metastatic melanoma [65], patients treated with ipilimumab prior to BRAF inhibitors experienced better outcomes than did patients treated with the opposite schedule. In the MM setting, a Phase III study of elotuzumab plus lenalidomide and dexamethasone as first-line therapy in previously untreated patients (ELOQUENT-1; NCT01335399) is under way. Further randomized controlled trials are needed to define the optimal sequencing of immuno-oncology therapies.

Conclusion

When considering immuno-oncology agents in MM, appropriate study designs, end points, statistical analyses and interpretation are needed, which can differ from those for cytotoxic chemotherapy and pathway-specific agents. Importantly, clinical benefit with immunooncology therapies is often delayed, so consideration of long-term (i.e., beyond the median) benefit is needed. In order to confirm early outcomes with surrogate end points, adequate long-term follow-up continues to be of utmost importance in immuno-oncology clinical trials, in order to capture the overall benefits of these new therapies. Clinical trials should take a longer-term view to understand the clinical benefit across the study population and ensure that potentially beneficial treatments are not disregarded too early.

Future perspective

Immuno-oncology therapies that harness the patient's immune system to target cancer cells are likely to have a significant impact on the treatment of MM, with improved clinical outcomes and changing treatment paradigms. Appropriate study designs, end points and statistical methods are essential for evaluating immuno-oncology therapies to achieve more favorable treatment outcomes. Currently, much focus is placed on median time-to-event end points, such as OS and PFS. As clinical benefit with immuno-oncology therapies is often delayed, a long-term approach to the assessment of immuno-oncology data should be taken, including assessment of HRs over time (which assess benefit across the entire course of a study and include all patients) and milestone PFS/OS analyses, as clinical trial results can be misleading if analyzed too early. In addition to the conventional efficacy end points, we propose selecting appropriate immune end points to determine early signs of an immune response, and adapting currently accepted end points to ensure the long-term evaluation of clinical benefit with immuno-oncology agents. Long-term data with promising immuno-oncology agents

may help us better predict future outcomes with immuno-oncology therapies in MM.

Financial & competing interests disclosure

B Durie has served as a consultant for Amgen, Takeda, Celgene, & Johnson and Johnson. Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number U10CA180819. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The authors have no other relevant affiliations or financial involvement with any organization or

entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Professional medical writing and editorial assistance was provided by Kate Jesien of Caudex, New York, USA, and funded by Bristol-Myers Squibb.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit http://creativecommons.org/ licenses/by-nc-nd/4.0/

Executive summary

- Long-term survival and delayed clinical benefit are common patterns of response to treatment with immuno-oncology agents.
- Alternate statistical methods and end points should be considered to allow treatment efficacy of immuno-oncology agents to be assessed appropriately.

Immuno-oncology therapies: response assessments & safety considerations

- Long-term follow-up and end points, such as, hazard ratios and 5-year progression-free survival and overall survival should be considered when designing immuno-oncology trials in the multiple myeloma (MM) setting.
- Immune-related response criteria were developed based on Response Evaluation Criteria In Solid Tumors and WHO criteria for evaluating immune response in patients with solid tumors treated with immuno-oncology therapies:
	- These criteria are not yet being used in hematologic malignancies and may require adaptation in the MM setting.
- Careful monitoring of safety and immune-related adverse events is required during immuno-oncology trials and long-term follow-up.

Study design issues & solutions

- Although statistical survival analysis models are typically nonparametric (e.g., Kaplan–Meier curves and log-rank tests), parametric survival models, such as the exponential or Weibull models, can reveal additional insights as they allow for the inclusion of covariates of survival times. In addition, compared with the exponential model with the underlying proportional assumption, the Weibull model can describe data from long-tailed distributions, which are common in immuno-oncology trials.
- Standard study designs and interim analyses allowing for early study termination for success or futility should be reassessed when considering immuno-oncology agents, so potentially beneficial treatments are not disregarded too early.

Data analysis issues & solutions

- The delayed clinical benefit and long-term survival often seen in immuno-oncology studies violate the fundamental assumption of proportional hazards, which can lead to a loss of statistical power.
- The weighted log-rank test is more powerful when nonproportional hazards are present and can be considered an appropriate alternative to traditional statistical methods.
- Milestone survival rates, such as 1- and 2-year overall survival rates and hazard ratios, are warranted in immuno-oncology trials.

