



Original Investigation | Oncology

# Evaluation of Continuous Tumor-Size–Based End Points as Surrogates for Overall Survival in Randomized Clinical Trials in Metastatic Colorectal Cancer

Tomasz Burzykowski, PhD; Elisabeth Coart, PhD; Everardo D. Saad, MD; Qian Shi, PhD; Dirkje W. Sommeijer, MD, PhD; Carsten Bokemeyer, MD, PhD; Eduardo Díaz-Rubio, MD, PhD; Jean-Yves Douillard, MD, PhD; Alfredo Falcone, MD; Charles S. Fuchs, MD, MPH; Richard M. Goldberg, MD; J. Randolph Hecht, MD; Paulo M. Hoff, MD; Herbert Hurwitz, MD; Fairouz F. Kabbinavar, MD; Miriam Koopman, MD, PhD; Timothy S. Maughan, MD; Cornelis J. A. Punt, MD, PhD; Leonard Saltz, MD; Hans-Joachim Schmoll, MD, PhD; Matthew T. Seymour, MD; Niall C. Tebbutt, MD, PhD; Christophe Tournigand, MD, PhD; Eric Van Cutsem, MD, PhD; Aimery de Gramont, MD, PhD; John R. Zalcberg, MBBS, PhD; Marc Buyse, ScD; for the Aide et Recherche en Cancerologie Digestive Group

## Abstract

**IMPORTANCE** Tumor measurements can be used to estimate time to nadir and depth of nadir as potential surrogates for overall survival (OS).

**OBJECTIVE** To assess time to nadir and depth of nadir as surrogates for OS in metastatic colorectal cancer.

**DESIGN, SETTING, AND PARTICIPANTS** Pooled analysis of 20 randomized clinical trials within the Aide et Recherche en Cancerologie Digestive database, which contains academic and industry-sponsored trials, was conducted. Three sets of comparisons were performed: chemotherapy alone, antiangiogenic agents, and anti-epidermal growth factor receptor agents in first-line treatment for patients with metastatic colorectal cancer.

**MAIN OUTCOMES AND MEASURES** Surrogacy of time to nadir and depth of nadir was assessed at the trial level based on joint modeling of relative tumor-size change vs baseline and OS. Treatment effects on time to nadir and on depth of nadir were defined in terms of between-arm differences in time to nadir and in depth of nadir, and both were assessed in linear regressions for their correlation with treatment effects (hazard ratios) on OS within each set. The strengths of association were quantified using sample-size-weighted coefficients of determination ( $R^2$ ), with values closer to 1.00 indicating stronger association. At the patient level, the correlation was assessed between modeled relative tumor-size change and OS.

**RESULTS** For 14 chemotherapy comparisons in 4289 patients, the  $R^2$  value was 0.63 (95% CI, 0.30-0.96) for the association between treatment effects on time to nadir and OS and 0.08 (95% CI, 0-0.37) for depth of nadir and OS. For 11 antiangiogenic agent comparisons (4854 patients), corresponding values of  $R^2$  were 0.25 (95% CI, 0-0.72) and 0.06 (95% CI, 0-0.35). For 8 anti-epidermal growth factor receptor comparisons (2684 patients), corresponding values of  $R^2$  were 0.24 (95% CI, 0-0.83) and 0.21 (95% CI, 0-0.78).

**CONCLUSIONS AND RELEVANCE** In contrast with early reports favoring depth of response as a surrogate, these results suggest that neither time to nadir nor depth of nadir is an acceptable surrogate for OS in the first-line treatment of metastatic colorectal cancer.

JAMA Network Open. 2019;2(9):e1911750. doi:10.1001/jamanetworkopen.2019.11750

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

JAMA Network Open. 2019;2(9):e1911750. doi:10.1001/jamanetworkopen.2019.11750

## Key Points

**Question** Can end points based on the kinetics of tumor size after treatment be used as surrogates for overall survival in metastatic colorectal cancer?

**Findings** In this pooled analysis of data from 20 randomized clinical trials, time to nadir and depth of nadir were modeled and assessed as potential surrogates for overall survival at the patient and trial levels. The associations found were weak or moderate; there were notable differences in tumor-size kinetics between antiangiogenic agents and anti-epidermal growth factor receptor agents.

**Meaning** The implications of these results for early drug development and clinical practice are unclear and warrant further studies; the findings of this study reinforce the need to develop more reliable end points that reflect tumor biology and patient benefit.

## + Supplemental content

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

## Introduction

The availability of active treatments for use in subsequent lines have called into question the use of overall survival (OS) as a primary end point in phase 3 trials on first-line therapy for metastatic colorectal cancer (mCRC).<sup>1</sup> As a result, there has been a long-standing interest in developing and validating surrogate end points for OS in this setting.<sup>2-3</sup> Such validation requires demonstration of a strong association between the surrogate and the final end point at the patient level (ie, patients with improvements in the surrogate end point also tend to have improvements in the final end point) and a strong association between the treatment effects on the surrogate end point and the final end point (the trial-level association).<sup>4</sup> Tumor-size-based end points have generated interest in the search for early treatment end points in mCRC.<sup>5-9</sup> These end points may be categorical or continuous and, among the latter type, the end point receiving the most attention has been the depth of response, defined as the maximum percent tumor shrinkage during treatment. In work published in abstract form, the depth of response was found to be associated with OS at the patient level in first-line cetuximab-based therapy.<sup>10</sup> That study was based on 2 randomized trials and did not assess the trial-level surrogacy. To obtain a more in-depth view of this question, we assessed the individual- and trial-level surrogacy for OS of 2 continuous tumor-size-based end points in first-line treatment of mCRC.

## Methods

### Trial Selection and Definition of Contrasts

Tumor measurements and OS data were available from 20 first-line randomized clinical trials in mCRC within the Aide et Recherche en Cancerologie Digestive (ARCAD) database (**Table 1**).<sup>11-30</sup> To evaluate the trial-level surrogacy, unbiased estimates of treatment effects are needed; hence, the clinical trial database was used. While our analysis used data from several randomized clinical trials, it is not a classic meta-analysis attempting to evaluate pooled estimates of treatment effects. As such, the study follows the recently published Reporting of Surrogate Endpoint Evaluation Using Meta-analyses (ReSEEM) Reporting Guidelines<sup>31</sup> rather than the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines.

Tumor measurements consisted of the longest diameters of target lesions, used in the original trials according to the Response Evaluation Criteria in Solid Tumors (RECIST) guideline, version 1.1.<sup>32</sup> Eight trials involved only chemotherapy; of the 12 trials that had at least 1 biological agent, 6 evaluated antiangiogenic (anti-ANG) agents as the only biological, 4 investigated an anti-epidermal growth factor receptor (anti-EGFR) agent as the only biological, and 2 trials had both an anti-ANG and anti-EGFR agent. The analysis was based on comparisons between 2 arms (henceforth termed *contrasts*) nested within trials, with control and experimental arms defined according to historical evolution. An exception to this rule was made for HORIZON III,<sup>24</sup> for which the cediranib arm was considered as control to have bevacizumab as the uniform experimental intervention for anti-ANG agents (Table 1). For 8 trials with more than 2 arms, each experimental arm was compared with a control arm created by randomly splitting the set of patients originally randomized to the control arm. This procedure was applied to avoid including each patient twice in the analysis, which would artificially induce a correlation that would confound the associations under investigation.

### Statistical Analysis

Target lesions measured up to 24 months after randomization were used, as 98% of the available postbaseline measurements were made within 24 months. Individual trials had tumor-assessment schedules that varied between 6 and 12 weeks, but this variation does not influence the models used here. Overall survival was defined as the time from randomization to death from any cause, with censoring of data from patients who were alive at the last contact date. Separate analyses were conducted for chemotherapy-only contrasts, anti-ANG-agent contrasts, and anti-EGFR-agent

contrasts. Because *KRAS* (OMIM \*190070) is a predictive biomarker for anti-EGFR treatment, only patients with wild-type *KRAS* were considered in contrasts evaluating the effects of such treatments. For trials of different treatment sequences, only contrasts for which the 2 arms testing different regimens at the beginning of the treatment sequence were analyzed. For the Bolus, Infusional, or Capecitabine With Camptosar-Celecoxib trial,<sup>12</sup> treatment arms with celecoxib were not analyzed.

