Original Research

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

Giandomenico Roviello [a,b,*], Fabrice Andre [c], Sergio Venturini [d], Barbara Pistilli [c], Giuseppe Curigliano [e], Massimo Cristofanilli [f], Pietro Rosellini [g], Daniele Generali [b,h]

[a] *Medical Oncology Unit, Department of Oncology, San Donato Hospital, Via Nenni 20, 52100 Arezzo, Italy*
[b] *Department of Medical, Surgery and Health Sciences, University of Trieste, Piazza Ospitale 1, 34129 Trieste, Italy*
[c] *Department of Medical Oncology, Gustave Roussy, Villejuif, France*
[d] *Centre for Research on Health and Social Care Management (CeRGAS), Bocconi University, Milan, Italy*
[e] *Early Drugs Development for Innovative Therapies Division, European Institute of Oncology, Milan, Italy*
[f] *Robert H Lurie Comprehensive Cancer Center, Feinberg School of Medicine, 710 North Fairbanks Ct, Suite 8-250a, Chicago, IL, USA*
[g] *Unit of Medical Oncology, Department of Medicine, Surgery and Neurosciences, University of Siena, Viale Bracci – Policlinico "Le Scotte", 53100 Siena, Italy*
[h] *Breast Cancer Unit and Translational Research Unit, ASST Cremona, Viale Concordia 1, 26100 Cremona, Italy*

KEYWORDS
Immune checkpoint;
Surrogate markers;
Survival;
Efficacy

Abstract ***Introduction:*** To assess the role of the tumour response rate (RR) after immune checkpoint inhibitors–based therapy as a potential surrogate end-point of progression-free survival (PFS) and overall survival (OS) in patients with solid tumours, we performed a trial-based meta-regression of randomised studies comparing different immune checkpoint inhibitors–based treatments.

Methods: The systematic literature search included the electronic databases and the proceedings of oncologic meetings. Treatment effects on PFS and OS were expressed as hazard ratios (HRs); treatment effects on RR were expressed as odds ratios (ORs). A weighted regression analysis was performed on log-transformed treatment effect estimates to test the association between treatment effects on the surrogate outcome and treatment effects on the clinical outcome.

* *Corresponding author*: Medical Oncology Unit, Department of Oncology, San Donato Hospital, Via Nenni 20, 52100 Arezzo, Italy.
E-mail address: giandomenicoroviello@hotmail.it (G. Roviello).

Results: Twenty-four trials, for a total of 11,894 patients, were included in the analysis. Using the complete set of data, the regression of either the log(HR) for PFS or the log(HR) for OS on the log(OR) for RR demonstrated weak associations ($R^2 = 0.47$; 95% confidence interval [CI], 0.03−0.77; $P = 0.001$; and $R^2 = 0.32$; 95% CI, 0.02−0.76; $P = 0.01$, respectively). The pre-planned analyses stratifying trials according to different type of disease and different mechanism of action of immune checkpoint inhibitors showed a very weak association of the RR with the OS for non−small cell lung cancer indicated and a modest association of the RR with the PFS for cytotoxic T lymphocyte−associated antigen 4 checkpoint inhibitors.
Conclusion: The results of the trial-based meta-regression analysis indicated a weak correlation between RR and OS, supporting future investigations to assess the surrogacy of RR in the patient treated with immune checkpoint inhibitors.

1. Introduction

Over the past years, the main role played by the host's immune system and the activation of the host's immune effectors against cancer was under a deep research focused on the discovery of the role of immunological agents that are able to directly or indirectly generate a potential immune response. Therefore, a new class of drugs targeting the tumour microenvironment has become available in clinic and has shown an unexpected efficacy [1]. The first identified target was the cytotoxic T lymphocyte−associated antigen 4 (CTLA-4) receptor, and ipilimumab is the first human monoclonal antibody against CTLA-4 [2]. Recently, the discovery of new antigens as the programmed cell death 1 (PD-1) in activated T cells and its ligand PD-L1 in tumour cells are the targets for the immunotherapeutic agent as nivolumab, already tested in patients with various types of advanced cancer [3]. Moreover, also pembrolizumab, another checkpoint inhibitor, blocking the binding of PD-1 and PD-L1 as well as PD-L2, showed a clear efficacy in several solid tumours [4]. Interestingly, with the exception of ipilimumab, nivolumab and pembrolizumab seem to demonstrate objective response in solid tumours other than melanoma or renal cancer [4]. Owing to the important clinical impact in several malignancies, new immune checkpoint inhibitors are in the advanced stages of clinical development [4]. Recent evidences indicated checkpoint-blocking monoclonal antibodies not only represent a promising means to induce robust and durable responses when employed as single agents [4−8] but can also be harnessed to boost the activity of several therapeutic regimens [9]. In this scenario, the identification of a potential surrogate marker for survival in patients under treatment with novel immune checkpoint inhibitors would represent an important advance in the early identification of patients' response/resistance to a potential active therapy.

The assessment of the tumour change induced by the treatment is an important issue for the clinical evaluation of activity and efficacy of the administered therapy. The Response Evaluation Criteria in Solid Tumours (RECIST) is a set of published rules defining when tumours in cancer patients respond or not during treatment [10]. According to the RECIST criteria, the complete response and the partial response defined the tumour response rate (RR) in phase III trials. Although, in many clinical trials the tumour RR has been used as a marker to guide treatment decisions on individual patients, the overall survival (OS) remained as the main end-point for phase III trials and for the evaluation of treatment efficacy. To assess the role of RR as a potential surrogate of the clinical outcomes in patients under treatment with the novel immune checkpoint inhibitors, we performed a systematic literature search and a trial-based meta-regression analysis of randomised controlled trials comparing different immune checkpoint inhibitors in different solid tumours and the survival outcomes. The aim of this study was to assess whether the treatment effects on tumour burden are able to predict the long-term effects on survival of patients receiving immune checkpoint inhibitors.