Perspectives on the changing treatment paradigm in MM predictive markers of anticancer immunity

● Further randomized controlled trials are needed to evaluate the efficacy of long-term maintenance therapy and to define the optimal sequencing of immuno-oncology therapies in the MM setting.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

- Finn OJ. Immuno-oncology: understanding the function and dysfunction of the immune system in cancer. *Ann. Oncol.* 23(Suppl. 8), viii6-viii9 (2012).
- 2 Hoos A, Britten C. The immuno-oncology framework: enabling a new era of cancer therapy. *Oncoimmunology* 1(3), 334–339 (2012).
- 3 Wolchok JD, Hoos A, O'Day S *et al.* Guidelines for the evaluation of immune therapy activity in solid tumors: immunerelated response criteria. *Clin. Cancer Res.* 15(23), 7412–7420 (2009).
- **Describes** guidelines for the evaluation of **the patterns of response seen with immuno-oncology therapies.**
- 4 Chen TT. Statistical issues and challenges in immuno-oncology. *J. Immunother. Cancer* 1, 18 (2013).
- 5 Hoos A, Eggermont AM, Janetzki S *et al.* Improved end points for cancer immunotherapy trials. *J. Natl Cancer Inst.* 102(18), 1388–1397 (2010).
- 6 Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J. Clin.* 65(1), 5–29 (2015).
- 7 Smith BD, Smith GL, Hurria A, Hortobagyi GN, Buchholz TA. Future of cancer incidence in the United States: burdens upon an aging, changing nation. *J. Clin. Oncol.* 27(17), 2758–2765 (2009).
- 8 Gay F, Larocca A, Wijermans P *et al.* Complete response correlates with long-term progression-free and overall survival in elderly myeloma treated with novel agents: analysis of 1175 patients. *Blood* 117(11), 3025–3031 (2011).
- 9 Harousseau JL. Ten years of improvement in the management of multiple myeloma: 2000–2010. *Clin. Lymphoma Myeloma Leuk.* 10(6), 424–442 (2010).
- 10 Kastritis E, Zervas K, Symeonidis A *et al.* Improved survival of patients with multiple myeloma after the introduction of novel agents and the applicability of the International Staging System (ISS): an analysis of the Greek Myeloma Study Group (GMSG). *Leukemia* 23(6), 1152–1157 (2009).
- 11 Morgan GJ, Walker BA, Davies FE. The genetic architecture of multiple myeloma. *Nat. Rev. Cancer* 12(5), 335–348 (2012).
- 12 International Myeloma Working Group. Criteria for the classification of monoclonal

gammopathies, multiple myeloma and related disorders: a report of the International Myeloma Working Group. *Br. J. Haematol.* 121(5), 749–757 (2003).

- 13 Usmani SZ, Crowley J, Hoering A *et al.* Improvement in long-term outcomes with successive Total Therapy trials for multiple myeloma: are patients now being cured? *Leukemia* 27(1), 226–232 (2013).
- 14 Barlogie B, Mitchell A, van Rhee F, Epstein J, Morgan GJ, Crowley J. Curing myeloma at last: defining criteria and providing the evidence. *Blood* 124(20), 3043–3051 (2014).
- 15 Clinical Trials. <www.Clinicaltrials.gov>
- 16 Hodi FS, O'Day SJ, McDermott DF *et al.* Improved survival with ipilimumab in patients with metastatic melanoma. *N. Engl. J. Med.* 363(8), 711–723 (2010).
- **• Evaluates the efficacy of the CTLA-4 inhibitor, ipilimumab, where for the first time, a method was developed to identify immune-related adverse events.**
- 17 Wolchok JD, Neyns B, Linette G *et al.* Ipilimumab monotherapy in patients with pretreated advanced melanoma: a randomised, double-blind, multicentre, Phase II, dose-ranging study. *Lancet Oncol.* 11(2), 155–164 (2010).
- 18 Small EJ, Schellhammer PF, Higano CS *et al.* Placebo-controlled Phase III trial of immunologic therapy with sipuleucel-T (APC8015) in patients with metastatic, asymptomatic hormone refractory prostate cancer. *J. Clin. Oncol.* 24(19), 3089–3094 (2006).
- 19 Finke LH, Wentworth K, Blumenstein B, Rudolph NS, Levitsky H, Hoos A. Lessons from randomized Phase III studies with active cancer immunotherapies-outcomes from the 2006 meeting of the Cancer Vaccine Consortium (CVC). *Vaccine* 25(Suppl. 2), B97–B109 (2007).
- 20 Fine G. Consequences of delayed treatment effects on analysis of time-to-event end points. *Drug Inf. J.* 41, 535–539 (2007).
- 21 Robert C, Thomas L, Bondarenko I *et al.* Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. *N. Engl. J. Med.* 364(26), 2517–2526 (2011).
- 22 Rea D, Vellenga E, Junghan C, Baccarani M, Kantarjian H. Six-year follow-up of patients with imatinib-resistant or imatinib-intolerant chronic-phase chronic myeloid leukemia (CP-CML) receiving dasatinib. *Haematologica* 97(Suppl. 1), Abstract 0199 (2012).
- 23 Kantarjian H, O'Brien S, Garcia-Manero G *et al.* Very long-term follow-up results of

imatinib mesylate therapy in chronic phase chronic myeloid leukemia after failure of interferon alpha therapy. *Cancer* 118(12), 3116–3122 (2012).

