

Change in Tumor Size by RECIST Correlates Linearly With Overall Survival in Phase I Oncology Studies

Rajul K. Jain, J. Jack Lee, Chaan Ng, David Hong, Jing Gong, Aung Naing, Jennifer Wheler, and Razelle Kurzrock

All authors: The University of Texas MD Anderson Cancer Center, Houston, TX.

Submitted April 11, 2011; accepted April 12, 2012; published online ahead of print at www.jco.org on June 11, 2012.

Supported in part by Grants No. RR024148 from the National Center for Research Resources (R.K.) and CA16672 from the National Cancer Institute (J.J.L.).

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

R.K.J. completed work on this article at The University of Texas MD Anderson Cancer Center.

Corresponding author: Rajul K. Jain, MD, Amgen, 1 Amgen Center Dr, MS 38-2-B, Thousand Oaks, CA 91320; e-mail: rajulj@amgen.com.

© 2012 by American Society of Clinical Oncology

0732-183X/12/3021-2684/\$20.00

DOI: 10.1200/JCO.2011.36.4752

ABSTRACT

Purpose

RECIST is used to quantify tumor changes during exposure to anticancer agents. Responses are categorized as complete response (CR), partial response (PR), stable disease (SD), or progressive disease (PD). Clinical trials dictate a patient's management options based on the category into which his or her response falls. However, the association between response and survival is not well studied in the early trial setting.

Patients and Methods

To study the correlation between response as quantified by RECIST and overall survival (OS, the gold-standard survival outcome), we analyzed 570 participants of 24 phase I trials conducted between October 2004 and May 2009, of whom 468 had quantifiable changes in tumor size. Analyses of Kaplan-Meier estimates of OS by response and null Martingale residuals of Cox models were the primary outcome measures. All analyses are landmark analyses.

Results

Kaplan-Meier analyses revealed strong associations between change in tumor size by RECIST and survival ($P = 4.5 \times 10^{-6}$ to $< 1 \times 10^{-8}$). The relationship was found to be near-linear ($R^2 = 0.75$ to 0.92) and confirmed by the residual analyses. No clear inflection points were found to exist in the relationship between tumor size changes and survival.

Conclusion

RECIST quantification of response correlates with survival, validating RECIST's use in phase I trials. However, the lack of apparent boundary values in the relationship between change in tumor size and OS demonstrates the arbitrary nature of the CR/PR/SD/PD categories and questions emphasis placed on this categorization scheme. Describing tumor responses as a continuous variable may be more informative than reporting categoric responses when evaluating novel anticancer therapies.

J Clin Oncol 30:2684-2690. © 2012 by American Society of Clinical Oncology

INTRODUCTION

For the last decade, the international cancer community has employed the RECIST to assess the response exhibited by a patient's tumor on exposure to both marketed and experimental antitumor therapies.¹⁻³ The criteria provide guidance on two issues: how to quantify the change in a patient's tumor burden, and how to categorize the quantified response. Specifically, the dimensions of select lesions, referred to as target lesions, are used to calculate the change in tumor burden between images from different time points. The calculated response is then categorized as complete response (CR), partial response (PR), stable disease (SD), or progressive disease (PD). CR is complete disappearance of tumor (-100%), PR is a change between -100% and -30% , SD is a change between -30% and

$+20\%$, and PD is an increase of 20% or greater. Of note, when scans show the appearance of new lesions the change in tumor burden cannot be quantified numerically, but the response is categorized as PD.

Although phase I trials are principally designed to study the dosing and safety profiles of investigational products and not drug efficacy,⁴⁻⁷ serial radiographic imaging is an important component in the evaluation of patients participating in most phase I studies of solid tumors.⁸ This is because though phase I trials in other disease types typically enroll healthy volunteers, phase I oncology studies are usually conducted in participants who are cancer patients,^{9,10} and serial imaging is an important part of their clinical management. Most trials allow participants whose tumors are not significantly advancing to continue on the trial, whereas those whose

Table 1. Patient Characteristics

Characteristic	Patients With Quantifiable Change by RECIST (n = 468)		Patients With Best Response of New Lesion (ie, unquantifiable; n = 102)	
	No. of Patients	%	No. of Patients	%
Age, years				
Median	57		56	
Range	2-88		33-82	
Sex				
Male	265	56.6	54	52.9
Female	203	43.4	48	47.1
Race				
White	367	78.4	79	77.5
Black	40	8.5	8	7.8
Hispanic	38	8.1	13	12.7
Asian	21	4.5	2	2.0
Other	2	0.5	0	0.0
Performance status				
0	155	33.1	33	32.5
1	295	63.0	61	59.8
≥ 2	18	3.8	8	7.8
Tumor type				
Breast, n = 47	32	6.8	15	14.7
GI, n = 177	137	29.3	40	39.2
Genitourinary, n = 51	43	9.2	8	7.8
Gynecologic, n = 17	13	2.8	4	3.9
Head and neck, n = 41	35	7.5	6	5.9
Lymphoma/myeloma, n = 19	8	1.7	1	1.0
Melanoma, n = 52	40	8.5	12	11.7
Sarcoma, n = 37	36	7.7	1	1.0
Thoracic, n = 45	38	8.1	7	6.7
All other, n = 157	86	18.4	8	7.8
Total No. of prior treatments				
Mean	5.8		6.2	
SD	2.7		3.1	
Systemic treatments				
Mean	3.6		4.1	
SD	2.3		2.6	
Radiation treatments				
Mean	0.6		0.7	
SD	0.8		1.0	
Surgical treatments				
Mean	1.4		1.2	
SD	1.3		0.9	
Other treatments*				
Mean	0.2		0.2	
SD	0.6		0.8	
Time from cancer diagnosis to C1D1 of study treatment, years†				
Median	3.7		2.8	
Range	0.1-36.7		0.1-26.8	
Time from end of last treatment to C1D1 of trial, months‡				
Median	2.0		1.7	
Range	0.0-269.9		0.0-20.2	
Time from C1D1 of study treatment to best RECIST response, days				
Median	59		57	
Range	25-854		25-79	

Abbreviation: SD, standard deviation.