**Table 1. Control and Experimental Arms for the 3 Treatment Classes Included in the Analysis**

Study	Contrast	Treatment (Sample Size, No.) <sup>a,b</sup>	
		Control	Experimental
<b>Chemotherapy Alone (n = 4289)</b>			
Díaz-Rubio et al, <sup>11</sup> 2007	O3-TTD-01	FUOX (136)	XELOX (137)
Fuchs et al, <sup>12</sup> 2007	BICC-C A	FOLFIRI (28)	Modified IFL (61)
Fuchs et al, <sup>12</sup> 2007	BICC-C B	FOLFIRI (27)	CAPIRI (54)
Tournigand et al, <sup>13</sup> 2004	C97-3	FOLFIRI → FOLFOX6 (79)	FOLFOX6 → FOLFIRI (86)
Koopman et al, <sup>14</sup> 2007	CAIRO1	Capecitabine → irinotecan → XELOX (295)	CAPIRI → XELOX (291)
Seymour et al, <sup>15</sup> 2007	FOCUS A	Fluorouracil/leucovorin → irinotecan (231)	FOLFIRI (231)
Seymour et al, <sup>15</sup> 2007	FOCUS B	Fluorouracil/leucovorin → I (227)	FOLFOX (235)
Seymour et al, <sup>16</sup> 2011	FOCUS2 A	Fluorouracil/leucovorin (74)	FOLFOX (80)
Seymour et al, <sup>16</sup> 2011	FOCUS2 B	Capecitabine (77)	XELOX (78)
Falcone et al, <sup>17</sup> 2007	GONO	FOLFIRI (33)	FOLFOXIRI (46)
Saltz et al, <sup>18</sup> 2008	N016966 A	FOLFOX4 (284)	XELOX (284)
Saltz et al, <sup>18</sup> 2008	N016966 B	FOLFOX4 (160)	XELOX (162)
Goldberg et al, <sup>19</sup> 2004	N9741 A	IFL (149)	FOLFOX (300)
Goldberg et al, <sup>19</sup> 2004	N9741 B	rIFL (171)	Irinotecan, oxaliplatin (273)
<b>Antiangiogenic Agents (n = 4854)</b>			
Tebbutt et al, <sup>20</sup> 2010	AGITG (MAX) A	Capecitabine (75)	Capecitabine + bevacizumab (140)
Tebbutt et al, <sup>20</sup> 2010	AGITG (MAX) B	Capecitabine (68)	Capecitabine + bevacizumab + mitomycin C (138)
Hurwitz et al, <sup>21</sup> 2004	AVF2107g A	IFL (187)	IFL + bevacizumab (363)
Hurwitz et al, <sup>21</sup> 2004	AVF2107g B	IFL (176)	Fluorouracil/leucovorin + bevacizumab (98)
Kabbinavar et al, <sup>22</sup> 2005	AVF2192g	Fluorouracil/leucovorin (80)	Fluorouracil/leucovorin + bevacizumab (95)
Hoff et al, <sup>23</sup> 2012	HORIZON II A	FOLFOX/XELOX (171)	FOLFOX/XELOX + cediranib (474)
Hoff et al, <sup>23</sup> 2012	HORIZON II B	FOLFOX/XELOX (170)	FOLFOX/XELOX + cediranib (198)
Schmoll et al, <sup>24</sup> 2012	HORIZON III A	FOLFOX + cediranib (654)	FOLFOX + bevacizumab (329)
Schmoll et al, <sup>24</sup> 2012	HORIZON III B	FOLFOX + cediranib (172)	FOLFOX + bevacizumab (330)
Saltz et al, <sup>18</sup> 2008	N016966 C	FOLFOX4 (161)	FOLFOX4 + bevacizumab (310)
Saltz et al, <sup>18</sup> 2008	N016966 D	XELOX (156)	XELOX + bevacizumab (309)
<b>Anti-EGFR Agents (n = 2684)</b>			
Tol et al, <sup>25</sup> 2009	CAIRO2	CAPOX + bevacizumab (126)	CAPOX + bevacizumab + cetuximab (128)
Maughan et al, <sup>26</sup> 2011	COIN A	Fluorouracil/leucovorin/oxaliplatin (99)	Fluorouracil/leucovorin/oxaliplatin + cetuximab (82)
Maughan et al, <sup>26</sup> 2011	COIN B	Capecitabine/oxaliplatin (189)	Capecitabine/oxaliplatin + cetuximab (184)
Van Cutsem et al, <sup>27</sup> 2009	CRYSTAL	FOLFIRI (324)	FOLFIRI + cetuximab (291)
Bokemeyer et al, <sup>28</sup> 2009	OPUS	FOLFOX (88)	FOLFOX + cetuximab (76)
Hecht et al, <sup>29</sup> 2009	PACCE (C249) A	Oxaliplatin based + bevacizumab (188)	Oxaliplatin based + bevacizumab + panitumumab (178)
Hecht et al, <sup>29</sup> 2009	PACCE (C249) B	Irinotecan based + bevacizumab (51)	Irinotecan based + bevacizumab + panitumumab (50)
Douillard et al, <sup>30</sup> 2010	PRIME (C203)	FOLFOX4 (318)	FOLFOX4 + panitumumab (312)

Abbreviations: AGITG, Australasian Gastro-Intestinal Cancer Trials Group; anti-EGFR, anti-epidermal growth factor receptor; BICC, Bolus, Infusional, or Capecitabine With Camptosar-Celecoxib; CAPIRI, capecitabine, irinotecan; CAPOX, capecitabine, oxaliplatin; FOCUS, Fluoxetine or Control Under Supervision; FOLFIRI, fluorouracil, leucovorin, irinotecan; FOLFOX, fluorouracil, leucovorin, oxaliplatin; FOLFOXIRI, fluorouracil, leucovorin, oxaliplatin, irinotecan; FUOX, fluorouracil, oxaliplatin; GONO, Gruppo Oncologico Nord Ovest; IFL, irinotecan, fluorouracil, leucovorin; MAX, Mitomycin C, Avastin and Xeloda; rIFL, reduced-dose irinotecan, fluorouracil, leucovorin; PACCE, Panitumumab Advanced Colorectal Cancer Evaluation; PRIME, Panitumumab

Randomized Trial in Combination With Chemotherapy for Metastatic Colorectal Cancer to Determine Efficacy; TTD, Spanish Cooperative Group for Gastrointestinal Tumor Therapy; XELOX, capecitabine, oxaliplatin; and →, subsequently.

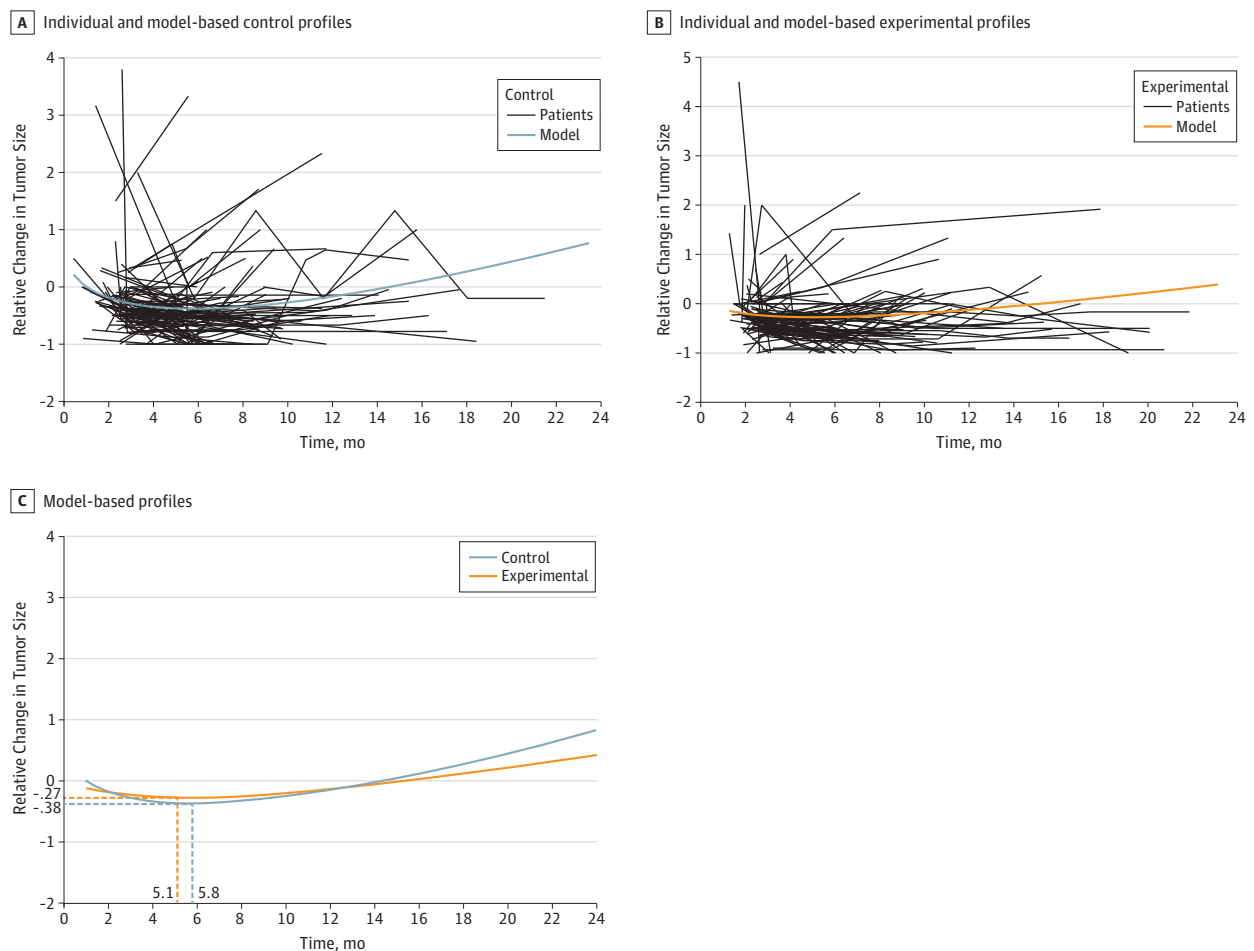
<sup>a</sup> Sample sizes may differ from those reported in the original publications owing to exclusion of patients in the present analysis (see Methods section for details).

