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

Georgia Ritchie, MBBS; Harry Gasper, MBBS; Johnathan Man, MBBS; Sally Lord, MSc;
Ian Marschner, PhD; Michael Friedlander, PhD; Chee Khoon Lee, PhD

➕ **Supplemental content**

IMPORTANCE Checkpoint inhibitors have a unique mechanism of action that differs from chemotherapy or targeted therapies. The validity of objective response rate (ORR) as a surrogate for progression-free survival (PFS) and overall survival (OS) in checkpoint-inhibitor trials is uncertain.

OBJECTIVE To determine the types of primary end points used in phase 2 checkpoint-inhibitor trials, and to assess the strength of associations for ORR with PFS and OS.

DATA SOURCES Trials listed in electronic databases from 2000 to 2017 (PREMEDLINE, MEDLINE, Embase, and the Cochrane Central Register of Controlled Trials).

STUDY SELECTION Advanced solid cancers in phase 2 and phase 3 trials.

DATA EXTRACTION AND SYNTHESIS Correlations between ORR odds ratios and hazard ratios (HRs) for PFS and OS were examined for randomized comparisons. Within checkpoint-inhibitor treatment arms, correlations for ORR with 6-month PFS and 12-month OS rates were examined. All analyses were weighted by trial size. Multivariable models to predict 6-month PFS and 12-month OS rates from ORR were developed and their performance validated in an independent sample of trials.

MAIN OUTCOMES AND MEASURES Correlation coefficient (r) of ORR with PFS and OS.

RESULTS Of 87 phase 2 trials identified, ORR was the most common (52 [60%]) primary end point. Twenty randomized clinical trials with 25 treatment comparisons were identified. Checkpoint-inhibitor therapy was associated with pooled ORR of 24% (95% CI, 18%-31%). For randomized comparisons, r between ORR odds ratio and PFS HR was 0.63 (95% CI, 0.35-0.89), ORR odds ratio and OS HR was 0.57 (95% CI, 0.23-0.89), and between PFS HR and OS HR was 0.42 (95% CI, 0.04-0.81). Within the checkpoint-inhibitor arms, r correlation coefficients between ORR with 6-month PFS, ORR with 12-month OS, and 6-month PFS with 12-month OS were 0.37 (95% CI, −0.06 to 0.95), 0.08 (95% CI, −0.17 to 0.70), and 0.74 (95% CI, 0.57-0.92), respectively. In validation, when 6-month PFS was used to predict 12-month OS, there was a good calibration between actual and predicted 12-month OS. When ORR was used to predict 6-month PFS and 12-month OS rates, respectively, the actual vs predicted rates calibrated poorly.

CONCLUSIONS AND RELEVANCE In checkpoint-inhibitor trials, ORR correlated poorly with OS. For future phase 2 studies, 6-month PFS rate is recommended as an end point.

Author Affiliations: Cancer Care Centre, St George Hospital, Sydney, Australia (Ritchie, Gasper, Man, Lee); School of Medicine, University of New South Wales, Sydney, Australia (Ritchie, Friedlander); National Health and Medical Research Council Clinical Trials Centre, University of Sydney, Sydney, Australia (Lord, Marschner, Friedlander, Lee); School of Medicine, University of Notre Dame, Sydney, Australia (Lord); Department of Statistics, Macquarie University, Sydney, Australia (Marschner); Nelune Comprehensive Cancer Centre, Prince of Wales Hospital, Sydney, Australia (Friedlander).

Corresponding Author: Chee Khoon Lee, PhD, NHMRC Clinical Trials Centre, University of Sydney, Locked Bag 77, Camperdown, NSW 1450, Australia (chee.lee@ctc.usyd.edu.au).

I mmune checkpoint inhibitors have become the standard of care for many patients with advanced solid cancers. These agents differ from chemotherapy by activating antitumor T-cells to detect and destroy tumor cells.[1] Agents currently in clinical use and being widely tested in clinical trials for multiple tumor types include monoclonal antibodies against anti–cytotoxic T-lymphocyte–associated antigen 4 (CTLA-4), anti–programmed cell death protein 1 (PD-1), and anti–programmed cell death ligand 1 (PDL-1).

Traditionally, new anticancer agents are evaluated for preliminary efficacy in phase 2 trials using objective response rate (ORR) as the primary trial end point. Typically, these agents will be examined in single-arm trials and their performance benchmarked against prespecified improvements in ORR, on the basis of historical control data, to identify promising agents for further testing in phase 3 randomized controlled trials (RCTs). With this approach, novel agents that result in predominantly stable disease are given lower priority for phase 3 testing than agents that have high ORR. Regulatory bodies are also more likely to provide preliminary approval of agents with high ORR.[2,3]

However, the validity of the ORR as a surrogate for progression-free survival (PFS) and overall survival (OS) is uncertain for checkpoint inhibitors because they have unique patterns of response and progression that differ from those of chemotherapy or molecular targeted agents. In particular, pseudoprogression has been recognized as an increase in size of target lesions or even the development of new lesions on imaging but not necessarily due to resistance to treatment but rather a consequence of treatment effect.[4] It has also been reported that some patients with apparent progressive disease on the basis of traditional response criteria such as Response Evaluation Criteria in Solid Tumors (RECIST)[5] have delayed deep and durable tumor responses when they continue treatment with a checkpoint inhibitor.[6] This paradoxical apparent increase in tumor burden preceding response might reflect continued tumor growth until a robust immune response develops[7] or transient immune-cell infiltrate with associated edema.[6,8]

