THERAPEUTIC ADVANCES in
Medical Oncology

Original Research

Tumor response assessment by measuring the single largest lesion per organ in advanced non-small cell lung cancer patients treated with PD-1/PD-L1 inhibitor

Ther Adv Med Oncol

2023, Vol. 15: 1–12

DOI: 10.1177/17588359231200463

Li-Na He*, Tao Chen*, Sha Fu*, Yongluo Jiang*, Xuanye Zhang, Chen Chen, Wei Du, Linfeng Luo, Anlin Li, Yixing Wang, Hui Yu, Yixin Zhou, Yuhong Wang, Yunpeng Yang, Yan Huang, Hongyun Zhao, Wenfeng Fang, Li Zhang and Shaodong Hong

Abstract

Background: For Response Evaluation Criteria in Solid Tumors version 1.1 (RECIST1.1), measuring up to two target lesions per organ is an arbitrary criterion.

Objectives: We sought to compare response assessment using RECIST1.1 and modified RECIST1.1 (mRECIST1.1, measuring the single largest lesion per organ) in advanced non-small cell lung cancer (aNSCLC) patients undergoing anti-PD-1/PD-L1 monotherapy.

Methods: Concordance of radiologic response categorization between RECIST1.1 and mRECIST1.1 was compared using the Kappa statistics. C-index was calculated to evaluate prognostic accuracy of radiologic response by the two criteria. The Kaplan–Meier method and Cox regression analysis were conducted for progression-free survival (PFS) and overall survival (OS).

Results: Eighty-seven patients who had at least two target lesions in any organ per the RECIST1.1 were eligible for comparison analysis. Tumor response showed excellent concordance when measured using the RECIST1.1 and mRECIST1.1 (Kappa = 0.961). C-index by these two criteria was similar for PFS (0.784 *versus* 0.785) and OS (0.649 *versus* 0.652). Responders had significantly longer PFS and OS *versus* non-responders ($p < 0.05$), whichever criterion adopted. Radiologic response remained a significant predictor of PFS and OS in multivariate analysis ($p < 0.05$).

Conclusion: The mRECIST1.1 was comparable to RECIST1.1 in response assessment among aNSCLC patients who received single-agent PD-1/PD-L1 inhibitor. The mRECIST1.1, with reduced number of lesions to be measured, may be sufficient and more convenient to assess antitumor activity in clinical practice.

Keywords: immunotherapy, modified RECIST 1.1, non-small cell lung cancer, RECIST 1.1, target lesion, tumor response

Received: 8 February 2023; revised manuscript accepted: 25 August 2023.

Correspondence to:
Shaodong Hong
State Key Laboratory of Oncology in South China, Guangzhou, China

Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

Department of Medical Oncology, Sun Yat-sen University Cancer Center, 651 Dongfeng East Road, Guangzhou, Guangdong 510060, China

Department of Oncology, The People's Hospital of Fengqing, No. 107, Qingyun Group 1, Pingcun Village Committee, Fengshan Town, Fengqing County, Lincang 675900, China
hongshd@sysucc.org.cn

Li Zhang
State Key Laboratory of Oncology in South China, Guangzhou, China

Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

Department of Medical Oncology, Sun Yat-sen University Cancer Center, 651 Dongfeng East Road, Guangzhou, Guangdong 510060, China
zhangli6@mail.sysu.edu.cn

Li-Na He
State Key Laboratory of Oncology in South China, Guangzhou, China

Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

Tao Chen
Yongluo Jiang
State Key Laboratory of Oncology in South China, Guangzhou, China

Introduction

Immune checkpoint inhibitors (ICIs) targeting programed cell death protein 1 (PD-1) or programed cell death ligand 1 (PD-L1) have revolutionized the therapeutic landscape of advanced non-small cell lung cancer (aNSCLC) and have become a standard treatment modality.[1–5] These promising anticancer agents have demonstrated response rates ranging from 14% to 20% in recent monotherapy trials.[2,3,5,6] Response Evaluation

Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

Department of Nuclear Medicine, Sun Yat-sen University Cancer Center, Guangzhou, China