<sup>b</sup> Numbers with the combination regimens (eg, FOLFOX6) are used by the original developers of these regimens to denote subsequent versions and improvements in the administration schedule.

Tumor-size measurements (the sum of all target lesions) were modeled using the relative tumor-size change (RTSC) vs baseline, defined (for time  $t$ ) as follows:  $RTSC(t) = (\text{tumor size at time } t - \text{tumor size at baseline}) / (\text{tumor size at baseline})$ .

Repeated values of RTSC and the time to death were analyzed in joint models.<sup>33,34</sup> In particular, RTSC measures were analyzed by linear mixed-effects models with contrast-specific fixed and random linear and square-root time effects. Overall survival was analyzed by proportional hazards models that included the random effects from the RTSC model to account for the association between RTSC and survival time. Based on the joint models, treatment effects on RTSC and OS were estimated. For OS, the effects were estimated using the natural logarithm of the hazard ratio (HR) obtained from the proportional hazards model (logHR). For RTSC, the outcomes were defined based on the mean treatment-specific time profiles estimated using the linear mixed-effects model. In particular, for each profile, the nadir (ie, the local minimum RTSC value) was obtained, together with the time at which the nadir took place. Treatment effects were then defined in terms of differences in time to nadir and differences in depth of nadir; the latter variable is analogous to depth of response but is estimated from the model rather than coming directly from patient data. **Figure 1** illustrates

Figure 1. Longitudinal Profiles



A, Relative tumor-size changes over time for individual patients and the model-based estimated profile for the control group. B, Relative tumor-size changes over time for individual patients and the model-based estimated profile for the experimental group. C, Based on the model-based profiles, the nadir for the control arm is estimated to occur at 5.8 months, with the depth of nadir  $-0.38$  (ie, a 38% reduction of the tumor mass relative to baseline). Corresponding figures for the experimental arm are 5.1 months for

the time of occurrence of the nadir and  $-0.27$  (ie, 27% reduction of the tumor mass relative to baseline) for the depth of nadir. Consequently, the effect of experimental treatment in terms of time to nadir and depth of nadir is equal to  $5.1 - 5.8 = -0.7$  months and  $-0.27 - (-0.38) = 0.11$ . That is, in the experimental arm, the nadir occurs earlier and is 11% smaller (ie, less deep) than in the control arm.

the calculation of longitudinal profiles for 1 of the contrasts. For differences in time to nadir, negative values indicate that the nadir occurs earlier with experimental treatment; for differences in depth of nadir, negative values indicate that the nadir is deeper with experimental treatment.

To assess the validity of time to nadir and depth of nadir as surrogates for OS, we applied the correlation approach.<sup>33</sup> Specifically, a linear regression was fitted to the estimated pairs of treatment effects on time to nadir or depth of nadir and OS. The regression was weighted by the contrast-specific sample size. The coefficient of determination ( $R^2$ ) was used to quantify the strength of association at the trial level between the treatment effects on time to nadir or depth of nadir and OS. An  $R^2$  value greater than 0.75 was considered an indicator of good surrogacy.<sup>35,36</sup> We also quantified the strength of association at the individual level between RTSC and OS. With this aim, we measured the correlation between the individual random effects included in the linear mixed-effects model for RTSC and the proportional hazards model for OS using a correlation coefficient, denoted by  $R(t)$ .<sup>33</sup> This correlation coefficient is a time-dependent measure, since the association between RTSC and the death process can be defined relative to any time over the course of tumor-size measurements. In the analysis, 2-sided 95% CIs were used. Analyses were conducted with SAS, version 9.4 (SAS Institute Inc) and Stata, version 13.1 (StataCorp LLC).

## Results

### Chemotherapy Alone

There were 6224 patients in the ARCAD database enrolled in 9 trials eligible for this analysis (8 trials involving only chemotherapy and 1 trial that included bevacizumab but provided chemotherapy-alone contrasts). After excluding patients without any tumor-size information or with tumor-size measurements available only more than 24 months after randomization, 4289 patients (68.9%) could be analyzed (Table 1). Such patients were grouped in 14 contrasts, with the median follow-up per trial ranging from 14 to 128 months. eFigure 1A in the [Supplement](#) presents the Kaplan-Meier OS curves for these 14 contrasts, with the corresponding HRs presented in **Table 2**.

eFigure 2A in the [Supplement](#) presents the estimated, model-based longitudinal profiles for each contrast in these trials. The corresponding estimates of treatment effects in terms of the differences in time to nadir and depth of nadir are presented in Table 2. There was large variability in the treatment effects, reflecting relatively small and inconsistent differences in the longitudinal profiles (eFigure 2A in the [Supplement](#)). For instance, for time to nadir, the estimated treatment effects varied (Table 2) from -4.53 months (BICC-C C) to 4.77 months (Gruppo Oncologico Nord Ovest). For depth of nadir, the range was from -0.49 (FOCUS B) to 0.17 (BICC-C C). For 1 comparison (FOCUS2 B), the effects could not be obtained because the estimated RTSC profile for the experimental arm did not reach a local minimum (the profile was a strictly increasing function of time).

The associations between the differences in time to nadir and logHRs for OS, as well as between the differences in depth of nadir and logHRs for OS, are presented in **Figure 2**, with a weighted regression line. The estimated value of  $R^2$  was 0.63 (95% CI, 0.30-0.96) for the association between the treatment effects on time to nadir and OS, and 0.08 (95% CI, 0-0.37) for the association between the treatment effects on depth of nadir and OS. eFigure 3A in the [Supplement](#) presents the estimated values of  $R(t)$  that quantify the association at the individual level between RTSC and OS at time  $t$ . At all considered time points,  $R(t)$  values were 0.9 or larger. Thus, the plot indicates that RTSC values provide much information on a patient's OS.

### Anti-ANG Agents

For anti-ANG agent contrasts, data on 5390 patients enrolled in 6 trials were available for analysis. After excluding patients with no tumor-size information or with tumor-size measurements available only more than 24 months after randomization, 4854 (90.1%) of the patients could be analyzed (Table 1). Eleven contrasts could be formed, with median follow-up in each trial ranging from 14 to 31

months. eFigure 1B in the [Supplement](#) shows the OS curves for each of these contrasts, and the corresponding HRs are presented in Table 2. eFigure 2B in the [Supplement](#) presents the longitudinal RTSC profiles for these contrasts, and the corresponding estimates of treatment effects on time to nadir and on depth of nadir are presented in Table 2. All effects on time to nadir were positive, suggesting that the nadir for the experimental treatments took place later than for the control treatments. At the same time, all but 2 (for HORIZON III A and N016966 C) effects on depth of nadir