The selection of the appropriate primary end point in phase 2 checkpoint-inhibitor trials is critical to inform the decision to proceed to phase 3 testing. In checkpoint-inhibitor trials, the validity of ORR, as determined by RECIST, and PFS as surrogates for OS remains unclear. We performed a systemic review and meta-analysis of trials of checkpoint inhibitors to determine the common types of primary outcome measures used in phase 2 checkpoint-inhibitor trials. We tested the strength of associations for ORR and PFS with OS for this new class of agents to assess the validity of ORR as a surrogate end point.

Methods

We performed a systematic search of electronic databases for trial results from January 2000 to January 2017. Published trials were identified through PREMEDLINE, MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials using the keywords: *ipilimumab, nivolumab, pembrolizumab, atezolizumab, avelumab, tremelimumab, CTLA, PD-L1, PD-1, checkpoint inhibitor, clinical trial, phase II, phase 2 clinical trial,*

Key Points

Question What is the most appropriate primary end point in phase 2 trials of checkpoint inhibitors?

Findings In this systematic review and meta-analysis of phase 2 and phase 3 trials of checkpoint inhibitors in advanced solid cancers, response rate correlated poorly with overall survival, but 6-month progression-free survival was a better predictor of 12-month overall survival.

Meaning Six-month progression-free survival is recommended in place of response rate as an end point in future phase 2 checkpoint-inhibitor trials.
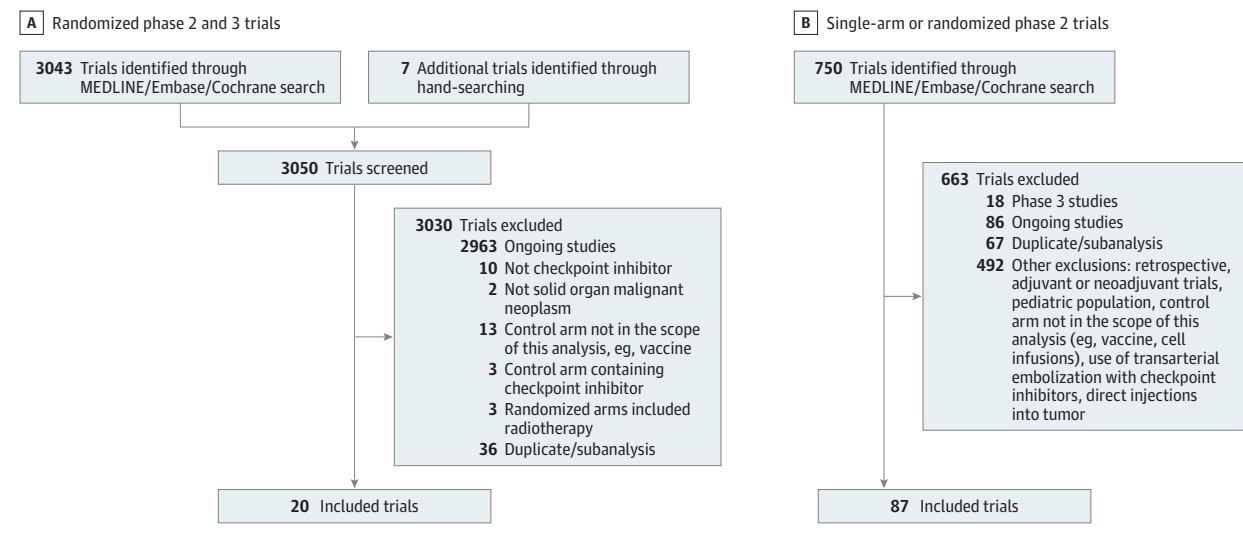
and variations on *randomi#ed and controlled and trial*. We also hand-searched abstracts and conference presentations on the European Society for Medical Oncology and American Society of Clinical Oncology websites. Where multiple publications and conference presentations were identified for the same trial, the most recent data were used.

Trials that met the following criteria were included: checkpoint inhibitors in advanced solid cancers in single-arm or RCTs of phase 2 and phase 3 designs. Eligible randomized comparisons were PD-1/PD-L1/CTLA4 inhibitor vs chemotherapy, placebo, or molecular targeted therapy; and combination of a checkpoint inhibitor with another checkpoint inhibitor, chemotherapy, or molecular targeted therapy vs non–checkpoint-inhibitor therapies. Trials that investigated checkpoint inhibitors with radiotherapy, direct local injections, transarterial embolization, vaccines, dendritic cell infusions, or granulocyte colony-stimulating factor were excluded.

Data on the baseline trial and patient characteristics and efficacy end point were extracted independently by 2 of us (G. R. and H. G.). Any discrepancies were resolved by consensus. The risk of bias for treatment-effect measures was assessed for each included study by examining the methods used for randomization and allocation concealment for the RCTs, outcome assessments, and handling of patient attrition and missing data.