Sha Fu
Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation of Sun Yat-sen University; Department of Cellular & Molecular Diagnostics Center, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China

Xuanye Zhang
Wei Du
Linfeng luo
Anlin Li
Yixing Wang
Hui Yu
Yunpeng Yang
Yan Huang
Wenfeng Fang
State Key Laboratory of Oncology in South China, Guangzhou, China

Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

Department of Medical Oncology, Sun Yat-sen University Cancer Center, Guangzhou, China

Chen Chen
State Key Laboratory of Oncology in South China, Guangzhou, China

Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

Department of Radiation Oncology, Sun Yat-sen University Cancer Center, Guangzhou, China

Yixin Zhou
State Key Laboratory of Oncology in South China, Guangzhou, China

Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

Department of VIP region, Sun Yat-sen University Cancer Center, Guangzhou, China

Yuhong Wang
State Key Laboratory of Oncology in South China, Guangzhou, China

Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

Department of Endoscopy, Sun Yat-sen University Cancer Center, Guangzhou, China

Criteria in Solid Tumors (RECIST) guideline, currently, still represent the most widely utilized and standardized criteria to assess antitumor activity of treatment across routine patient management and clinical trial settings. It appears optimal to measure all lesions in order to capture accurate antitumor activity, which, in practice, is time-consuming and laborious.

The RECIST Working Group introduced in 2000 the RECIST version 1.0 (RECIST1.0)[7] and published in 2009 the revised RECIST version1.1 (RECIST1.1), with the latter based in part on assessment of tumor measurements from a data warehouse consisting of over 6500 patients treated with chemotherapy.[8–10] One of the major revisions of RECIST1.1 *versus* the RECIST1.0 guideline is that the number of target lesions has been reduced from 10 to 5 in total per patient and from 5 to 2 per organ (Supplemental Table 1). The codification of total 10 lesions in RECIST1.0 lacked extensive supporting data and has raised multiple issues in practical application. This could be mainly reflected by substantial cost of collecting, processing, and auditing imaging data and increased difficulty of response assessment due to ambiguous thus hard-to-evaluate products, which were characterized as target lesions at baseline. Contrarily, the RECIST1.1 defined a total of five targets according to continuous test and validation across an expanded data warehouse[9] and statistical simulation studies[10]; however, the criterion of up to two targets per organ is considered to be an arbitrary one and devoid of sufficient supporting evidence.

As reported in NSCLC,[11] metastatic colorectal cancer (CRC),[12] advanced gastric cancer (AGC),[13] and small cell lung cancer (SCLC)[14] patients who received first-line chemotherapy or chemoradiation, measuring the single largest lesion per organ (modified RECIST1.1, hereafter referred to as mRECIST1.1) could produce highly concordant response categorizations compared with measuring up to two target lesions per organ (RECIST1.1). These observations indicate that it remains to establish the ideal number of target lesions per organ to accurately assess tumor response to therapy. In terms of the emerging immunotherapy, atypical response patterns have been noted, as is the case in dissociated response (DR), which is denoted as some lesions shrinking while others growing.[15,16] The RECIST1.1 may miss out on information about some organs and misjudge a DR case, due to limited number of lesions measured. Furthermore, since RECIST criteria were initially developed and validated for chemotherapy, it requires continuous re-evaluation as a response assessment instrument with the emergence of new treatment paradigms.[17]

No previous studies we know of has illuminated response assessment using the mRECIST1.1 in immunotherapy landscape. We sought to investigate whether measuring the single largest target lesion per organ (mRECIST1.1; Supplemental Table 1) could produce similar response classifications and survival prediction with measuring up to two lesions per organ (RECIST1.1) in aNSCLC patients who received single-agent PD1/PD-L1 inhibitors.

