Original Study

A Quality-adjusted Time Without Symptoms or Toxicity (Q-TWiST) Analysis of Nivolumab Versus Everolimus in Advanced Renal Cell Carcinoma (aRCC)

Ruchitbhai Shah,[1] Marc Botteman,[1] Caitlyn T. Solem,[1] Linlin Luo,[1] Justin Doan,[2] David Cella,[3] Robert J. Motzer[4]

Abstract

This study assessed the net health benefits of treatment with nivolumab versus everolimus among patients with advanced renal cell carcinoma by assessing the quality (ie, patient preferences) and quantity of survival (ie, time spent with significant toxicities, in progression, or before progression and without significant toxicities). Nivolumab resulted in a 3.3-month quality-adjusted survival gain versus everolimus that was statistically significant and clearly clinically meaningful.

Background: This analysis compared quality-adjusted time without symptoms of disease progression or toxicity (Q-TWiST) between nivolumab and everolimus among previously treated patients with advanced renal cell carcinoma enrolled in the phase III CheckMate 025 trial (NCT01668784). **Materials and Methods:** At 45-month follow-up, overall survival (OS) was partitioned into 3 health states: TWiST, time with grade $\geq$ 3 toxicity (TOX), and time after progression (REL). Mean Q-TWiST was determined by multiplying each state's duration with its utility (TWiST, 1.0; TOX, 0.5; REL, 0.5). Relative Q-TWiST gains (calculated as Q-TWiST difference divided by everolimus OS) of $\geq$ 10% were predefined as clinically important. Immuno-oncology-specific sensitivity analyses considered 4 alternative progression definitions: Tumor size increase $\geq$ 25% from nadir; treatment discontinuation; $\geq$ 2-point reduction from baseline in Functional Assessment of Cancer Therapy-Kidney Symptom Index Disease-Related Symptoms scores; and a composite definition. A scenario incorporating grade $\geq$ 2 toxicities was tested. **Results:** Compared with everolimus, nivolumab was associated with a significant Q-TWiST improvement of 3.3 months ($P < .001$). In all sensitivity analyses, nivolumab was associated with Q-TWiST gains (relative gain %) ranging from 3.3 months (14.4%) to 4.8 months (20.9%). **Conclusions:** Nivolumab is associated with a statistically significant and clinically meaningful gain in quality-adjusted OS versus everolimus among previously treated patients with advanced renal cell carcinoma.

Clinical Genitourinary Cancer, Vol. 17, No. 5, 356-65 © 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: CheckMate 025, Immunotherapy, Kidney cancer, Quality-adjusted survival, Risk-benefit

Introduction

Renal cell carcinoma (RCC) is the most common type of kidney cancer and one of the most common cancers in the United States.[1,2] Globally, approximately 338,000 new RCC cases are diagnosed annually, including 30% at advanced stage.[3]

Several targeted therapies are indicated for advanced RCC (aRCC), including vascular endothelial growth factor/platelet-derived growth factor pathway inhibitors and mammalian target of rapamycin inhibitors.[4] Everolimus, a mammalian target of rapamycin inhibitor, is recommended for patients with aRCC previously treated with a first-line therapy (ie, sunitinib or pazopanib).[5]

Nivolumab is a humanized immunoglobulin G4 monoclonal antibody targeting programmed cell death 1 (PD-1) receptor. It is approved by the United States Food and Drug Administration (FDA) for patients with advanced/metastatic RCC whose disease

[1]Pharmerit International, Bethesda, MD
[2]Bristol-Myers Squibb, Princeton, NJ
[3]Northwestern University Feinberg School of Medicine, Chicago, IL
[4]Memorial Sloan Kettering Cancer Center, New York, NY

Submitted: Feb 25, 2019; Revised: May 10, 2019; Accepted: May 20, 2019; Epub: May 31, 2019

Address for correspondence: Marc Botteman, MSc, MA, Partner, Pharmerit International LP, 4350 East West Hwy #1100, Bethesda, MD 20814
E-mail contact: mbotteman@pharmerit.com

https://doi.org/10.1016/j.clgc.2019.05.010

progressed while on an antiangiogenic therapy. Nivolumab blocks the interaction between PD-1 and PD-1 ligand, which inhibits T-cell activation.[6] This could restore antitumor immunity and lead to improvements in overall survival (OS).[6-8] A phase III trial (CheckMate 025) comparing nivolumab with everolimus among patients with aRCC whose previous treatments had failed found nivolumab resulted in a better overall objective response rate (ORR, 25% vs. 5%; odds ratio, 5.98; $P < .001$), longer OS (hazard ratio, 0.73; $P = .002$), and fewer grade 3 or 4 treatment-related adverse events (AEs, 19% vs. 37%).[9]

Regulators and clinicians are paying increasing attention to understanding the clinical risk/benefit of various oncology therapies from a patient perspective. This includes the European Medicine Agency[10-12] and the United States FDA,[13] as well as the American Society of Clinical Oncology,[14] which formally defines a net health benefit score based on clinical benefits and toxicities. The quality-adjusted time without symptoms or toxicity (Q-TWiST)[15,16] method has been used since the mid-1980s to assess the net benefits of oncology treatments across 50 cancers,[17] including RCC,[18,19] in terms of the quantity (ie, OS, progression-free survival [PFS], and AEs) and quality (ie, patient health utilities) of survival gained.[20,21]

The present analysis estimated the Q-TWiST of nivolumab versus everolimus based on the CheckMate 025 trial, while accounting for immunotherapy-relevant definitions of progression.

Materials and Methods

Data Source/Study Population

This was a post-hoc analysis of the patient-level data from the phase III, open-label CheckMate 025 trial (ClinicalTrials.gov number, NCT01668784), in which patients with advanced/metastatic RCC with a clear-cell component who had received ≤ 2 prior antiangiogenic therapies were randomized (1:1) to either nivolumab (at biweekly intravenous doses of 3 mg/kg of body weight) or everolimus (daily oral 10-mg tablet dose). Patients were treated until progression, unacceptable toxicity, withdrawal of consent, or study end. The primary endpoint of the trial was OS, defined as the duration from randomization to death (if occurred). The secondary endpoints were ORR, PFS, and incidence of AEs.[9]

This analysis used the intent-to-treat cohort from the CheckMate 025 trial. Trial inclusion and exclusion criteria can be found elsewhere.[22]

Statistical Analysis

Q-TWiST Method. The OS time was partitioned into 3 health states: TOX, time after randomization and before progression/censoring during which patients experienced any grade ≥ 3 AEs; TWiST, time after randomization and before progression/censoring without any toxicity; and REL, time after disease progression until death/censoring (assessed based on Response Evaluation Criteria in Solid Tumors [RECIST] 1.1 criteria in the base case). RECIST 1.1 qualifies progression with the appearance of new lesions and/or an absolute increase in tumor size of $\geq 20\%$ and 5 mm versus nadir.[23] Each state's restricted mean duration was obtained by calculating the area under the Kaplan-Meier (K-M) curve. The following steps were used to build the K-M curves and calculate the restricted mean duration for each health state using the proc lifetest procedure in SAS 9.4:

- The K-M curves for TOX, PFS, and OS were built separately.
- Following this, the area under each K-M curve was calculated. This gave us the restricted mean durations for TOX, PFS, and OS.
- TOX time was calculated as the area under the TOX curve.
- TWiST time was calculated as the difference between the area under K-M the curve for the PFS and TOX curves.
- REL time was calculated as the difference between the area under the K-M curves for the OS and the PFS curves.