*Concurrent chemotherapy with radiation therapy is categorized as other treatments.

†C1D1 is cycle 1, day 1 of phase I treatment.

‡Two patients (one in each group) had no therapy before phase I treatment.

tumors show any substantial growth are required to discontinue the investigational treatment. This is because of the commonsense relationships between tumor shrinkage and possible benefit, and between tumor growth and risk, including potential impacts on survival.^{11,12} However, correlations between response and survival are not universal in oncology.¹³⁻¹⁶ Additionally, they have been most studied in patients receiving initial lines of therapy.¹⁷⁻¹⁹ It is unclear if the correlations are maintained in the phase I population, which is composed of patients who have usually exhausted standard therapies and whose tumors may fundamentally differ from tumors in patients with earlier stages of disease. These differences manifest clinically with more aggressive tumor behavior and biologically with increased genetic and cellular alterations.

Because the relationship between survival and tumor response as described by RECIST is not well studied in the phase I setting, despite the widespread use of RECIST in phase I oncology trials,²⁰⁻²² we studied the relationship between overall survival (OS) and change in tumor burden as quantified and categorized by RECIST in phase I trial participants. We also studied the relationship between progression-free survival (PFS, another common end point in clinical trials¹⁸) and change in tumor burden in the same group of participants.

PATIENTS AND METHODS

Data Sources

All participants enrolled onto any phase I trial of systemic anticancer therapy in the Department of Investigational Cancer Therapeutics (phase I department) at The University of Texas MD Anderson Cancer Center between October 10, 2004 and May 1, 2008 who underwent ≥ 1 reimaging evaluation after initiation of study treatment were eligible for inclusion. A total of 581 consecutive participants in 24 trials met the criteria. Among them, 570 patients (98.1%) are included (Table 1; Appendix Table A1 [online-only]); 11 patients (1.9%) were excluded owing to inability to evaluate changes in pre- and post-treatment imaging. Of the 570 patients, 468 had changes quantifiable by RECIST 1.0. Patient outcomes through May 1, 2009 are included in the analysis.

This study was conducted in accordance with our institutional review board's guidelines. Data were obtained from the electronic patient record system and from clinical data management tools used in the department.

Imaging

All patients had baseline imaging ≤ 4 weeks before the initiation of investigational drug (median, 12 days; Q1-3, 6 to 21 days). Repeat scans were performed after every one to two cycles of treatment (4 to 6 weeks between scans) depending on individual protocol guidelines. All measurements were taken from computed tomography or magnetic resonance imaging scans interpreted by at least one MD Anderson staff radiologist and were evaluated using RECIST 1.0.¹ Each patient's best response was defined as the smallest ratio in [tumor burden on any postbaseline imaging] to [tumor burden at baseline]; tumor burden is calculated using RECIST 1.0.

Statistical Analysis

PFS and OS curves were calculated using the Kaplan-Meier method.²³ PFS is defined as the time from the first dose of study drug to either disease progression per RECIST 1.0 or death. If a patient was diagnosed with progressive disease based on clinical progression without a RECIST-specified assessment (radiographic or other) at the time of clinical progression, then the last date that a RECIST-specified assessment showed no evidence of progressive disease is taken as a censored end date for PFS calculations. OS is defined as the time from the first dose of study drug to time of death. Log-rank test was applied to compare PFS and OS between groups. The median PFS and OS for each group of patients in each Kaplan-Meier analysis was plotted against the mean change in tumor size for the group, and best fit lines were calculated

using least-squares fit (Figs 1B, 1D, 1F, and 2B and Appendix Figs A1B, A1D, A1F, and A2B). To assess the RECIST response categorization boundary values, Cox models with no covariates were fit to each set of survival data and best response was plotted versus the null Martingale residual.²⁴ Similar to the usual residual plot for checking the model fit for a linear regression analysis, the Martingale residual plot is useful for evaluating the model fit for time-to-event data. The resultant scatterplots are evaluated with locally weighted scatterplot smoothing (loess curves in blue, Figs 2A to 2C and Fig 3C). This local-fitting methodology provides an exploratory graphic tool, to provide insight into the data distribution for model checking.²⁵

Because the Kaplan-Meier and null Martingale residual analyses compare time-to-event distributions among patient subsets defined by a treatment outcome variable, all Kaplan-Meier and residual analyses reported in this article are performed exclusively with the landmark method.^{26,27} Accordingly, the tumor size changes used for both the Kaplan-Meier and null Martingale analyses are the best change observed by the specified landmark time point (compared with baseline or pretreatment imaging), and survival is calculated with the landmark time point taken as time 0. For OS analyses, the landmark time points employed were the median PFS (58 days or 1.9 months), the median OS (250 days or 8.2 months), and a third intermediate time point (4 months). For PFS analyses, the landmark time point employed was the median PFS. Any patients who had not had their first restaging imaging evaluations by

the landmark time points were excluded from the Kaplan-Meier and null Martingale residual analyses (1.9-month landmark, n = 19; 4-month landmark, n = 0; 8.2-month landmark, n = 0).