2. Methods

We planned a literature-based meta-regression analysis of randomised controlled trials based on the use of the novel immune checkpoint inhibitors. The primary objective of the meta-analysis was the validation of the RR as a potential surrogate end-point for efficacy and survival meant as progression-free survival (PFS) or OS in patients with solid tumours receiving the new immune checkpoint inhibitors.

2.1. Identification of randomised trials

A systematic search of the literature was conducted to identify all randomised trials based on the use of the novel immune checkpoint inhibitors in solid tumours. Relevant publications from PubMed and the Central Register of Controlled Trials of the Cochrane Library were identified. The search was focused on terms describing novel immune checkpoint inhibitors; therefore, following medical subject heading terms were used:

'nivolumab' OR 'ipilimumab' OR 'pembrolizumab' OR 'atezolizumab' AND 'solid tumours' AND 'prospective' OR 'clinical' OR 'random'. After the first search, article types were chosen as follows: 'clinical trial' and 'humans' was chosen in PubMed, and no restrictions were imposed in the Cochrane library. Publications available in these databases up to February 28, 2017 were analysed. To minimise the risk of selection or information bias, the search criteria were limited to articles reporting the results of phase II or phase III randomized controlled trials (RCT). The computer search was supplemented with a manual search of the primary studies referenced in all the retrieved review articles. When the results of a study were reported in subsequent analysis, only the most recent and complete version was included in this meta-analysis. The detailed flow diagram is reported in Fig. 1s Supplementary Data. No restriction for language was set.

2.2. Eligibility criteria

Two authors (BP and PR) screened the studies according to specific selection and exclusion criteria. Studies eligible for the meta-analysis had to report RR and the survival or efficacy outcomes in the different treatment arms. The inclusion/exclusion decisions regarding contentious studies were made in consultation with a third author (GR). The studies were identified according to the following inclusion criteria: 1) participants treated with novel immune checkpoint inhibitors; 2) the presence of a control arm for comparison without immune checkpoint inhibitors (chemotherapy or placebo) and 3) a primary outcome of OS and PFS. The following exclusion criteria were used: 1) insufficient data were available to estimate the outcomes; 2) the size of each arm was <10 participants and 3) non-randomised studies.

2.3. Data extraction

Two authors independently extracted the relevant data including the name of the first author, publication year, patient demographics (i.e. age, number, drug administered and type of tumours), median follow-up, median treatment duration, study design (i.e. the type of blinding, the type of control and the methods for randomisation), median OS and PFS and the number of patients who experienced a RR. For each trial, the arm with novel immune checkpoint inhibitors was considered to be the experimental arm and chemotherapy or the placebo control.

2.4. Statistical methods

The primary clinical outcomes were the PFS and OS. The hazard ratio (HR) of disease progression and death between the experimental arm and the control arm was used as the treatment effect estimated on the clinical outcome. The odds ratio (OR) of RR between the experimental and the control arms was used as the treatment effect estimated on the surrogate outcome. The proportion of patients with an RR per treatment arm was extracted from each single trial as the surrogate endpoint for the analysis. The statistical analysis consisted of a weighted regression on a logarithmic scale between treatment effects on the clinical outcomes and treatment effects on the RR. For trials with multiple arms, a contrast with the control arm was studied for any experimental arm (i.e. the treatment effect was estimated for any experimental arm versus the control arm).

A weighted regression analysis was performed on log-transformed treatment effect estimates [i.e. log(HR)] to test the association between treatment effects on the surrogate outcome (RR) and treatment effects on the clinical outcome (PFS or OS). The coefficient of determination (R2) was used to assess the goodness of fit for each model and to quantify the surrogacy level of RR [11–13]. 95% confidence intervals (CIs) of the R^2 values were estimated by bootstrap techniques. Moreover, due to the different classes of immune checkpoint inhibitors and the different types of considered tumours, a subgroup analysis was performed according to these last two subgroups. Data were acquired and analysed using the R software (version 3.3.2) and the 'boot' package (version 1.3–18) for the R^2 bootstrapping.

3. Results

3.1. Trials included in the analysis

We identified 910 citations after the systematic search (Fig. 1s). A total of 701 studies were excluded as duplicates. After viewing the titles and abstracts of the 109 remaining studies, the full texts of 46 studies were retrieved and 17 studies [14–30] were included in the analysis. The final set included a total of 8994 patients (4947 in the experimental arm and 4047 in the control arm) in the study. Three trials had two arms for a total of 20 analysed arms. Main characteristics of included trials are listed in Table 1. Five studies were on non–small cell lung cancer, four on melanoma, three on small-cell lung cancer, one on renal cancer, one on prostate cancer and one on head and neck cancer. Among the new experimental immune checkpoint inhibitors, nivolumab was the most represented drug. The RECIST criteria were the most used standard for classifying tumour responses in the included trials (Table 1). Data on PFS and OS along the median follow-up of the included studies are reported in Table 2.

3.2. Analysis

The RR ranged from 11% to 62% in the experimental group. Meanwhile, the RR ranged from 11% to 62% in the control arm. Notably, in the experimental arm, the

Table 1
Trials included in the analysis.