- 24 Brahmer J, Reckamp KL, Baas P *et al.* Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. *N. Engl. J. Med.* 373(2), 123–135 (2015).
- 25 Lebbe C, Weber JS, Maio M *et al.* Survival follow-up and ipilimumab retreatment of patients with advanced melanoma who received ipilimumab in prior Phase II studies. *Ann. Oncol.* 25(11), 2277–2284 (2014).
- 26 Gahrton G, Krishnan A. Allogeneic transplantation in multiple myeloma. *Expert Rev. Hematol.* 7(1), 79–90 (2014).
- 27 Barlogie B, Kyle RA, Anderson KC *et al.* Standard chemotherapy compared with high-dose chemoradiotherapy for multiple myeloma: final results of Phase III US Intergroup Trial S9321. *J. Clin. Oncol.* 24(6), 929–936 (2006).
- 28 Gahrton G, Iacobelli S, Bjorkstrand B *et al.* Autologous/reduced-intensity allogeneic stem cell transplantation vs autologous transplantation in multiple myeloma: long-term results of the EBMT-NMAM2000 study. *Blood* 121(25), 5055–5063 (2013).
- 29 Barlogie B, Tricot G, Anaissie E *et al.* Thalidomide and hematopoietic-cell transplantation for multiple myeloma. *N. Engl. J. Med.* 354(10), 1021–1030 (2006).
- 30 Barlogie B, Pineda-Roman M, van RF *et al.* Thalidomide arm of Total Therapy 2 improves complete remission duration and survival in myeloma patients with metaphase cytogenetic abnormalities. *Blood* 112(8), 3115–3121 (2008).
- 31 Lonial S, Dimopoulos M, Palumbo A *et al.* Elotuzumab therapy for relapsed or refractory multiple myeloma. *N. Engl. J. Med.* 373(7), 621–631 (2015).
- 32 Eisenhauer EA, Therasse P, Bogaerts J *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* 45(2), 228–247 (2009).
- 33 Therasse P, Arbuck SG, Eisenhauer EA *et al.* New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J. Natl Cancer Inst.* 92(3), 205–216 (2000).
- 34 Hoos A, Wolchok JD, Humphrey RW, Hodi FS. CCR 20th Anniversary Commentary: immune-related response criteria – capturing clinical activity in immuno-

oncology. *Clin. Cancer Res.* 21(22), 4989–4991 (2015).

- P **B** Describes the novel set of antitumor **assessment criteria evaluating immune response in patients treated with immunooncology therapies, developed based on the Response Evaluation Criteria In Solid Tumors and WHO criteria.**
- 35 Rajkumar SV, Dimopoulos MA, Palumbo A *et al.* International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. *Lancet Oncol.* 15(12), e538–e548 (2014).
- 36 Durie BG, Harousseau JL, Miguel JS *et al.* International uniform response criteria for multiple myeloma. *Leukemia* 20(9), 1467–1473 (2006).
- **•• Describes the International Myeloma Working Group response criteria for multiple myeloma.**
- 37 Kumar S, Paiva B, Anderson KC *et al.* International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. *Lancet Oncol.* 17(8), e328–e346 (2016).
- 38 Iwama K, Chihara D, Tsuda K *et al.* Normalization of free light chain kappa/ lambda ratio is a robust prognostic indicator of favorable outcome in patients with multiple myeloma. *Eur. J. Haematol.* 90(2), 134–141 (2013).
- 39 Carroll KJ. On the use and utility of the Weibull model in the analysis of survival data. *Control Clin. Trials* 24(6), 682–701 (2003).
- **• Describes the use and utility of the Weibull model for the analysis of survival data from clinical trials.**
- 40 Ying GS, Heitjan DF. Weibull prediction of event times in clinical trials. *Pharm. Stat.* 7(2), 107–120 (2008).
- 41 Farewell VT. The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics* 38(4), 1041–1046 (1982).
- 42 Perperoglou A, Keramopoullos A, van Houwelingen HC. Approaches in modelling long-term survival: an application to breast cancer. *Stat. Med.* 26(13), 2666–2685 (2007).
- 43 Othus M, Barlogie B, Leblanc ML, Crowley JJ. Cure models as a useful statistical tool for analyzing survival. *Clin. Cancer Res.* 18(14), 3731–3736 (2012).
- 44 Petersdorf SH, Kopecky KJ, Slovak M *et al.* A Phase III study of gemtuzumab ozogamicin during induction and postconsolidation therapy in younger patients with acute myeloid leukemia. *Blood* 121(24), 4854–4860 (2013).
- 45 Han G, Schell MJ, Kim J. Improved survival modeling in cancer research using a reduced piecewise exponential approach. *Stat. Med.* 33(1), 59–73 (2014).
- **• Describes the benefits of using parametric survival modeling in cancer clinical studies.**
- 46 Hoos A, Britten CM, Huber C, O'Donnell-Tormey J. A methodological framework to enhance the clinical success of cancer immunotherapy. *Nat. Biotechnol.* 29(10), $867 - 870(2011)$.
- 47 Kalbfleisch JD, Prentice RL. Marginal likelihoods based on Cox's regression and life model. *Biometrika* 60(2), 267–278 (1973).
- 48 Maio M, Grob JJ, Aamdal S *et al.* Five-year survival rates for treatment-naive patients with advanced melanoma who received ipilimumab plus dacarbazine in a Phase III trial. *J. Clin. Oncol.* 33(10), 1191–1196 (2015).
- 49 Uno H, Claggett B, Tian L *et al.* Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. *J. Clin. Oncol.* 32(22), 2380–2385 (2014).
- 50 Royston P, Parmar MK, Altman DG. Visualizing length of survival in time-to-event studies: a complement to Kaplan-Meier plots. *J. Natl Cancer Inst.* 100(2), 92–97 (2008).
- 51 Moser BK, McCann MH. Reformulating the hazard ratio to enhance communication with clinical investigators. *Clin. Trials* 5(3), 248–252 (2008).
- 52 Buyse M. Generalized pairwise comparisons of prioritized outcomes in the two-sample problem. *Stat. Med.* 29(30), 3245–3257 (2010).
- 53 Harrington DP, Fleming TR. A class of rank test procedures for censored survival data. *Biometrika* 69(3), 533–566 (1982).
- 54 Yuan J, Adamow M, Ginsberg BA *et al.* Integrated NY-ESO-1 antibody and CD8+ T-cell responses correlate with clinical benefit in advanced melanoma patients treated with ipilimumab. *Proc. Natl Acad. Sci. USA* 108(40), 16723–16728 (2011).
- 55 Kirkwood JM, Lee S, Moschos SJ *et al.* Immunogenicity and anti tumor effects of vaccination with peptide vaccine+/−

