Critical Reviews in Oncology / Hematology 137 (2019) 35–42

Contents lists available at ScienceDirect

Critical Reviews in Oncology / Hematology

journal homepage: www.elsevier.com/locate/critrevonc

Immune-checkpoint inhibitors and candidate surrogate endpoints for overall survival across tumour types: A systematic literature review

Sébastien Branchoux[a,b,c,*], Carine Bellera[c,d,e], Antoine Italiano[f], Denis Rustand[b], Anne-Françoise Gaudin[a], Virginie Rondeau[b,c]

[a] Department of Health Economics & Outcomes Research, Bristol-Myers Squibb, Rueil-Malmaison, France
[b] Biostatistic Team, Bordeaux Population Health Center, ISPED, Centre INSERM U1219, INSERM, Bordeaux, France
[c] Epicene Team (Cancer & Environnement), Bordeaux Population Health Center, ISPED, Centre INSERM U1219, INSERM, Bordeaux, France
[d] Clinical and Epidemiological Research Unit, Institut Bergonié, Comprehensive Cancer Center, Bordeaux, France
[e] Department of Clinical Epidemiology, INSERM CIC-EC 14.01, Bordeaux, France
[f] Department of Early Phase Trial Unit, Institut Bergonié Comprehensive Cancer Centre, Bordeaux, France

ARTICLE INFO

Keywords:
Immune-checkpoint inhibitors
Surrogate endpoints
Overall survival
Literature review

ABSTRACT

Background: Surrogate endpoints (SEs) for overall survival (OS) are specific to therapeutic class. The objective of this review was to document all alternative endpoints studied for their association with OS in Immune-Checkpoint Inhibitors (ICI)-treated patients.

Methods: We searched PubMed and Embase for publications reporting the association between a clinical endpoint and OS in ICI-treated populations from 01/01/2003 to 03/31/2018.

Results: Out of 6,335 references identified, 24 were selected. Only 3 studies assessed surrogacy at both the patient and trial levels. The main traditional alternative endpoints included progression-free survival (N = 10) and objective response rate (N = 8). New alternative endpoints, such as durable response rate (N = 1) and intermediate response endpoint (N = 1) statistically better correlate with OS in the cancer types analysed.

Conclusion: Based on the published evidence, there is insufficient data to support validated SE for OS. Adequate surrogacy assessment of promising composite endpoints which consider a duration component is encouraged.

1. Introduction

Recently, cancer treatment has been revolutionized by the development of therapeutic agents targeting the immune system, considered as the third generation in this field after cytotoxic drugs and molecular agents acting at oncogene-related targets (Martin-Liberal et al., 2017). Among the different types of newly available immunotherapy treatments, immune-checkpoint inhibitors (ICI) constitute the most established and clinically promising therapy (Galluzzi et al., 2014). These immunomodulatory monoclonal antibodies are designed to elicit a novel or revitalise an existing antitumoral immune response.3 Antibodies which bind to the programmed cell death-1 (PD-1) receptor and block their interaction with ligands PD-L1 and PD-L2 are the most widely used therapies at the advanced stage of numerous cancers (Fujii et al., 2018; Rolfo et al., 2017). To date, six ICIs have received marketing authorization (ipilimumab, pembrolizumab, nivolumab, atezolizumab, avelumab and durvalumab) in several types of cancer. Some of them are still under investigation as monotherapy or in combination in other types of cancer or disease settings.

In oncology, several clinical endpoints, defined as characteristics or variables that reflect how a patient feels, functions, or survives in response to a treatment, have been used to assess treatment efficacy (Fleming and Powers, 2012; Mathoulin-Pelissier et al., 2008). Overall survival (OS), defined as the time from randomization to the date of death (any cause), is considered as the gold standard for the primary efficacy endpoint in randomized clinical trials (RCT) (Pazdur, 2008). However, in some limited cases, progression-free survival (PFS) may be used as primary endpoint to measure efficacy of the drug, such as, for example, in trials when there is cross-over because of early demonstration of activity of the new treatment and lack of alternatives (Ocana and Tannock, 2001).

OS may require extensive follow-up to demonstrate significant and clinically relevant differences compared to standard care, and may thus delay patient access to promising new drugs. The development of alternative endpoints, such as recurrence-free survival (RFS - adjuvant setting) or PFS (metastatic setting), that could capture treatment benefit

* Corresponding author at: Bristol-Myers Squibb, 3, rue Joseph Monier, 92500, Rueil-Malmaison, France.
E-mail address: Sebastien.branchoux@bms.com (S. Branchoux).

https://doi.org/10.1016/j.critrevonc.2019.02.013
Received 4 December 2018; Received in revised form 26 February 2019; Accepted 27 February 2019

