



Original Investigation | Oncology

Correlation of Milestone Restricted Mean Survival Time Ratio With Overall Survival Hazard Ratio in Randomized Clinical Trials of Immune Checkpoint Inhibitors

A Systematic Review and Meta-analysis

Zi-Xian Wang, MD; Hao-Xiang Wu, MD; Li Xie, MD; Ying-Nan Wang, MD; Lu-Ping Yang, MD; Ming-Ming He, MD; Hui-Yan Luo, MD; Pei-Rong Ding, MD, PhD; Dan Xie, MD, PhD; Gong Chen, MD, PhD; Yu-Hong Li, MD, PhD; Feng Wang, MD, PhD; Rui-Hua Xu, MD, PhD

Abstract

IMPORTANCE Immune checkpoint inhibitors (ICIs) have unique patterns of response and survival that differ from conventional chemotherapies. Novel intermediate end points are urgently required to detect the early signals of ICI activity.

OBJECTIVE To evaluate milestone rate (Kaplan-Meier estimates of survival probabilities at given time points) and milestone restricted mean survival time (RMST, the area under survival curves up to given time points) as potential intermediate end points for ICI trials.

DATA SOURCES Electronic databases (pre-MEDLINE, MEDLINE, Embase, and the Cochrane Central Register of Controlled Trials) were searched for randomized clinical trials published between January 1, 2000, and December 31, 2017.

STUDY SELECTION Phase 2 and phase 3 randomized clinical trials evaluating ICIs in advanced solid tumors.

DATA EXTRACTION AND SYNTHESIS Two investigators extracted the data and reconstructed individual patient data to estimate the milestone rate or milestone RMST from the published Kaplan-Meier curves.

MAIN OUTCOMES AND MEASURES Trial-level milestone rates or milestone RMSTs were estimated for 6-month and 9-month progression-free survival (PFS) and 9-month and 12-month overall survival (OS). A weighted linear regression model evaluated the correlations of ratios of milestone rates or milestone RMSTs with OS hazard ratios (HRs).

RESULTS Twenty-six trials examining 8 different tumor types were identified, including 12 892 patients. Overall survival HR was correlated with the ratio of 9-month OS milestone rate ($R^2 = 0.45$; 95% CI, 0.27-0.74), and with the ratio of 12-month OS milestone rate ($R^2 = 0.40$; 95% CI, 0.22-0.70). The ratio of 9-month OS milestone RMST ($R^2 = 0.60$; 95% CI, 0.28-0.74) and ratio of 12-month OS milestone RMST were correlated with OS HR ($R^2 = 0.64$; 95% CI, 0.42-0.78). No correlations were observed between OS HR and the ratio of 6-month or 9-month PFS milestone rates or milestone RMSTs.

CONCLUSIONS AND RELEVANCE Ratios of OS milestone RMSTs had a stronger correlation with OS HRs than ratios of OS milestone rates, whereas ratios of PFS milestone rates and ratios of PFS

(continued)

Key Points

Question What intermediate end points could be useful in randomized clinical trials studying immune checkpoint inhibitors?

Findings In this systematic review and meta-analysis of 26 trials studying immune checkpoint inhibitors in 12 892 participants, the ratio of milestone restricted mean survival time for overall survival was more strongly correlated with the overall survival hazard ratio than the ratio of overall survival milestone rates.

Meaning Milestone restricted mean survival time could be studied as a potential intermediate end point for overall survival in future trials of immune checkpoint inhibitors.

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

Open Access. This is an open access article distributed under the terms of the CC-BY License.

Abstract (continued)

milestone RMSTs were not correlated with OS HRs. The OS milestone RMST could be further studied as an intermediate end point in future ICI trials.

JAMA Network Open. 2019;2(5):e193433. doi:10.1001/jamanetworkopen.2019.3433

Introduction

There have been major advances in cancer immunotherapies in the last 10 years, among which the most promising approach is the use of immune checkpoint inhibitors (ICIs) to activate self-immunity toward tumors.¹ The broad efficacy of ICIs for various malignant tumors has led to unprecedented levels of research and development of these agents.² Meanwhile, a unique pattern of response, progression, and survival among patients who received ICIs that differs from conventional chemotherapy or targeted therapy has stimulated concern.³ The delayed clinical effects and long-term survival demonstrated by ICI treatments could lead to substantially extended study duration and loss of statistical power if these 2 unique features are not accounted for in the study design and statistical analyses.⁴ Additionally, several previous studies have shown no association of the odds ratio for the objective response rate or the progression-free survival (PFS) hazard ratio (HR) with the overall survival (OS) HR, raising questions for the use of these traditional surrogate end points in ICI trials.⁵⁻⁷

With the statistical issues and challenges posed by ICIs, researchers are investigating potential intermediate end points for ICI trials, including milestone survival, defined as the Kaplan-Meier estimate of survival probability at a given time point.⁸ A 2017 study of milestone survival⁹ suggested that the 12-month OS rate could be used as a potential intermediate end point in ICI trials for non-small cell lung cancer (NSCLC). However, that investigation was based on the analysis of trials pooled by different therapeutic classes, including conventional therapy, targeted therapy, and immunotherapy. Furthermore, drawbacks of milestone survival include the cross-sectional assessment at a given time point, the inability to account for the totality of the survival period, and the effect of censoring before the milestone time point.

The restricted mean survival time (RMST) is an alternative treatment outcome measure that can be estimated as the area under the survival curve up to a prespecified time horizon and hence can account for all survival information before that time horizon.^{10,11} In this study, we used RMST to measure milestone treatment effect and assessed ratios of milestone RMSTs against ratios of milestone rates as a potential intermediate end point for ICI trials.

Methods

This study was conducted in compliance with the recommendations of the Cochrane Handbook for Systematic Reviews of Interventions and reported based on Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline.

Selection of Randomized Trials

We searched pre-MEDLINE, MEDLINE, Embase, and the Cochrane Central Register of Controlled Trials for randomized clinical trial results published from January 1, 2000, to December 31, 2017. Analysis began in June 2018. We combined Medical Subject Headings and free-text terms to identify relevant studies. The full search strategy is detailed in the eMethods in the Supplement. Inclusion criteria included reports of randomized clinical trials of ICIs that included a Kaplan-Meier curve for OS. Phase 1, single-arm phase 2, dose-finding, and adjuvant setting trials were excluded. News, editorials, letters, commentaries, retrospective studies, review articles, and secondary analyses of randomized clinical trials were also excluded. For multiple-arm trials, all comparisons were included. The intention-to-treat population was included except for 1 trial, in which the population with

programmed cell death 1 (PDCD1) ligand 1 expression in 5% or more of tumor-infiltrating immune cells was selected according to the trial design.¹²

Data Extraction and Reconstruction of Individual Patient Data

Two of us (Z.-X.W. and H.-X.W.) screened the trials independently for eligibility and extracted the following information from each included trial: first author's name, year of publication, tumor type, treatment regimen in both arms, phase and treatment line of the trial, primary or coprimary end point, minimum follow-up, and sample size. Any discrepancies were resolved by consensus.