A 45-month cutoff for the maximum follow-up was used to estimate restricted means. Differences in mean health state durations between treatment arms were tested using log-rank tests. The Q-TWiST was then calculated summing up the time in each state multiplied by its respective utility weight (U) ranging from 0 to 1, to reflect patient preferences for time spent in each state, as per the following equation:

$$\text{Q-}TWiST = (TOX * U_{TOX}) + (TWiST * U_{TWiST}) + (REL * U_{REL})$$

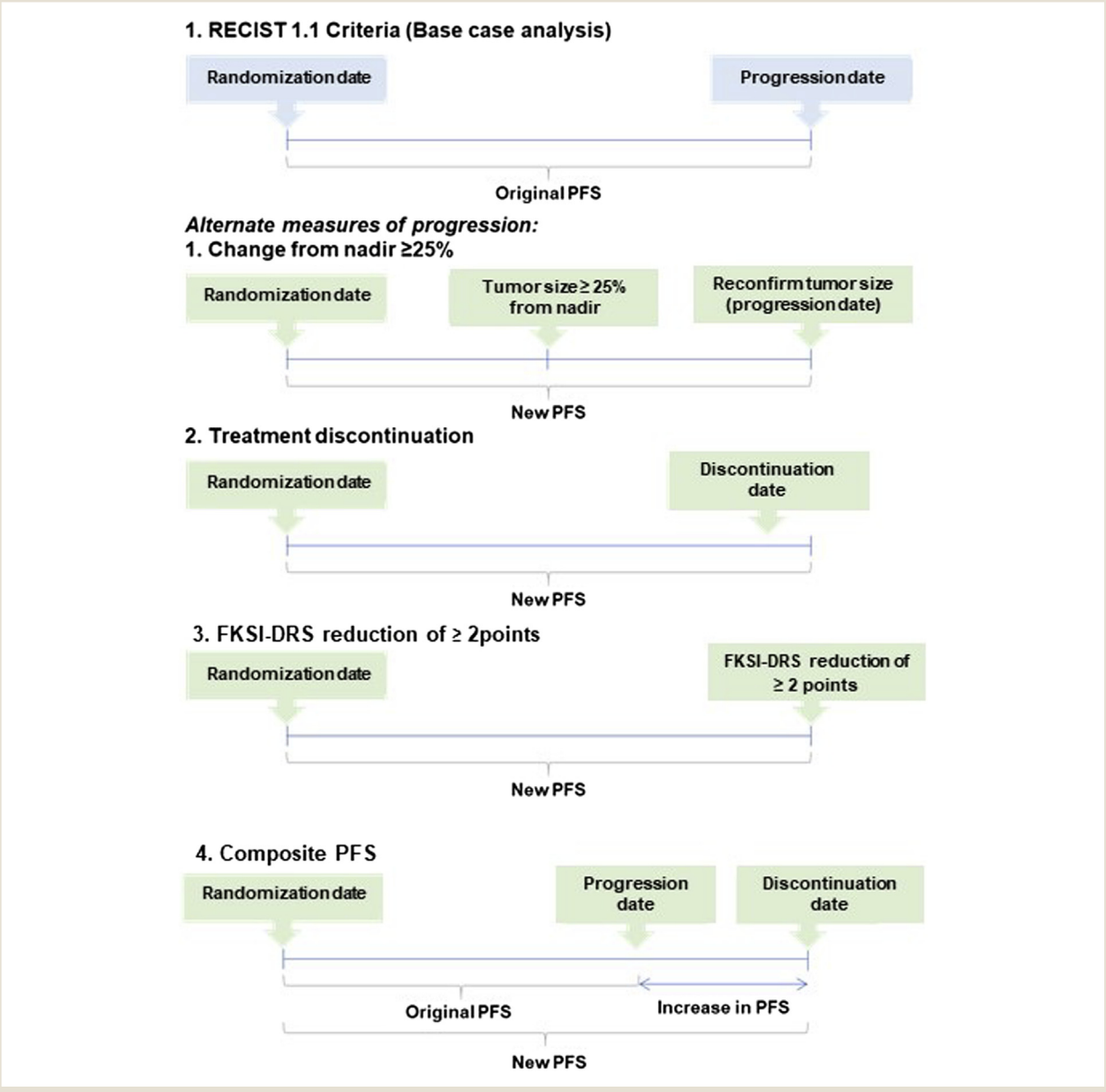
To assess the precision of the mean restricted time in each health state, overall Q-TWiST, and difference in Q-TWiST, 95% confidence intervals (CIs) were computed using 1000 bootstrapped (with replacement) samples of trial patients. Finally, we calculated the relative Q-TWiST gain (ie, difference in Q-TWiST between arms divided by mean OS of everolimus). Using the Revicki et al criteria,[24] relative Q-TWiST gains of $\geq 10\%$ and $\geq 15\%$ were considered clinically important and clearly clinically important, respectively.

Six key assumptions were considered in the base case:

- Fixed utility values were adopted ($U_{TWiST} = 1$, $U_{TOX} = 0.5$, $U_{REL} = 0.5$). These utilities have been commonly used in the Q-TWiST literature. In a systematic review of 51 Q-TWiST studies, 30 (60.8%) used $U_{TOX} = 0.5$, $U_{REL} = 0.5$, and $U_{TWiST} = 1.0$.[19] This set of utilities was also suggested in the original 1995 article by Gelber et al that introduced the Q-TWiST method.[25]
- Regardless of the type, severity, or extent of AE-related symptoms, U_{TOX} was set at 0.5.
- Both treatment-related and unrelated grade ≥ 3 AEs were included for TOX calculation.
- Each AE had a start and end date. The end date for each AE was the resolution date, disease progression date, death date, or end of follow-up, whichever occurred first. TOX duration was calculated as the number of days spent with grade ≥ 3 AEs before disease progression, death, or end of follow-up, whichever occurred first.
- A day with multiple AEs was counted only once.
- All days with AEs before progression were grouped together to calculate total time in TOX irrespective of whether AEs occurred consecutively or not.

Threshold Analysis and Q-TWiST Gain Function. A threshold analysis illustrates the results for the treatment comparison (nivolumab vs. everolimus) whereby the TOX and REL utility values were varied between 0 and 1 (ie, to cover the entire range of possible values). This was presented graphically in a Q-TWiST "plane" indicating scenarios in which the benefits are statistically significant

Figure 1 Alternative Progression Measures Considered for REL



Abbreviations: FSKI-DRS = Functional Assessment of Cancer Therapy-Kidney Symptom Index-Disease Related Symptoms; PFS = progression-free survival; RECIST = Response Evaluation Criteria in Solid Tumors; REL = time after progression.