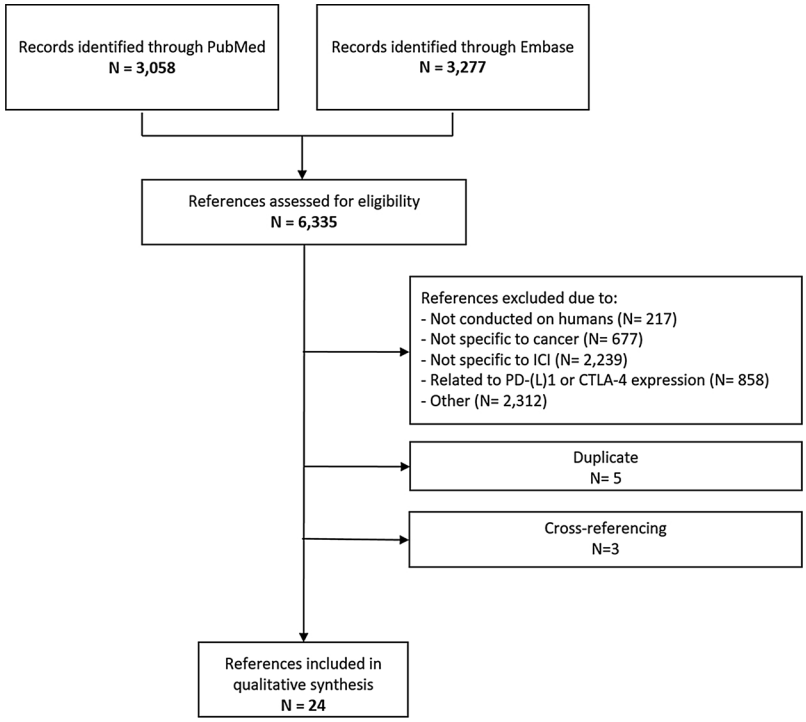

Fig. 1. Flow of information through the different phases of the systematic review of studies reporting on association between a clinical endpoint and overall survival in ICI-treated patients, as per PRISMA guidelines. (ICI = Immune-Checkpoint Inhibitors).

appropriately and be measurable earlier, has thus become central to clinical oncology. Such an endpoint, if clinically and statistically validated, could be considered as a surrogate endpoint (SE) for OS.

An SE is defined as 'a biomarker intended to substitute for a clinical endpoint' (Fleming and Powers, 2012). SEs have been extensively studied within the molecular targeted therapy era as they may reduce the need for large sample sizes, and thus decrease the duration and cost of trials (Alonso et al., 2017). Assessments of surrogacy should be performed at both the patient and trial levels, the most robust statistical method for validating a SE being the meta-analytic approach (Burzykowski et al., 2005). Since the validity of an SE depends both on the mechanism of action of the treatment and on the disease setting, it is recommended to validate all potential SEs on a case-by-case basis.

As of today, there is evidence for good surrogacy for certain SEs, as highlighted in a recent systematic literature review by Savina et al. (Savina et al., 2018) For example, PFS may be an appropriate SE for OS in the context of advanced colorectal cancer treated with chemotherapy and/or radiotherapy, or in locally advanced lung cancer treated with chemotherapy. As ICIs had not been developed extensively at the time of this review, information on SEs for OS in ICI-treated patients are still lacking.

In this context, we conducted a systematic literature review (i) to summarize current evidence on clinical alternative endpoints associated with OS in ICI-treated patients by cancer type and (ii) to identify validated SE for OS for this drug class, if any.

2. Methods

2.1. Search strategy and selection criteria

We searched PubMed and Embase for English language publications reporting the association between a clinical endpoint and OS in ICI-treated population from 01/01/2003 (year of the first clinical trial of the first ICI, ipilimumab) to 03/31/2018 (cut-off date). The computerized algorithm designed for extracting all references is available as Supplementary Material. Publications were first classified as eligible if the abstract presented at least one of the three following characteristics: conducted in humans, specific to cancer and specific to ICI. The retained publications were then classified as related to ICI expression (such as prognostic role of PD-1 expression), which were excluded, or as related to an association between a clinical endpoint and OS, which were included in the qualitative synthesis. Other publications on subjects such as adverse events related to ICIs or practice guidelines were excluded. When studies led to multiple communications (i.e. conference abstract and then peer-review article), we only included the principal publication. We reported the results of the selection process following PRISMA guidelines for the reporting of systematic reviews. (Moher et al., 2009) This study is registered in the PROSPERO database (identification number: CRD42018097434).

2.2. Data analysis

Two reviewers (SB, DR) independently collected information on the format of the communication (abstract or article), type of publication (research study or discussion article), cancer type, treatment setting (adjuvant or metastatic), type of treatment, type of data (patient-level or aggregated data), number of trials, number of patients, alternative clinical endpoints analysed, statistical method and results. Disagreements between reviewers were resolved through discussions with senior authors (VR, CB).

3. Results

Out of 6,335 references retrieved through the algorithms, 24 were selected (Fig. 1). The most common reasons for exclusion were references not related to ICI (N = 2,239) and description of RCT results, practice guidelines or related to immune-related adverse events, main contributors of the "other" category (N = 2,312). PD-(L)1 or CTLA-4 expression was described in 15% of the publications (N = 858).

Of the references selected, twenty were primary research studies and four were discussion articles such as literature reviews (Table 1). Key characteristics of the research studies are summarized in Table 2.

Table 1
Characteristics of the references identified in the systematic literature review (N = 24).