granulocyte-monocyte colony-stimulating factor and/or IFN-alpha2b in advanced metastatic melanoma: Eastern Cooperative Oncology Group Phase II Trial E1696. *Clin. Cancer Res.* 15(4), 1443–1451 (2009).

- 56 Attal M, Lauwers-Cances V, Marit G *et al.* Lenalidomide maintenance after stem-cell transplantation for multiple myeloma. *N. Engl. J. Med.* 366(19), 1782–1791 (2012).
- 57 McCarthy PL, Owzar K, Hofmeister CC *et al.* Lenalidomide after stem-cell transplantation for multiple myeloma. *N. Engl. J. Med.* 366(19), 1770–1781 (2012).
- 58 Palumbo A, Bringhen S, Rossi D *et al.* Bortezomib-melphalan-prednisonethalidomide followed by maintenance with bortezomib-thalidomide compared with bortezomib-melphalan-prednisone for initial treatment of multiple myeloma: a randomized controlled trial. *J. Clin. Oncol.* 28(34), 5101–5109 (2010).
- 59 Palumbo A, Hajek R, Delforge M *et al.* Continuous lenalidomide treatment for newly diagnosed multiple myeloma. *N. Engl. J. Med.* 366(19), 1759–1769 (2012).
- 60 Dimopoulos MA, Richardson PG, Brandenburg N *et al.* A review of second primary malignancy in patients with relapsed or refractory multiple myeloma treated with lenalidomide. *Blood* 119(12), 2764–2767 (2012).
- 61 Pembrolizumab (KEYTRUDA). Prescribing information, December 2015. [www.merck.com/product/usa/pi_circulars/](www.merck.com/product/usa/pi_circulars/k/keytruda/keytruda_pi.pdf)
- 62 Nivolumab (OPDIVO). Prescribing information, January 2016. [http://packageinserts.bms.com/pi/](http://packageinserts.bms.com/pi/pi_opdivo.pdf)
- 63 Clave E, Douay C, Coman T *et al.* Lenalidomide consolidation and maintenance therapy after autologous stem cell transplant for multiple myeloma induces persistent changes in T-cell homeostasis. *Leuk. Lymphoma* 55(8), 1788–1795 (2014).
- 64 Usmani S, Binder G, Hu XH *et al.* Lenalidomide dose modifications and associated patient outcomes in newly diagnosed multiple myeloma (NDMM). *Clin. Lymphoma Myeloma Leuk.* 15, e257–e258 (2015).
- 65 Ackerman A, Klein O, McDermott DF *et al.* Outcomes of patients with metastatic melanoma treated with immunotherapy prior to or after BRAF inhibitors. *Cancer* 120(11), 1695–1701 (2014).