($P < .05$ based on 1000 bootstraps). The range of Q-TWiST differences were also calculated at different restricted follow-up times to create a Q-TWiST gain function over time. The results of a Q-TWiST analysis depend on the values of the utility coefficients used for each health state. The threshold utility analysis is key because it allows one to determine the preferred treatment (ie, Q-TWiST gain) given any choice of values for U_{TOX} and U_{REL}.

Subgroup Analysis. Analyses were performed for the following prespecified subgroups: age (≥ 65 years, < 65 years), gender, region (United States/Canada, Western Europe, other), prior anti-angiogenic therapy regimens in the advanced or metastatic setting (1, 2), and Memorial Sloan Kettering Cancer Center (MSKCC) prognostic risk group (favorable, intermediate, poor).[26]

Immuno-Oncology-Specific Sensitivity Analyses

In a traditional Q-TWiST, the transition of a patient from TOX/TWiST to REL is governed by PFS. However, traditional measures to assess cancer progression in non-immunotherapies may not be suitable when evaluating treatment effectiveness for immunotherapies. For the current analysis, in the base case, REL was assessed using the RECIST 1.1 criteria, which defines progression with the appearance of new lesions and/or an absolute increase in tumor size of ≥ 20% and 5 mm versus nadir.[27] Using this definition may lead

Table 1 Restricted Mean Duration of Health States at 45 Months (Base Case)

Health State	Mean Duration (95% CI), mo			
	Nivolumab (n = 410)	Everolimus (n = 411)	Difference	*P* Value
TOX	0.1 (0.1-0.1)	0.5 (0.3-0.6)	−0.4 (−0.5 to −0.2)	<.001
TWiST	9.7 (8.5-11.0)	7.0 (6.2-7.7)	2.7 (1.3-4.2)	<.001
REL	17.0 (15.5-18.5)	15.5 (14.2-16.9)	1.5 (-0.5-3.5)	.142
PFS	9.8 (8.6-11.1)	7.4 (6.5-8.2)	2.4 (0.9-3.9)	.002
OS	26.8 (25.2-28.4)	23.0 (21.3-24.4)	3.8 (1.7-6.2)	.001

Base case refers to the scenario where TOX included grade ≥ 3 AEs and disease progression was measured using the Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 criteria.
Abbreviations: AEs = adverse events; CI = confidence interval; OS = overall survival; PFS = progression-free survival; REL = time after progression; TOX = time with grade ≥ 3 toxicity; TWiST = time without symptoms of disease progression or toxicity.

to a premature declaration of progression when the treatment effects from immunotherapies are not fully realized.[23,28,29] Specifically, the RECIST 1.1 criteria neglects the importance of the 'flare effect' (ie, the pseudo-progression), resulting from inflammation of the cancer tissue from the immunotherapy because of which there may be an increase in tumor size before eventual shrinkage. This flare effect may lead to the false impression on radiographic assessments that the tumor has grown or that new lesions have emerged when in fact this is not the case.[30,31]

Therefore, in addition to declaring progression using RECIST 1.1 criteria (base case scenario), the following alternative definitions of progression were considered (Figure 1): tumor burden increase ≥ 25% from nadir (reconfirmed at a subsequent visit) adapted from the immune-related response criteria developed by Wolchok et al[30,32]; treatment discontinuation; reduction from baseline in Functional Assessment of Cancer Therapy-Kidney Symptom Index-Disease Related Symptoms (FKSI-DRS) score of ≥ 2 to 3 points (ie, the minimally important difference for FKSI-DRS[33]); and meeting ≥ 2 of the following criteria (ie, composite criteria): traditional progression, treatment discontinuation, or FKSI-DRS reduction of ≥ 2 points from baseline.

The European Society for Medical Oncology clinical guidelines on management of immunotherapy-related toxicities recommend interrupting treatment upon development of grade ≥ 2 immune-related skin toxicity, endocrinopathies, hepatotoxicity, pneumonitis, or gastrointestinal hepatotoxicity.[34] Therefore, a sensitivity analysis was conducted where TOX included grade ≥ 2 AEs.

Results

Base Case Q-TWiST Analysis (Using RECIST 1.1 Criteria)

In the intent-to-treat population, the median OS was 19.7 months (95% CI, 17.6-22.1 months) for everolimus (n = 411) versus 25.8 months (95% CI, 22.2-29.8 months) for nivolumab (n = 410). Deaths occurred in 273 (66.59%) and 298 (72.51%) patients in the nivolumab and everolimus trial arms, respectively. The median PFS was 4.5 months (95% CI, 3.7-5.5 months) for everolimus versus 4.2 months (95% CI, 3.7-5.4 months) for nivolumab. Progression was observed in 349 (85.12%) and 346 (84.18%) patients in the nivolumab and everolimus trial arm, respectively. Fifty-eight (14.15%) and 120 (29.20%) patients in the nivolumab and everolimus trial arm experienced a grade ≥ 3 AE. The median follow-up time for the nivolumab arm was 45.04 months and the everolimus arm was 44.19 months. Therefore, the 45-month cutoff was used to conduct the analysis.

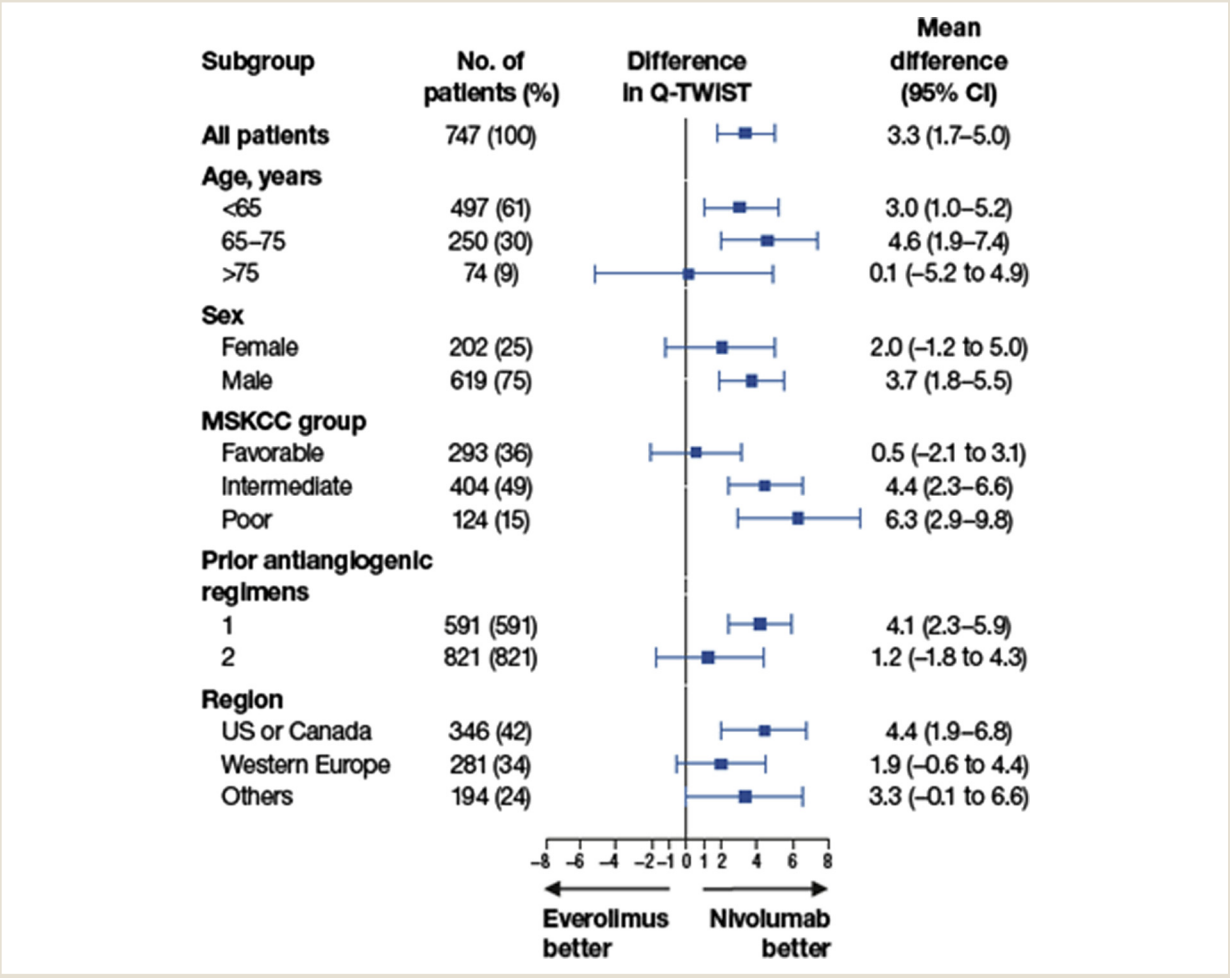
Nivolumab patients had a significantly longer mean restricted TWiST time (difference, 2.7 months; 95% CI, 1.3-4.2 months; $P < .001$), significantly shorter TOX time (difference, 0.4 months; 95% CI, −0.5 to −0.2 months; $P < .001$), and longer REL time (difference, 1.5 months; 95% CI, −0.5 to 3.5 months; $P = .142$)