Type of cancer	Treatment setting	Type of treatment	Authors	Type of publication	Type of data	Number of trials	Number of patients	Method to assess the association	Alternative clinical endpoints	Results
Melanoma	adjuvant	Interferon (main analysis) Anti-CTLA-4 (secondary predictive analysis)	(Suciu et al., 2018)	article	individual patient-level	13	6,708	2-stage approach: - ILS: Copula function with 2 Cox models (1 for RFS and 1 for OS), Spearman coefficient (r) - TLS: R^2 from linear regression with adjustment for the treatment effect estimates	RFS	r = 0.89 (95%CI [0.88–0.89]) R^2 = 0.91 (95%CI [0.81–1.01])
	metastatic	Mix of treatment (including ICI)	(Flaherty et al., 2013)	abstract	Aggregated	10	4,215	weighted linear regressions	PFS	r = 0.74 (95%CI [0.26–0.93]) r = 0.90 (95%CI [0.66–0.97])
		Anti-CTLA-4	(Lee et al., 2013)	Abstract	Individual patient-level	1	676	Principal component analysis & Cox model	HRQoL (QLQC-30 questionnaire)	OS was associated with most of the HRQoL domain scores
		Mix of ICI	(Petrelli et al., 2016)	article	aggregated	13	3,373	weighted linear regressions (Pearson coefficient)	median PFS OS milestone (1-y) OS milestone (2-y)	by treatment arm : r(mPFS) = 0.45 (0.12–0.78) r(1y OS) = 0.93 (0.84–0.96) r(2y OS) = 0.79 (0.51–0.91) between treatment effects on SE and OS r(1-y OS) = −0.86 (0.3–0.97) r(2-y OS) = −0.83 (0.07–0.97)
		Oncolytic virus (talimogene laherparepvec)	(Kaufman et al., 2017)	article	individual patient-level	NA	436	Cox model	durable response rate	durable responders vs non-durable responders: at 9 months: HR = 0.07; (95% CI [0.01–0.48]) at 12 months: HR = 0.05 ; (95% CI [0.01–0.33]) at 18 months: HR = 0.11; (95% CI [0.03–0.44])
		Anti-PD-1	(Nishino et al., 2017a)	article	individual patient-level	NA	96	Cox model	tumour burden increase from baseline	patients with < 20% tumour burden increase during therapy had significantly reduced risk of death (HR = 0.19 95% CI [0.08–0.43])
		ICI	(Chen, 2015)	article	NA	NA	NA	NA	OS milestone rate	General discussion
		ICI	(Izar et al., 2017)	article	NA	NA	NA	NA	OS milestone rate PFS milestone rate Treatment-free interval Treatment-free survival	General discussion
Non-small cell lung cancer	metastatic / advanced	Mix of treatment (including ICI)	(Blumenthal et al., 2017)	article	aggregated	25	20,013	weighted linear regressions	12-month OS milestone ratio 9-month OS milestone ratio 9-month PFS milestone ratio 6-month ORR milestone ratio	R^2 = 0.80 (95%CI [0.63–0.91]) R^2 = 0.67 (95% CI [0.49–0.82]) R^2 = 0.19 (95%CI [0.03–0.49]) R^2 = 0.04 (95% CI [0.0002–0.28])
		Anti-PD(L)-1	(Shukuya et al., 2016)	article	aggregated	10	NR	Spearman correlation coefficient	median PFS Response rate	r = 0.473 r = 0.452
		Mix of treatment (including ICI)	(McCoach et al., 2017)	article	individual patient-level	4	660	Cox model	DepOR	For PD-1 treated patients: tumour shrinkage > 50% is associated with OS improvement
		Mix of treatment (including ICI)	(Saad and Buyse, 2017)	article	NA	NA	NA	NA	DepOR	Discussion article on the article of McCoach et al.
		Anti-PD-1	(Kazandjian et al., 2015)	abstract	individual patient-level	1	117	Cox model	objective response rate	patients who achieved a best response of complete or partial response had the longest survival.
		Anti-PD-1	(Nishino et al., 2017b)	article	individual patient-level	NR	160	Cox model	tumour burden increase from baseline	patients with < 20% tumour burden increase during therapy had significantly reduced hazards of death (HR = 0.24 ; p < 0.001)

(*continued on next page*)

Table 1 (*continued*)

Type of cancer	Treatment setting	Type of treatment	Authors	Type of publication	Type of data	Number of trials	Number of patients	Method to assess the association	Alternative clinical endpoints	Results
		Anti-PD-1	(Nishino et al., 2017c)	abstract	individual patient-level	NR	134	Cox model	tumour burden increase from baseline	median OS when ≥50% increase of tumour burden = 12·7 months [8.5–14.7] / median OS when < 50% = 4.5 [1.3–4.9] logrank test p-value = 0.0003 Tumour burden increase of 50% at 8 weeks of therapy was associated with significantly shorter OS
Urinary tract cancer	metastatic	Anti-PD(L)-1	(Abdel-Rahman, 2018)	article	aggregated	13	2,792	Spearman or Pearson correlation coefficient	Overall response rate PFS OS milestone (1 y)	Uro. carcinoma: r = −0.120; p = 0.758 ; RCC: r = −0.397; p = 0.440 Uro. carcinoma: r = −0.024; p = 0.955 ; RCC: r = 0.394; p = 0.440 Uro. carcinoma: r = 0.806; p = 0.016 ; RCC: r = 0.800; p = 0.104
Multi-tumours	Advanced	Anti-CTLA-4 Anti-PD(L)-1	(Ritchie et al., 2018)	article	aggregated	20	10,828	weighted linear regression, correlation coefficient	ORR PFS 6-month PFS	r = 0.57 (95%CI [0.23–0.89]) r = 0.42 (95%CI [0.04–0.81]) r = 0.55 (95%CI [0.14–0.92])
	Metastatic	Anti-CTLA4 Anti-PD (L)-1	(Roviello et al., 2017)	article	aggregated	17	8,994	weighted linear regression	Response rate	R^2 = 0.47 (95%CI [0.03–0.77])
		Anti-PD(L)-1	(Gao et al., 2018)	article	individual patient-level	9	5,806	2-stage approach: - ILS: Cox model - TLS: weighted linear regression (R^2)	IME	IME responders vs non responders: HR = 0.09 (95% CI [0.07–0.11]) R^2 = 0.68 (95% CI [0.02–0.91])
		Anti-PD(L)-1	(Mushti et al., 2018)	article	individual patient-level	13	6,722	2-stage approach: - ILS: Cox model, Spearman correlation coefficient (r) - TLS: weighted linear regression (R^2)	ORR PFS modified PFS	r = NA ; R^2 = 0.1277 r = 0.61 ; R^2 = 0.1303 r = 0.60 ; R^2 = 0.07 to 0.10
		Anti-PD(L)-1	(Korn and Freidlin, 2018)	article	NA	NA	NA	NA	ORR PFS Modified PFS	Discussion article on the article of Mushti et al.
		Anti-PD(L)-1 Anti-CTLA-4	(Kaufmann et al., 2017)	abstract	aggregated	18	7,140	weighted regression model	ORR DCR PFS	anti-CTLA4: anti-PD(L)1: R^2 = 0.016 R^2 = 0.066 R^2 = 0.160 R^2 = 0.038 R^2 = 0.000 R^2 = 0.432
		Mix of treatment	(Tan et al., 2017)	article	aggregated	51	NR	Spearman correlation coefficient, weighted linear regression model	PFS	r = 0.62 R^2 = 0.38
		Anti-PD-1	(Gyawali et al., 2017)	abstract	aggregated	9	NR	correlation coefficient regression model	PFS	r = 0.676 R^2 = 0.457