P values are calculated from Cox proportional hazards models. Two-sided P values ≤ .05 are considered significant. All statistical analyses were performed using R.²⁸ Data were plotted using R or Microsoft Excel (Microsoft, Redmond, WA).

RESULTS

Patient and Trial Characteristics

Patient characteristics are listed in Table 1. Patients' median age was 57 years and 56% of patients were male. The majority of patients were white (78.2%) and had an Eastern Cooperative Oncology Group performance status of 0 to 1 (95.8%). Not surprisingly for an aggregate of patients from 24 phase I studies, a variety of tumor types are represented, with GI malignancies representing the single largest group owing to disease prevalence and referral patterns. Patients had on average almost six prior therapies, including an average of 3.7

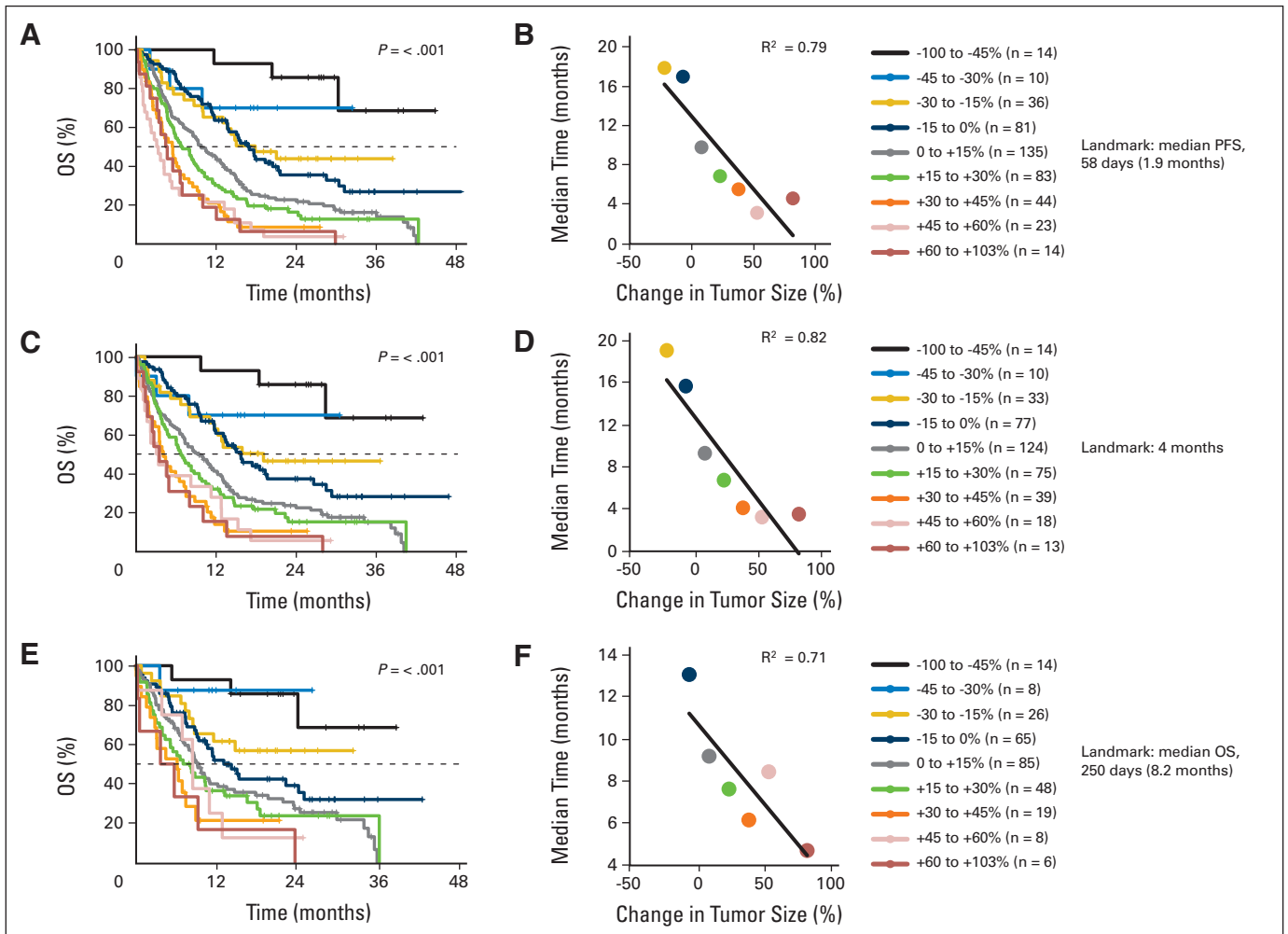


Fig 1. (A), (C), and (E) Kaplan-Meier overall survival (OS) based on tumor response using landmark analyses. Patient cohorts are separated by 15% changes in best tumor response (except at the two extremes where cohorts are enlarged due to small patient numbers). (B), (D), and (F) For cohorts that reached a median OS, each cohort's median is plotted against its mean change in tumor burden. Linearity is assessed by using least-squares fit and calculating the correlation coefficient R^2 . Landmark in A-B, 1.9 months; C-D, 4 months; E-F, 8.2 months. PFS, progression-free survival.