Phase 2 trial design and primary end points were tabulated. Using the data from the checkpoint-inhibitor treatment arms of randomized phase 2 and phase 3 trials only, we used the fixed-effects inverse-variance-weighted method to obtain the pooled ORR and 95% CI for checkpoint inhibitors. Using data from the randomized treatment comparisons, we examined correlations between ORR odds ratio with PFS hazard ratio (HR) and OS HR. The correlation between 6-month PFS and 12-month OS rate ratios were also evaluated. Odds ratios and HRs of less than 1 denote a favorable result for checkpoint-inhibitor therapy vs non–checkpoint-inhibitor control therapy. Linear regression models were fitted by use of ordinary least-squares regression to examine the associations between the surrogates (ORR or PFS) and OS. Tumor origin was included as a covariate in these models. We also performed milestone analysis[9-11] by examining correlations within checkpoint-inhibitor treatment arms of the RCTs between ORR and 6-month PFS rate and 12-month OS rate. All analyses were weighted by trial size. The strength of associations was expressed as *r*, the correlation coefficient between the treat-

Figure 1. Flow Diagram of Study Inclusions and Exclusions

A Randomized phase 2 and 3 trials

3043 Trials identified through MEDLINE/Embase/Cochrane search

7 Additional trials identified through hand-searching

3050 Trials screened

3030 Trials excluded
  2963 Ongoing studies
  10 Not checkpoint inhibitor
  2 Not solid organ malignant neoplasm
  13 Control arm not in the scope of this analysis, eg, vaccine
  3 Control arm containing checkpoint inhibitor
  3 Randomized arms included radiotherapy
  36 Duplicate/subanalysis

20 Included trials

B Single-arm or randomized phase 2 trials

750 Trials identified through MEDLINE/Embase/Cochrane search

663 Trials excluded
  18 Phase 3 studies
  86 Ongoing studies
  67 Duplicate/subanalysis
  492 Other exclusions: retrospective, adjuvant or neoadjuvant trials, pediatric population, control arm not in the scope of this analysis (eg, vaccine, cell infusions), use of transarterial embolization with checkpoint inhibitors, direct injections into tumor

87 Included trials

ment effects on surrogates and on OS, with values close to 1 indicating strong associations.[12] The 95% CIs of *r* were obtained by bootstrap method with 1000 replications.

We validated our findings by applying these predictive linear regression models to the single-arm or multiarm phase 2 trials with adequate outcome data. These multiarm trials did not have a concurrent control standard of care or placebo arm. With the reported ORR and 6-month PFS rate, we used these models to generate the predicted 12-month OS rates. We also estimated predicted 12-month OS rates based on observed 6-month PFS rates. We assessed visually for calibration using plots of actual vs model-predicted 12-month OS rates.

We performed 4 sensitivity analyses, in which (1) included studies were restricted to only non–small cell lung cancer (NSCLC), the most common tumor type in the included RCTs, because ORR, PFS, and OS associations might potentially differ from other tumors; (2) the analysis was limited to PD-1/PD-L1 inhibitor studies because we recognized that there are differences in the mechanism of action between PD-1/PD-L1 and CTLA4 inhibitors; (3) exclusion of Keynote 024 because only patients with tumor of high PD-L1 expression were enrolled in this study; and (4) the analysis was limited to phase 3 checkpoint-inhibitor studies that were of larger sample size with potentially more robust results.

Results

Of 87 phase 2 trials identified (**Figure 1**), most were single-arm design (59 [68%]). Nineteen (22%) were multi-arm studies with no concurrent standard-care arms; and 9 (10%) were RCTs with concurrent standard-care arms. OBJECTIVE response rate was the most common (52 [60%]) primary end point, followed by PFS (11 [13%]), toxicity (10 [11%]), OS (6 [7%]), disease control rate (2 [2%]), and other molecular biomarker end points (6 [7%]). Immune-related response criteria, either as a primary or secondary end point, were considered in only 5 trials (6%).

We identified 20 eligible RCTs, including 4 with phase 2 designs (eTable 1 in the Supplement; Figure 1), comprising 25 treatment comparisons and 10 828 patients. Ten trials examined PD-1 inhibitor monotherapy (nivolumab,[13-19] 7; pembrolizumab,[20-22] 3); 2 trials, PD-L1 inhibitor monotherapy (atezolizumab,[23,24] 2); 3 trials, CTLA4 inhibitor monotherapy (tremelimumab,[25,26] 2; ipilimumab,[27] 1); and 5, a checkpoint inhibitor and chemotherapy combination (ipilimumab and chemotherapy,[28-31] 3; nivolumab and chemotherapy,[32] 1, pembrolizumab and chemotherapy,[33] 1). These trials examined 8 different tumor types, predominantly NSCLC (9 [45%]) and melanoma (4 [20%]). Other advanced cancers (7 [35%]) included mesothelioma, small cell lung, renal cell, prostate, gastric, and squamous cell head and neck cancers. Risk of bias was assessed as unclear in 3 unpublished trials[14,19,26] and low in all others.