CI: confident interval; DepOR: depth of response; DCR: disease control rate; HR: hazard ratio; HRQoL: health-related quality of life; ICI: immune-checkpoint inhibitors; ILS: individual-level surrogacy; IME: intermediate response endpoint; NA: not applicable; NR : not reported ; ORR: objective response rate; OS: overall survival; PFS: progression-free survival; r: correlation coefficient; RCC : renal cell carcinoma; RFS: recurrence-free survival; R2: determination coefficient ; SE: surrogate endpoint ; TLS: trial-level surrogacy; Uro: urothelial Y: year.

Table 2
Summary of the key characteristics of the publications (N = 20 research studies) assessing association between a clinical endpoint and overall survival in ICI-treated patients.

Format of the communication, n (%)	
Article	14 (70)
Conference abstract	6 (30)
Type of cancer, n (%)	
Pooled cancer types	7 (35)
Non-small cell lung cancer	6 (30)
Melanoma	6 (30)
Urinary tract cancer	1 (5)
Type of treatment, n (%)	
anti-PD-1 / anti-PDL-1 / anti-CTLA-4	8 (40)
anti-PD-1 or anti-CTLA-4 alone	6 (30)
Mix of treatment types (including ICI)	4 (20)
Interferon	1 (5)
Oncolytic virus	1 (5)
Treatment setting, n (%)	
Advanced / metastatic	19 (95)
Adjuvant	1 (5)
Type of data, n (%)	
Aggregated data	10 (50)
Individual patient data	10 (50)
Statistical method[†], n (%)	
Weighted linear regression model	8 (40)
Spearman or Pearson correlation coefficient	5 (25)
Cox model	7 (35)
Two-stage approach (individual-level surrogacy and trial-level surrogacy)	3 (15)
Alternative clinical endpoints[†], n (%)	
Time-to-event endpoints	
Progression-free survival	10 (50)
OS milestone rate/ratio	3 (15)
Recurrence-free survival	1 (5)
Categorical or continuous endpoints	
Objective response rate	8 (40)
Disease control rate	1 (5)
Tumour burden increase from baseline	3 (15)
Depth of response	1 (5)
Durable response rate	1 (5)
Intermediate response endpoint	1 (5)
Health-related quality of life	1 (5)

[†] Multiple approaches or endpoints could be considered in the same study.

Most of the research studies were peer-reviewed articles (70%) focusing on a single cancer type (65%). Studies were primarily performed on pooled ICIs (40%) or on anti-PD-1 or anti-CTLA-4 only (30%), almost exclusively in the advanced/metastatic setting (95%). One study was performed on resected stage II–III melanoma interferon-treated patients and the predictive model was applied to an ICI-treated population, in the same disease setting. (Suciu et al., 2018)

For half of the publications, analyses on the association with OS were performed on individual patient data (N = 10). From a methodological perspective, multiple statistical approaches were considered, depending on the type of data available. Of the ten references presenting analyses performed on individual patient data, one third evaluated the association between the alternative endpoint and OS at both the patient and trial levels (meta-analytic approach). Only one of them used a two-stage approach with a bivariate joint distribution function. (Suciu et al., 2018)

Association between alternative endpoints and OS has been assessed for ten different types of clinical endpoints, classified as “time-to-event” or “categorical or continuous” endpoints. Definition of each alternative criterion is presented in Table 3. Half of the publications analysed the association between PFS and OS (N = 10). Correlation coefficients (r) between PFS and OS ranged from −0.024 to 0.90. Determination coefficients, while performed, were also low ($R^2 \leq 0.38$). For OS milestone rate, defined as the Kaplan-Meier survival probability at a time point defined *a priori* (milestone rate: ratio of them between treatment arms), a moderate to strong association was highlighted in three publications (r: from 0.79 to 0.93; R^2: from 0.67 to 0.93) (Petrelli et al., 2016; Izar et al., 2017; Abdel-Rahman, 2018).