**Table 2. Estimated Time to Nadir and Depth of Nadir**

Contrast	Time to Nadir, mo <sup>a</sup>			Depth of Nadir, m <sup>b</sup>			HR for OS <sup>d</sup>
	Control	Experimental	Treatment Effect <sup>c</sup>	Control	Experimental	Treatment Effect <sup>c</sup>	
<b>Chemotherapy Alone</b>							
O3-TTD-01	5.81	5.09	-0.72	-0.38	-0.27	0.11	1.06
BICC-C A	5.89	5.26	-0.64	-0.32	-0.36	-0.04	1.07
BICC-C C	11.04	6.52	-4.52	-0.46	-0.29	0.17	1.57
C97-3	5.32	4.82	-0.50	-0.40	-0.34	0.06	0.83
CAIRO1	2.72	3.73	1.01	-0.06	-0.24	-0.18	0.80
FOCUS A	0.66	3.05	2.39	-0.12	-0.09	0.02	0.88
FOCUS B	2.97	2.70	-0.27	0.23	-0.26	-0.49	0.93
FOCUS2 A	1.17	2.00	0.83	0.40	-0.07	-0.47	1.01
FOCUS2 B	0.05	NA	NA	-0.01	NA	NA	0.99
GONO	6.37	11.14	4.77	-0.43	-0.66	-0.23	0.78
N016966 A	5.25	4.82	-0.43	-0.38	-0.40	-0.02	0.89
N016966 B	6.17	4.59	-1.58	-0.43	-0.36	0.07	1.16
N9741 A	4.66	7.31	2.65	-0.28	-0.44	-0.16	0.68
N9741 B	4.57	4.75	0.18	-0.24	-0.27	-0.01	0.90
<b>Antiangiogenic Agents</b>							
AGITG (MAX) A	3.42	4.13	0.70	-0.15	-0.26	-0.11	0.88
AGITG (MAX) B	3.05	4.82	1.78	-0.11	-0.28	-0.17	1.07
AVF2107g A	4.02	6.34	2.32	-0.26	-0.37	-0.11	0.73
AVF2107g B	3.66	6.66	2.99	-0.21	-0.27	-0.06	0.80
AVF2192g	3.88	4.51	0.63	-0.24	-0.26	-0.02	0.91
HORIZON II A	4.98	5.28	0.30	-0.38	-0.39	-0.01	0.88
HORIZON II B	4.56	5.93	1.37	-0.32	-0.41	-0.09	0.96
HORIZON III A	5.70	6.64	0.93	-0.35	-0.35	0.00	1.09
HORIZON III B	5.38	5.92	0.54	-0.30	-0.34	-0.04	1.00
N016966 C	5.37	6.79	1.42	-0.37	-0.36	0.01	0.85
N016966 D	4.93	6.04	1.10	-0.33	-0.36	-0.03	0.82
<b>Anti-EGFR Agents</b>							
CAIRO2	6.78	5.21	-1.57	-0.26	-0.33	-0.07	1.13
COIN A	6.37	8.34	1.97	-0.31	-0.40	-0.09	0.76
COIN B	5.82	2.97	-2.85	-0.03	-0.30	-0.27	1.09
CRYSTAL	6.28	8.26	1.98	-0.31	-0.46	-0.16	0.74
OPUS	7.83	10.23	2.40	-0.34	-0.55	-0.22	0.86
PACCE (C249) A	7.40	7.77	0.37	-0.37	-0.31	0.06	1.48
PACCE (C249) B	171.1	7.99	-163.1	-0.78	-0.37	0.41	1.76
PRIME (C203)	8.36	9.22	0.86	-0.40	-0.48	-0.08	0.81

Abbreviations: AGITG, Australasian Gastro-Intestinal Cancer Trials Group; anti-EGFR, anti-epidermal growth factor receptor; BICC, Bolus, Infusional, or Capecitabine With Camptosar-Celecoxib; FOCUS, Fluoxetine or Control Under Supervision; GONO, Gruppo Oncologico Nord Ovest; HR, hazard ratio; OS, overall survival; PACCE, Panitumumab Advanced Colorectal Cancer Evaluation; PRIME, Panitumumab Randomized Trial in Combination With Chemotherapy for Metastatic Colorectal Cancer to Determine Efficacy; TTD, Spanish Cooperative Group for Gastrointestinal Tumor Therapy.

<sup>a</sup> For differences in time to nadir, negative values indicate that the nadir occurred earlier with experimental treatment.

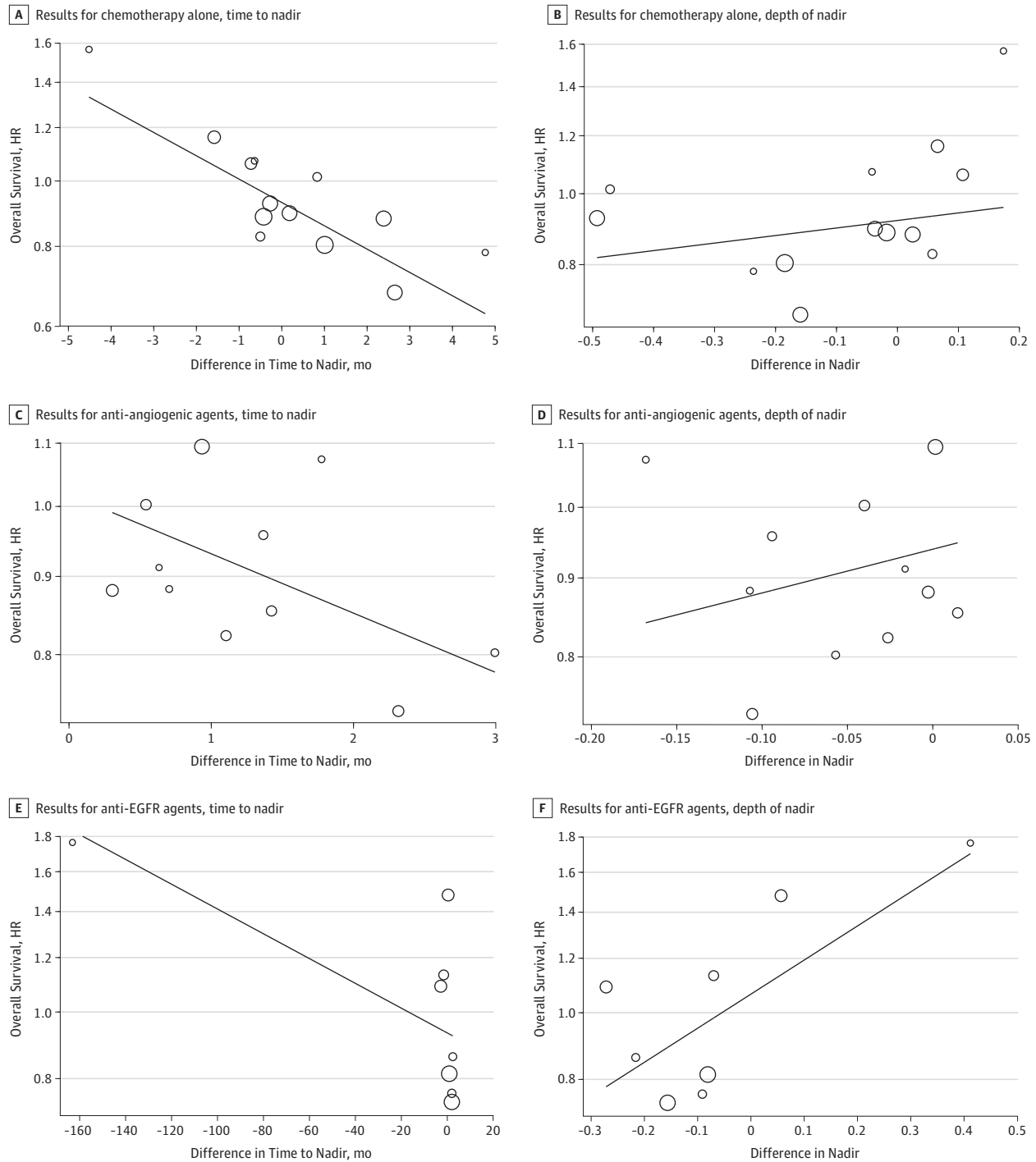
<sup>b</sup> For differences in depth of nadir, negative values indicate that the nadir was deeper with experimental treatment.

<sup>c</sup> Experimental minus control.

<sup>d</sup> Hazard ratios may differ from those reported in the original publications owing to exclusion of patients in the present analysis and the use of a different modeling framework (a joint model for relative tumor-size change and OS).

were negative, suggesting that the experimental treatments led to a larger relative reduction in tumor size than the control treatments. This finding reflects that the RTSC profiles for the control arms exhibited a higher curvature than the profiles for the experimental arms (eFigure 2B in the Supplement).

Figure 2. Trial-Level Associations Between Treatment Effects



Hazard ratios (HRs) of overall survival associated with time to nadir and depth of nadir in the chemotherapy-alone (A and B), antiangiogenic agent (C and D), and anti-epidermal growth factor receptor agent (E and F) groups. The difference in nadir is the difference

between the model-estimated mean relative tumor-size change at nadir (relative to baseline) in each contrast. The line indicates weighted regression; the sizes of the circles are proportional to the total sample sizes of the corresponding contrasts.

The associations between treatment effects on time to nadir and depth of nadir and on OS are shown in Figure 2B. The estimated value of  $R^2$  was 0.25 (95% CI, 0-0.72) for the association between the treatment effects on time to nadir and OS and 0.06 (95% CI, 0-0.35) for the association between the treatment effects on depth of nadir and OS. eFigure 3B in the [Supplement](#) depicts the association at the individual level between RTSC and OS at time  $t$ . Values of  $R(t)$  become larger than 0.9 for  $t$  of approximately 6 months. Thus, the plot suggests that, initially, RTSC values provided relatively little information on a patient's OS. However, as additional information on tumor size was gathered over time during the first year of treatment, RTSC achieved a better predictive strength for OS, with no further gain in the subsequent year.