To estimate the milestone survival rate and milestone RMST, we reconstructed individual patient data for each group from the published Kaplan-Meier curves.¹³ Digitizelt software version 2.2 (Digitizelt) was used to measure the time and survival probability coordinates on the Kaplan-Meier curves. The number of patients at risk and the total number of events were extracted when available. Next, the data were entered into an algorithm on the basis of iterative numerical methods to solve the inverted Kaplan-Meier equations. All data were reconstructed by one of us (H.-X.W.) and validated by another (Z.-X.W.).

Outcome Measurement

Four trial-level milestones were chosen for analysis: 6-month PFS, 9-month PFS, 9-month OS, and 12-month OS. We calculated the milestone rates using Kaplan-Meier estimates. The milestone ratio was defined as the ratio of milestone rates between 2 treatment arms. We assessed the milestone RMSTs in the experimental and control groups by prespecifying the time horizon as 6, 9, or 12 months.^{14,15} The milestone RMST from the Kaplan-Meier estimate of the survival function was determined, and the ratio of milestone RMST was estimated. The associated variances were estimated using the δ method.¹⁶

Statistical Analysis

Nonproportional hazards were tested using the Grambsch-Therneau test, and a 2-tailed *P* less than .10 indicated a statistically significant violation of the proportional hazard assumption.¹⁷ By pooling the reconstructed individual patient data of the included ICI trials, Kaplan-Meier analyses of the treatment arm vs the control were performed to investigate the survival kinetics among the pooled cohort.

The correlations of treatment effects measured by the ratios of the milestone rate or milestone RMST with the HR were evaluated using weighted linear regression models, with weights equal to the sample size of each randomized comparison. The coefficient of determination (R^2) and 95% CIs from the weighted linear regression model were used to measure strength of the correlations. The 95% CIs of R^2 were obtained using the bootstrap method with 1000 replications. R^2 equal to 0.80 was chosen prospectively as the cutoff value to establish the milestone rate or RMST as validated intermediate end points for the OS HR.⁵

To assess whether any trial was more influential in the trial-level correlation of the OS HR with the ratio of 9-month or 12-month OS milestone RMST, a leave-1-out cross-validation was performed by excluding 1 comparison at a time. Statistical analyses were performed using R statistical software version 3.5.1 (R Project for Statistical Computing), and the survRM2 package was used to derive the milestone RMST.

Results

Twenty-six eligible randomized clinical trials studying ICIs were identified (**Figure 1** and **Table**), including 31 treatment comparisons and 12 892 patients. Twenty trials (77%) were phase 3 studies, and 6 trials (23%) were phase 2 studies. The 26 trials examined 8 tumor types, including 9 on NSCLC (35%) and 8 on melanoma (30%). There were 12 trials (46%) that examined PDCD1 inhibitor monotherapy (8 with nivolumab and 4 with pembrolizumab), 3 trials (12%) of PDCD1 ligand 1

inhibitor monotherapy (atezolizumab), 3 trials (12%) of cytotoxic T lymphocyte-associated antigen 4 inhibitor monotherapy (2 with ipilimumab and 1 with tremelimumab), 7 trials (27%) of a checkpoint inhibitor and chemotherapy or vaccine combination (4 with ipilimumab and chemotherapy, 2 with ipilimumab and vaccine, and 1 with pembrolizumab and chemotherapy), and 1 trial of a PD1 inhibitor and cytotoxic T lymphocyte-associated antigen 4 inhibitor combination (nivolumab and ipilimumab). Twelve trials (46%) were first-line studies, and the remaining 14 trials (54%) were second-line studies or beyond. Most of the studies set OS as the primary or coprimary end point (20 [77%]). In most of the trials, PFS was determined by Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1.⁴³ At the time of the database lock of each trial, the minimum follow-up duration range was 4.63 months to 39.60 months with a median (interquartile range) of 12.17 (7.80-18.22) months. Detailed regimen information and other trial characteristics are summarized in eTable 1 in the [Supplement](#). eTable 2 and eTable 3 in the [Supplement](#) present the reconstructed individual patient data demonstrating estimates that are close to the HRs or median survival times reported in the original articles.

A delayed clinical effect of the ICIs was observed in most trials. When visually assessed, 25 OS curves (81%) and 26 PFS curves (74%) were inseparable in the first 3 months of the 31 examined treatment comparisons. In the pooled cohort of 12 892 patients, the OS curve and the PFS curve were intertwined at the beginning but began to separate between 3 to 6 months (eFigure in the [Supplement](#)).

Figure 2 shows the scatterplots of the treatment effects, illustrating trial-level correlations among the end points. The OS HR was weakly correlated with the ratio of the 6-month PFS milestone rate ($R^2 = 0.12$; 95% CI, 0-0.30) (Figure 2A) and with the ratio of the 9-month PFS milestone rate ($R^2 = 0.19$; 95% CI, 0-0.44) (Figure 2B). However, the OS HR was correlated with the ratio of the 9-month OS milestone rate ($R^2 = 0.45$; 95% CI, 0.27-0.74) (Figure 2C) and the ratio of the 12-month OS milestone rate ($R^2 = 0.40$; 95% CI, 0.22-0.70) (Figure 2D).