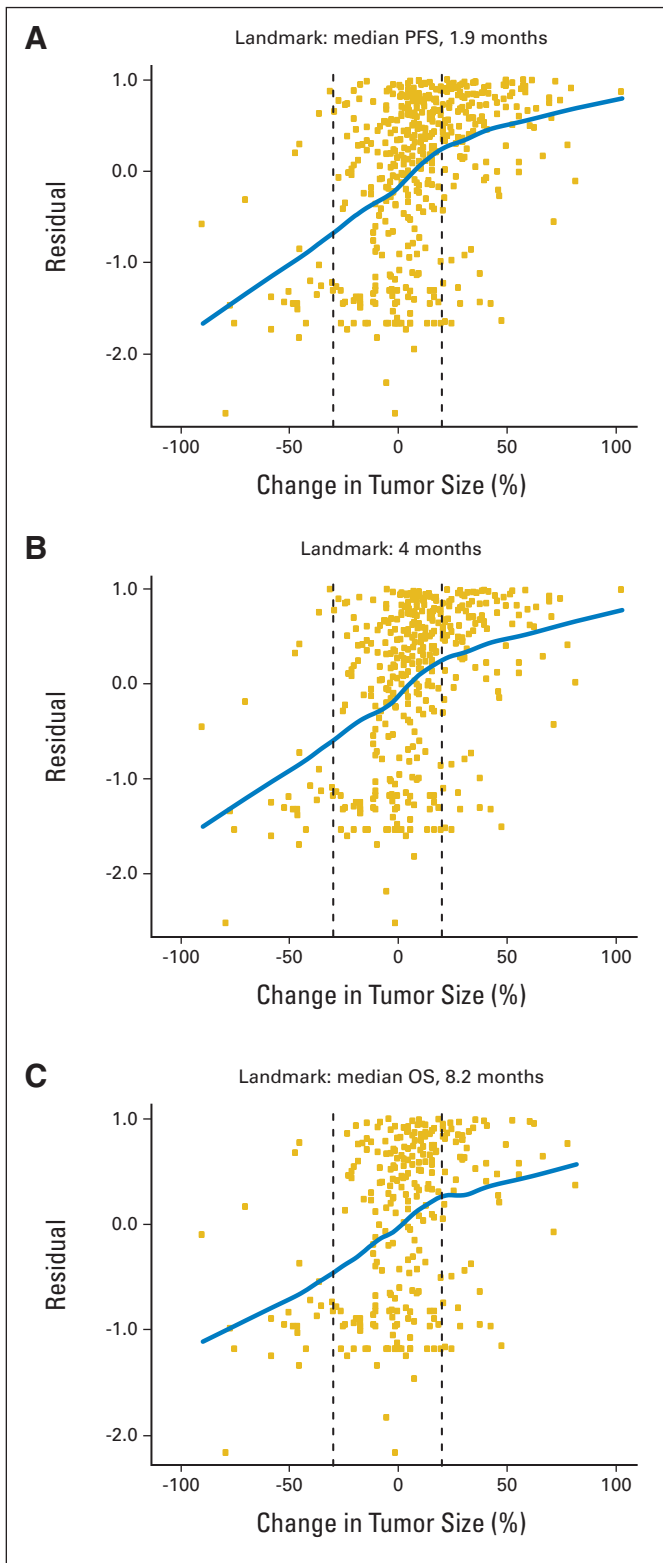


Fig 2. Null Martingale residual analyses of overall survival (OS) using indicated landmark time points. Squares represent individual patients. Blue line (loess line) represents best fit by local regression as described in Patients and Methods. PFS, progression-free survival.