### Anti-EGFR Agents

Of 3081 eligible patients enrolled in 6 trials involving anti-EGFR agents, 2684 patients (87.1%) could be analyzed after excluding those without any tumor-size information or with tumor-size measurements available only more than 24 months after randomization (Table 1). These patients were grouped into 8 contrasts, and the median follow-up in each trial ranged from 10 to 47 months. eFigure 1C in the [Supplement](#) presents the OS curves for these contrasts. The corresponding HRs are reported in Table 2. eFigure 2C in the [Supplement](#) presents the longitudinal RTSC profiles for these contrasts, and the corresponding estimates of the treatment effects on time to nadir and depth of nadir are given in Table 2. Although the effects on time to nadir show some heterogeneity (range from -2.85 for COIN B to 2.40 for OPUS, excluding PACCE [C249] B), once again, all but 2 (for PACCE [C249] A and B) of the effects on depth of nadir were negative, suggesting that the experimental treatments led to larger tumor shrinkage than the control treatments. This finding reflects that the RTSC profiles for the experimental arms seem to be shifted down as compared with the control-arm profiles, while exhibiting roughly a similar curvature (eFigure 2C in the [Supplement](#)). An exception was the PACCE B comparison, for which the estimated RTSC profile for the control arm decreased, unlike for the experimental arm. As a consequence, the estimated time to nadir for the control arm was long (equal to 171.1 months) and resulted in treatment effects on time to nadir (-163.1) and depth of nadir (0.41) that were markedly different from the other comparisons (Table 2).

The associations between treatment effects are depicted in Figure 2E and F. All comparisons were taken into account, and the estimated value of  $R^2$  was 0.24 (95% CI, 0-0.83) for the association between the treatment effects on time to nadir and OS and 0.21 (95% CI, 0-0.78) for the association between the treatment effects on depth of nadir and OS. When the PACCE B comparison was excluded from the analysis, the estimates of  $R^2$  were 0.36 (95% CI, 0-0.97) for depth of nadir and 0.18 (95% CI, 0-0.74) for OS. eFigure 3C in the [Supplement](#) depicts the individual-level association between RTSC and OS at time  $t$ . At all considered time points, values of  $R(t)$  are smaller than 0.4, suggesting that RTSC provided little information on a patient's OS.

## Discussion

Given the continuum of care in mCRC, it becomes increasingly difficult to demonstrate gains in OS in first-line treatment trials. This difficulty has heightened interest in alternative strategies, such as adaptive designs<sup>37</sup> and the use of surrogate end points, including those based on tumor measurements. The latter approach is contrary to the key finding from the present study that neither time to nadir nor depth of nadir can be considered a valid surrogate for OS using contemporary regimens for first-line therapy of mCRC. At best, time to nadir appears to display a moderate association with OS at the trial level with chemotherapy alone or combined with an anti-ANG agent, while depth of nadir appears to display a weak association with OS in all treatment classes. Another finding from this study is the apparent difference between the response kinetics of regimens that include an anti-ANG agent and those that involve an anti-EGFR agent.

The difference in tumor-growth kinetics between anti-ANG and anti-EGFR agents may warrant further exploration. Data presented in Table 2 and eFigure 2 in the [Supplement](#) suggest that the



addition of an anti-ANG agent to chemotherapy is associated with a later, although not often deeper, nadir. Conversely, the addition of an anti-EGFR agent often produces a deeper nadir, with less-conclusive results about its timing of occurrence. These exploratory observations are based on a relatively small number of contrasts, but they may support the clinical impression that the addition of an anti-EGFR agent produces a larger influence on the depth of responses than the addition of an anti-ANG agent. Albeit subject to bias owing to the above-mentioned reasons, the often-divergent slopes after nadir between control and experimental arms as shown in eFigure 2 in the Supplement suggest that the tumor-growth kinetics with both classes of agents are not marked by a rebound effect after progression. The differences in tumor-growth kinetics among different classes of agents are also reflected on the individual-level associations between the RTSC and OS processes. For chemotherapy, it seems that RTSC may provide a strong prediction of a patient's survival. For anti-ANG agents, a strong correlation might be inferred after the initial half-year of treatment. However, for anti-EGFR agents, the correlation appeared to be weak. These individual-level estimates depend largely on the form of the models applied and should be interpreted with caution.

### Strengths and Limitations

Strengths of this study are the large sample size and representativeness in terms of contemporary first-line therapy. Moreover, results of this study suggest that the dimensions of measurable tumor lesions can be modeled to provide information on tumor-growth kinetics. In this sense, our approach differs from the one used by Mansmann et al,<sup>10</sup> who did not model tumor size as a function of time and did not estimate trial-level associations, which is a current requirement for surrogacy validation.<sup>38</sup>

This study has limitations. The chief limitation of this study is the absence of tumor measurements for all patients, which is a potential source of bias through exclusion of individuals with features that may differ systematically from those of included patients. Likewise, extended RAS testing was not available at the time that these trials were conducted, leading to a predictably small percentage of patients being falsely considered as having wild-type tumors. Moreover, no data were available on tumor sidedness or other potential prognostic or predictive molecular markers, such as the status of microsatellite instability, *BRAF*, or *HER2*. Limitations also apply to the model building, which is affected by the absence of postprogression measurements. Moreover, if progression is due to new lesions before the sum of target lesions has reached the nadir, there is increased uncertainty in the estimation of time to nadir and depth of nadir. Also, new lesions could not be included in the definition of RTSC, because the size of such lesions was not reported. In addition, the strength of the association between treatment effects on time to nadir or depth of nadir and on OS was assessed by using a linear regression model weighted by the sample size to account for the uncertainty in the estimated treatment effects. A methodologically more appropriate approach would be to take into account estimates of the SEs and correlation of the estimated treatment effects.<sup>39</sup> However, obtaining such estimates for the joint model used in our analysis was not possible, because the model was fitted by using the expectation-maximization algorithm.

### Conclusions

Neither time to nadir nor depth of nadir appears to be an acceptable surrogate for OS. These findings are not surprising, given the weak trial-level association between conventional response rates and OS in mCRC, despite their association with OS at the patient level, both in mCRC and advanced breast cancer.<sup>40,41</sup> This distinction indicates that achieving response may convey prognostic information for patients in clinical practice, but at the same time suggests that response-based end points cannot replace OS in clinical trials. In none of the treatment classes analyzed was the association between treatment effects strong enough to warrant reasonable precision of the prediction of the treatment effect on OS from the effect on time to nadir or depth of nadir. Such a reasonable precision of the prediction is currently considered the key requirement for a surrogate end point.<sup>38</sup> Nevertheless, at

least for chemotherapy and targeted agents, the use of response-based end points in early-phase trials has been helpful in selecting regimens for further testing in phase 3 trials. Moreover, in clinical practice, a deeper response may help in controlling symptoms and increase the chance of performing secondary resections. Therefore, the implications of these results for early drug development and clinical practice are unclear and warrant further studies. In addition, the findings of this study reinforce the need to develop more reliable end points that reflect tumor biology and patient benefit.

## ARTICLE INFORMATION

**Accepted for Publication:** July 28, 2019.

**Published:** September 20, 2019. doi:10.1001/jamanetworkopen.2019.11750

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

**Corresponding Author:** Tomasz Burzykowski, PhD, International Drug Development Institute, Avenue Provinciale 30-1340, Louvain-la-Neuve, Belgium ([tomasz.burzykowski@iddi.com](mailto:tomasz.burzykowski@iddi.com)).

**Author Affiliations:** International Drug Development Institute, Louvain-la-Neuve, Belgium (Burzykowski, Coart, Saad); Hasselt University, Diepenbeek, Belgium (Burzykowski, Buysse); Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota (Shi); The University of Sydney, Camperdown, New South Wales, Australia (Sommeijer); Academic Medical Centre, Amsterdam, the Netherlands (Sommeijer); Flevohospital, Almere, the Netherlands (Sommeijer); Department of Internal Medicine II and Clinic, University of Hamburg, Hamburg, Germany (Bokemeyer); Hospital Clinico San Carlos and Centro de Investigación Biomédica en Red Cáncer, CIBERONC, Madrid, Spain (Díaz-Rubio); Centre René Gauducheau, St Herblain, France (Douillard); University Hospital S Chiara, Pisa, Italy (Falcone); Dana-Farber Cancer Institute, Boston, Massachusetts (Fuchs); West Virginia University Cancer Institute, Morgantown (Goldberg); David Geffen School of Medicine, University of California, Los Angeles (Hecht, Kabbinnavar); Instituto de Câncer do Estado de São Paulo, São Paulo, Brazil (Hoff); Genentech, South San Francisco, California (Hurwitz); Department of Medical Oncology, University Medical Centre Utrecht, Utrecht University, Utrecht, the Netherlands (Koopman); Cancer Research UK and the Medical Research Council Oxford Institute for Radiation Oncology, Oxford, United Kingdom (Maughan); Amsterdam University Medical Centrum, Department of Medical Oncology, University of Amsterdam, Amsterdam, the Netherlands (Punt); Memorial Sloan-Kettering Cancer Center, New York, New York (Saltz); Martin-Luther University, Halle, Germany (Schmoll); St James's Hospital, University of Leeds, Leeds, United Kingdom (Seymour); Austin Health, Heidelberg, Victoria, Australia (Tebbutt); Hôpital Henri Mondor, Creteil, France (Tournigand); Division of Digestive Oncology, University Hospitals Gasthuisberg Leuven, Leuven, Belgium (Van Cutsem); Katholieke Universiteit, Leuven, Belgium (de Gramont); Franco-British Institute, Levallois-Perret, France (de Gramont); School of Public Health and Preventative Medicine, Monash University, Melbourne, Australia (Zalcberg); International Drug Development Institute Inc, San Francisco, California (Buysse).