When the RMST was incorporated as a milestone treatment effect measurement, the OS HR correlated weakly with the 6-month PFS milestone RMST ratio ($R^2 = 0.08$; 95% CI, 0-0.34)

Figure 1. PRISMA Flowchart of Study Inclusions and Exclusions

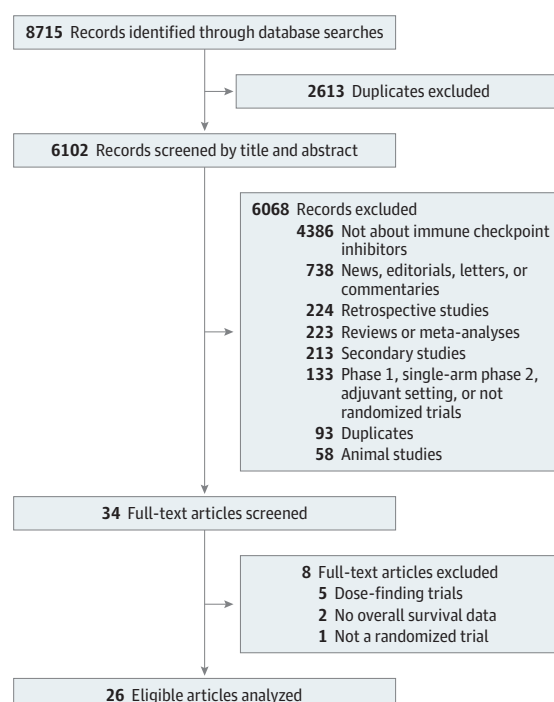


Table. Summary of Trials Included

Trial	Tumor Type	Experimental Arm(s)	Control Arm	Primary End Point	Phase	Line	Minimum Follow-up, mo	No. of Patients
Robert et al, ¹⁸ 2015	Melanoma	Nivolumab	Dacarbazine	OS	3	First	6.17	418
Hodi et al, ¹⁹ 2016	Melanoma	Nivolumab + ipilimumab	Ipilimumab	ORR	2	First	25.1	142
Larkin et al, ²⁰ 2018	Melanoma	Nivolumab	Physician's choice chemotherapy	OS/ORR	3	≥Second	26.97	405
Robert et al, ²¹ 2015	Melanoma	A: Pembrolizumab once every 2 wk; B: pembrolizumab once every 3 wk	Ipilimumab	OS/PFS	3	First/second	12.17	834
Hodi et al, ²² 2010	Melanoma	A: Ipilimumab + glycoprotein 100; B: ipilimumab	Glycoprotein 100	OS	3	≥Second	8.43	676
Robert et al, ²³ 2011	Melanoma	Ipilimumab + dacarbazine	Dacarbazine	OS	3	First	37.1	502
Hodi et al, ²⁴ 2014	Melanoma	Ipilimumab + sargramostim	Ipilimumab	OS	2	≥Second	17.37	245
Ribas et al, ²⁵ 2013	Melanoma	Tremelimumab	Physician's choice chemotherapy	OS	3	First	39.6	655
Fehrenbacher et al, ²⁶ 2016	NSCLC	Atezolizumab	Docetaxel	OS	2	≥Second	13.43	287
Rittmeyer et al, ²⁷ 2017	NSCLC	Atezolizumab	Docetaxel	OS	3	≥Second	19.57	850
Borghaei et al, ²⁸ 2015	NSCLC	Nivolumab	Docetaxel	OS	3	Second	15.73	582
Brahmer et al, ²⁹ 2015	NSCLC	Nivolumab	Docetaxel	OS	3	Second	12.63	272
Carbone et al, ³⁰ 2017	NSCLC	Nivolumab	Physician's choice chemotherapy	PFS	3	First	16.3	423
Herbst et al, ³¹ 2016	NSCLC	A: Pembrolizumab 2 mg/kg; B: pembrolizumab 10 mg/kg	Docetaxel	OS/PFS	2-3	≥Second	7.17	1034
Langer et al, ³² 2016	NSCLC	Pembrolizumab + carboplatin + pemetrexed	Carboplatin + pemetrexed	ORR	2	First	6.53	123
Reck et al, ³³ 2016	NSCLC	Pembrolizumab	Physician's choice chemotherapy	PFS	3	First	6.43	305
Lynch et al, ³⁴ 2012	NSCLC	A: Concurrent ipilimumab + carboplatin + paclitaxel; B: phased ipilimumab + carboplatin + paclitaxel	Paclitaxel + carboplatin	irPFS	2	First	19.07	204
Kang et al, ³⁵ 2017	EG	Nivolumab	Placebo	OS	3	Third	5.63	493
Ferris et al, ³⁶ 2016	HNSCC	Nivolumab	Physician's choice chemotherapy	OS	3	≥Second	4.63	361
Kwon et al, ³⁷ 2014	Prostate cancer	Ipilimumab	Placebo	OS	3	Second	11.9	799
Beer et al, ³⁸ 2017	Prostate cancer	Ipilimumab	Placebo	OS	3	First	24	602
Motzer et al, ³⁹ 2015	Renal cell carcinoma	Nivolumab	Everolimus	OS	3	Second/third	15.23	821
Reck et al, ⁴⁰ 2013	SCLC	A: Concurrent ipilimumab + carboplatin + paclitaxel; B: phased ipilimumab + carboplatin + paclitaxel	Paclitaxel + carboplatin	irPFS	2	First	11.1	130
Reck et al, ⁴¹ 2016	SCLC	Ipilimumab + etoposide + platinum	Etoposide + platinum	OS	3	First	9.4	954
Powles et al, ⁴² 2018 ^a	Urothelial carcinoma	Atezolizumab	Physician's choice chemotherapy	OS	3	Second/third	13.07	234
Bellmunt et al, ⁴² 2017	Urothelial carcinoma	Pembrolizumab	Physician's choice chemotherapy	OS/PFS	3	Second	9.97	542

Abbreviations: EG, gastric or gastroesophageal junction cancer; HNSCC, squamous cell carcinoma of the head and neck; irPFS, immune-related progression-free survival; NSCLC, non-small cell lung cancer; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; SCLC, small cell lung cancer