Table 2 Threshold Utility Analysis (Base Case)

Utility		Mean Q-TWiST (95% CI), mo				
TOX	REL	Nivolumab (n = 410)	Everolimus (n = 411)	Difference	Improvement	*P* Value
0	0	9.7 (8.5-11.0)	7.0 (6.2-7.7)	2.7 (1.3-4.2)	11.8%	<.001
0	0.5	18.2 (16.9-19.4)	14.7 (13.6-15.7)	3.5 (1.9-5.1)	15.3%	<.001
0	1	26.7 (25.1-28.3)	22.5 (20.9-23.9)	4.2 (2.0-6.5)	18.3%	<.001
0.5	0	9.7 (8.5-11.1)	7.2 (6.4-8.0)	2.5 (1.1-4.0)	10.9%	.001
0.5	0.5	18.2 (17.0-19.5)	15.0 (13.8-15.9)	3.3 (1.7-5.0)	14.4%	<.001
0.5	1	26.7 (25.1-28.3)	22.7 (21.0-24.1)	4.0 (1.9-6.4)	17.4%	.001
1	0	9.8 (8.6-11.1)	7.4 (6.5-8.2)	2.4 (0.9-3.9)	10.5%	.002
1	0.5	18.3 (17.0-19.5)	15.2 (14.1-16.1)	3.1 (1.5-4.8)	13.5%	<.001
1	1	26.8 (25.2-28.4)	23.0 (21.3-24.4)	3.8 (1.7-6.2)	16.6%	<.001

Base case refers to the scenario where TOX included grade ≥ 3 adverse events and disease progression was measured using the Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 criteria.
Abbreviations: CI = confidence interval; Q-TWiST = quality-adjusted time without symptoms of disease progression or toxicity; REL = time after progression; TOX = time with grade ≥ 3 toxicity.

Figure 2 Subgroup Analyses of Q-TWiST Difference



Abbreviations: CI = confidence interval; MSKCC = Memorial Sloan Kettering Cancer Center; Q-TWiST = quality-adjusted time without symptoms of disease progression or toxicity; US = United States.

(Table 1) versus everolimus patients, as well as statistically significantly greater quality-adjusted OS (difference, 3.3 months; 95% CI, 1.7-5.0 months; $P < .001$), representing a 14.4% relative Q-TWiST gain (Table 2).

Threshold analyses showed significantly positive Q-TWiST differences for nivolumab versus everolimus across the full range of TOX and REL utility values (Table 2 and Supplemental Figure 1 [in the online version]). Specifically, mean Q-TWiST differences ranged from 2.4 months (95% CI, 0.9-3.9 months; $P = .002$) to 4.2 months (95% CI, 2.0-6.5 months; $P < .001$), and relative Q-TWiST gains ranged from 10.5% to 18.3%. The relative gains in Q-TWiST increased over time, from 3.7% at 6 months' follow up to 14.4% at 45 months' follow-up (see Supplemental Figure 2 in the online version). When stratified by subgroups, Q-TWiST value differences varied from 0.1 to 4.6 months (Figure 2) and were significant in patients who were male, aged < 75 years at baseline, from the United States/Canada, in poor or intermediate MSKCC risk groups at baseline, or had received 1 prior antiangiogenic regimen (all $P < .01$).