**Author Contributions:** Dr Burzykowski 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:** Burzykowski, Coart, Saad, Shi, Sommeijer, Douillard, Hurwitz, Maughan, Saltz, Zalcberg, Buysse.

**Acquisition, analysis, or interpretation of data:** Burzykowski, Coart, Saad, Shi, Sommeijer, Bokemeyer, Díaz-Rubio, Douillard, Falcone, Fuchs, Goldberg, Hecht, Hoff, Hurwitz, Kabbinnavar, Koopman, Maughan, Punt, Schmoll, Seymour, Tebbutt, Tournigand, Van Cutsem, de Gramont, Zalcberg, Buysse.

**Drafting of the manuscript:** Burzykowski, Saad, Sommeijer, Zalcberg.

**Critical revision of the manuscript for important intellectual content:** Burzykowski, Coart, Shi, Sommeijer, Bokemeyer, Díaz-Rubio, Douillard, Falcone, Fuchs, Goldberg, Hecht, Hoff, Hurwitz, Kabbinnavar, Koopman, Maughan, Punt, Saltz, Schmoll, Seymour, Tebbutt, Tournigand, Van Cutsem, de Gramont, Zalcberg, Buysse.

**Statistical analysis:** Burzykowski, Shi, Buysse.

**Obtained funding:** Maughan, de Gramont.

**Administrative, technical, or material support:** Coart, Saad, Bokemeyer, Douillard, Goldberg, Hecht, Hoff, Kabbinnavar, Tournigand, Van Cutsem.

**Supervision:** Burzykowski, Sommeijer, Fuchs, Goldberg, Punt, Schmoll, Zalcberg.

**Conflict of Interest Disclosures:** Dr Coart reported being an employee of IDDI from Aide et Recherche en Cancérologie Digestive (ARCAD) during the conduct of the study. Dr Saad reported receiving grants from

Fondation ARCAD during the conduct of the study and other support from IDDI outside the submitted work. Dr Bokemeyer reported receiving grants and personal fees from Merck KGa and personal fees from Sanofi Aventis, Roche, Bayer Healthcare, Bristol-Myers Squibb, AstraZeneca, Lilly/Imclone, Merck Sharp & Dohme, AOK Health Insurance, and Pfizer outside the submitted work. Dr Díaz-Rubio reported receiving personal fees from Roche, Merck Serono, Amgen, Bayer, MSD, Genomica, Servier, Merck Sharp & Dohme, and Amgen; and grants from Roche, Merck Serono, Amgen, and AstraZeneca outside the submitted work. Dr Falcone reported receiving grants, personal fees, and nonfinancial support from Amgen, Bayer, Roche, Bristol-Myers Squibb, Servier, Sanofi, Lilly, Merck, and MSD outside the submitted work. Dr Fuchs reported receiving personal fees from Agios, Bain Capital, Bayer, Celgene, Dicerna, Five Prime Therapeutics, Gilead Sciences, Eli Lilly, Entrinsic Health, KEW, Merck, Merrimack Therapeutics, Pfizer, Sanofi, Taiho, Unum Therapeutics, CytomX Therapeutics, and Genentech outside the submitted work; serving as a director for CytomX Therapeutics; and owning unexercised stock options for CytomX and Entrinsic Health. Dr Goldberg reported receiving personal fees from Merck Sharp & Dohme, KGA, Taiho, Novartis, EMD Serono, and Amgen outside the submitted work. Dr Hoff reported receiving grants from Roche during the conduct of the study. Dr Hurwitz reported other support from Genentech Roche during the conduct of the study and grants and personal fees from Genentech Roche outside the submitted work. Dr Koopman reported receiving grants from KWF, DCCG, and MLDS outside the submitted work. Dr Maughan reported receiving grants from Merck Serono and Cancer Research UK during the conduct of the study. Dr Van Cutsem reported receiving grants from Bayer, Boehringer Ingelheim, Celgene, Ipsen, Lilly, Roche, Merck Sharp & Dohme, Merck KGaA, Novartis, Roche, and Servier; and personal fees from Array, AstraZeneca, Bayer, Biocartis, Bristol-Myers Squibb, Celgene, Daiichi, Pierre-Fabre, Incyte, Ipsen, Lilly, Merck Sharp & Dohme, Merck KGaA, Novartis, Roche, Servier, and Sirtex outside the submitted work. Dr Zalberg reported receiving grants and personal fees from Pfizer, MerckSerono, Specialized Therapeutics Australia, and MSD; personal fees from Targovax, Halozyme, Gilead, Sirtex, and Bayer; and grants from Roche, Lilly, Shire, Boehringer-Ingelheim, BMS, and AstraZeneca outside the submitted work. Dr Buyse reported receiving personal fees and other support from IDDI during the conduct of the study as well as outside the submitted work. No other disclosures were reported.

**Group Information:** The Aide et Recherche en Cancerologie Digestive Colorectal Cancer Group members are as follows: René Adam, Centre Hépatobiliaire, Hôpital AP-HP Paul Brousse, Université Paris-Sud, Inserm U935, Villejuif, France; Richard Adams, Cardiff University and Velindre Cancer Centre, Cardiff, United Kingdom; Jaffer Ajani, The University of Texas MD Anderson Cancer Center, Houston; Carmen Joseph Allegra, Division of Hematology and Oncology, University of Florida Medical Center, Gainesville; Thierry Andre, Department of Medical Oncology, Hôpital Saint-Antoine, Paris, France; Dirk Arnold, CUF Infante Santo Hospital, Lisboa, Portugal; Jean-Baptiste Bachet, Department of Hepato-Gastroenterology, Groupe Hospitalier Pitié Salpêtrière, Paris, France; Al Bowen Benson, Robert H. Lurie Comprehensive Cancer Center of Northwestern University, Chicago, Illinois; Jordan Berlin, Vanderbilt University, Nashville, Tennessee; Harry Bleiberg, Department of Medicine, Digestive Oncology Unit, Medical Oncology, Institut Jules Bordet, ULB, Brussels, Belgium; György Bodoky, Department of Oncology, St László Teaching Hospital, Budapest, Hungary; Marc Buyse, International Drug Development Institute Inc, San Francisco, California; Docteur Benoist Chibaudel, Department of Oncology, Institut Franco-Britannique, Levallois-Perret, France; Eduardo Diaz-Rubio, Laboratorio de Investigación Traslacional, IdISSC, Hospital Clínico San Carlos, Centro de Investigación Biomédica en Red Cáncer, Madrid, Spain; Jean-Yves Douillard, Scientific and Medical Division, European Society for Medical Oncology, Lugano, Switzerland, and Institut de Cancérologie de l'Ouest René Gauducheau, Saint-Herblain, France; Lee Ellis, Division of Surgery, Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston; Cathy Eng, Department of Gastrointestinal Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston; Alfredo Falcone, Unit of Medical Oncology, Department of Translational Research and New Technologies in Medicine, University of Pisa, Pisa, Italy; Jan Franko, Department of Surgery, Mercy Medical Center, Des Moines, Iowa; Charles S. Fuchs, Yale University School of Medicine, Yale Cancer Center, New Haven, Connecticut, and Dana-Farber Cancer Institute, Boston, Massachusetts; Masashi Fujii, Department of Digestive Surgery, Nihon University School of Medicine, Itabashi, Tokyo, Japan; Bruce J. Giantonio, Division of Hematology and Oncology, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston; Richard M. Goldberg, West Virginia University Cancer Institute and the Mary Babb Randolph Cancer Center, Morgantown; Aimery de Gramont, Department of Oncology, Institut Hospitalier Franco-Britannique, Levallois-Perret, France; Axel Grothey, West Cancer Center, Memphis, Tennessee; Daniel Haller, Abrahamson Cancer Center, Philadelphia, Pennsylvania; Stan R. Hamilton, University of Texas MD Anderson Cancer, Houston; Petr F. Hausner, University of Maryland Greenebaum Comprehensive Cancer Center, Baltimore; J. Randolph Hecht, David Geffen School of Medicine, University of California, Los Angeles; Volker Heinemann, Department of Medical Oncology and Comprehensive Cancer Center, University Hospital Grosshadern, Ludwig Maximilian University of Munich, Munich, Germany; Alain Herrera, Alain Oncologie Consulting, Paris, France; Howard S. Hochster, Rutgers Cancer Institute of New Jersey, New Brunswick; Paulo M. Hoff, Instituto de Câncer do Estado de São Paulo, Universidade de São Paulo, São Paulo, Brazil; Derek J. Jonker, The Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Ontario, Canada; Rick Kaplan, MRC Clinical Trials Unit at UCL, University College London, London, United Kingdom; Dieter Koeberle, Department of Nuclear