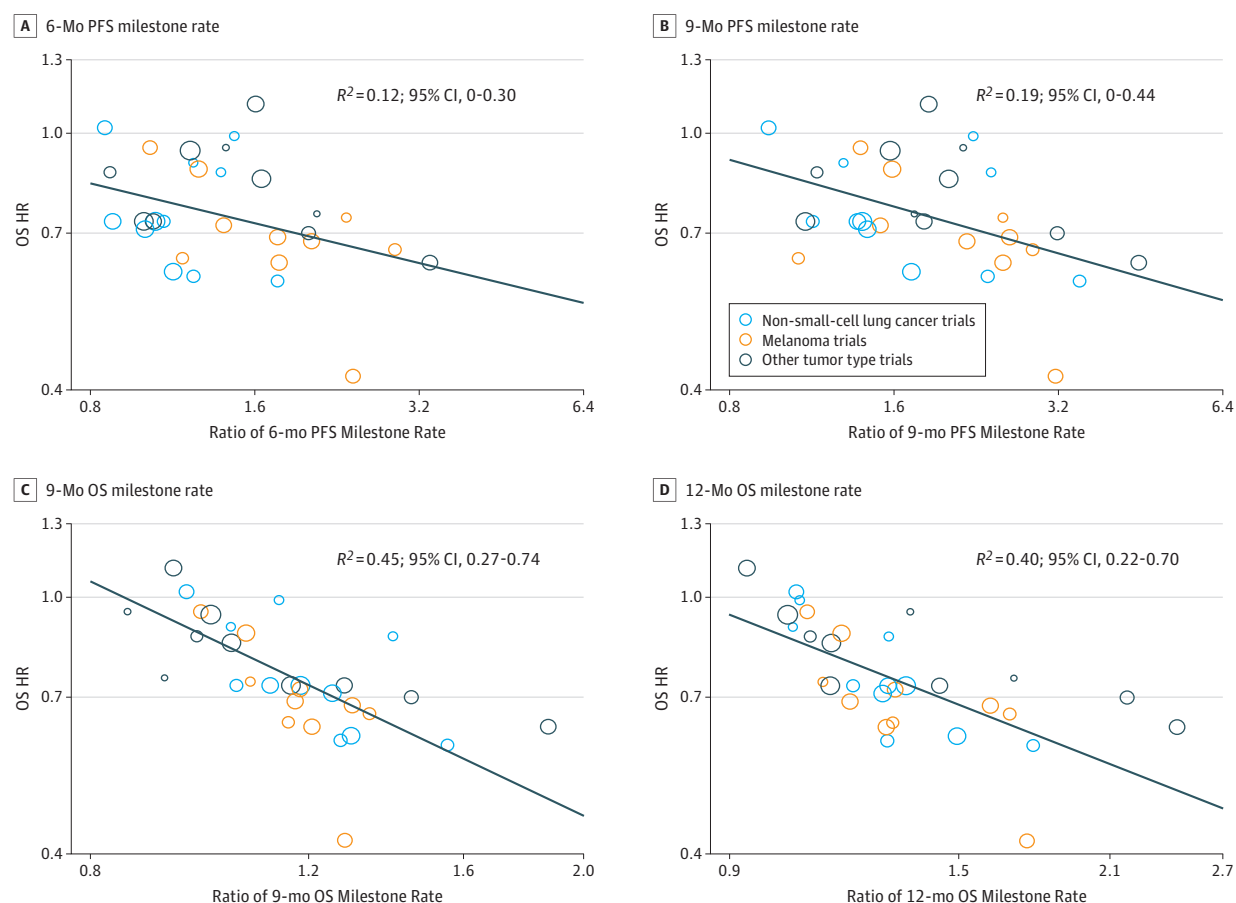
^a The population with programmed cell death 1 ligand expression in 5% or less of tumor-infiltrating immune cells was included according to the trial design.

(Figure 3A) and with the 9-month PFS milestone RMST ratio ($R^2 = 0.13$; 95% CI, 0-0.37) (Figure 3B). However, compared with the ratio of the 9-month or 12-month OS milestone rates, the ratio of the 9-month OS milestone RMST exhibited a strong correlation with the OS HR ($R^2 = 0.60$; 95% CI, 0.28-0.74) (Figure 3C), and the correlation of the OS HR with the 12-month milestone RMST ratio was even stronger ($R^2 = 0.64$; 95% CI, 0.42-0.78) (Figure 3D). In the leave-1-out cross-validation analysis, the median (range) R^2 was 0.59 (0.57-0.64) for the 9-month OS milestone RMST and 0.64 (0.62-0.71) for the 12-month OS milestone RMST.

To provide a more reliable calculation of the OS milestone RMST, we assessed the correlation of the OS HR with the OS milestone RMST ratio by limiting the analysis to trials that set OS as the primary or coprimary end point and that had a minimum follow-up longer than 9 months (for the 9-month OS milestone RMST calculation) or 12 months (for the 12-month OS milestone RMST calculation). There were 16 comparisons that met criteria for the 9-month OS milestone RMST calculation and 13 comparisons that met the criteria for the 12-month OS milestone RMST calculation. In the 9-month comparison group, there was a slight improvement in the correlation of the OS HR with the ratio of the 9-month OS milestone RMST ($R^2 = 0.62$; 95% CI, 0.09-0.76) (Figure 4A). Among the 12-month comparison group, the correlation of the OS HR with the ratio of the 12-month OS milestone RMST was significantly improved ($R^2 = 0.74$; 95% CI, 0.43-0.84) (Figure 4B).

The proportional hazard assumption was violated in 10 of 31 comparisons (eTable 1 in the Supplement), which introduced uncertainty into the treatment effect measured by HR. To address

Figure 2. Correlation of Treatment Effects on the Overall Survival (OS) Hazard Ratio (HR) With the Ratios of OS or Progression-Free Survival (PFS) Milestone Rates



A, Correlation of treatment effects on OS HR with the ratio of 6-month PFS milestone rate. B, Correlation of treatment effects on OS HR with the ratio of 9-month PFS milestone rate. C, Correlation of treatment effects on OS HR with the ratio of 9-month