Study	Experimental regimen (number)	Control regimen (number)	Site	Design	Primary end-point	System for classifying response
Borghaei *et al.*, 2015	292	290	NSCLC	Nivolumab versus docetaxel	OS	RECIST 1.1
Brahmer *et al.*, 2015	135	137	NSCLC	Nivolumab versus docetaxel	OS	RECIST 1.1
Ferris *et al.*, 2016	240	121	Head & neck	Nivolumab versus methotrexate/ docetaxel/cetuximab	OS	RECIST 1.1
Herbst *et al.*, 2016	344	343	NSCLC	Pembrolizumab versus docetaxel	OS & PFS	RECIST 1.1
Herbst *et al.*, 2016 (2)	346	343	NSCLC	Pembrolizumab versus docetaxel	OS & PFS	RECIST 1.1
Know *et al.*, 2014	399	400	Prostate	Ipilimumab versus placebo	OS	Prostate cancer clinical trials working group's recommendations & RECIST 1.0
Langer *et al.*, 2016	60	63	NSCLC	Pembrolizumab + CHT versus CHT	Objective response	RECIST 1.1
Lynch *et al.*, 2012	68	66	NSCLC	Ipilimumab + CHT versus CHT	irPFS	mWHO & irRC
Lynch *et al.*, 2012 (2)	70	66	NSCLC	Ipilimumab + CHT versus CHT	irPFS	mWHO & irRC
Motzer *et al.*, 2015	410	411	RENAL	Nivolumab versus everolimus	OS	RECIST 1.1
Reck *et al.*, 2013	43	45	SCLC	Ipilimumab + CHT versus CHT	irPFS	mWHO & irRC
Reck *et al.*, 2013 (2)	42	45	SCLC	Ipilimumab + CHT versus CHT	irPFS	mWHO & irRC
Reck *et al.*, 2016	478	476	SCLC	Ipilimumab + CHT versus placebo + CHT	OS	mWHO
Reck *et al.*, 2016 NEJM	154	151	NSCLC	Pembrolizumab versus CHT	PFS	RECIST 1.1
Ribas *et al.*, 2013	328	327	Melanoma	Tremelimumab versus CHT	OS	RECIST 1.1
Ribas *et al.*, 2015	180	179	Melanoma	Pembrolizumab versus CHT	PFS	RECIST 1.1
Ribas *et al.*, 2015 (2)	181	179	Melanoma	Pembrolizumab versus CHT	PFS	RECIST 1.1
Robert *et al.*, 2015 NEJM	210	208	Melanoma	Nivolumab versus dacarbazine	OS	RECIST 1.1
Weber *et al.*, 2015	272	133	Melanoma	Nivolumab versus CHT	Objective response	RECIST 1.1
Bellmunt *et al.*, 2017	270	272	Urothelial	Pembrolizumab versus CHT	OS & PFS	RECIST 1.1
Rittmeyer *et al.*, 2017	425	425	NSCLC	Atezolizumab versus docetaxel	OS	RECIST 1.1

NSCLC, non−small cell lung cancer; CHT, chemotherapy; PFS, progression-free survival; OS, overall survival; WHO, World Health Organisation; RECIST, Response Evaluation Criteria in Solid Tumours.

RR was generally higher than that of the control arm. The higher RR reported in the experimental group was 62% for ipilimumab plus chemotherapy [24]; the lower RR was 4% with the use of tremelimumab [26].

We estimated the relationship between treatment effects on OS and treatment effects on RR and estimated a regression equation. We regressed the log(HR) for OS on the log(OR) for RR using the total number of events as the weights in the weighted regression, demonstrating a weak association between the two effects ($R^2 = 0.47$, $P = 0.001$). The equation of the resulting line (Fig. 1A) is the following: $\log(HR) = -0.1329 - 0.2575 \times \log(OR)$.

The R^2 value of the weighted regression line was 0.47 (95% CI, 0.03−0.77; $P = 0.001$), indicating that the 47% of the variability among the effects on OS can be explained by the observed effects on RR.

We estimated the relationship between treatment effects on PFS versus treatment effects on RR and estimated a regression equation. We regressed the log(HR) for PFS on the log(OR) for RR using the total number of events as the weights in the weighted regression, demonstrating a weak association between the two effects ($R^2 = 0.32$, $P = 0.01$). The equation of the resulting line (Fig. 1B) is the following: $\log(HR) = -0.1281 - 0.2384 \times \log(OR)$.

The R^2 value of the weighted regression line was 0.32 (95% CI, 0.02−0.76; $P = 0.01$), indicating that the 32% of the variability among the effects on PFS can be explained by the observed effects on RR.

The surrogacy of RR was also explored in the preplanned analyses that are stratifying trials according to different type of disease and different mechanism of action of the immune checkpoint inhibitors. Regarding the disease site, we performed the analysis separately by non−small cell lung cancer (NSCLC) only because of the limited samples available for the other type. As for the mechanism of action of the considered checkpoint inhibitors, we performed the analysis for the group of anti-CTLA-4 and the group of anti-PD-1/PD-L1 pathway.