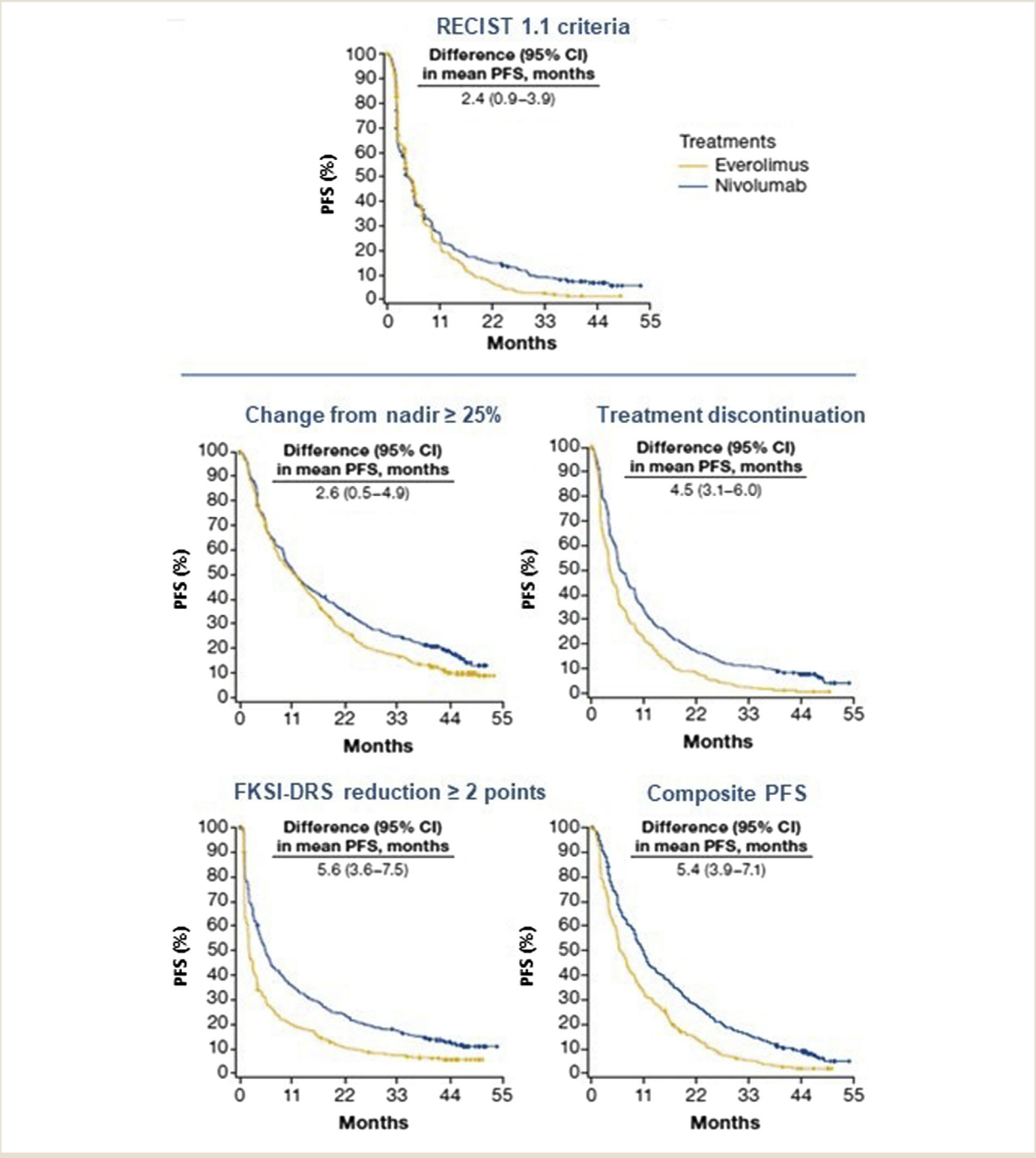
Immuno-oncology Specific Sensitivity Analyses

Mean PFS improved when alternate measures of progression were considered (Figure 3). Q-TWiST gains and relative gains across all scenarios using alternate measures of progression have been reported in Table 3. Compared with the base case (Q-TWiST gain of 3.3 months, relative Q-TWiST gain of 14.4%), the alternate measures of progression resulted in higher Q-TWiST gains varying from 3.5 to 5.6 months (all $P < .01$) and higher relative Q-TWiST gains ranging from 15.3% to 24.4%, with the highest improvement observed for the definition based on a FKSI-DRS reduction of ≥ 3 points.

Discussion

Q-TWiST is a well-established and generally accepted method that was originally developed specifically to assess quality-adjusted survival in oncology. It can be a useful tool in making treatment choices for patients and physicians when they consider trade-offs between the clinical benefits and toxicity of various treatments. Our analysis demonstrated that nivolumab is associated with a

Figure 3 Kaplan-Meier Curve of PFS Based on Different Relapse Definitions



Abbreviations: CI = confidence interval; FKSI-DRS = Functional Assessment of Cancer Therapy-Kidney Symptom Index-Disease Related Symptoms; PFS = progression-free survival; RECIST = Response Evaluation Criteria in Solid Tumors.

quality-adjusted OS benefit of 3.3 months (corresponding to a 14.4% relative Q-TWiST gain) versus everolimus in previously treated patients with aRCC. In the threshold utility analysis, the relative Q-TWiST gain ranged from 10.5% to 16.6%, all above the threshold of clinical importance.[24] Such findings are also consistent with prior trial analysis results that have demonstrated the clinical and quality-of-life benefits of nivolumab versus everolimus as a second-line treatment in aRCC.[9,35]

This is the first Q-TWiST analysis conducted among previously treated patients with aRCC. Prior studies conducted among treatment-naive patients with aRCC have reported relative Q-TWiST gains of 14.57% (sunitinib vs. interferon alpha [IFN-α][20]),

Table 3 Sensitivity Analyses of Q-TWiST Gains

Health State	Mean (95% CI), mo			
	Nivolumab (n = 410)	Everolimus (n = 411)	Difference	*P* Value
Traditional RECIST criteria (base case analysis)				
TOX	0.1 (0.1-0.1)	0.5 (0.3-0.6)	−0.4 (−0.5 to −0.2)	<.001
TWiST	9.7 (8.5-11.0)	7.0 (6.2-7.7)	2.7 (1.3-4.2)	<.001
REL	17.0 (15.5-18.5)	15.5 (14.2-16.9)	1.5 (−0.5 to 3.5)	.049
Q-TWiST	18.2 (17.0-19.5)	15.0 (13.8-15.9)	3.3 (1.7-5.0)	<.001
Relative Q-TWiST gain	14.4%			
Alternate measures of progression				
1. Change from nadir ≥ 25%				
TOX	0.2 (0.1-0.3)	0.8 (0.5-1.1)	−0.6 (−0.9 to −0.3)	.028
TWiST	18.2 (16.7-19.9)	15.1 (13.7-16.5)	3.1 (1.0-5.3)	.005
REL	8.3 (7.1-9.6)	7.1 (5.9-8.3)	1.3 (−0.4 to 3.0)	.134
Q-TWiST	22.5 (21.1-23.9)	19.0 (17.6-20.3)	3.5 (1.6-5.6)	<.001
Relative Q-TWiST gain	15.3%			
2. Treatment discontinuation				
TOX	0.1 (0.1-0.2)	0.5 (0.3-0.6)	−0.4 (−0.5 to −0.2)	<.001
TWiST	11.7 (10.5-12.9)	6.9 (6.1-7.6)	4.8 (3.4-6.3)	<.001
REL	14.9 (13.7-16.2)	15.6 (14.1-16.8)	−0.6 (−2.6 to 1.3)	.547
Q-TWiST	19.3 (18.0-20.6)	14.9 (13.9-15.9)	4.3 (2.8-6.0)	<.001
Relative Q-TWiST gain	18.7%			
3. FKSI-DRS reduction ≥ 2 points				
TOX	0.1 (0.0-0.1)	0.2 (0.1-0.3)	−0.2 (−0.3 to −0.1)	.028
TWiST	13.0 (11.6-14.5)	7.2 (6.1-8.4)	5.7 (3.8-7.7)	<.001
REL	13.7 (12.1-15.3)	15.5 (13.8-17.2)	−1.8 (−4.2 to 0.4)	.173
Q-TWiST	19.9 (18.6-21.2)	15.1 (13.9-16.1)	4.8 (3.1-6.6)	<.001
Relative Q-TWiST gain	20.9%			
4. FKSI-DRS reduction ≥ 3 points				
TOX	0.1 (0.1-0.1)	0.3 (0.2-0.5)	−0.2 (−0.4 to −0.1)	.009
TWiST	15.9 (14.3-17.6)	8.6 (7.2-9.8)	7.4 (5.2-9.6)	<.001
REL	10.8 (9.3-12.2)	14.1 (12.5-15.8)	−3.3 (−5.6 to −1.1)	.004
Q-TWiST	21.4 (19.9-22.9)	15.8 (14.5-16.9)	5.6 (3.8-7.5)	<.001
Relative Q-TWiST gain	24.4%			
5. Composite PFS				
TOX	0.2 (0.1-0.2)	0.6 (0.4-0.8)	−0.4 (−0.7 to −0.2)	.002
TWiST	15.6 (14.3-16.9)	9.7 (8.7-10.7)	5.8 (4.3-7.4)	<.001
REL	11.1 (9.8-12.4)	12.6 (11.2-14.1)	−1.6 (−3.5 to 0.4)	.043
Q-TWiST	21.2 (19.9-22.4)	16.3 (15.1-17.3)	4.8 (3.2-6.6)	<.001
Relative Q-TWiST gain	20.9%			
6. Including grade ≥ 2 adverse events				
TOX	1.5 (1-1.9)	2.8 (2.2-3.3)	−1.3 (−2 to −0.6)	<.001
TWiST	8.3 (7.2-9.5)	4.6 (4-5.2)	3.7 (2.5-5.1)	<.001
REL	17.0 (15.5-18.5)	15.5 (14.2-16.9)	1.5 (−0.5 to 3.5)	.141
Q-TWiST	17.6 (16.4-18.7)	13.8 (12.8-14.7)	3.8 (2.3-5.4)	<.001
Relative Q-TWiST gain	16.6%			