Of the 10 828 patients, 6144 (57%) had been randomly assigned to checkpoint-inhibitor therapy, and 4684 (43%) to standard-care treatment or placebo. Data were unavailable for ORR in 1 trial. For the remaining 19 trials, with 23 comparisons, treatment with checkpoint-inhibitor therapy was associated with a pooled ORR of 24% (95% CI, 18%-31%; Cochrane Q χ^2 = 617.29 ($P < .001$); I^2 = 96%). Of these, only 8 trials (42%) or 9 checkpoint-inhibitor arms (47%) (monotherapy, or in combination with another agent) had an ORR of 30% or higher.

For the 24 treatment comparisons, the correlation coefficients for treatment effects between randomized arms, *r*, between ORR odds ratio and PFS HR was 0.63 (95% CI, 0.35-0.89), between ORR odds ratio and OS HR was 0.57 (95% CI, 0.23-0.89), and between PFS HR and OS HR was 0.42 (95% CI, 0.04-0.81). For the association of 6-month PFS and 12-month OS rate ratios, *r* was 0.55 (95% CI, 0.14-0.92). eFigure 1A in the Supplement shows the linear regression line used to predict the effect of treatment on PFS from the observed effect on ORR. For **Figure 2**, the regression lines predict the effects of treatment on OS from the observed effects on ORR and PFS, respectively.

Among the 24 checkpoint-inhibitor treatment arms only, *r* between ORR and 6-month PFS rate was 0.37 (95% CI, −0.06

Figure 2. Correlations in Relative Treatment Effect
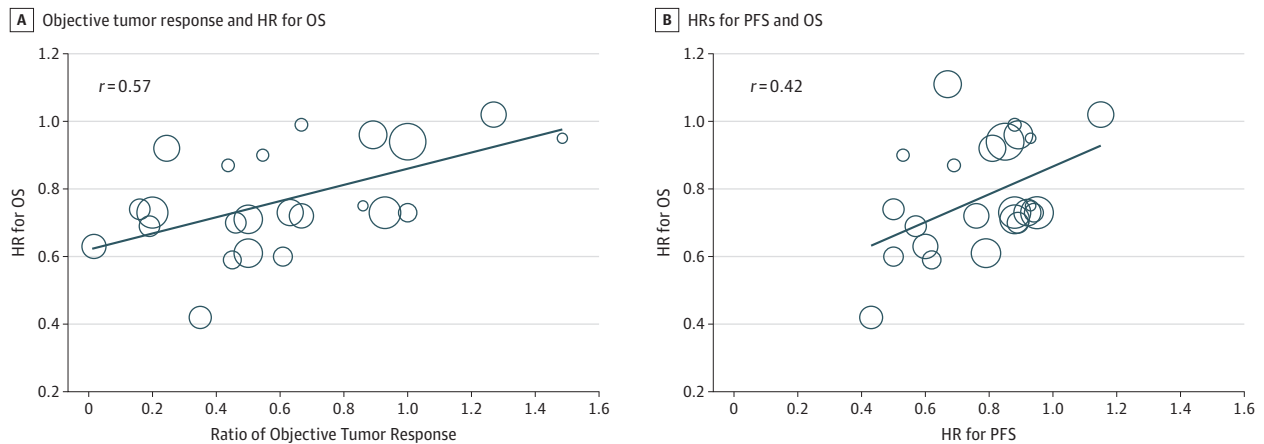
A, Ratio of objective tumor response and hazard ratio (HR) for overall survival (OS). B, Hazard ratio for progression-free survival (PFS) and HR for OS. Each circle represents a trial or treatment comparison. The circle size is proportional to the number of patients.

Figure 3. Correlations in Absolute Treatment Effect Within the Checkpoint-Inhibitor Arms
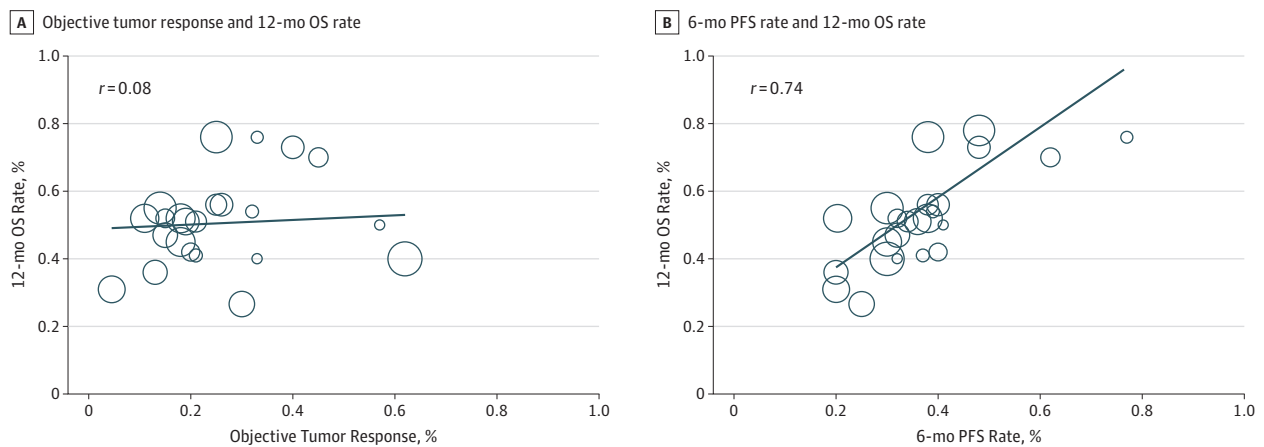
A, Objective tumor response and 12-month overall survival (OS) rate. B, Six-month progression-free survival (PFS) rate and 12-month OS rate. Each circle represents a trial or treatment comparison. The circle size is proportional to the number of patients.