Medicine and Positron Emission Tomography/Computed Tomography Center North-West Switzerland, St Claraspital, Basel, Switzerland; Scott Kopetz, Department of Gastrointestinal Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston; Roberto F. Labianca, Center of Oncology, Papa Giovanni XXIII Hospital, Bergamo, Italy; Annette K. Larsen, Cancer Biology and Therapeutics, Centre de Recherche Saint-Antoine, Sorbonne University, INSERM, Paris, France; Heinz-Joseph Lenz, Division of Medical Oncology, Norris Comprehensive Cancer Center, University of Southern California, Los Angeles; Christopher Lieu, University of Colorado Cancer Center, Aurora; Christophe Louvet, Department of Oncology, Institut Mutualiste Montsouris, Paris, France; Fotios Loupakis, Department of Oncology, Veneto Institute of Oncology IOV-IRCCS, Padua, Italy; John Marshall, Department of Pharmacy, Massachusetts General Hospital, Boston; Timothy S. Maughan, St James's Hospital and University of Leeds, Leeds, United Kingdom; Robert J. Mayer, Dana-Farber Cancer Institute, Boston, Massachusetts; Neal J. Meropol, University Hospitals Seidman Cancer Center, Cleveland, Ohio; Edith P. Mitchell, Medical Oncology, Thomas Jefferson University, Philadelphia, Pennsylvania; Michael J. O'Connell, Allegheny Center, Pittsburgh, Pennsylvania; Marc Peeters, Department of Oncology, Antwerp University Hospital/Antwerp University, Edegem, Belgium; Rainer Porschen, Klinikum Bremen-Ost Klinik für Innere Medizin, Bremen, Germany; Timothy Price, Queen Elizabeth Hospital and University of Adelaide, Adelaide, Australia; Cornelis J. A. Punt, Department of Medical Oncology, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands; Mohamed E. Salem, MD, Levine Cancer Institute, Atrium Health, Charlotte, North Carolina; Leonard Saltz, Memorial Sloan-Kettering Cancer Center, New York, New York; Richard Schilsky, American Society of Clinical Oncology, Alexandria, Virginia; Hans-Joachim Schmoll, Martin-Luther-University, Halle, Germany; Matthew T. Seymour, Cancer Research UK Clinical Centre, Leeds, United Kingdom; Einat Shacham Shmueli, Department of Oncology, Sheba Medical Center, Tel-Hashomer, Israel, and Tel-Aviv University, Tel-Aviv, Israel; Qian Shi, Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota; Alberto Sobrero, IRCCS Azienda Ospedaliera Universitaria San Martino, Genova, Italy; John Souglakos, University of Crete, Heraklion, Greece; Josep Tabernero, Vall d'Hebron University Hospital and Institute of Oncology, Centro de Investigación Biomédica en Red Cáncer, Universitat Autònoma de Barcelona, Barcelona, Spain; Julien Taieb, Department of Oncology Digestive, Hôpital Européen Georges Pompidou, Paris, France; Niall C. Tebbutt, Austin Health, Heidelberg, Victoria, Australia; Sabine Tejpar, Universitair Ziekenhuis Leuven, Leuven, Belgium; Margaret Tempero, Division of Hematology and Oncology, Department of Medicine, UCSF Medical Center, San Francisco, California; Christophe Tournigand, Department of Oncology, CHU Henri Mondor, Créteil, France; Yasushi Tsuji, Department of Clinical Oncology, Tonan Hospital, Sapporo, Hokkaido, Japan; Eric Van Cutsem, University Hospital Leuven, Leuven, Belgium; Alan P. Venook, Department of Medicine, University of California, San Francisco; Takayuki Yoshino, Department of Urology, Kyoto University Graduate School of Medicine, Kyoto, Japan; Benjamin A. Weinberg, The Ruesch Center for the Cure of Gastrointestinal Cancers, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC; Norman Wolmark, Allegheny Health Network Cancer Institute, Pittsburgh, Pennsylvania; John R. Zalcberg, School of Public Health and Preventative Medicine, Monash University, Melbourne, Australia.

**Funding/Support:** Fondation ARCAD provided data for analysis and financial support that allowed analysis and interpretation of the data and the preparation of the manuscript.

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

## REFERENCES

1. Di Leo A, Bleiberg H, Buysse M. Overall survival is not a realistic end point for clinical trials of new drugs in advanced solid tumors: a critical assessment based on recently reported phase III trials in colorectal and breast cancer. *J Clin Oncol*. 2003;21(10):2045-2047. doi:10.1200/JCO.2003.99.089
2. Burzykowski T, Buysse M, Yothers G, Sakamoto J, Sargent D. Exploring and validating surrogate endpoints in colorectal cancer. *Lifetime Data Anal*. 2008;14(1):54-64. doi:10.1007/s10985-007-9079-4
3. Shi Q, Sargent DJ. Meta-analysis for the evaluation of surrogate endpoints in cancer clinical trials. *Int J Clin Oncol*. 2009;14(2):102-111. doi:10.1007/s10147-009-0885-4
4. Buysse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics*. 2000;1(1):49-67. doi:10.1093/biostatistics/1.1.49
5. Piessevaux H, Buysse M, Schlichting M, et al. Use of early tumor shrinkage to predict long-term outcome in metastatic colorectal cancer treated with cetuximab. *J Clin Oncol*. 2013;31(30):3764-3775. doi:10.1200/JCO.2012.42.8532
6. Claret L, Gupta M, Han K, et al. Evaluation of tumor-size response metrics to predict overall survival in Western and Chinese patients with first-line metastatic colorectal cancer. *J Clin Oncol*. 2013;31(17):2110-2114. doi:10.1200/JCO.2012.45.0973