Abbreviations: CI = confidence interval; FSKI-DRS = Functional Assessment of Cancer Therapy-Kidney Symptom Index-Disease Related Symptoms; PFS = progression-free survival; Q-TWiST = quality-adjusted time without symptoms of disease progression or toxicity; RECIST = Response Evaluation Criteria in Solid Tumors; REL = time after progression; TOX = time with grade ≥ 3 toxicity; TWiST = time without symptoms of disease progression or toxicity.

15.7% (temsirolimus vs. IFN-α[18]), < 5% (pazopanib vs. sunitinib[21]), and 11.47% (cabozanitib vs. sunitinib[36]). Our results may be benchmarked against a recently published systematic review of 81 previously conducted Q-TWiST comparisons.[19] Specifically, the relative Q-TWiST gains of nivolumab reported herein are greater than reported across prior immuno-oncology therapy assessments (14.4% vs. 8.9%).

A key goal of RCC treatment is tumor control while maintaining, if not improving, quality of life.[37] The Q-TWiST method takes into consideration 2 key drivers of quality of life: toxicities and disease progression. Disease progression endpoints are included because they are routinely reported in clinical trials and are thought to correlate with quality of life. However, the traditional progression definition (using RECIST criteria) is potentially less correlated with quality of life in solid tumor immunotherapies (owing to pseudo-progression) than traditional chemotherapies.[38,39] Therefore, to more fully account for nivolumab's treatment benefits, 4 alternative progression definitions beyond the base case (RECIST 1.1) were considered. Importantly, these alternate measures do not refer to PFS in a clinical sense but rather are proxies for symptoms or milestones in the patient experience that could correlate with quality of life.

In the first scenario analysis, progression was defined as an increase in tumor size of ≥ 25% from nadir. This measure was adapted from the immune-related response criteria developed by Wolchok et al[30,32] specifically to address the shortcomings of the RECIST 1.1 criteria for immunotherapies.[27,40] The resulting relative Q-TWiST gain of 15.3% for nivolumab versus everolimus in this scenario may be considered clearly clinically important.[24] The second scenario analysis accounted for the pseudo-progression effect by measuring progression only after treatment discontinuation. The post-hoc analysis of the CheckMate 025 trial revealed that there was a longer median time to discontinuation among nivolumab than everolimus patients, despite similar median PFS.[41] This means that treatment discontinuation at first occurrence of progression may be considered premature, as it underestimates an immunotherapy's benefit over time and post-progression. Of note, in a subgroup analysis of a phase II study of nivolumab versus everolimus, George et al demonstrated clinical benefits in those who continued nivolumab beyond first progression, in terms of lower incidence of all-grade treatment-related AEs compared with patients who discontinued.[42] The results from the second alternative measure were consistent with the findings by George et al, and the relative Q-TWiST gain of 18.7% in our analysis may be considered clearly clinically important.

Patient-reported outcomes have shown associations with survival endpoints and have been increasingly utilized as prognostic tools for predicting survival in solid cancers.[43] Cella et al revealed a robust relationship between baseline quality-of-life scores and median PFS/OS among patients with RCC receiving sunitinib and IFN-α.[44] In their analysis, the total baseline FKSI 15-item (FSKI-15) score was a significant predictor of survival outcomes ($P < .0001$). Therefore, for the third alternative progression measure scenario, a change of ≥ 2 to 3 points in the FKSI-DRS score from baseline (consistent with the FKSI-DRS minimally important difference[33]) was considered to be indicative of progression. Relative Q-TWiST gains of 20.9% and 24.4% (for ≥ 2- and ≥ 3-point reductions, respectively) were observed under this scenario, the highest across all 4 scenarios. A similar result was also observed for the composite PFS measure, which incorporated all alternative progression measures.

Another key aspect of immunotherapy treatment is a consideration of drug-related toxicities. The European Society for Medical Oncology guidelines on management of immunotherapy-related toxicities suggest interrupting treatment upon occurrence of grade ≥ 2 AEs.[34] Therefore, a sensitivity analysis included all grade ≥ 2 AEs as a part of TOX, resulting in a relative Q-TWiST gain of 16.6%, which suggests that the time spent with grade ≥ 2 toxicities was significantly longer ($P < .0001$) for everolimus versus nivolumab patients. To fully quantify the benefit of immunotherapies, it is important to include lower-grade toxicities that could impact patient quality of life.

The present analysis has certain limitations. First, established definitions of progression for immunotherapy, such as iRECIST,[45] could not be applied in this study because the CheckMate 025 trial's patient data were collected using scheduled assessments per the RECIST 1.1 framework. Clearly, the use of the iRECIST criteria in the present analysis would have been of interest. Second, multiple AEs co-occurring were counted only once to avoid duplication, and AEs were not differentiated by type. These assumptions were made to ensure the current analysis remains consistent with previous Q-TWiST analyses.[46] However, one might expect that the inclusion of all AEs would have affected the assessment of net benefits.

Conclusion

In conclusion, this study presents a Q-TWiST analysis of nivolumab versus everolimus in the CheckMate 025 trial population, consisting of patients with aRCC who had received prior anti-angiogenic therapy. The base case analysis showed significant TWiST benefit for nivolumab as well as clinically meaningful relative gains in Q-TWiST, which highlights the survival and quality-of-life benefits of nivolumab as a second-line treatment of aRCC. Moreover, the relative Q-TWiST improvement increased when using alternate definitions of progression that better reflect the treatment potential of immunotherapies, suggesting that traditional progression definitions may need to be reconsidered when assessing immunotherapy benefits in solid tumor cancers such as RCC.