The regression lines estimated for NSCLC indicated a very weak association of the RR with the OS ($R^2 = 0.0007$; 95% CI, 0.09−0.91; $P = 0.94$; Fig. 2A) and a weak association of the RR with the PFS ($R^2 = 0.42$; 95% CI, 0.003−0.85; $P = 0.06$; Fig. 2B). The regression lines estimated for the subgroup related to the use of the CTLA-4 checkpoint inhibitors indicated a very weak association of the RR with the OS ($R^2 = 0.00$; 95% CI, 0.00−0.97; $P = 0.96$; Fig. 2C) and a modest association of the RR with the PFS

Table 2
Value of overall survival, progression-free survival and response rate of the included studies.

Study	Median OS (months)		Median PFS (months)		Response rate (%)	
	Exp	Cont	Exp	Cont	Exp	Cont
Borghaei *et al.*, 2015	12.2 (9.7−15.0)	9.4 (8.1−10.7)	2.3 (2.2−3.3)	4.2 (3.5−4.9)	19 (15−24)	12 (9−17)
Brahmer *et al.*, 2015	9.2 (7.3−13.3)	6.0 (5.1−7.3)	3.5 (2.1−4.9)	2.8 (2.1−3.5)	20 (14−28)	9 (5−15)
Ferris *et al.*, 2016	7.5 (5.5−9.1)	5.1 (4.0−6.0)	2.0 (1.9−2.1)	2.3 (1.9−3.1)	13.3 (9.3−18.3)	5.8 (2.4−11.6)
Herbst *et al.*, 2016	10.4 (9.4−11.9)	8.5 (7.5−9.8)	3.9 (3.1−4.1)	4.0 (3.1−4.2)	18	9
Herbst *et al.*, 2016 (2)	12.7 (10.0−17.3)	8.5 (7.5−9.8)	4.0 (2.7−4.3)	4.0 (3.1−4.2)	18	9
Know *et al.*, 2014	11.2 (9.5−12.7)	10.0 (8.3−11.0)	4.0 (3.6−4.3)	3.1 (2.9−3.4)	13.1 (9.5−17.5)	5.2 (3.0−8.4)
Langer *et al.*, 2016	NR	NR	NR	NR	55 (42−68)	29 (18−41)
Lynch *et al.*, 2012	12.22 (9.26−14.39)	8.28 (6.80−12.39)	5.68 (4.76−7.79)	4.63 (4.14−5.52)	32 (22−45)	14 (6−24)
Lynch *et al.*, 2012 (2)	9.69 (7.59−12.48)	8.28 (6.80−12.39)	5.52 (4.17−6.74)	4.63 (4.14−5.52)	21 (13−33)	14 (6−24)
Motzer *et al.*, 2015	25.0 (21.8-ne)	19.6 (17.6−23.1)	4.6 (3.7−5.4)	4.4 (3.7−5.5)	25	5
Reck *et al.*, 2013	9.13 (6.67−12.98)	9.92 (8.64−11.73)	3.89 (2.89−5.85)	5.19 (4.40−5.59)	33 (19−49)	49 (34−64)
Reck *et al.*, 2013 (2)	12.94 (7.89−16.46)	9.92 (8.64−11.73)	5.22 (4.14−6.57)	5.19 (4.40−5.59)	57 (41−72)	49 (34−64)
Reck *et al.*, 2016	11.0 (10.5−11.3)	10.9 (10.0−11.5)	4.6 (4.5−5.0)	4.4 (4.4−4.6)	62 (58−67)	62 (58−67)
Reck *et al.*, 2016 NEJM	NR	NR	10.3 (6.7-ne)	6.0 (4.2−6.2)	44.8 (38.6−53.0)	27.8 (20.8−35.7)
Ribas *et al.*, 2013	12.6 (10.8−14.3)	10.7 (9.36−11.96)	20.3 (15.9−24.6) at 6 mo	18.1 (13.9−22.3) at 6 mo	10.7 (7.8−14.9)	9.8 (6.8−13.5)
Ribas *et al.*, 2015	NR	NR	2.9 (2.8−3.8)	2.7 (2.5−2.8)	21 (15−28)	4 (2−9)
Ribas *et al.*, 2015 (2)	NR		2.9 (2.8−4.7)	2.7 (2.5−2.8)	25 (19−32)	4 (2−9)
Robert *et al.*, 2015 NEJM	NR	10.8 (9.3−12.1)	5.1 (3.5−10.8)	2.2 (2.1−2.4)	40.0 (33.3−47.0)	13.9 (9.5−19.4)
Weber *et al.*, 2015	NR	NR	4.7 (2.3−6.5)	4.2 (2.1−6.3)	31.7 (23.5−40.8)	10.6 (3.5−23.1)
Bellmunt *et al.*, 2017	10.3 (8.0−11.8)	7.4 (6.1−8.3)	2.1 (2.0)	3.3 (2.2)	21.1 (16.4−26.5)	11.4 (7.9−15.8)
Rittmeyer *et al.*, 2017	13.8 (11.8−15.7)	9.6 (8.6−11.2)	2.8 (2.6−3.0)	4.0 (3.3−4.2)	14	13

OS, overall survival; PFS, progression-free survival; NR, not reported.

(R^2 = 0.67; 95% CI, 0.02; 1.00; P = 0.05; Fig. 2D). Finally, the regression lines estimated for the subgroup of PD-1/PD-L1 checkpoint inhibitors indicated a weak association of the RR with both OS (R^2 = 0.18; 95% CI, 0.00−0.97; P = 0.17; Fig. 2E) and PFS (R^2 = 0.25; 95% CI, 0.02−1.00; P = 0.08; Fig. 2F).