to 0.95), between ORR and 12-month OS rate was 0.08 (95% CI −0.17 to 0.70), and between the 6-month PFS rate and 12-month OS rate was 0.74 (95% CI, 0.57-0.92) (eFigure 1B in the Supplement; **Figure 3**). The regression equation for association of the 6-month PFS rate with the 12-month OS rate, accounting for different tumor types, was $HR_{OS} = 1.10 \times HR_{PFS} + 0.16 + 0.04 \times$ melanoma $−0.04 \times$ NSCLC $+ 0 \times$ other tumors.

Nineteen single-arm or multicheckpoint-inhibitor arm phase 2 trials comprising 3023 patients were used for validation of the OS predictive models (eTable 2 in the Supplement). These trials were conducted in 7 types of advanced tumor (34% in melanoma,[34-41] 21% in NSCLC,[42-46] 17% in urothelial cancers,[47-49] 28% in other cancers[50-52]). For the analysis of 6-month PFS rate as predictor of 12-month OS rate, data were available from 12 trials or treatment arms. For the analysis of ORR as a predictor of 6-month PFS rate and 12-month OS rate, data were available from 16 and 25 trials or treatment arms, respec-

tively. Model-predicted 12-month OS rates based on 6-month PFS rates showed good calibration (**Figure 4**A). However, there was poor calibration between the actual and ORR-predicted 12-month OS rates (Figure 4B), and actual and ORR-predicted 6-month PFS rate (eFigure 1C in the Supplement).

When only NSCLC trials were considered, there were 8 trials[14,15,18,20,21,23,28,33] with 11 treatment comparisons. For treatment effects between randomized arms, r for ORR odds ratio with PFS HR was 0.74 (95% CI, 0.38-1.08), ORR odds ratio with OS HR was 0.68 (95% CI, 0.08-1.10), and PFS HR with OS HR was 0.63 (95% CI, 0.12-1.06) (eFigure 2 in the Supplement). For treatment effects within the checkpoint-inhibitor arms, r for ORR with 6-month PFS rate was 0.85 (95% CI, 0.63-1.06), ORR with 12-month OS rate was 0.66 (95% CI, 0.17-1.08), and PFS rate with OS rate was 0.76 (95% CI, 0.29-1.15) (eFigure 3 in the Supplement).

eTable 3 in the Supplement provides the sensitivity analyses for the correlations between different end points when

Figure 4. Observed vs Predicted 12-Month Overall Survival (OS) Rate
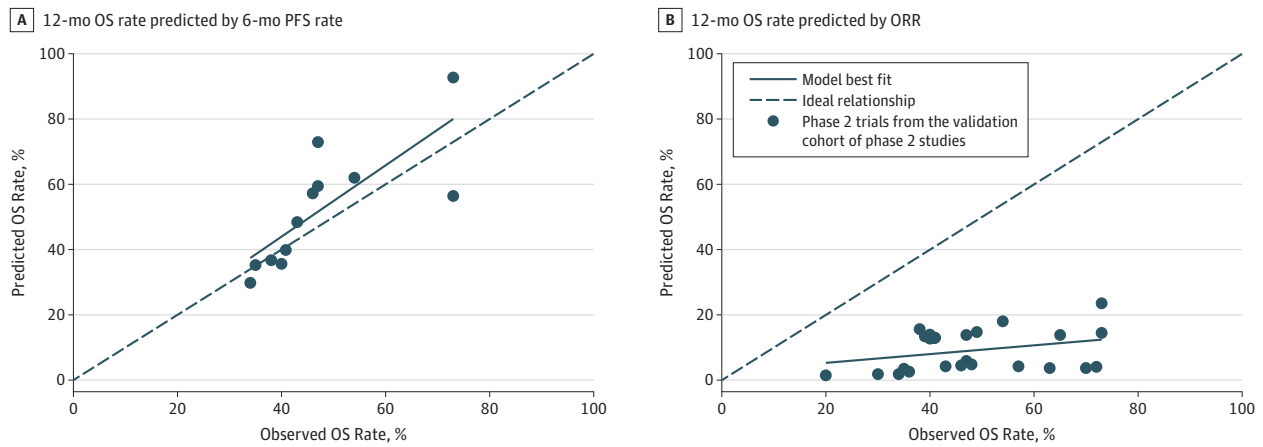
A, Observed 12-month OS rate and 12-month OS rate predicted by 6-month progression-free survival (PFS). B, Observed 12-month OS rate and 12-month OS rate predicted by objective response rate (ORR). Each data point represents a phase 2 trial from the validation cohort of phase 2 studies.