Clinical Practice Points

- This study examined the net clinical benefits of nivolumab versus everolimus among previously treated patients with aRCC enrolled in the CheckMate 025 trial.
- We divided the patient's survival time into 3 health states: time before progression spent with grade 3/4 toxicities (TOX), time after progression (REL), and time without toxicity or progression (TWiST). Time in each health state was weighted by its respective utility to calculate the Q-TWiST gain.
- Given the limitations of the RECIST 1.1 criteria in correctly assessing progression for immuno-oncology treatments, 4 alternate progression (REL) definitions were used as sensitivity analyses. These alternate scenarios included assessing an increase in tumor size (reconfirmed at a subsequent diagnosis), treatment discontinuation, a decrease in patient-reported FKSI-DRS score by ≥ 2 points, and a composite scenario integrating the base case definition and the other 3 alternate scenarios.

- Nivolumab was associated with a statistically significant and clinically meaningful gain in quality-adjusted OS versus everolimus in the base case and across all sensitivity analyses.
- Nivolumab improves both the quantity and quality of survival gain among previously treated patients with aRCC.

Acknowledgments

This study was funded by Bristol-Myers Squibb. The authors would like to acknowledge the following Pharmerit project team members in providing medical writing assistance for this manuscript: Youngmin Kwon and Catherine Mirvis.

Disclosure

R.S. is employed by Pharmerit International. R.S.'s institution has received consulting or advisory role and research funding from Bristol-Myers Squibb and numerous other sponsors from the pharmaceutical industry. M.F.B. is employed by Pharmerit International and has stock or ownership interests in Pharmerit. M.F.B.'s institution has received consulting or advisory role and research funding from Bristol-Myers Squibb and numerous other sponsors from the pharmaceutical industry. C.T.S. is employed by Pharmerit International. C.T.S.'s institution has received consulting or advisory role and research funding from Bristol-Myers Squibb and numerous other sponsors from the pharmaceutical industry. L.L. is employed by Pharmerit International. L.L.'s institution has received consulting or advisory role and research funding from Bristol-Myers Squibb and numerous other sponsors from the pharmaceutical industry. J.D. is an employee of Bristol-Myers Squibb and owns Bristol-Myers Squibb stock. D.C. is a consultant and reports personal fees from Bristol-Myers Squibb, Pfizer, Abbvie, and Novartis, outside the submitted work. R.M. reports grants from Bristol-Myers Squibb during the conduct of the study; grants and personal fees from Pfizer, Novartis, Eisai, Exelixis, and Genentech/Roche outside the submitted work.

Supplemental Data

Supplemental figures accompanying this article can be found in the online version at https://doi.org/10.1016/j.clgc.2019.05.010.

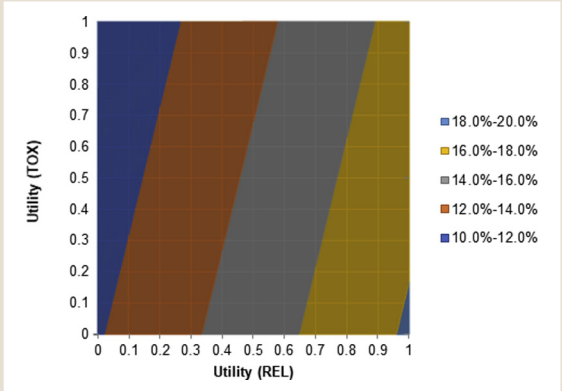
References

1. Chow WH, Dong LM, Devesa SS. Epidemiology and risk factors for kidney cancer. *Nat Rev Urol* 2010; 7:245-57.
2. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin* 2012; 62:10-29.
3. Fisher R, Gore M, Larkin J. Current and future systemic treatments for renal cell carcinoma. *Semin Cancer Biol* 2013; 23:38-45.
4. Motzer RJ, Agarwal N, Beard C, et al. NCCN clinical practice guidelines in oncology: kidney cancer. *J Natl Compr Canc Netw* 2009; 7:618-30.
5. Motzer RJ, Jonasch E, Agarwal N, et al. Kidney cancer, version 2.2017, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw* 2017; 15: 804-34.
6. Nurieva RI, Liu X, Dong C. Molecular mechanisms of T-cell tolerance. *Immunol Rev* 2011; 241:133-44.
7. Hamanishi J, Mandai M, Iwasaki M, et al. Programmed cell death 1 ligand 1 and tumor-infiltrating CD8+ T lymphocytes are prognostic factors of human ovarian cancer. *Proc Natl Acad Sci U S A* 2007; 104:3360-5.
8. Thompson RH, Kuntz SM, Leibovich BC, et al. Tumor B7-H1 is associated with poor prognosis in renal cell carcinoma patients with long-term follow-up. *Cancer Res* 2006; 66:3381-5.
9. Motzer RJ, Escudier B, McDermott DF, et al; CheckMate 025 Investigators. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med* 2015; 373:1803-13.
10. Salvatore L. Management of metastatic colorectal cancer patients: guidelines of the Italian Medical Oncology Association (AIOM). *ESMO Open* 2017; 2:e000147.
11. Postmus D, Richard S, Bere N, et al. Individual trade-offs between possible benefits and risks of cancer treatments: results from a stated preference study with patients with multiple myeloma. *Oncologist* 2018; 23:44-51.
12. Postmus D, Mavris M, Hillege HL, et al. Incorporating patient preferences into drug development and regulatory decision making: results from a quantitative pilot study with cancer patients, carers, and regulators. *Clin Pharmacol Ther* 2016; 99: 548-54.
13. U.S. Food and Drug Administration. *Plan for issuance of patient-focused drug development guidance under 21st Century Cures Act Title III Section 3002* 2017. Available at: https://www.fda.gov/media/105979/download. Accessed: September 5, 2017.
14. Schnipper LE, Davidson NE, Wollins DS, et al; American Society of Clinical Oncology. American Society of Clinical Oncology Statement: a conceptual framework to assess the value of cancer treatment options. *J Clin Oncol* 2015; 33: 2563-77.
15. Goldhirsch A, Gelber RD, Simes RJ, Glasziou P, Coates AS. Costs and benefits of adjuvant therapy in breast cancer: a quality-adjusted survival analysis. *J Clin Oncol* 1989; 7:36-44.
16. Gelber RD, Goldhirsch A. A new endpoint for the assessment of adjuvant therapy in postmenopausal women with operable breast cancer. *J Clin Oncol* 1986; 4:1772-9.
17. Husson O, Jones RL. Q-TWiST: what really matters to the cancer patient? *Cancer* 2017; 123:2200-2.
18. Zbrozek AS, Hudes G, Levy D, et al. Q-TWiST analysis of patients receiving temsirolimus or interferon alpha for treatment of advanced renal cell carcinoma. *Pharmacoeconomics* 2010; 28:577-84.
19. Solem CT, Kwon Y, Shah RM, Aly A, Botteman MF. Systematic review and benchmarking of quality-adjusted time without symptoms or toxicity (Q-TWiST) in oncology. *Expert Rev Pharmacoecon Outcomes Res* 2018; 18:245-53.
20. Patil S, Figlin RA, Hutson TE, et al. Q-TWiST analysis to estimate overall benefit for patients with metastatic renal cell carcinoma treated in a phase III trial of sunitinib vs interferon-alpha. *Br J Cancer* 2012; 106:1587-90.
21. Beaumont JL, Salsman JM, Diaz J, et al. Quality-adjusted time without symptoms or toxicity analysis of pazopanib versus sunitinib in patients with renal cell carcinoma. *Cancer* 2016; 122:1108-15.
22. Escudier B, Sharma P, McDermott DF, et al; CheckMate 025 Investigators. CheckMate 025 randomized phase 3 study: outcomes by key baseline factors and prior therapy for nivolumab versus everolimus in advanced renal cell carcinoma. *Eur Urol* 2017; 72:962-71.
23. Hoos A, Eggermont AMM, Janetzki S, et al. Improved endpoints for cancer immunotherapy trials. *J Natl Cancer Inst* 2010; 102:1388-97.
24. Revicki DA, Feeny D, Hunt TL, et al. Analyzing oncology clinical trial data using the Q-TWiST method: clinical importance and sources for health state preference data. *Qual Life Res* 2006; 15:411-23.
25. Gelber RD, Cole BF, Gelber S, Goldhirsch A. Comparing treatments using quality-adjusted survival: the Q-TWiST method. *Am Statist* 1995; 49:161-9.
26. Motzer RJ, Bacik J, Schwartz LH, et al. Prognostic factors for survival in previously treated patients with metastatic renal cell carcinoma. *J Clin Oncol* 2004; 22:454-63.
27. Hodi FS, Hwu WJ, 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; 34:1510-7.
28. Chen TT. Statistical issues and challenges in immuno-oncology. *J Immunother Cancer* 2013; 1:18.
29. Quinn DI, Shore ND, Egawa S, Gerritsen WR, Fizazi K. Immunotherapy for castration-resistant prostate cancer: progress and new paradigms. *Urol Oncol* 2015; 33:245-60.
30. 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; 15:7412-20.
31. Chow LQ. Exploring novel immune-related toxicities and endpoints with immune-checkpoint inhibitors in non-small cell lung cancer. *Am Soc Clin Oncol Educ Book* 2013.
32. Wolchok J. How recent advances in immunotherapy are changing the standard of care for patients with metastatic melanoma. *Ann Oncol* 2012; 23(Suppl 8): viii15-21.
33. Cella D, Yount S, Brucker PS, et al. Development and validation of a scale to measure disease-related symptoms of kidney cancer. *Value Health* 2007; 10:285-93.
34. Haanen J, Carbonnel F, Robert C, et al; ESMO Guidelines Committee. Management of toxicities from immunotherapy: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2017; 28:iv119-i142.
35. Mazza C, Escudier B, Albiges L. Nivolumab in renal cell carcinoma: latest evidence and clinical potential. *Ther Adv Med Oncol* 2017; 9:171-81.
36. Chen RC, Feuilly M, Meng J, et al. Quality-adjusted time without symptoms or toxicity (Q-TWiST): analysis of cabozantinib (Cabo) vs sunitinib (Sun) in patients with advanced renal cell carcinoma (aRCC) of intermediate or poor risk (Alliance A031203). *J Clin Oncol* 2018; 36(15 Suppl):4556.
37. Rini BI. Metastatic renal cell carcinoma: many treatment options, one patient. *J Clin Oncol* 2009; 27:3225-34.
38. Sundar R, Cho BC, Brahmer JR, Soo RA. Nivolumab in NSCLC: latest evidence and clinical potential. *Ther Adv Med Oncol* 2015; 7:85-96.
39. Chiou VL, Burotto M. Pseudoprogression and immune-related response in solid tumors. *J Clin Oncol* 2015; 33:3541-3.