Among the categorical or continuous alternative endpoints, objective response rate (ORR) was the most frequently analysed (N = 8). Similarly to PFS, associations between ORR and OS were weak ($r \leq 0.57$; $R^2 \leq 0.47$). Disease control rate (DCR), defined as the sum of complete or partial response and stable disease, was tested in only one study (Kaufmann et al., 2017). No association between DCR and OS was highlighted ($R^2 \leq 0.16$). Finally, two new composite endpoints, especially designed for immunotherapy treatment were identified. Durable response rate (DRR), a combination of standard response criteria and a prospective duration dimension of 6 months, was highly associated with OS (Kaufman et al., 2017). In melanoma patients, achieving a durable response was associated with a statistically significant improvement in OS in all the landmark analyses (Hazard ratio (HR) ranged from 0.05 to 0.11). Another more complex binary composite endpoint, called “Intermediate response endpoint” (IME) has also been described (Gao et al., 2018). The IME response is defined on the basis of non-target lesion progression, new lesion, and target lesion information determined by baseline tumour burden, tumour reduction depth, and tumour change dynamic within one year after randomisation (Table 3). At the patient-level, IME responders had improved OS compared with non-responders (HR = 0.09 (95% CI [0.07–0.11])). At the trial-level, association between OS and IME was moderate ($R^2 = 0.68$ between the HR for OS and the odds ratio for IME).

In the four review articles selected, (Chen, 2015; Izar et al., 2017; Saad and Buyse, 2017; Korn and Freidlin, 2018) only two additional alternative endpoints were discussed for patients with advanced melanoma, namely treatment-free interval (TFI) and treatment-free survival (TFS). TFI, defined as the time since treatment interruption to disease progression and the need for further treatment, is used for the management of hematologic malignancies.

4. Discussion

To our knowledge, this is the first systematic literature review of alternative endpoints associated with OS in ICI-treated patients with cancer. Few references were identified (N = 24), even though the search algorithms were broad. This might be explained by the limited data available from RCT of ICIs due to the recent approval of these drugs. In light of the large number of publications identified on the prognostic role of PD-(L)1 or CTLA-4 expression (N = 858), research in this field has focused primarily on identification of subsets of responders rather than on surrogacy for OS. From our perspective, the surrogacy issue deserves as much interest, and could definitely be addressed by at least reporting specific data on OS and alternative endpoints.

Traditional alternative endpoints (PFS, ORR) as well as new promising composite ones (DRR, IME) were retrieved. PFS was the most studied alternative endpoint in ICI-treated patients with cancer. Historically, this has been demonstrated to be a valid SE for OS in some tumour types for other drug classes (Savina et al., 2018). Overall, we noticed a weak correlation between PFS and OS in ICI-treated patients. In a meta-analysis on the association between PFS and OS in the same population, no correlation between OS and PFS was found in terms of medians or gains in medians (Gyawali et al., 2018). Recently, Kaufman et al. concluded as well that PFS is an imperfect surrogate of OS according to the results of their meta-analysis performed on RCT of ICI (Kaufman et al., 2018). Pseudo-progression, even though relatively rare for certain tumour types (Wang et al., 2018), may partially explain this weak correlation between PFS and OS. Indeed, the effect of ICI as opposed to chemotherapy, is not on tumour cells, but on immune cells. Being treated by ICI, some patients experience immune-related

Table 3
Definition of the clinical endpoints identified in the systematic literature review.

Clinical alternative endpoint	Definition
Time to event endpoints	
Progression-free survival (PFS)	Time from randomization until the minimum between a disease progression or death (any cause)
OS Milestone rate	Kaplan-Meier survival probability at a time point defined a priori
OS Milestone ratio	Ratio of milestone rates between 2 treatment arms
Treatment-free interval (TFI)	Time since treatment interruption to disease progression and the need for re-initiation of the same treatment or initiation of another therapy
Treatment-free survival (TFS)	Time from end of therapy until need for next line treatment or death (the minimum)
Recurrence-free survival (RFS)	Time from randomization until recurrence of tumour or death from any cause (the minimum)
Categorical or continuous endpoints	
Objective response rate (ORR)	Proportion of patients with tumour size reduction of a predefined amount and for a minimum time period. ORR is composed as the sum of partial plus complete responses.
Disease control rate (DCR)	Proportion of patients with partial or complete responses to therapy or in stable disease and for a minimum time period
Tumour burden increase from baseline	Proportion of patients with < X% tumour burden increase from baseline, after a specific period of treatment and at a specific landmark time point
Depth of response (DepOR)	Percent tumour shrinkage at nadir, in comparison with baseline
Durable response rate (DRR)	Continuous response (complete response or partial response) beginning in the first 12 months of treatment and lasting 6 months or longer
Intermediate response endpoint (IME)	IME response is a binary endpoint (response or non-response) defined as satisfying all of the three following criteria: 1. The patient needed to be a target lesion responder, meaning the patient's target lesion score was less than an optimal cut-off value; 2. The patient had no unequivocal non-target lesion progression as determined per RECIST 1.1 criteria within 1 year; 3. The patient had no new unequivocal lesion as determined per RECIST 1.1 criteria within 1 year
Health-related quality of life (HRQoL)	Individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept affected in a complex way by the person's physical health, psychological state, personal beliefs, social relationships and their relationship to salient features of their environment *(World Health Organization definition)*.