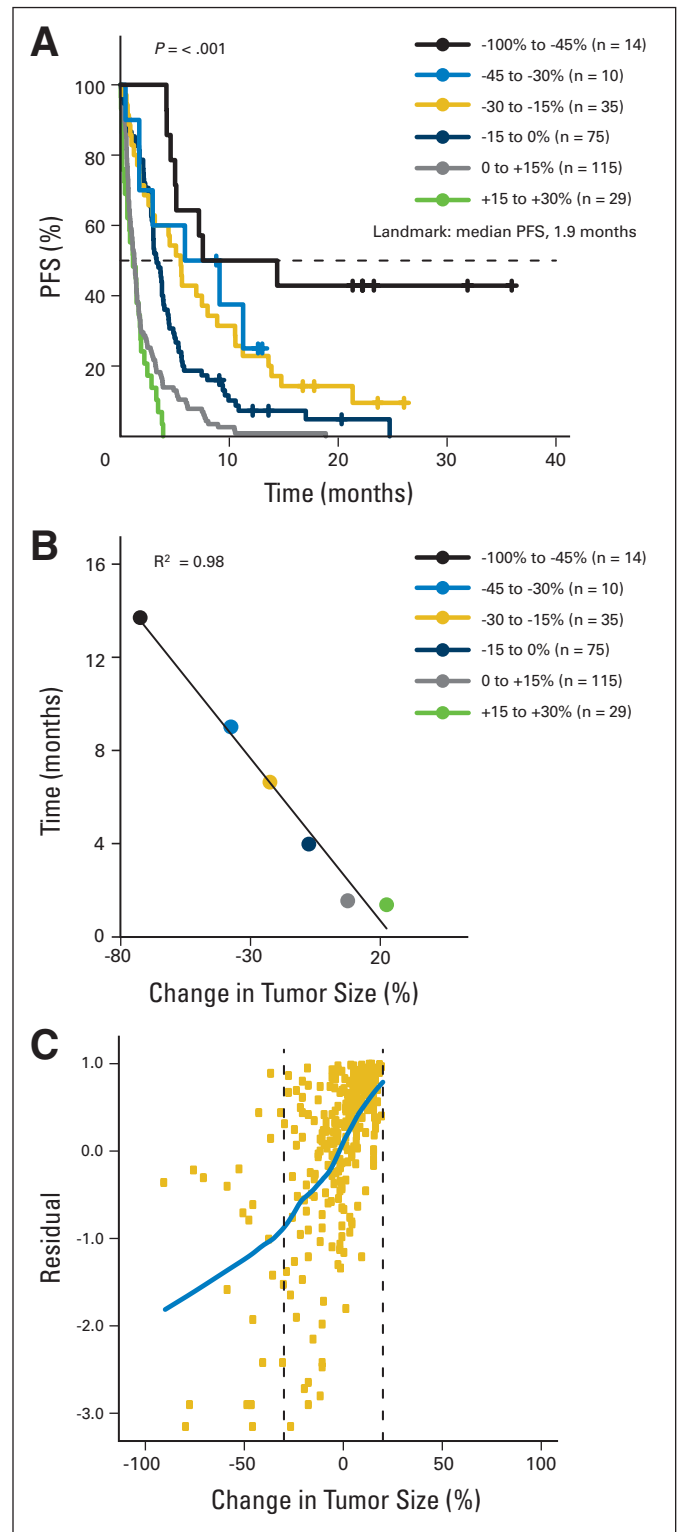


Fig 3. (A) Kaplan-Meier analysis of progression-free survival (PFS) based on tumor response using landmark analysis. Cohorts are separated by 15% changes in best tumor response (except at the extreme representing the greatest shrinkage, where cohort is enlarged because of small patient numbers). (B) Each cohort's median PFS is plotted against its mean change in tumor burden. Linearity is assessed by least-squares fit and calculating the correlation coefficient R^2 . (C) Null Martingale residual analysis. Squares represent individual patients. Blue line (loess line) represents best fit by local regression as described in Patients and Methods. In (A) and (B), the green line is labeled +15% to 30% to coincide with Figure 1 (although no patients with change > 20% are included owing to all having progressive disease by the 1.9-month landmark).

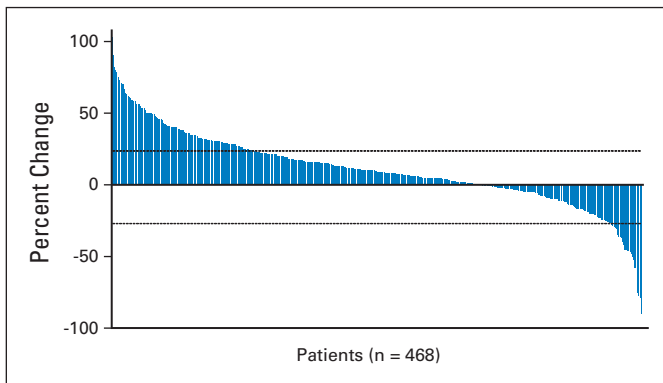


Fig 4. Waterfall plot of best response by RECIST. Four hundred sixty-eight patients had quantifiable changes that are illustrated in the figure. The remaining 102 patients had new lesions at first restaging and are therefore unquantifiable and not shown in the figure.

systemic treatments. Characteristics of patients who had tumor responses quantifiable by RECIST ($n = 468$) were similar to those with unquantifiable responses because of the appearance of new lesions on first restaging ($n = 102$; Table 1).

Most of the trials included in this analysis were of single agents (18 trials); five combined two agents, one combined three, and the agents span a broad range of antitumor mechanisms (Appendix Table A1). One trial (14 participants, 2.5%) was a study of a classically cytotoxic agent, two trials (42 participants, 7.4%) combined cytotoxic with biologic or targeted agents, and 21 trials (514 participants, 90.2%) studied solely biologic and targeted agents.

Waterfall Plot

Changes in tumor burden were quantifiable by RECIST for 468 patients included in this study (Fig 4). The other 102 patients had new lesions on first restaging, which are unquantifiable by RECIST.^{1,2} The 468 patients with measurable changes had a range of best responses from a -90% decrease in tumor to a $+103\%$ increase, with 141 patients (30.1%) showing at least some decrease by RECIST.

Association Between Response and OS

The association between response as calculated by RECIST and OS was studied using Kaplan-Meier analyses for patients with quantifiable changes in tumor burden on reimaging ($n = 468$; Fig 1 and Appendix Fig A1). Data for the patients with nonquantifiable lesions ($n = 102$) were also studied with Kaplan-Meier analyses and results are presented in Overall Survival Outcomes for Patients With New Lesions (Fig 5). All Kaplan-Meier estimates were calculated exclusively using the landmark method (see Patients and Methods). Participants were grouped by two schemes. In one, patients were divided into groups separated by 15% increments in tumor size change (Fig 1). In the second, groups were divided so that there were approximately equal numbers of patients per group (Appendix Fig A1). Regardless of how patients were grouped, the trend for increased survival with better tumor response is clear, as evidenced by the observed separation in the Kaplan-Meier curves with P values for the log-rank test ranging from $P \leq 4.5 \times 10^{-6}$ for the 8.2-month landmark analyses to $P < 1 \times 10^{-8}$ for the 4-month and 1.9-month landmark analyses.