4. Discussion

To date, there is no consensus on the definition for valid surrogate end-points for therapeutic benefits in patients under treatment with immune checkpoint inhibitors. While objective response is usually end-points in patients treated with chemotherapy and/or targeted therapies, its value is not documented in patients treated with immune therapies. The purpose of this analysis was to identify potential surrogate end-points of response to the new immune-modulating drug to predict the long-term effect of the intervention on true end-point without having to wait and observe the true end-point through a strict and demonstrably correlation [31]. Unfortunately, the discovery of the novel immune checkpoint

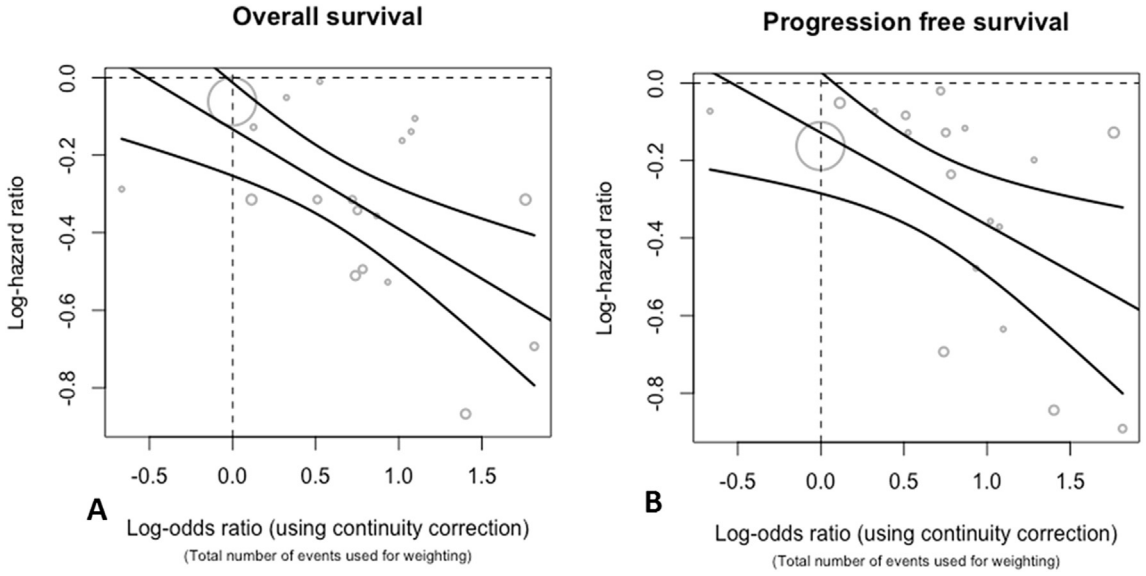

Fig. 1. Treatment effect on tumour response versus overall survival (**A**) and progression-free survival (**B**).