responses such as initial increase in the size of tumours or appearance of new lesions, before a subsequent and sustained reduction in tumour burden occurs. This immunotherapy-specific phenomenon has stimulated the development of immune-related response criteria, the iRECIST (Seymour et al., 2017). Another explanation might be the residual efficacy of ICI for a longer duration (delayed treatment effect), these drugs affecting OS more than PFS even after treatment discontinuation (Gyawali et al., 2018). The poor correlation between ORR or DCR and OS may be due to these ICI-specific phenomena as well. Lastly, potential information provided by other known metrics, such as tumour burden increase, depth of response, health-related quality of life or treatment-free interval merit further investigation in this population. To date, few studies have investigated these alternative endpoints.

This work also highlights new clinical criteria, such as OS milestone rate/ratio, DRR and IME, not assessed for previous generations of therapy. A moderate to strong correlation between 1-year OS milestone ratio and OS was observed in metastatic non-small cell lung cancer patients (Blumenthal et al., 2017). Unfortunately, several type of treatments were aggregated in this study due to the low number of RCT of individual ICIs, which precluded formal validation of this endpoint as an SE. Notably, Petrelli et al. focused on ICI-treated melanoma patients and described a strong correlation between 1-year OS milestone rate and median OS (Petrelli et al., 2016). However, we face methodological concerns such as the impossibility of assessing surrogacy at the patient-level with the milestone rate, which is an aggregate measure. For this reason, this metric cannot be considered an adequate candidate SE. Moreover, even though milestone rates may present some advantages, another major limitation lies in the challenge of selecting the optimal milestone time point which may differ between cancer types and tumour stage (Hoss et al., 2013).

Based on the current published evidence, new composite endpoints, such as DRR and IME, statistically better correlate with OS in the metastatic cancer types considered. It seems to be more appropriate for capturing the unique pattern of antitumor response and survival with ICI treatments in the advanced setting. The main advantage of the suggested exploratory metrics is the inclusion of a duration component, which allows the persistence of the response to be taken into account, even though pre-specification of the time dimension is challenging. It also necessitates the use of specific statistical methods such as landmark analysis. However, duration of response should not be considered as the sole criterion for surrogacy. Emens et al. have argued that a composite endpoint which includes ORR and duration of response might best predict the effect of immunotherapy on long-term survival (Emans et al., 2017). Recently, Pfeiffer et al. considered both ORR and duration of response in their surrogacy assessment for OS in advanced non-small cell lung cancer patients (Peiffer et al., 2018). Combination of ORR and duration of response performed better as a surrogate for OS than duration of response alone.

Considerable heterogeneity in the statistical approaches used to assess surrogacy was noted in the studies evaluated. Only three research studies applied a statistical method for evaluating both patient-level and trial-level surrogacy. This two-stage approach is the only technique to validate an SE adequately, and is most robust when performed with a joint distribution function, as in the study of Suciu et al. (Shi et al., 2011; Suciu et al., 2018). This method ensures that correlation between endpoints measured on the same patient is taken into account. Only half of the studies identified in this review considered individual patient data. Difficulty in gaining access to individual patient data might be the main reason for not considering this type of analysis. Another major limitation of the available research is the pooling of tumour types, since validation of an SE should be performed within a given disease setting and for a specific drug class with a shared mechanism of action (Burzykowski et al., 2005). Thus, based on the literature published and from a methodological point of view, no validated SEs for OS are available today for the study of ICIs in clinical trials.

This study presents some limitations inherent to any systematic literature review. Publication bias stressed by Moher et al. may occur in this type of review as non-significant associations are less likely to be reported (Moher et al., 2009). Unfortunately, the impact of such bias could not be assessed. Two databases were consulted and one can assume that some references might not be included in these databases. We selected these databases as they are the two largest medical ones. Embase has the additional attractive characteristic to include major oncology conference proceedings (such as ESMO and ASCO). Keeping in mind these limitations, this work however is a valuable step forward in the context of assessing surrogacy in ICI-treated patients. To our knowledge, we have provided the first comprehensive list of alternative endpoints analysed for their association with OS in ICI-treated

populations, relying on a strong methodological approach. Results were reported according to international guidelines (Moher et al., 2009), and we did not restrict our investigations to primary research studies, but also considered general discussions of alternative endpoints. Repeating this study in upcoming years would be of interest, when more clinical studies of ICIs have been performed in the same or other cancer types and published.

As previously noted, endpoints other than OS are essential for regulatory approval of anticancer agents and it is of importance to identify novel surrogate for efficacy for ICI (Kaufman et al., 2018; Saad and Buyse, 2016). However, based on the current literature published, there is no sufficient data to support validated SE for OS. In ICI-treated patients at the metastatic stage, adequate surrogacy assessment of promising composite endpoints which take into account a duration component is encouraged.

Funding

The study was funded by Bristol-Myers Squibb. The funder contributed to the study design, the collection, the analysis, and the interpretation of the data in collaboration with the authors of this manuscript.

Declaration of interests

SB and AFG are employees of Bristol-Myers Squibb.

CB declares no competing interests.

AI has advisory board consulting with Epizyme, Lilly, Merck Sharp & Dohme, Novartis, Pharmamar and Roche. He has received research grants from Astra-Zeneca, Bayer, Bristol-Myers Squibb, Chugai, Merck Sharp & Dohme, Novartis, Pharmamar, Pfizer and Roche.