To further evaluate the trends observed in the Kaplan-Meier graphs, we plotted the median OS versus the average change in tumor

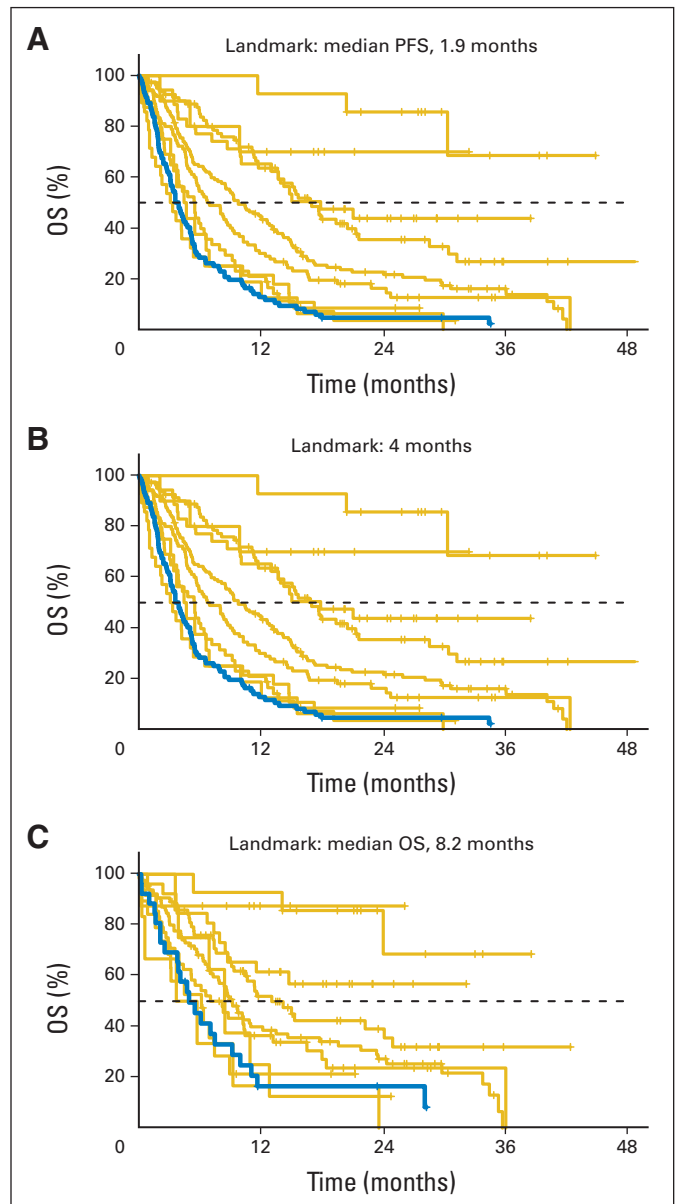


Fig 5. Kaplan-Meier analysis of overall survival (OS) for patients with new lesions on first restaging using the landmark analysis. Gold curves represent patients with measurable lesions and are identical to the curves in Figs 1A, 1C, and 1E. Blue curves represent survival rates for patients with best response of new lesion(s). For patients with new lesions, (A) OS, 3.9; $n = 93$; (B) OS, 3.2 months; $n = 64$; (C) OS, 4.9 months; $n = 26$. PFS, progression-free survival.

size for each group of patients in each Kaplan-Meier analysis (Figs 1B, 1D, and 1F and Appendix Figs A1B, A1D, and A1F; see Patients and Methods). A linear relationship is found for all landmark dates in both patient grouping schemes, with the correlation coefficient R^2 from least-squares fits ranging from 0.71 to 0.93, indicating a strong association between change in tumor size and median survival.

Evaluation of Boundaries in the Response-OS Relationship

To further evaluate the linearity observed in the Kaplan-Meier analyses in Association Between Response and OS as well as to test

for boundary values in the relationship between best tumor response and OS, we evaluated Cox models of survival compared with best tumor response (see Patients and Methods). The relationship between best response and OS seems to be generally linear (blue lines, Fig 2), consistent with the Kaplan-Meier results described in Association Between Response and OS. Regarding the evaluation of boundary conditions, Figure 2 illustrates a slight change in the relationship between best change in tumor size and OS in the vicinity of 20% tumor growth (change in curvature of blue lines, Fig 2). Specifically, the slopes of the loess lines in the portions of the plots representing less than 20% tumor growth are greater than the slopes of the lines in the portions representing more than 20% tumor growth. This may suggest that changes in tumor size correspond to greater changes in OS for patients with CR, PR, or SD as a best tumor response than for patients with PD as a best response. No significant changes to the slope of the line occur at -30% change in tumor size, the RECIST boundary between PR and SD. Overall, across the entire range from -100% to $+100\%$ changes in tumor size, the trend is for linearly increasing survival with more favorable tumor responses.

Association Between Response and PFS

Although OS is generally regarded as the gold-standard survival outcome in oncology studies, PFS is an end point commonly used in clinical studies and can serve as a regulatory end point, for example in accelerated approval in the United States²⁹ or conditional approval in the European Union.³⁰ For this reason, we conducted the same analyses performed to study the relationship between response and OS to study the relationship between response and PFS. Kaplan-Meier estimates of PFS using median PFS (1.9 months) as the landmark time point showed a strong correlation between response and PFS, with *P* values for the log-rank test less than 1×10^{-8} when cohorts are separated by an equal spacing in tumor size change (Fig 3A) and when separated by equal numbers of patients per group (Appendix Fig A2A). These relationships were also found to be highly linear ($R^2 = 0.97$ to 0.98 ; Fig 3B; Appendix Fig A2B). Finally, null Martingale residual analysis showed the relationship between PFS and change in tumor size to be nearly linear from approximately -100% to $+20\%$ changes in tumor size (Fig 3C). There are no patients in these PFS analyses who had a best response of more than 20% because all patients who had a best change in tumor size of more than 20% by definition reached a PFS end point, and because all had restaging scans before the 1.9-month landmark (all patients were required to have first restaging by week 4 to 6 depending on individual study guidelines, see Patients and Methods).