(1) the analysis limited to only PD-1/PD-L1 inhibitor studies, (2) exclusion of Keynote 024, and (3) the analysis limited to phase 3 RCTs only.

Discussion

We found that phase 2 checkpoint-inhibitor trials typically use a single-arm design with ORR as the primary trial end point. The ORRs reported are modest (pooled ORR, 23%). With the exception of NSCLC, ORR correlated poorly with 6-month PFS rate or 12-month OS rate within treatment arms. In contrast, the 6-month PFS rate correlated moderate strongly with the 12-month OS rate.

The results of the 20 RCTs included in this meta-analysis have led to 13 unique indications approved by the US Food and Drug Administration (FDA). It is noteworthy that where checkpoint inhibitors were approved, 58% had an ORR less than 30%. A recent study[3] suggests that single-arm trials in advanced solid cancers with ORR exceeding 30% were associated with greater likelihood of "breakthrough therapy" designation. However, drug approval involves multifactorial considerations, including unmet need and lack of effective treatment options. If recommendation for drug approval is based only on a high ORR, many potentially efficacious agents could be overlooked for further evaluation in phase 3 trials or may not obtain conditional regulatory approval.

In the phase 2 setting, surrogate end points are frequently used as an early signal of drug activity and assist in go or no-go decision making for phase 3 testing. Response and progression by conventional RECIST poorly reflect the treatment efficacy on OS for checkpoint inhibitors, and immune-related response criteria have been proposed.[6,53,54] Despite the availability of such criteria, our review has shown that only 6% of recently conducted phase 2 trials used them as a trial end point. Further research is required to assess the validity of these modified criteria as a surrogate end point for OS.

Recently, it has also been proposed that milestone analysis, which looks at OS for a given time point, such as at 12 months, could potentially be a better surrogate end point for checkpoint-inhibitor trials.[9,10] Using individual patient data (IPD) from 14 RCTs of checkpoint inhibitors, targeted therapy, and chemotherapy, analysis in advanced NSCLC by the FDA demonstrated that OS at 12 months had the strongest association with OS HR, but this association was modest.[11] This study also found no association between 6-month ORR and the OS HR. Interestingly, there was also a poor association between the 9-month PFS milestone ratio and OS HR. There are multiple possible reasons for differences in the findings between this study and our analysis. First, these analyses, based on milestone ratio and HR, focused on relative differences between the experimental and control therapy. Second, the result was based on only 6 checkpoint-inhibitor trials out of the total of 25 included studies. Only 17% of the patients were treated with checkpoint-inhibitor therapies and the remainder with chemotherapy and targeted therapies. Subgroup analysis comparing OS HR vs PFS HR for trials comparing immunotherapy with chemotherapy demonstrated better correlation than trials comparing targeted therapies with chemotherapy. We speculate that high rates of crossover and/or long postprogression survival are possible reasons for these observations. Our study is unique because we included only checkpoint-inhibitor trials, and we examined the within-arm association between 6-month PFS and 12-month OS, particularly in the phase 2 setting, where most trials used a single-arm design. Moreover, the appropriateness of OS, including milestone analysis, as an end point in phase 2 checkpoint-inhibitor trials, remains an area of ongoing debate because one of the important objectives of these studies is to rapidly screen them for preliminary drug activity using an earlier intermediate end point before definitive testing in phase 3 trials.

Strengths and Limitations

This study has several strengths. We have performed a comprehensive review to include all reported RCTs of immuno-

therapy across different advanced solid cancers involving more than 10 000 patients. Because a wide variety of checkpoint inhibitors was investigated in a heterogeneous population of patients with advanced cancer, we were able to examine for variability in treatment outcomes and hence improved generalizability of our study. Our generation and validation of a prediction model to describe the association between 6-month PFS rate and 12-month OS rate is unique.

There are also several limitations. We had no access to IPD and were unable to examine patient subgroups in detail. Our analysis is also limited for trial-level surrogacy analysis but is unable to provide any information on patient-level association. Only 1 RCT reported immune-related response criteria as an end point, and hence we were unable to assess whether it represents a more valid surrogate than RECIST ORR for PFS and OS. We assumed that the best ORR, as reported in the included studies, occurred before 6 months, but we were unable to confirm whether this assumption was correct without access to IPD. Furthermore, reliable and comprehensive data on crossover between treatment arms and salvage checkpoint-inhibitor therapy for control-arm patients are poorly reported in most studies, and it is possible that postprogression therapies could have an impact on OS. Another important confounder is that in some patients, who came off study because of progressive disease by RECIST, subsequent responses may be incorrectly attributed to the crossover therapy rather than to a delayed response to the immunotherapy, particularly for studies that did not use immune-related response criteria. Although the poor correlation between ORR and 12-month OS treatment effect may be in part be due to treatment crossover subsequent to progression, this phenomenon cannot explain the poor correlation between ORR and 6-month PFS treatment effect. With the use of 6-month PFS rate as a clinical trial end point,

we were unable to account for the totality of the PFS times or the effect of censoring before this milestone time point. However, in the setting of phase 2 trials, 6-month PFS rate is advantageous as it is simple, clinically meaningful, and predictable, because it is a time-driven rather than event-driven end point.