DR declares activity as independent methodological expert for Servier Laboratories.

VR declares activities as independent methodological expert for advisory boards for AstraZeneca, GlaxoSmithKline, Novartis and Medtronic.

Acknowledgments

We wish to sincerely thank Adam Doble from Foxymed company for having reviewed this manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.critrevonc.2019.02.013.

References

Abdel-Rahman, O., 2018. Surrogate end points for overall survival in trials of PD-(L)1 inhibitors for urinary cancers: a systematic review. Immunotherapy 10 (2), 139–148.

Alonso, A., Bigirumurame, T., Burzykowski, T., et al., 2017. Applied Surrogate Endpoint Evaluation Method with SAS and R. CRC Press, Boca Raton.

Blumenthal, G.M., Zhang, L., Zhang, H., et al., 2017. Milestone analyses of immune checkpoint inhibitors, targeted therapy, and conventional therapy in metastatic non–small cell lung cancer trials: a meta-analysis. JAMA Oncol. 3 (8), e171029. https://doi.org/10.1001/jamaoncol.2017.1029.

Burzykowski, T., Molenberghs, G., Buyse, M., 2005. The Evaluation of Surrogate Endpoints. Springer-Verlag, New-York.

Chen, T.-T., 2015. Milestone survival: a potential intermediate endpoint for immune checkpoint inhibitors. J. Natl. Cancer Inst. 107 (9). https://doi.org/10.1093/jnci/djv156.

Emans, L.A., Ascierto, P.A., Darcy, P.K., et al., 2017. Cancer immunotherapy: opportunities and challenges in the rapidly evolving clinical landscape. Eur. J. Cancer 81, 116–129.

Flaherty, K.T., Lee, S.J., Dummer, R., et al., 2013. A meta-analysis of randomized, controlled trials in metastatic melanoma establishes progression-free survival as a surrogate for overall survival. Eur. J. Cancer 49 (2), S856.

Fleming, T.R., Powers, J.H., 2012. Biomarkers and surrogate endpoints in clinical trials. Stat. Med. 31 (25), 2973–2984.

Fujii, T., Naing, A., Rolfo, C., Hajjar, J., 2018. Biomarkers of response to immune checkpoint blockade in cancer treatment. Crit. Rev. Oncol. Hematol. 130, 108–120.

Galluzzi, L., Vacchelli, E., Bravo-San Pedro, J.-M., et al., 2014. Classification of current anticancer immunotherapies. Oncotarget 5 (24), 12472–12508.

Gao, X., Zhang, L., Sridhara, R., 2018. Exploration of a novel intermediate response endpoint in immunotherapy clinical studies. Clin. Cancer Res. 24 (10), 2262–2267.

Gyawali, B., Shimokata, T., Ando, Y., 2017. Correlation and differences in effect sizes between progression free survival (PFS) and overall survival (OS) among PD-1 inhibitors. Ann. Oncol. 28 (5S), v403–v427.

Gyawali, B., Hey, S.P., Kesselheim, A.S., 2018. A comparison of response patterns for progression-free survival and overall survival following treatment for cancer with PD-1 Inhibitors. A meta-analysis of correlation and differences in effect sizes. JAMA Network Open 1 (2), e180416.

Hoss, A., Topalian, S., Chen, T.-T., et al., 2013. Facilitating the development of immunotherapies: intermediate endpoints for immune checkpoint modulators. Conference on Clinical Research. Available from: https://www.pharmamedtechbi.com/~/media/Supporting%20Documents/The%20Pink%20Sheet/75/46/FOCR_immunotherapy_issue_brief.pdf.

Izar, B., Regan, M.M., McDermott, D.F., 2017. Clinical trial design and endpoints for stage IV melanoma in the modern era. Cancer J. 23 (1), 63–67.

Kaufman, H.L., Andtbacka, R., Collichio, F.A., et al., 2017. Durable response rate as an endpoint in cancer immunotherapy: insights from oncolytic virus clinical trials. J. Immunother. Cancer 5 (72). https://doi.org/10.1186/s40425-017-0276-8.

Kaufman, H., Schwartz, L., William Jr, W., et al., 2018. Evaluation of classical endpoints as surrogates for overall survival in patients treated with immune checkpoint blockers: a systematic review and meta-analysis. J. Cancer Res. Clin. Oncol. https://doi.org/10.1007/s00432-018-2738-x.

Kaufmann, H., Schwartz, L.H., William, W.N., et al., 2017. Evaluation of clinical endpoints as surrogates for overall survival in patients treated with immunotherapies. J. Clin. Oncol. 35 (15s), e14557.

Kazandjian, D., Blumenthal, G., Khozin, S., et al., 2015. An exploratory responder analysis of best RECIST response and survival in patients with metastatic squamous NSCLC treated with nivolumab. J. Thorac. Oncol. 10 (9), S234.

Korn, E.L., Freidlin, B., 2018. Surrogate and intermediate endpoints in randomized trials: What's the goal? Clin. Cancer Res. 24 (10), 2239–2240.

Lee, D., Harvey, B., Gaudin, A.-F., et al., 2013. An exploratory analysis of the association between EORT C-QLQ C30 domains and progression free/overall survival in advanced melanoma after 12 weeks of treatment on ipilimumab compared to GP100 in a Phase III clinical trial. Value Health 16 (7), A422.