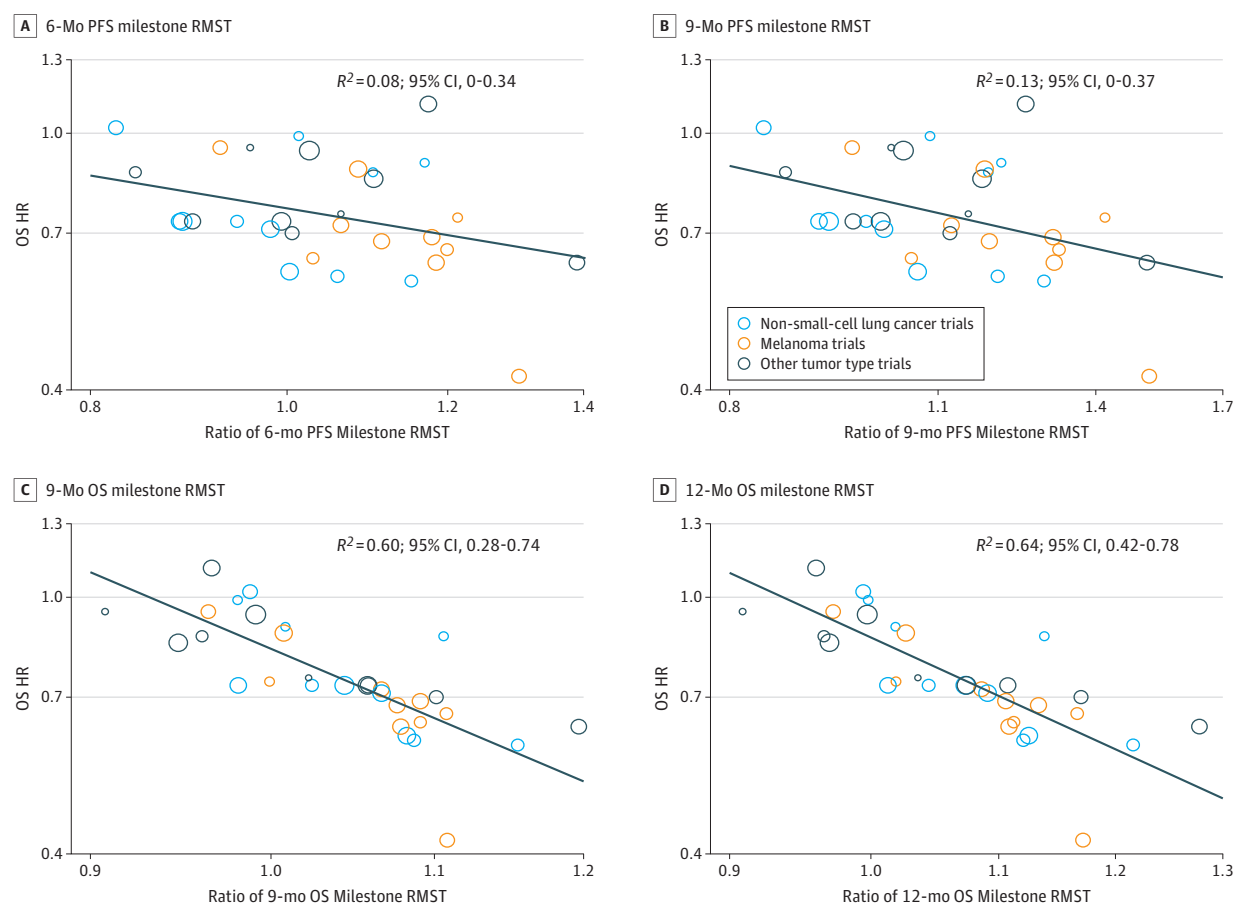
OS milestone rate. D, Correlation of treatment effects on OS HR with ratio of the 12-month OS milestone rate. Size of circles indicates size of trial (ie, larger circle indicates more participants).

this issue, we performed a sensitivity analysis in which comparisons with nonproportionality were excluded. In this analysis, a stronger correlation was observed of the OS HR with the ratios of the 9-month milestone RMST ($R^2 = 0.70$; 95% CI, 0.38-0.90) or 12-month milestone RMST ($R^2 = 0.67$; 95% CI, 0.39-0.87); however, there was no improvement in the correlation of the OS HR with the ratios of the 9-month milestone rate ($R^2 = 0.45$; 95% CI 0.21-0.85) or 12-month milestone rate ($R^2 = 0.35$; 95% CI 0.08-0.55). Moreover, the findings remained robust in additional sensitivity analyses in which the included trials were limited to trials with OS as the primary or coprimary end point, trials examining PDCD1 or PDCD1 ligand 1 inhibitors, phase 3 trials, trials examining only NSCLC, trials examining only melanoma, trials examining only cancers other than NSCLC and melanoma, first-line trials, and second-line or beyond trials (eTable 4 in the [Supplement](#)).

Discussion

The results of this study demonstrate that the ratios of the OS milestone RMSTs had a stronger correlation with the OS HRs compared with the ratios of the OS milestone rates. In contrast, there was a poor correlation between the ratios of the PFS milestone rates or PFS milestone RMSTs and OS HRs. Notably, the correlations of the OS HR with the ratio of the OS milestone rate at 12 months or 9 months were much weaker than those reported in the study by Blumenthal et al⁹ ($R^2 = 0.45$ vs

Figure 3. Correlation of Treatment Effects on Overall Survival (OS) Hazard Ratio (HR) With the Ratios of OS and Progression-Free Survival (PFS) Milestone Restricted Mean Survival Times (RMSTs)



A, Correlation of treatment effects on OS HR with the ratio of 6-month PFS milestone RMST. B, Correlation of OS HR with the ratio of 9-month PFS milestone RMST. C, Correlation of treatment effects on OS HR with the ratio of 9-month OS milestone RMST.

D, Correlation of treatment effects on OS HR with the ratio of 12-month OS milestone RMST. Size of circles indicates size of trial (ie, larger circle indicates more participants).

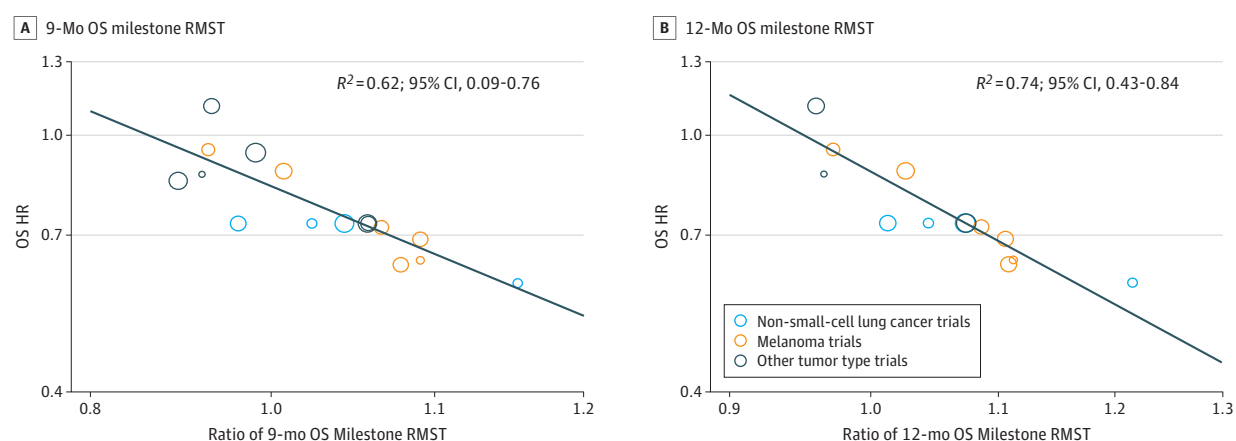
$R^2 = 0.67$ at 9 months and $R^2 = 0.40$ vs $R^2 = 0.80$ at 12 months). One possible explanation is that our analysis only included trials studying ICIs, whereas only 6 of the 25 included trials in the study by Blumenthal et al⁹ were ICI trials: 17% of the patients were treated with checkpoint inhibitor therapies, and the remainder were treated with chemotherapy and targeted therapies. Considering that these 3 therapeutic classes have different response patterns, the stronger correlations observed in the study by Blumenthal et al⁹ might reflect a mixed effect.