This study has several important implications. With the use of a 6-month PFS rate, future phase 2 trials might require a larger sample size, more resources, and take longer to report on this result than RECIST ORR. However, given the current poor success rate of phase 3 RCTs in oncology,[55] we believe that this overall approach will still be worthwhile and cost-effective because more robust preliminary data of efficacy could be generated with smaller phase 2 studies to better guide the selection of agents for a significantly more expensive and time-consuming phase 3 testing. Furthermore, most phase 2 studies do not collect information on patient-reported outcomes (PROs) that could potentially provide patient relevant information about treatment benefit. However, if ORR is shown to be associated with improvement of symptoms and PROs, it may still be an important a secondary end point in phase 2 trials. This question, and the potential inclusion of PROs as a co-primary end point for future phase 2 trials, warrants further research.

Conclusions

Objective response rate correlated poorly with OS. In future phase 2 checkpoint-inhibitor trials, 6-month PFS is recommended as an end point over RECIST ORR. Further research is required to assess the validity of milestone analysis with 6-month PFS as a potential surrogate for OS in treatment comparisons between checkpoint inhibitors and standard of care therapy.

REFERENCES

1. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer*. 2012;12(4):252-264.

2. Johnson JR, Ning Y-M, Farrell A, Justice R, Keegan P, Pazdur R. Accelerated approval of oncology products: the Food and Drug Administration experience. *J Natl Cancer Inst*. 2011; 103(8):636-644.

3. Oxnard GR, Wilcox KH, Gonen M, Polotsky M, Hirsch BR, Schwartz LH. Response rate as a regulatory end point in single-arm studies of advanced solid tumors. *JAMA Oncol*. 2016;2(6):772-779.

4. Chiou VL, Burotto M. Pseudoprogression and immune-related response in solid tumors. *J Clin Oncol*. 2015;33(31):3541-3543.

5. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. *J Natl Cancer Inst*. 2000;92(3):205-216.

6. 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.

7. Hodi FS, Oble DA, Drappatz J, et al. CTLA-4 blockade with ipilimumab induces significant clinical benefit in a female with melanoma metastases to the CNS. *Nat Clin Pract Oncol*. 2008; 5(9):557-561.

8. Hodi FS, Butler M, Oble DA, et al. Immunologic and clinical effects of antibody blockade of cytotoxic T lymphocyte–associated antigen 4 in previously vaccinated cancer patients. *Proc Natl Acad Sci U S A*. 2008;105(8):3005-3010.

9. Chen T-T. Milestone survival: a potential intermediate endpoint for immune checkpoint inhibitors. *J Natl Cancer Inst*. 2015;107(9):djv156.

10. Hellmann MD, Kris MG, Rudin CM. Medians and milestones in describing the path to cancer cures: telling "tails". *JAMA Oncol*. 2016;2(2):167-168.

11. 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.

12. Burzykowski T, Molenberghs G, Buyse M. *Evaluation of Surrogate Endpoints*. Heidelberg, Germany: Springer Verlag; 2005.

13. 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.

14. Socinski M, Creelan B, Horn L, et al. NSCLC, metastaticCheckMate 026: a phase 3 trial of nivolumab vs investigator's choice (IC) of platinum-based doublet chemotherapy (PT-DC) as first-line therapy for stage IV/recurrent programmed death ligand 1 (PD-L1)–positive

NSCLC. *Ann Oncol*. 2016;27(suppl 6): LBA7_PR-LBA7_PR.

15. 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.

16. 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.

17. Weber JS, D'Angelo SP, Minor D, 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(4):375-384.

18. 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.

19. Kang Y-K, Satoh T, Ryu M-H, et al. Nivolumab (ONO-4538/BMS-936558) as salvage treatment after second or later-line chemotherapy for advanced gastric or gastro-esophageal junction cancer (AGC): a double-blinded, randomized, phase III trial. *J Clin Oncol*. 2017;35(4 suppl):2-2.

20. Herbst RS, Baas P, Kim D-W, 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.

21. 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.

22. Ribas A, Puzanov I, Dummer R, et al. Pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory melanoma (KEYNOTE-002): a randomised, controlled, phase 2 trial. *Lancet Oncol*. 2015;16(8):908-918.

23. 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.

24. 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.

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.

26. Kindler HL, Scherpereel A, Calabrò L, et al. Tremelimumab as second- or third-line treatment of unresectable malignant mesothelioma (MM): results from the global, double-blind, placebo-controlled DETERMINE study. *J Clin Oncol*. 2016;34(15 suppl):8502-8502.

27. 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-naive castration-resistant prostate cancer. *J Clin Oncol*. 2017;35(1):40-47.

28. 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.

29. 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.

30. 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.

31. 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.

32. 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.