40. Wallace T, Torre T, Grob M, et al. Current approaches, challenges and future directions for monitoring treatment response in prostate cancer. *J Cancer* 2014; 5:3-24.
41. Johal S, Santi I, Doan J, George S. Is RECIST-defined progression free-survival a meaningful endpoint in the era of immunotherapy? *J Clin Oncol* 2017; 35(6 Suppl):488.
42. George S, Motzer RJ, Hammers HJ, et al. Safety and efficacy of nivolumab in patients with metastatic renal cell carcinoma treated beyond progression: a subgroup analysis of a randomized clinical trial. *JAMA Oncol* 2016; 2:1179-86.
43. Secord AA, Coleman RL, Havrilesky LJ, Abernethy AP, Samsa GP, Cella D. Patient-reported outcomes as end points and outcome indicators in solid tumours. *Nat Rev Clin Oncol* 2015; 12:358-70.
44. Cella D, Bushmakin AG, Cappelleri JC, Charbonneau C, Michaelson MD, Motzer RJ. Baseline quality of life as a prognostic survival tool in patients receiving sunitinib for metastatic renal cell carcinoma. *Br J Cancer* 2012; 106: 646-50.
45. 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:e143-52.
46. Berry V, Basson L, Bogart E, et al. REGOSARC: Regorafenib versus placebo in doxorubicin-refractory soft-tissue sarcoma: a quality-adjusted time without symptoms of progression or toxicity analysis. *Cancer* 2017; 123: 2294-302.

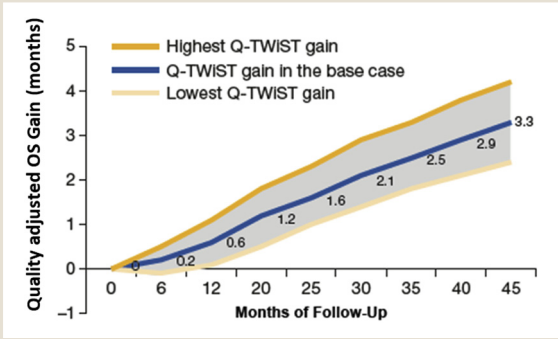
Supplemental Data

Supplemental Figure 1 Q-TWiST Threshold Utility Analysis. The y-axis Represents the Utility for the TOX Health State (U_{TOX}) and the x-axis Represents the Utility for the REL Health State (U_{REL}). Both Vary Across Their Full Range From 0 to 1, Whereas U_{TWiST} Is Fixed at 1. The Diagonal Bands of Different Colors or Shading Represent Relative Q-TWiST Gains. To Assess the Relative Q-TWiST Gain Associated With a Given Combination of U_{TOX} and U_{REL}, One Must Select a Pair of Values for U_{TOX} and U_{REL} on the y-axis and x-axis, Respectively. The Intersection of the U_{TOX} and U_{REL} Values Inside the Plot Indicates the Band to Which the Relative Q-TWiST Gain Belongs for the Given U_{TOX} and U_{REL} Combination. For Instance, in the Base Case $U_{TOX} = 0.5$ and $U_{REL} = 0.5$, Which Intersect in the 14% to 16% Relative Q-TWiST Gain Band and are Consistent with the Point Estimate of 14.4%. Note That All Q-TWiST Gains in This Plot Are Statistically Significant



Abbreviations: Q-TWiST = quality-adjusted time without symptoms of disease progression or toxicity; REL = time after progression; TOX = time with grade ≥ 3 toxicity; U_{REL} = utility of time after progression; U_{TOX} = utility of time with grade ≥ 3 toxicity. All values significant at $P < .01$.

Supplemental Figure 2 Q-TWiST Gain Function Over Time



Abbreviations: OS = overall survival; Q-TWiST = quality-adjusted time without symptoms of disease progression or toxicity.