The ratio of the OS milestone rate at a given time point is unable to account for the survival information before that particular time point.⁴⁴ Immune checkpoint inhibitors have delayed clinical benefits; therefore, it is likely that the ratio of the OS milestone rate may overestimate the treatment effect. The ratio of the 9-month OS milestone rate was even more closely correlated with the OS HR than the ratio of the 12-month OS milestone rate ($R^2 = 0.45$ vs $R^2 = 0.40$), which might have been caused by a more prominent overestimation of the treatment effect using the latter end point.

With the further introduction of the RMST for milestone treatment effect measurement, a substantial improvement was observed in the correlation of the OS HR with the ratio of the OS milestone RMST. Unfortunately, as the R^2 did not reach the predefined validated cutoff value of 0.80, it may be too early to recommend the use of the OS milestone RMST as an intermediate end point for ICI trials. However, the results suggest that the OS milestone RMST outperforms the OS milestone rate as a potential intermediate end point for ICI trials and is worthy of further investigation. In addition, the OS milestone RMST may also be explored prospectively as a secondary end point in future ICI trials.

The use of the milestone RMST in the analyses not only improved the strength of the correlation with the OS HRs but also retained the advantage of milestone analysis, ie, the predictability and simplicity of the analysis as a time-driven end point. A mature time point is of great importance and should be predefined for milestone analysis, which is different from event-driven end points. Among trials with a minimum follow-up duration longer than 12 months, the correlation of the OS HR with the ratio of the 12-month OS milestone RMST was further improved, suggesting that data maturity is indispensable for the milestone RMST in terms of treatment effect measurement. Considering that future trials may increasingly use ICIs as controls in front-line design, in biomarker-enrichment strategies, and in combinational ICIs therapy,⁹ a substantially prolonged follow-up duration may be required to obtain the expected number of events. In contrast, the milestone RMST as a time-driven end point would have a well-defined maximum follow-up duration according to the prespecified milestone time point. Thus, while there should be a continuous optimization of the choice of

Figure 4. Correlation of Overall Survival (OS) Hazard Ratio (HR) With the Ratio of OS Milestone Restricted Mean Survival Times (RMSTs) Among Trials With Adequate Follow-up



A, Correlation of treatment effects on OS HR with the ratio of 9-month OS milestone RMST among comparisons with a minimum follow-up longer than 9 months. B, Correlation of treatment effects on OS HR with the ratio of 12-month OS milestone RMST

among comparisons with a minimum follow-up longer than 12 months. Size of circles indicates size of trial (ie, larger circle indicates more participants).

milestones depending on increasing amounts of trial data, it is reasonable to further investigate the 12-month OS milestone RMST as an intermediate end point in future ICI trials. Moreover, even the 9-month OS milestone RMST ratio outperformed the ratio of the 12-month OS rate in terms of strength of correlation with OS HR, suggesting that OS milestone RMST ratio-based end points may require a shorter follow-up duration to detect early signals of the efficacy of ICI-based therapy.

The ratios of the PFS milestone rate and PFS milestone RMST showed weak correlation with OS HR. In most of the included trials, disease progression was defined using the standard RECIST version 1.1 criteria,⁴³ which may not be able to fully capture the treatment effect of ICIs because of unique patterns of patient response.⁴⁵ In addition, the correlation of PFS with OS was found to be weak in a previous study of 13 trials containing ICIs submitted to the US Food and Drug Administration, which suggests that PFS may not capture the survival benefit of patients with long-term survival.⁵ In 2018, Ritchie et al⁶ recommended the use of the 6-month PFS milestone rate as the preferred end point for phase 2 trials studying ICIs, primarily based on its correlation with the 12-month OS milestone rate. Our study found weak correlations of the ratio of 12-month OS milestone rate or 6-month PFS milestone rate with the OS HR, suggesting that it might be inappropriate to adopt the 6-month PFS rate as the primary end point in phase 2 ICI trials.

Limitations

Several limitations should be acknowledged. Although the individual patient survival data were reconstructed accurately for milestone analysis, there was no access to detailed individual patient data, and information on patient-level associations with the examined end points could not be provided. In one trial,²³ the difference between the reconstructed HR and the reported HR was greater than 5%, which may introduce potential bias to the findings. In addition, although only trials studying ICIs were included, the different study designs and patient populations among the included trials may have led to potential bias. The proportional hazard assumption was violated in some trials, which may have compromised the accuracy of HR as measurement of treatment effectiveness. Comparisons with nonproportionality were excluded and revealed correlations of the OS HR with the ratio of the 9-month or 12-month milestone RMST. Moreover, although assessment at the patient level is of value to identify appropriate surrogate end points, neither the milestone survival rate nor the milestone RMST are measurable at the individual patient level.

Conclusions

In the present study, the results demonstrated that the ratios of the OS milestone RMST were more strongly correlated with the OS HRs than the ratios of the OS milestone rate, whereas there was weak correlation of OS HRs with the ratios of the PFS milestone rates and with the PFS milestone RMSTs. Although the OS milestone RMST should not yet be used as an intermediate end point for future ICI trials, it could be prospectively incorporated as a secondary end point for further testing in future studies with front-line strategies involving ICI combinations or biomarker-enrichment designs. To validate the OS milestone RMST as an intermediate end point for ICI trials and to optimize the choice of milestone time points, further studies are required.

ARTICLE INFORMATION

Accepted for Publication: March 20, 2019.

Published: May 3, 2019. doi:10.1001/jamanetworkopen.2019.3433

Open Access: This is an open access article distributed under the terms of the [CC-BY License](#). © 2019 Wang Z-X et al. *JAMA Network Open*.

Corresponding Author: Rui-Hua Xu, MD, PhD, Department of Medical Oncology, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, 651 Dong Feng Road East, Guangzhou, 510060, China (xurh@sysucc.org.cn).

Author Affiliations: Department of Medical Oncology, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, Guangzhou, China (Z.-X. Wang, Wu, Y.-N. Wang, Yang, He, Luo, Li, F. Wang, Xu); Faculty of Medical Sciences, Sun Yat-Sen University, Guangzhou, China (Wu); Clinical Research Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China (L. Xie); Department of Colorectal Surgery, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, Guangzhou, China (Ding, Chen); Department of Pathology, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, Guangzhou, China (D. Xie).