7. Heinemann V, Stintzing S, Modest DP, Giessen-Jung C, Michl M, Mansmann UR. Early tumour shrinkage (ETS) and depth of response (DpR) in the treatment of patients with metastatic colorectal cancer (mCRC). *Eur J Cancer*. 2015;51(14):1927-1936. doi:10.1016/j.ejca.2015.06.116
8. Petrelli F, Pietrantonio F, Cremolini C, et al. Early tumour shrinkage as a prognostic factor and surrogate end-point in colorectal cancer: a systematic review and pooled-analysis. *Eur J Cancer*. 2015;51(7):800-807. doi:10.1016/j.ejca.2015.02.011
9. Cremolini C, Loupakis F, Antoniotti C, et al. Early tumor shrinkage and depth of response predict long-term outcome in metastatic colorectal cancer patients treated with first-line chemotherapy plus bevacizumab: results from phase III TRIBE trial by the Gruppo Oncologico del Nord Ovest. *Ann Oncol*. 2015;26(6):1188-1194. doi:10.1093/annonc/mdv112
10. Mansmann UR, Laubender RP, Sartorius U, Giessen CA, Graser A, Heinemann V. Improved early prediction of individual prognosis for patients with mCRC: joint modeling of tumor shrinkage with volume data for PFS and OS [abstract]. *J Clin Oncol*. 2012;(suppl):30.
11. Díaz-Rubio E, Tabernero J, Gómez-España A, et al; Spanish Cooperative Group for the Treatment of Digestive Tumors Trial. Phase III study of capecitabine plus oxaliplatin compared with continuous-infusion fluorouracil plus oxaliplatin as first-line therapy in metastatic colorectal cancer: final report of the Spanish Cooperative Group for the Treatment of Digestive Tumors Trial. *J Clin Oncol*. 2007;25(27):4224-4230. doi:10.1200/JCO.2006.09.8467
12. Fuchs CS, Marshall J, Mitchell E, et al. Randomized, controlled trial of irinotecan plus infusional, bolus, or oral fluoropyrimidines in first-line treatment of metastatic colorectal cancer: results from the BICC-C Study. *J Clin Oncol*. 2007;25(30):4779-4786. doi:10.1200/JCO.2007.11.3357
13. Tournigand C, André T, Achille E, et al. FOLFIRI followed by FOLFOX6 or the reverse sequence in advanced colorectal cancer: a randomized GERCOR study. *J Clin Oncol*. 2004;22(2):229-237. doi:10.1200/JCO.2004.05.113
14. Koopman M, Antonini NF, Douma J, et al. Sequential versus combination chemotherapy with capecitabine, irinotecan, and oxaliplatin in advanced colorectal cancer (CAIRO): a phase III randomised controlled trial. *Lancet*. 2007;370(9582):135-142. doi:10.1016/S0140-6736(07)61086-1
15. Seymour MT, Maughan TS, Ledermann JA, et al; FOCUS Trial Investigators; National Cancer Research Institute Colorectal Clinical Studies Group. Different strategies of sequential and combination chemotherapy for patients with poor prognosis advanced colorectal cancer (MRC FOCUS): a randomised controlled trial. *Lancet*. 2007;370(9582):143-152. doi:10.1016/S0140-6736(07)61087-3
16. Seymour MT, Thompson LC, Wasan HS, et al; FOCUS2 Investigators; National Cancer Research Institute Colorectal Clinical Studies Group. Chemotherapy options in elderly and frail patients with metastatic colorectal cancer (MRC FOCUS2): an open-label, randomised factorial trial. *Lancet*. 2011;377(9779):1749-1759. doi:10.1016/S0140-6736(11)60399-1
17. Falcone A, Ricci S, Brunetti I, et al; Gruppo Oncologico Nord Ovest. Phase III trial of infusional fluorouracil, leucovorin, oxaliplatin, and irinotecan (FOLFOXIRI) compared with infusional fluorouracil, leucovorin, and irinotecan (FOLFIRI) as first-line treatment for metastatic colorectal cancer: the Gruppo Oncologico Nord Ovest. *J Clin Oncol*. 2007;25(13):1670-1676. doi:10.1200/JCO.2006.09.0928
18. Saltz LB, Clarke S, Díaz-Rubio E, et al. Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase III study. *J Clin Oncol*. 2008;26(12):2013-2019. doi:10.1200/JCO.2007.14.9930
19. Goldberg RM, Sargent DJ, Morton RF, et al. A randomized controlled trial of fluorouracil plus leucovorin, irinotecan, and oxaliplatin combinations in patients with previously untreated metastatic colorectal cancer. *J Clin Oncol*. 2004;22(1):23-30. doi:10.1200/JCO.2004.09.046
20. Tebbutt NC, Wilson K, GebSKI VJ, et al. Capecitabine, bevacizumab, and mitomycin in first-line treatment of metastatic colorectal cancer: results of the Australasian Gastrointestinal Trials Group randomized phase III MAX Study. *J Clin Oncol*. 2010;28(19):3191-3198. doi:10.1200/JCO.2009.27.7723
21. Hurwitz H, Fehrenbacher L, Novotny W, et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N Engl J Med*. 2004;350(23):2335-2342. doi:10.1056/NEJMoa032691
22. Kabbinar FF, Schulz J, McCleod M, et al. Addition of bevacizumab to bolus fluorouracil and leucovorin in first-line metastatic colorectal cancer: results of a randomized phase II trial. *J Clin Oncol*. 2005;23(16):3697-3705. doi:10.1200/JCO.2005.05.112
23. Hoff PM, Hochhaus A, Pestalozzi BC, et al. Cediranib plus FOLFOX/CAPOX versus placebo plus FOLFOX/CAPOX in patients with previously untreated metastatic colorectal cancer: a randomized, double-blind, phase III study (HORIZON II). *J Clin Oncol*. 2012;30(29):3596-3603. doi:10.1200/JCO.2012.42.6031

24. Schmol HJ, Cunningham D, Sobrero A, et al. Cediranib with mFOLFOX6 versus bevacizumab with mFOLFOX6 as first-line treatment for patients with advanced colorectal cancer: a double-blind, randomized phase III study (HORIZON III). *J Clin Oncol*. 2012;30(29):3588-3595. doi:10.1200/JCO.2012.42.5355
25. Tol J, Koopman M, Cats A, et al. Chemotherapy, bevacizumab, and cetuximab in metastatic colorectal cancer. *N Engl J Med*. 2009;360(6):563-572. doi:10.1056/NEJMoa0808268
26. Maughan TS, Adams RA, Smith CG, et al; MRC COIN Trial Investigators. Addition of cetuximab to oxaliplatin-based first-line combination chemotherapy for treatment of advanced colorectal cancer: results of the randomised phase 3 MRC COIN trial. *Lancet*. 2011;377(9783):2103-2114. doi:10.1016/S0140-6736(11)60613-2
27. Van Cutsem E, Köhne CH, Hitre E, et al. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med*. 2009;360(14):1408-1417. doi:10.1056/NEJMoa0805019
28. Bokemeyer C, Bondarenko I, Makhson A, et al. Fluorouracil, leucovorin, and oxaliplatin with and without cetuximab in the first-line treatment of metastatic colorectal cancer. *J Clin Oncol*. 2009;27(5):663-671. doi:10.1200/JCO.2008.20.8397
29. Hecht JR, Mitchell E, Chidiac T, et al. A randomized phase IIIB trial of chemotherapy, bevacizumab, and panitumumab compared with chemotherapy and bevacizumab alone for metastatic colorectal cancer. *J Clin Oncol*. 2009;27(5):672-680. doi:10.1200/JCO.2008.19.8135
30. Douillard JY, Siena S, Cassidy J, et al. Randomized, phase III trial of panitumumab with infusional fluorouracil, leucovorin, and oxaliplatin (FOLFOX4) versus FOLFOX4 alone as first-line treatment in patients with previously untreated metastatic colorectal cancer: the PRIME study. *J Clin Oncol*. 2010;28(31):4697-4705. doi:10.1200/JCO.2009.27.4860
31. Xie W, Halabi S, Tierney JF, et al. A systematic review and recommendation for reporting of surrogate endpoint evaluation using meta-analyses. *J Natl Cancer Inst Cancer Spectr*. 2019;3(1):pkz002. doi:10.1093/jncics/pkz002
32. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45(2):228-247. doi:10.1016/j.ejca.2008.10.026
33. Renard D, Geys H, Molenberghs G, et al. Validation of a longitudinally measured surrogate marker for a time-to-event endpoint. *J Appl Stat*. 2003;30:235-247. doi:10.1080/0266476022000023776
34. Henderson R, Diggle P, Dobson A. Joint modelling of longitudinal measurements and event time data. *Biostatistics*. 2000;1(4):465-480. doi:10.1093/biostatistics/1.4.465
35. Lassere MN. The Biomarker-Surrogacy Evaluation Schema: a review of the biomarker-surrogate literature and a proposal for a criterion-based, quantitative, multidimensional hierarchical levels of evidence schema for evaluating the status of biomarkers as surrogate endpoints. *Stat Methods Med Res*. 2008;17(3):303-340. doi:10.1177/0962280207082719
36. Buyse M, Burzykowski T, Carroll K, et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. *J Clin Oncol*. 2007;25(33):5218-5224. doi:10.1200/JCO.2007.11.8836
37. Mauer M, Collette L, Bogaerts J; European Organisation for Research and Treatment of Cancer Statistics Department. Adaptive designs at European Organisation for Research and Treatment of Cancer (EORTC) with a focus on adaptive sample size re-estimation based on interim-effect size. *Eur J Cancer*. 2012;48(9):1386-1391. doi:10.1016/j.ejca.2011.12.024
38. Biomarkers Definitions Working G; Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001;69(3):89-95. doi:10.1067/mcp.2001.113989
39. Burzykowski T, Molenberghs G, Buyse M, Geys H, Renard D. Validation of surrogate endpoints in multiple randomized clinical trials with failure-time endpoints. *J Royal Stat Soc C*. 2001;50(4):405-422. doi:10.1111/1467-9876.00244
40. Buyse M, Thirion P, Carlson RW, Burzykowski T, Molenberghs G, Piedbois P; Meta-analysis Group in Cancer. Relation between tumour response to first-line chemotherapy and survival in advanced colorectal cancer: a meta-analysis. *Lancet*. 2000;356(9227):373-378. doi:10.1016/S0140-6736(00)02528-9
41. Burzykowski T, Buyse M, Piccart-Gebhart MJ, et al. Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. *J Clin Oncol*. 2008;26(12):1987-1992. doi:10.1200/JCO.2007.10.8407

#### SUPPLEMENT.

**eFigure 1.** Kaplan-Meier Overall Survival Curves for Each Contrast

**eFigure 2.** Model-Based Estimated Longitudinal Profiles for Each Contrast

**eFigure 3.** Individual-Level  $R_{ind}$  (Referred to as  $R(t)$  in the manuscript)