33. Langer CJ, Gadgeel 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.

34. Di Giacomo AM, Ascierto PA, Queirolo P, et al. Three-year follow-up of advanced melanoma patients who received ipilimumab plus fotemustine in the Italian Network for Tumor Biotherapy (NIBIT)-M1 phase II study. *Ann Oncol*. 2015;26(4):798-803.

35. Kirkwood JM, Lorigan P, Hersey P, et al. Phase II trial of tremelimumab (CP-675,206) in patients with advanced refractory or relapsed melanoma. *Clin Cancer Res*. 2010;16(3):1042-1048.

36. Zimmer L, Eigentler TK, Kiecker F, et al. Open-label, multicenter, single-arm phase II DeCOG-study of ipilimumab in pretreated patients with different subtypes of metastatic melanoma. *J Transl Med*. 2015;13:351.

37. O'Day SJ, Ibrahim R, DePril V, et al. Efficacy and safety of ipilimumab induction and maintenance dosing in patients with advanced melanoma who progressed on one or more prior therapies. *J Clin Oncol*. 2008;26(15 suppl):9021.

38. Atkins MB, Kudchadkar RR, Sznol M, et al. Phase 2, multicenter, safety and efficacy study of pidilizumab in patients with metastatic melanoma. *J Clin Oncol*. 2014;32(15 suppl):9001.

39. 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.

40. Patel SP, Kim DW, Bassett RL, et al. A phase II study of ipilimumab plus temozolomide in patients with metastatic melanoma. *Cancer Immunol Immunother*. 2017;66(10):1359-1366.

41. Wolchok JD, Neyns B, Linette G, et al. Ipilimumab monotherapy in patients with pretreated advanced melanoma: a randomised, double-blind, multicentre, phase 2, dose-ranging study. *Lancet Oncol*. 2010;11(2):155-164.

42. Rizvi NA, Mazières J, Planchard D, et al. Activity and safety of nivolumab, an anti-PD-1 immune checkpoint inhibitor, for patients with advanced, refractory squamous non-small-cell lung cancer (CheckMate 063): a phase 2, single-arm trial. *Lancet Oncol*. 2015;16(3):257-265.

43. Gettinger S, Rizvi NA, Chow LQ, et al. Nivolumab monotherapy for first-line treatment of advanced non-small-cell lung cancer. *J Clin Oncol*. 2016;34(25):2980-2987.

44. Antonia SJ, López-Martin JA, Bendell J, et al. Nivolumab alone and nivolumab plus ipilimumab in recurrent small-cell lung cancer (CheckMate 032): a multicentre, open-label, phase 1/2 trial. *Lancet Oncol*. 2016;17(7):883-895.

45. Besse B, Johnson M, Janne PA, et al. 16LBA Phase II, single-arm trial (BIRCH) of atezolizumab as first-line or subsequent therapy for locally advanced or metastatic PD-L1-selected non-small cell lung cancer (NSCLC). *Eur J Cancer*. 51(suppl 3):S717-S718.

46. Lena H, Rizvi NA, Wolf J, et al. 137O: Nivolumab in patients (pts) with advanced refractory squamous (SQ) non-small cell lung cancer (NSCLC): 2-year follow-up from CheckMate 063 and exploratory cytokine profiling analyses. *J Thorac Oncol*. 2016;11(4)(suppl):S115-S116.

47. Balar AV, Galsky MD, Rosenberg JE, et al; IMvigor210 Study Group. Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. *Lancet*. 2017;389(10064):67-76.

48. Rosenberg JE, Hoffman-Censits J, Powles T, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *Lancet*. 2016;387(10031): 1909-1920.

49. Sharma P, Callahan MK, Bono P, et al. Nivolumab monotherapy in recurrent metastatic urothelial carcinoma (CheckMate 032): a multicentre, open-label, two-stage, multi-arm, phase 1/2 trial. *Lancet Oncol*. 2016;17(11):1590-1598.

50. Kaufman HL, Russell J, Hamid O, et al. Avelumab in patients with chemotherapy-refractory metastatic Merkel cell carcinoma: a multicentre, single-group, open-label, phase 2 trial. *Lancet Oncol*. 2016;17(10):1374-1385.

51. Janjigian YY, Bendell JC, Calvo E, et al. CheckMate-032: phase I/II, open-label study of safety and activity of nivolumab (nivo) alone or with ipilimumab (ipi) in advanced and metastatic (A/M) gastric cancer (GC). *J Clin Oncol*. 2016;34(15 suppl): 4010.

52. Plimack ER, Hammers HJ, Rini BI, et al. Updated survival results from a randomized, dose-ranging phase II study of nivolumab (NIVO) in metastatic renal cell carcinoma (mRCC). *J Clin Oncol*. 2015;33(15 suppl):4553.

53. Seymour L, Bogaerts J, Perrone A, et al; RECIST Working Group. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol*. 2017;18(3):e143-e152.

54. Hodi FS, Hwu W-J, Kefford R, et al. Evaluation of immune-related response criteria and RECIST v1.1 in patients with advanced melanoma treated with pembrolizumab. *J Clin Oncol*. 2016;34(13):1510-1517.

55. Harrison RK. Phase II and phase III failures: 2013-2015. *Nat Rev Drug Discov*. 2016;15(12):817-818.