Author Contributions: Drs Z.-X. Wang, Wu, L. Xie, and Y.-N. Wang contributed equally to this study and are considered co-first authors. Drs F. Wang and Xu contributed equally to the study and are considered co-senior authors. Dr Xu had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Z.-X. Wang, Wu, L. Xie, Xu.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Z.-X. Wang, Wu, L. Xie, Y.-N. Wang, Xu.

Critical revision of the manuscript for important intellectual content: Wu, L. Xie, Y.-N. Wang, Yang, He, Luo, Ding, D. Xie, Chen, Li, F. Wang, Xu.

Statistical analysis: Z.-X. Wang, Wu, L. Xie, Xu.

Obtained funding: Xu.

Administrative, technical, or material support: Wu, Chen, Xu.

Supervision: Xu.

Conflict of Interest Disclosures: None reported.

Funding/Support: This study was supported by grants from the National Key Technologies R&D Program of China (2018YFC1313300), the Natural Science Foundation of Guangdong Province (2014A030312015), the Science and Technology Program of Guangdong (2015B020232008), and the Science and Technology Program of Guangzhou (15570006, 201508020250, and 201604020003).

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

REFERENCES

1. Khalil DN, Smith EL, Brentjens RJ, Wolchok JD. The future of cancer treatment: immunomodulation, CARs and combination immunotherapy. *Nat Rev Clin Oncol*. 2016;13(5):273-290. doi:10.1038/nrclinonc.2016.25
2. Brawley L. With 20 agents, 803 trials, and 166,736 patient slots, is pharma investing too heavily in PD-1 drug development? http://cancerletter.com/articles/20161007_1/. Accessed August 8, 2018.
3. Chen TT. Statistical issues and challenges in immuno-oncology. *J Immunother Cancer*. 2013;1:18. doi:10.1186/2051-1426-1-18
4. Chen TT. Milestone survival: a potential intermediate endpoint for immune checkpoint inhibitors. *J Natl Cancer Inst*. 2015;107(9):djv156. doi:10.1093/jnci/djv156
5. Mushti SL, Mulkey F, Sridhara R. Evaluation of overall response rate and progression-free survival as potential surrogate endpoints for overall survival in immunotherapy trials. *Clin Cancer Res*. 2018;24(10):2268-2275. doi:10.1158/1078-0432.CCR-17-1902
6. Ritchie G, Gasper H, Man J, et al. 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*. 2018;4(4):522-528. doi:10.1001/jamaoncol.2017.5236
7. Nie RC, Chen FP, Yuan SQ, et al. Evaluation of objective response, disease control and progression-free survival as surrogate end-points for overall survival in anti-programmed death-1 and anti-programmed death ligand 1 trials. *Eur J Cancer*. 2019;106:1-11. doi:10.1016/j.ejca.2018.10.011
8. Chen TT. Designing late-stage randomized clinical trials with cancer immunotherapy: can we make it simpler? *Cancer Immunol Res*. 2018;6(3):250-254. doi:10.1158/2326-6066.CIR-17-0465
9. Blumenthal GM, Zhang L, Zhang H, et al. Milestone analyses of immune checkpoint inhibitors, targeted therapy, and conventional therapy in metastatic non-small cell lung cancer trials: a meta-analysis. *JAMA Oncol*. 2017;3(8):e171029. doi:10.1001/jamaoncol.2017.1029
10. 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*. 2014;32(22):2380-2385. doi:10.1200/JCO.2014.55.2208

11. Uno H, Wittes J, Fu H, et al. Alternatives to hazard ratios for comparing the efficacy or safety of therapies in noninferiority studies. *Ann Intern Med*. 2015;163(2):127-134. doi:[10.7326/M14-1741](https://doi.org/10.7326/M14-1741)
12. Powles T, Durán I, van der Heijden MS, et al. Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. *Lancet*. 2018;391(10122):748-757. doi:[10.1016/S0140-6736\(17\)33297-X](https://doi.org/10.1016/S0140-6736(17)33297-X)
13. Guyot P, Ades AE, Ouwers MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol*. 2012;12:9. doi:[10.1186/1471-2288-12-9](https://doi.org/10.1186/1471-2288-12-9)
14. Royston P, Parmar MK. The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. *Stat Med*. 2011;30(19):2409-2421. doi:[10.1002/sim.4274](https://doi.org/10.1002/sim.4274)
15. Royston P, Parmar MK. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Med Res Methodol*. 2013;13:152. doi:[10.1186/1471-2288-13-152](https://doi.org/10.1186/1471-2288-13-152)
16. Trinquart L, Jacot J, Conner SC, Porcher R. Comparison of treatment effects measured by the hazard ratio and by the ratio of restricted mean survival times in oncology randomized controlled trials. *J Clin Oncol*. 2016;34(15):1813-1819. doi:[10.1200/JCO.2015.64.2488](https://doi.org/10.1200/JCO.2015.64.2488)
17. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*. 1994;81(3):515-526. doi:[10.1093/biomet/81.3.515](https://doi.org/10.1093/biomet/81.3.515)
18. Robert C, Long GV, Brady B, et al. Nivolumab in previously untreated melanoma without BRAF mutation. *N Engl J Med*. 2015;372(4):320-330. doi:[10.1056/NEJMoa1412082](https://doi.org/10.1056/NEJMoa1412082)
19. Hodi FS, Chesney J, Pavlick AC, et al. Combined nivolumab and ipilimumab versus ipilimumab alone in patients with advanced melanoma: 2-year overall survival outcomes in a multicentre, randomised, controlled, phase 2 trial. *Lancet Oncol*. 2016;17(11):1558-1568. doi:[10.1016/S1470-2045\(16\)30366-7](https://doi.org/10.1016/S1470-2045(16)30366-7)
20. Larkin J, Minor D, D'Angelo S, et al. Overall survival in patients with advanced melanoma who received nivolumab versus investigator's choice chemotherapy in CheckMate 037: a randomized, controlled, open-label phase III trial. *J Clin Oncol*. 2018;36(4):383-390. doi:[10.1200/JCO.2016.71.8023](https://doi.org/10.1200/JCO.2016.71.8023)
21. Robert C, Schachter J, Long GV, et al; KEYNOTE-006 Investigators. Pembrolizumab versus ipilimumab in advanced melanoma. *N Engl J Med*. 2015;372(26):2521-2532. doi:[10.1056/NEJMoa1503093](https://doi.org/10.1056/NEJMoa1503093)
22. Hodi FS, O'Day SJ, McDermott DF, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med*. 2010;363(8):711-723. doi:[10.1056/NEJMoa1003466](https://doi.org/10.1056/NEJMoa1003466)
23. Robert C, Thomas L, Bondarenko I, et al. Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. *N Engl J Med*. 2011;364(26):2517-2526. doi:[10.1056/NEJMoa1104621](https://doi.org/10.1056/NEJMoa1104621)
24. Hodi FS, Lee S, McDermott DF, et al. Ipilimumab plus sargramostim vs ipilimumab alone for treatment of metastatic melanoma: a randomized clinical trial. *JAMA*. 2014;312(17):1744-1753. doi:[10.1001/jama.2014.13943](https://doi.org/10.1001/jama.2014.13943)
25. Ribas A, Kefford R, Marshall MA, et al. Phase III randomized clinical trial comparing tremelimumab with standard-of-care chemotherapy in patients with advanced melanoma. *J Clin Oncol*. 2013;31(5):616-622. doi:[10.1200/JCO.2012.44.6112](https://doi.org/10.1200/JCO.2012.44.6112)
26. Fehrenbacher L, Spira A, Ballinger M, et al; POPLAR Study Group. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet*. 2016;387(10030):1837-1846. doi:[10.1016/S0140-6736\(16\)00587-0](https://doi.org/10.1016/S0140-6736(16)00587-0)
27. Rittmeyer A, Barlesi F, Waterkamp D, et al; OAK Study Group. 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(10066):255-265. doi:[10.1016/S0140-6736\(16\)32517-X](https://doi.org/10.1016/S0140-6736(16)32517-X)
28. Borghaei H, Paz-Ares L, Horn L, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N Engl J Med*. 2015;373(17):1627-1639. doi:[10.1056/NEJMoa1507643](https://doi.org/10.1056/NEJMoa1507643)
29. Brahmer J, Reckamp KL, Baas P, et al. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. *N Engl J Med*. 2015;373(2):123-135. doi:[10.1056/NEJMoa1504627](https://doi.org/10.1056/NEJMoa1504627)
30. Carbone DP, Reck M, Paz-Ares L, et al; CheckMate 026 Investigators. First-line nivolumab in stage IV or recurrent non-small-cell lung cancer. *N Engl J Med*. 2017;376(25):2415-2426. doi:[10.1056/NEJMoa1613493](https://doi.org/10.1056/NEJMoa1613493)
31. Herbst RS, Baas P, Kim DW, 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(10027):1540-1550. doi:[10.1016/S0140-6736\(15\)01281-7](https://doi.org/10.1016/S0140-6736(15)01281-7)