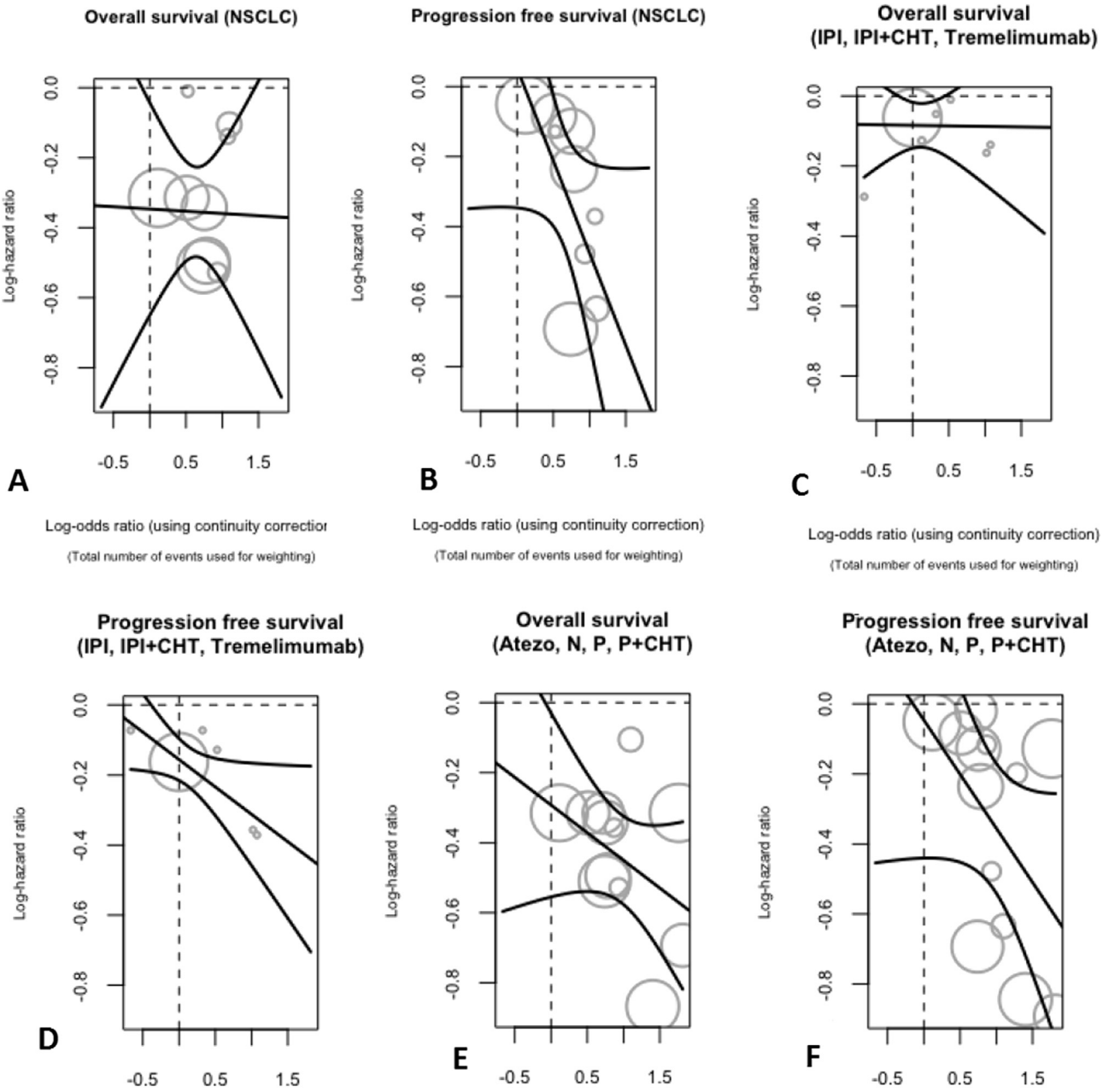

Fig. 2. Treatment effect on tumour response versus overall survival and progression-free survival for trials with non−small cell lung cancer (NSCLC) (**A** and **B**), anti-CTLA-4 (**C** and **D**), and anti-PD-L1−PD pathway (**E** and **F**). IPI, ipilimumab; N, nivolumab; CHT, chemotherapy.

inhibitors, which are revolutionising the treatment of solid tumours, is very recent and therefore few data are available for the prediction of response to novel checkpoint inhibitors. Historically, the World Health Organisation (WHO) and the RECIST criteria are an accepted approach for defining the disease response or progression to systemic therapies [10,32]. Therefore, we conducted a systematic search of all published randomised controlled trials on novel immune checkpoint inhibitors reporting information on both RR and survival to define the role of RR as a predictor of outcome for patients exposed to novel immune checkpoint inhibitors. With data extracted from a total of 17 trials, we estimated the regression equation of the log(HR) for the clinical outcome on the log(OR) for the RR. The estimated regression equation showed a weak relationship between the treatment effect on the RR and the treatment effect on the clinical outcomes as the PFS and the OS, suggesting that the activity of immune checkpoint inhibitors on RR is able to explain almost 50% of the effects detected on survival in patients with solid tumours. Considering the heterogeneity of the patients' cohort as the different type of solid tumours or treatment and follow-up in the considered trials for the

analysis, the results of the meta-regression seem to support the use of RR as a valid surrogate end-point to be used in place of the true clinical end-points only in a small subgroup of patients. Therefore, it might be possible that patients' responders to immune checkpoint inhibitors likely will achieve some disease control but, there are other factors that may help predicting the impact of survival (e.g. immune biomarkers? mutational tumour burden? and other clinical or biological factors).

The identification of the RR as surrogate end-point capturing the treatment effect on survival or efficacy for the checkpoint inhibitors–based treatment would be important for several reasons. First, the use of RR as a surrogate end-point may facilitate the earlier analysis of trial data, permitting to plan quicker and less expensive studies. Second, although novel immune checkpoint inhibitors have shown good efficacy in several tumours, no data are available on the optimal time and duration of these treatments and thus, the early evaluation of response may be useful to individuate those patients who are not responsive to immune checkpoint inhibitors in an earlier phase and address them towards other treatments.

Unfortunately, we are still waiting for a predictive biomarker of response. Until now, one of the most intriguing markers to identify potential responder to novel immune checkpoint inhibitors is the evaluation of the PD-L1 expression on formalin-fixed, paraffin-embedded tumour section by immunohistochemistry (IHC) [33]. However, the use of PD-L1 stained by IHC is confounded by multiple unresolved issues such as the variable detection antibodies, the differing cut-offs, the tissue preparation, the processing variability, the primary versus metastatic biopsies, the oncogenic versus induced PD-L1 expression and staining of tumour versus immune cells [34]. Nevertheless, emerging data are suggesting that patients, whose tumours are PD-L1+, have improved the clinical outcomes with anti-PD-1–directed therapy. On the counterpart, the presence of robust responses in some patients with low levels of PD-L1 expression put in doubt the role of PD-L1 as a predictive biomarker of response to the new immune-modulating agents [35]. Moreover, a sensitivity analysis aiming the evaluation of potential differential activity of three different immune checkpoint inhibitors according to the PD-L1 expression showed a significant interaction between tumour PD-L1 expression and overall sample with an RR of 34.1% in the PD-L1-positive and 19.9% in the PD-L1-negative population. In particular, the predictive value of PD-L1 on tumour cells seems to be more robust with the anti-PD-1 antibody (nivolumab and pembrolizumab), and with regards to the advanced melanoma and NSCLC [35]. Other surrogate end-points have been investigated such as baseline neutrophil-to-lymphocyte ratio or peripheral CD8 effector-memory type 1 T-cells [36,37] with an uncertainty on their use as surrogate markers for OS and prospective validations of these potential surrogates are required. Considering all these data together, the 50% rate of prediction on survival detected by our analysis seems to be relevant to define future studies on immune checkpoint inhibitors. In the near future, it might be possible that our data, along with mutational tumour burden and/or PD-L1 expression, can help build a predictive model for longitudinal evaluation of patients treated with immune therapies.