Martin-Liberal, J., Hierro, C., Ochoa de Olza, M., Rodon, J., 2017. Immuno-oncology: the third paragdim in early drug developement. Target. Oncol. 12 (2), 125–138.

Mathoulin-Pelissier, S., Gourgou-Bourgade, S., Bonnetain, F., Kramar, A., 2008. Survival end point reporting in randomized cancer clinical trials: a review of major journals. J. Clin. Oncol. 26 (22), 3721–3726.

McCoach, C.E., Blumenthal, G.M., Zhang, L., et al., 2017. Exploratory analysis of the association of depth of response and survival in patients with metastatic non-small-cell lung cancer treated with a targeted therapy or immunotherapy. Ann. Oncol. 28 (11), 2707–2714.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. PLoS Med. 6 (7), e1000097.

Mushti, S.L., Mulkey, F., Sridhara, R., 2018. Evaluation of overall response rate and progression-free survival as potential surrogate endpoints for overall survival in immunotherapy trials. Clin. Cancer Res. 24 (10), 2268–2275.

Nishino, M., Giobbie-Hurder, A., Manos, M.P., et al., 2017a. Immune-related tumor response dynamics in melanoma patients treated with pembrolizumab: identifying markers for clinical outcome and treatment decisions. Clin. Cancer Res. 23 (16), 4671–4679.

Nishino, M., Dahlberg, S.E., Adeni, A.E., et al., 2017b. Tumor response dynamics of advanced non–small cell lung cancer patients treated with PD-1 inhibitors: imaging markers for treatment outcome. Clin. Cancer Res. 23 (19), 5737–5744.

Nishino, M., Adeni, A., Lydon, C., et al., 2017c. Spectrum of early progression in advanced NSCLC patients treated with PD-1 inhibitors: identifying markers for poor outcome. J. Thorac. Oncol. 12 (11), S1915–16.

Ocana, A., Tannock, I.F., 2001. When are "positive" clinical trials in oncology truly positive? JNCI 103, 16–20.

Pazdur, R., 2008. Endpoints for assessing drug activity in clinical trials. Oncologist 13 (S2), 19–21.

Peiffer, B.M., Kulakova, M., Hashim, M., Heeg, B., 2018. Can duration of response be used as a surrogate endpoint for overall survival in advanced non-small cell lung cancer? J. Clin. Oncol. 36 (15), S9082.

Petrelli, F., Coinu, A., Cabiddu, A., et al., 2016. Early analysis of surrogate endpoints for metastatic melanoma in immune checkpoint inhibitor trials. Medicine 95 (26), e3997.

Ritchie, G., Gasper, H., Man, J., et al., 2018. Defining the most appropriate primary end point in phase 2 trials of immune checkpoint inhibitors for advanced solid cancers. A systematic review and meta-analysis. JAMA Oncol. 4 (4), 522–528.

Rolfo, C., Caglevic, C., Santarpia, M., et al., 2017. Immunotherapy in NSCLC: a promising and revolutionary weapon. Adv. Exp. Med. Biol. 995, 97–125.

Roviello, G., Andre, F., Venturini, S., et al., 2017. Response rate as a potential surrogate for survival and efficacy in patients treated with novel immune checkpoint inhibitors: a meta-regression of randomised prospective studies. Eur. J. Cancer 86, 257–265.

Saad, E.D., Buyse, M., 2016. Statistical controversies in clinical research: end points other than overall survival are vital for regulatory approval of anticancer agents. Ann. Oncol. 27 (3), 373–378.

Saad, E.D., Buyse, M., 2017. Exploratory analysis of the association of depth of response and survival in patients with metastatic non-small-cell lung cancer treated with a

targeted therapy or immunotherapy. Ann. Oncol. 28 (11), 2629–2630.
Savina, M., Gourgou, S., Italiano, A., et al., 2018. Meta-analyses evaluating surrogate endpoints for overall survival in cancer randomized trials: a critical review. Crit. Rev. Oncol. Hematol. 123, 21–41.
Seymour, L., Bogaerts, J., Perrone, A., et al., 2017. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. Lancet Oncol. 18 (3), e143–52.
Shi, Q., Renfro, L.A., Bot, B.M., Burzykowski, T., Buyse, M., Sargent, D.J., 2011. Comparative assessment of trial-level surrogacy measures for candidate time-to-event surrogate endpoints in clinical trials. Comput. Stat. Data Anal. 55, 2748–2757.
Shukuya, T., Mori, K., Amann, J.M., et al., 2016. Relationship between overall survival and response or progression-free survival in advanced non–small cell lung cancer patients treated with anti–PD-1/PD-L1 antibodies. J. Thorac. Oncol. 11 (11), 1927–1939.
Suciu, S., Eggermont, A., Lorigan, P., et al., 2018. Relapse-Free Survival as a surrogate for overall survival in the evaluation of stage II–III melanoma adjuvant therapy. J. Natl. Cancer Inst.
Tan, A., Porcher, R., Crequit, P., Ravaud, P., Dechartres, A., 2017. Differences in treatment effect size between overall survival and progression-free survival in immunotherapy trials: a meta-epidemiologic study of trials with results posted at ClinicalTrials.gOv. J. Clin. Oncol. 35 (15), 1686–1694.
Wang, Q., Gao, J., Wu, X., 2018. Pseudoprogression and hyperprogression after checkpoint blockade. Int. Immunopharmacol. 58, 125–135.