32. Langer CJ, Gadgil SM, Borghaei H, et al; KEYNOTE-021 Investigators. 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(11):1497-1508. doi:10.1016/S1470-2045(16)30498-3
33. Reck M, Rodríguez-Abreu D, Robinson AG, et al; KEYNOTE-024 Investigators. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N Engl J Med*. 2016;375(19):1823-1833. doi:10.1056/NEJMoa1606774
34. Lynch TJ, Bondarenko I, Luft A, 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(17):2046-2054. doi:10.1200/JCO.2011.38.4032
35. Kang YK, Boku N, Satoh T, et al. Nivolumab in patients with advanced gastric or gastro-oesophageal junction cancer refractory to, or intolerant of, at least two previous chemotherapy regimens (ONO-4538-12, ATTRACTION-2): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet*. 2017;390(10111):2461-2471. doi:10.1016/S0140-6736(17)31827-5
36. Ferris RL, Blumenschein G Jr, Fayette J, et al. Nivolumab for recurrent squamous-cell carcinoma of the head and neck. *N Engl J Med*. 2016;375(19):1856-1867. doi:10.1056/NEJMoa1602252
37. Kwon ED, Drake CG, Scher HI, et al; CA184-043 Investigators. 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(7):700-712. doi:10.1016/S1470-2045(14)70189-5
38. Beer TM, Kwon ED, Drake CG, et al. Randomized, double-blind, phase III trial of ipilimumab versus placebo in asymptomatic or minimally symptomatic patients with metastatic chemotherapy-naïve castration-resistant prostate cancer. *J Clin Oncol*. 2017;35(1):40-47. doi:10.1200/JCO.2016.69.1584
39. Motzer RJ, Escudier B, McDermott DF, et al; CheckMate 025 Investigators. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med*. 2015;373(19):1803-1813. doi:10.1056/NEJMoa1510665
40. Reck M, Bondarenko I, Luft A, 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(1):75-83. doi:10.1093/annonc/mds213
41. Reck M, Luft A, Szczesna A, 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(31):3740-3748. doi:10.1200/JCO.2016.67.6601
42. Bellmunt J, de Wit R, Vaughn DJ, et al; KEYNOTE-045 Investigators. Pembrolizumab as second-line therapy for advanced urothelial carcinoma. *N Engl J Med*. 2017;376(11):1015-1026. doi:10.1056/NEJMoa1613683
43. Eisenhauer EA, Therasse P, Boagaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45:228-247. doi:10.1016/j.ejca.2008.10.026
44. Seruga B, Pond GR, Hertz PC, Amir E, Ocana A, Tannock IF. Comparison of absolute benefits of anticancer therapies determined by snapshot and area methods. *Ann Oncol*. 2012;23(11):2977-2982. doi:10.1093/annonc/mds174
45. Wolchok JD, Hoos A, O'Day S, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. *Clin Cancer Res*. 2009;15(23):7412-7420. doi:10.1158/1078-0432.CCR-09-1624

SUPPLEMENT.

eMethods. Search Strategy

eTable 1. Detailed Regimen and Characteristics of Trials Included

eTable 2. Comparison of Reported Overall Survival Results With Those Calculated Based on Reconstructed Individual Patient Data

eTable 3. Comparison of Reported Progression-Free Survival Results With Those Calculated Based on Reconstructed Individual Patient Data

eTable 4. Sensitivity Analyses in Which A, Only Trials With Primary or Coprimary End Point of Overall Survival Were Included; B, Only PDCD1 or PDCD1 Ligand 1 Inhibitors Trials Were Included; C, Only Phase 3 Trials Were Included; D, Only NSCLC Trials Were Included; E, Only Melanoma Trials Were Included; F, Only Trials Other Than NSCLC and Melanoma Were Included; G, Only First-Line Trials Were Included; and H, Only Second-Line or Beyond Trials Were Included

eFigure. Kaplan-Meier Survival Estimates

eReferences