Immunotherapy induces different patterns of response compared with those related to chemotherapy, usually evaluated with the RECIST or WHO criteria [10,32]. To try to avoid possible differences, a consortium of approximately 200 oncologists, immunotherapists and regulatory experts developed the immune-related response criteria (irRC) [38]. The most important difference between the irRC and RECIST or WHO criteria is related to the concept of tumour burden, which replaces the concept of target lesion. Therefore, the changes are evaluated in all lesions, and the possible appearance of new lesions are considered in the context of the whole disease and not considered as a disease progression (the thresholds to define progression are different for irRC compared with RECIST or WHO criteria). Unfortunately, the use of irRC is not well standardised or largely used, and the majority of existing trials used irRC as a corollary to the RECIST criteria. However, in the absence of a consensus on this issue, the RECIST criteria together with clinical conditions are the most used criteria in clinical practice to guide treatment decisions for patients treated with immune checkpoint inhibitors.

Overall, analysing all the included studies, we observed a non-negligible significant association between either OS or PFS and response [the former measured using the log(HR) and the latter using the log(OR)]. This implies that the response can be considered as a reliable potential surrogate for overall or PFS. The results are qualitatively similar even if the total number of events observed or with the total sample sizes are used as the weights in the analysis.

However, splitting the study arms by site (NSCLC versus all other sites) or by treatment (Atezo, nivolumab, pembrolizumab (P), P + chemotherapy (CHT) versus ipilimumab (IPI), IPI + CHT, tremelimumab), very little, if any, emerges about the validity of the response as a potential surrogate for overall or PFS.

Our analysis presents a few limitations we should report: it was based on the literature rather than on individual patients' data; on different chemotherapeutic regimens, dosing and schedules reported in the various analysed trials, on mainly different locations of metastatic tumours and on different immune checkpoint inhibitors. Therefore, although our total sample size is large, we cannot yield safe and definitive conclusions. In addition, one of the mayor criticism is that some patients treated

with immune checkpoint inhibitors experienced an initial increased size of tumour lesions with subsequent decreased tumour burden. These findings of pseudoprogression would have been classified prematurely as progressive disease by RR in our analysed trials. However, the overall reported incidence of pseudoprogression in solid tumours is low [39]; therefore, we deem that it only marginally affects our analysis.

In conclusion, the results of this trial-based meta-regression analysis indicated a weak correlation between RR and OS; therefore, high and stringent criteria are necessary to validate a surrogate end-point, which can then be considered as a replacement end-point in future investigations.

Conflict of interest statement

No author has actual or potential conflicts of interest, including any financial, personal or other relationships with other people or organisations within three years of beginning the submitted work that could influence, or be perceived to influence, their work.

Appendix A. Supplementary data

References

[1] Wolchok JD, Kluger H, Callahan MK, Postow MA, Rizvi NA, Lesokhin AM, et al. Nivolumab plus ipilimumab in advanced melanoma. N Engl J Med 2013;369:122–33.

[2] Tarhini AA, Iqbal F. CTLA-4 blockade: therapeutic potential in cancer treatments. Onco Targets Ther 2010;3:15–25.

[3] Guo L, Zhang H, Chen B. Nivolumab as programmed death-1 (PD-1) inhibitor for targeted immunotherapy in tumor. J Cancer 2017 Feb 10;8(3):410–6.

[4] Martin-Liberal J, Ochoa de Olza M, Hierro C, Gros A, Rodon J, Tabernero J. The expanding role of immunotherapy. Cancer Treat Rev 2017 Mar;54:74–86.

[5] Zitvogel L, Kroemer G. Targeting PD-1/PD-L1 interactions for cancer immunotherapy. Oncoimmunology 2012;1:1223–5.

[6] Mavilio D, Lugli E. Inhibiting the inhibitors: checkpoints blockade in solid tumors. Oncoimmunology 2013;2, e26535.

[7] Lipson EJ, Sharfman WH, Drake CG, Wollner I, Taube JM, Anders RA, et-al. Durable cancer regression off-treatment and effective reinduction therapy with an anti-PD-1 antibody. Clin Cancer Res 2013;19:462–8.

[8] Cooper ZA, Frederick DT, Ahmed Z, Wargo JA. Combining checkpoint inhibitors and BRAF-targeted agents against metastatic melanoma. Oncoimmunology 2013;2, e24320.

[9] Vacchelli E, Prada N, Kepp O, Galluzzi L. Current trends of anticancer immunochemotherapy. Oncoimmunology 2013;2, e25396.

[10] Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45:228–47.

[11] Burzykowski T, Molenberghs G, Buyse M. The evaluation of surrogate endpoints. New York, NY: Springer; 2005.

[12] Gail MH, Pfeiffer R, Van Houwelingen HC, Carroll RJ. On meta-analytic assessment of surrogate outcomes. Biostatistics 2000;1:231–46.

[13] Chambers JM. Linear models. In: Chambers JM, Hastie TJ, editors. Statistical models in S. Pacific Grove, CA: Wadsworth & Brooks/Cole; 1992.

[14] Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, et al. Nivolumab versus docetaxel in advanced non-squamous non–small-cell lung cancer. N Engl J Med 2015;373:1627–39.

[15] Brahmer J, Reckamp KL, Baas P, Crinò L, Eberhardt WE, Poddubskaya E, et al. Nivolumab versus docetaxel in advanced squamous-cell non–small-cell lung cancer. N Engl J Med 2015;373:123–35.

[16] Ferris RL, Blumenschein Jr G, Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab for recurrent squamous-cell carcinoma of the head and neck. N Engl J Med 2016;375:1856–67.