Overall Survival Outcomes for Patients With New Lesions

As noted, the analyses presented so far have excluded patients found to have new lesions on first reimaging evaluation because RECIST cannot quantify change in tumor size when new lesions appear.^{1,2} However, to obtain a sense of how well RECIST criteria categorize these patients' responses (new lesions are categorized as PD), we estimated their overall survival with Kaplan-Meier analyses using the same landmark time points as in the analyses of the other patients (1.9-, 4.0-, and 8.2-months). Long-term follow-up data for these patients show that they do as poorly in terms of overall survival as patients who are found to have large quantifiable increases in tumor burden as a best response (Fig 5). For example, using a 1.9-month landmark, the median OS for patients with new lesions on first restaging is 3.9 months compared with a median OS of 4.6 months for increases in tumor burden of 60% to 103%. And using a 8.2-month landmark, the median OS for patients with new lesions is 4.9 months compared with 4.7 months for patients with increases of 60% to 103%.

DISCUSSION

Considering the importance radiologic evaluations have in the management of oncology clinical trial participants,^{6,11,12} as well as the role imaging data plays in decisions related to advancing drugs into further stages of clinical development^{31,32} and in the granting of marketing indications by regulatory bodies,^{18,30,33} we studied the correlation between survival and response in phase I oncology subjects. Although other investigators have evaluated the relationship between response classification (CR, PR, SD, or PD) and survival outcomes,^{17,34-36} to our knowledge this is the first evaluation of the association between survival and tumor response as a continuous variable and the first time the association has been reported for a phase I population. Our findings suggest that RECIST quantification of changes in tumor burden correlate well and in a linear fashion with OS and PFS in the phase I population studied. Our results also suggest that the CR, PR, SD, and PD response categorizations are more akin to arbitrarily chosen points on a continuum than inflection points in a step function. For this reason, degree of tumor response shown as a continuous variable (for example, as in waterfall plots) may be more informative than aggregate CR, PR, SD, and PD rates when evaluating novel anticancer therapies.

One criticism that could be made of these analyses that correlate quantified response with overall survival is that patients with more aggressive disease will have more rapidly enlarging tumors and will therefore have larger changes in tumor size as a best response than patients with less aggressive disease because clinical trial protocols specify the same imaging frequency for all patients (every 4 to 6 weeks, see Patients and Methods). Continuing with this logic, those with more aggressive disease should also die sooner and therefore the linearity we demonstrate between change in tumor size and survival could be a consequence of imaging frequency. However, this does not seem to be the case. Patients with a best response of PD that was quantifiable ($n = 184$; 39.3%) had a median best response of $+30\%$ at a median of 59 days. Patients with a best response of SD ($n = 260$; 55.5%) had a median best response of $+2\%$ at a median of 58 days, only 1 day sooner. The correlation coefficient between best quantified response and time to that response is -0.339 for all 468 subjects with quantifiable tumor changes. Therefore, overall survival seems to be associated with best tumor response, but best response does not seem to be associated with imaging frequency.

Regarding data analyses, in this study survival data are interpreted using landmark methods. As has been previously described, comparing responders to nonresponders without employing such methodologies biases the data in favor of responders and can lead to the misinterpretation that a study treatment is effective because response may be a surrogate of favorable prognostic factors and not of drug efficacy.^{26,27}

Potential limitations of this study include its retrospective nature and that it includes trials of multiple agents with differing mechanisms of action as well as multiple tumor types. In reference to the latter, 9.8% of patients in this analysis received at least one cytotoxic agent as part of their treatment regimen, and 31% had GI malignancies (the most common tumor type in this study). If either the 9.8% of patients who received a cytotoxic agent or the 31% with GI malignancies are excluded from the study, the resulting Kaplan-Meier and Martingale residual analyses yield identical trends to those found on analysis of all patients (data not shown), making it unlikely this study's results are driven by any particular agent, study, or patient subset.

In conclusion, in this age of targeted therapy it has become increasingly clear that matching patients with therapies based on underlying molecular profiles may yield better outcomes. Many articles describing such studies use waterfall plots. The data presented in our article suggest that use of these plots is justified in that the correlation of tumor regression or progression on the continuous scale with survival outcome is linear, rather than polytomous, with artificial cutoffs to define PR, SD, and PD. Further study of RECIST is warranted to refine the use of this tool that is widely implemented in drug development and in the everyday management of patients.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) and/or an author's immediate family member(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships marked

with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Employment or Leadership Position: Rajul K. Jain, Amgen (C) **Consultant or Advisory Role:** None **Stock Ownership:** None **Honoraria:** None **Research Funding:** None **Expert Testimony:** None **Other Remuneration:** None

AUTHOR CONTRIBUTIONS

Conception and design: Rajul K. Jain, J. Jack Lee, Razelle Kurzrock

Financial support: J. Jack Lee, Razelle Kurzrock

Administrative support: Razelle Kurzrock

Provision of study materials or patients: Razelle Kurzrock

Collection and assembly of data: Rajul K. Jain, David Hong, Jing Gong, Aung Naing, Jennifer Wheler

Data analysis and interpretation: Rajul K. Jain, J. Jack Lee, Chuan Ng, Razelle Kurzrock