[17] Herbst RS, Baas P, Kim DW, Felip E, Pérez-Gracia JL, Han JY, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. Lancet 2016;387:1540–50.

[18] Kwon ED, Drake CG, Scher HI, Fizazi K, Bossi A, van den Eertwegh AJ, et al. Ipilimumab versus placebo after radiotherapy in patients with metastatic castration-resistant prostate cancer that had progressed after docetaxel chemotherapy (CA184-043): a multicentre, randomised, double-blind, phase 3 trial. Lancet Oncol 2014;15:700–12.

[19] Langer CJ, Gadgeel SM, Borghaei H, Papadimitrakopoulou VA, Patnaik A, Powell SF, et al. Carboplatin and pemetrexed with or without pembrolizumab for advanced, non-squamous non-small-cell lung cancer: a randomised, phase 2 cohort of the open-label KEYNOTE-021 study. Lancet Oncol 2016;17:1497–508.

[20] Lynch TJ, Bondarenko I, Luft A, Serwatowski P, Barlesi F, Chacko R, et al. Ipilimumab in combination with paclitaxel and carboplatin as first-line treatment in stage IIIB/IV non–small-cell lung cancer: results from a randomized, double-blind, multicenter phase II study. J Clin Oncol 2012;30:2046–54.

[21] Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. N Engl J Med 2015;373(19):1803–13.

[22] Lynch TJ, Bondarenko I, Luft A, Serwatowski P, Barlesi F, Chacko R, et al. Ipilimumab in combination with paclitaxel and carboplatin as first-line therapy in extensive disease-small-cell lung cancer: results from a randomized, double-blind, multicenter phase 2 trial. Ann Oncol 2013;24:75–83.

[23] Reck M, Luft A, Szczesna A, Havel L, Kim SW, Akerley W, et al. Phase III randomized trial of ipilimumab plus etoposide and platinum versus placebo plus etoposide and platinum in extensive-stage small-cell lung cancer. J Clin Oncol 2016;34:3740–8.

[24] Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, et al. Pembrolizumab versus chemotherapy for PD-L1–positive non–small-cell lung cancer. N Engl J Med 2016;375:1823–33.

[25] Ribas A, Kefford R, Marshall MA, Punt CJ, Haanen JB, Marmol M, et al. Phase III randomized clinical trial comparing tremelimumab with standard-of-care chemotherapy in patients with advanced melanoma. J Clin Oncol 2013;31:616–22.

[26] Ribas A, Puzanov I, Dummer R, Schadendorf D, Hamid O, Robert C, et al. Pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory melanoma (KEYNOTE-002): a randomised, controlled, phase 2 trial. Lancet Oncol 2015;16:908–18.

[27] Robert C, Long GV, Brady B, Dutriaux C, Maio M, Mortier L, et al. Nivolumab in previously untreated melanoma without BRAF mutation. N Engl J Med 2015;372:320–30.

[28] Weber JS, D'Angelo SP, Minor D, Hodi FS, Gutzmer R, Neyns B, et al. Nivolumab versus chemotherapy in patients with

advanced melanoma who progressed after anti-CTLA-4 treatment (CheckMate 037): a randomised, controlled, open-label, phase 3 trial. Lancet Oncol 2015;16:375–84.
[29] Bellmunt J, de Wit R, Vaughn DJ, Fradet Y, Lee JL, Fong L, et al. Pembrolizumab as second-line therapy for advanced urothelial carcinoma. N Engl J Med 2017 Mar 16;376(11):1015–26. https://doi.org/10.1056/NEJMoa1613683. Epub 2017 Feb 17.
[30] Rittmeyer A, Barlesi F, Waterkamp D, Park K, Ciardiello F, von Pawel J, et al. Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial. Lancet 2017;389:255–65.
[31] Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Stat Med 1989;8:431–40.
[32] Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, 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 2000;92:205–16.
[33] Mahoney KM, Atkins MB. Prognostic and predictive markers for the new immunotherapies. Oncology 2014;28(Suppl. 3):39–48.
[34] Patel SP, Kurzrock R. PD-L1 expression as a predictive biomarker in cancer immunotherapy. Mol Cancer Ther 2015;14(4):847–56.
[35] Carbognin L, Pilotto S, Milella M, Vaccaro V, Brunelli M, Caliò A, et al. Differential activity of nivolumab, pembrolizumab and MPDL3280A according to the tumor expression of programmed death-ligand-1 (PD-L1): sensitivity analysis of trials in melanoma, lung and genitourinary cancers. PLoS One 2015 Jun 18;10(6), e0130142.
[36] Ferrucci PF, Gandini S, Battaglia A, Alfieri S, Di Giacomo AM, Giannarelli D, et al. Baseline neutrophil-to-lymphocyte ratio is associated with outcome of ipilimumab-treated metastatic melanoma patients. Br J Cancer 2015 Jun 9;112(12):1904–10.
[37] Wistuba-Hamprecht K, Martens A, Heubach F, Romano E, Geukes Foppen M, Yuan J, et al. Peripheral CD8 effector-memory type 1 T-cells correlate with outcome in ipilimumab-treated stage IV melanoma patients. Eur J Cancer 2017 Mar;73:61–70.
[38] Fojo AT, Noonan A. Why RECIST works and why it should stay: counterpoint. Cancer Res 2012;72:5151–7.
[39] Chiou VL, Burotto M. Pseudoprogression and immune-related response in solid tumors. J Clin Oncol 2015 Nov 1;33(31):3541–3.