Manuscript writing: All authors

Final approval of manuscript: All authors

REFERENCES

1. Therasse P, Arbuck SG, Eisenhauer EA, et al: New guidelines to evaluate the response to treatment in solid tumors: European Organisation for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 92:205-216, 2000
2. Eisenhauer EA, Therasse P, Bogaerts J, et al: New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur J Cancer* 45:228-247, 2009
3. Shanbhogue AK, Karnad AB, Prasad SR: Tumor response evaluation in oncology: Current update. *J Comput Assist Tomogr* 34:479-484, 2010
4. Horstmann E, McCabe MS, Grochow L, et al: Risks and benefits of phase 1 oncology trials, 1991 through 2002. *N Engl J Med* 352:895-904, 2005
5. Markman M: The needs of science vs the needs of patients: Ethical concerns in cancer clinical trials. *Cleve Clin J Med* 70:1008-1009, 2003; additional discussion: 1013-1014; 1016
6. Eisenhauer EA, O'Dwyer PJ, Christian M, et al: Phase I clinical trial design in cancer drug development. *J Clin Oncol* 18:684-692, 2000
7. Lipsett MB: On the nature and ethics of phase I clinical trials of cancer chemotherapies. *JAMA* 248:941-942, 1982
8. LoRusso PM, Boerner SA, Seymour L: An overview of the optimal planning, design, and conduct of phase I studies of new therapeutics. *Clin Cancer Res* 16:1710-1718, 2010
9. Daugherty CK: Ethical issues in the development of new agents. *Invest New Drugs* 17:145-153, 1999
10. Pasqualetti G, Gori G, Blandizzi C, et al: Healthy volunteers and early phases of clinical experimentation. *Eur J Clin Pharmacol* 66:647-653, 2010
11. Kelly K, Halabi S: *Oncology Clinical Trials: Successful Design, Conduct, and Analysis*. New York, NY, Demos Medical Publishing, 2010
12. Eisenhauer E, Twelves C, Buyse M: *Phase I Cancer Trials: A Practical Guide*. New York, NY, Oxford University Press, 2006
13. Henderson IC, Hayes DF, Gelman R: Dose-response in the treatment of breast cancer: A critical review. *J Clin Oncol* 6:1501-1515, 1988
14. Wilkerson J, Fojo T: Progression-free survival is simply a measure of a drug's effect while administered and is not a surrogate for overall survival. *Cancer J* 15:379-385, 2009
15. Benjamin RS, Choi H, Macapinlac HA, et al: We should desist using RECIST, at least in GIST. *J Clin Oncol* 25:1760-1764, 2007
16. Oye RK, Shapiro MF: Reporting results from chemotherapy trials. Does response make a difference in patient survival? *JAMA* 252:2722-2725, 1984
17. Buyse M, Thirion P, Carlson RW, et al: Relation between tumour response to first-line chemotherapy and survival in advanced colorectal cancer: A meta-analysis—Meta-Analysis Group in Cancer. *Lancet* 356:373-378, 2000
18. Pazdur R: Response rates, survival, and chemotherapy trials. *J Natl Cancer Inst* 92:1552-1553, 2000
19. Michaelis LC, Ratain MJ: Measuring response in a post-RECIST world: From black and white to shades of grey. *Nat Rev Cancer* 6:409-414, 2006
20. O'Shaughnessy J, Osborne C, Pippen JE, et al: Iniparib plus chemotherapy in metastatic triple-negative breast cancer. *N Engl J Med* 364:205-214, 2011
21. Scher HI, Beer TM, Higano CS, et al: Antitumor activity of MDV3100 in castration-resistant prostate cancer: A phase 1-2 study. *Lancet* 375:1437-1446, 2010
22. Von Hoff DD, LoRusso PM, Rudin CM, et al: Inhibition of the hedgehog pathway in advanced basal-cell carcinoma. *N Engl J Med* 361:1164-1172, 2009
23. Jiang H, Fine JP: Survival analysis. *Methods Mol Biol* 404:303-318, 2007
24. Grambsch PM, Therneau TM, Fleming TR: Diagnostic plots to reveal functional form for covariates in multiplicative intensity models. *Biometrics* 51:1469-1482, 1995
25. Cleveland W, Devlin SJ: Locally weighted regression: An approach to regression analysis by local fitting. *J Am Stat Assoc* 83:596-610, 1988
26. Anderson JR, Cain KC, Gelber RD: Analysis of survival by tumor response. *J Clin Oncol* 1:710-719, 1983
27. Anderson JR, Cain KC, Gelber RD: Analysis of survival by tumor response and other comparisons of time-to-event by outcome variables. *J Clin Oncol* 26:3913-3915, 2008
28. R Development Core Team: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2005. <http://www.R-project.org>
29. US Food and Drug Administration: *Reinventing the Regulation of Cancer Drugs: Accelerating Approval and Expanding Access*. Washington, DC, National Performance Review, 1996
30. Llinares J: A regulatory overview about rare diseases. *Adv Exp Med Biol* 686:193-207
31. Gelmon KA, Eisenhauer EA, Harris AL, et al: Anticancer agents targeting signaling molecules and cancer cell environment: Challenges for drug development? *J Natl Cancer Inst* 91:1281-1287, 1999
32. Roberts TG Jr, Lynch TJ Jr, Chabner BA: The phase III trial in the era of targeted therapy: Unraveling the "go or no go" decision. *J Clin Oncol* 21:3683-3695, 2003
33. Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research: *Guidance for industry: Clinical trial endpoints for the approval of cancer drugs and biologics*. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071590.pdf>
34. Choi H, Charnsangavej C, Faria SC, et al: Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: Proposal of new computed tomography response criteria. *J Clin Oncol* 25:1753-1759, 2007
35. Chun YS, Vauthey JN, Boonsirikamchai P, et al: Association of computed tomography morphologic criteria with pathologic response and survival in patients treated with bevacizumab for colorectal liver metastases. *JAMA* 302:2338-2344, 2009
36. Paesmans M, Sculier JP, Libert P, et al: Response to chemotherapy has predictive value for further survival of patients with advanced non-small cell lung cancer: 10 years' experience of the European Lung Cancer Working Party. *Eur J Cancer* 33:2326-2332, 1997