Materials and methods

Patients

We retrospectively reviewed the electronic medical records of all histologically confirmed aNSCLC patients who received ICI monotherapy ($N=172$) at Sun Yat-sen University Cancer Center (SYSUCC) between August 2016 and June 2018. This study was conducted in accordance with the Declaration of Helsinki, and the study protocol was approved by the Institutional Review Board of SYSUCC with a waiver of patients' informed consent due to the retrospective design. Patients were eligible for inclusion in the study if they had radiologically or histologically confirmed advanced disease (stage IIIB or IV), available baseline and at least one follow-up radiologic tumor examinations, and no locoregional anti-cancer treatments during ICI therapy. Among the eligible patients for response assessment, those with at least two measurable lesions in any organ by RECIST1.1 were further included in the comparison analysis between mRECIST1.1 and RECIST1.1.

Computed tomography examinations

Computed tomography (CT) examinations were performed using a 64-slice spiral CT system (Aquilion TSX-101A; Toshiba Medical System; Osaka, Japan), a 128-slice spiral CT system (Discovery CT750 HD; GE System), a 256-slice spiral CT system (Brilliance iCT; Philips System; Orlando, USA), or a dual-source spiral CT system (SOMATOM Force;

Hongyun Zhao
State Key Laboratory of Oncology in South China, Guangzhou, China
Collaborative Innovation Center for Cancer Medicine, Guangzhou, China
Department of Clinical Research, Sun Yat-sen University Cancer Center, Guangzhou, China
*These authors contributed equally

Siemens Medical System; Erlangen, Germany). CT studies were performed using the following guidelines: slice thickness, 5 mm; slice interval, 1 mm; tube voltage, 80–140 kVp; tube current, automatic tube current modulation (maximum 450) mAs; and sagittal and coronal reconstruction thicknesses, 2 mm with 2 mm intervals. Contrast-enhanced CT was performed after an intravenous bolus dose of 1.5–2 mL/kg body weight of a non-ionic iodinated contrast agent (iopromide; Ultraist 300; Schering; Berlin, Germany) that was administered into the antecubital vein at a rate of 3.0 mL/s via a high-pressure syringe. The images were uploaded on a workstation (Advantage version 4.2; GE Healthcare; Chicago, USA). Imaging examinations were performed at the discretion of the physicians without a pre-determined interval.

Tumor measurements

Tumor measurements of each patient were evaluated manually from the original CT images using the calipers of a measurement tool on the workstation. According to RECIST1.1 and mRECIST1.1, two experienced oncologists (TC and YJ), who were blinded to patients' information, independently reviewed the CT images and recorded data including dates of CT examinations, target lesion description, CT size measurement, sum of the longest diameters of the target lesions (SLD), descriptions of any non-target lesions, occurrence of new lesions and overall response at each assessment time point, and the best overall response for each patient. Measurement of lymph node (LN) was performed in its short axis, considering LN of $\geqslant$15 mm to be a target lesion. LNs that measured $\geqslant$10 mm and $<$15 mm were considered to be non-target lesions, and LNs with a short axis of $<$10 mm were regarded as normal. The maximum number of target lesions to be assessed was five in total, with a maximum of two per organ (RECIST1.1) or a single largest lesion in each organ (mRECIST1.1).

Response and outcome assessment

Treatment responses according to RECIST1.1[8] and mRECIST1.1 were evaluated by the forementioned two reviewers, and discrepancy was resolved by consensus. Response at each evaluation time point comprehensively takes into account the changes of target, non-target, and new lesions. The best overall response (BOR) is the best response recorded from treatment initiation until the disease progression. Patients underwent tumor assessments until radiological-defined disease progression, study termination by physicians, or death. Durable clinical benefit (DCB) was defined as complete response (CR) plus partial response (PR) plus stable disease (SD) that lasted for 6 months or more from baseline, and overall response rate (ORR) as percentage of patients with a CR or PR. Progression-free survival (PFS) was calculated from the date of treatment onset until disease progression or death from any cause, whichever came first. Overall survival (OS) was calculated from the date of immunotherapy initiation until death due to any cause or last follow-up. Follow-up was completed on January 20, 2021.

Statistical analysis

Wilcoxon matched-pairs signed-ranks test was used for comparison of changes in the number of target lesions at baseline between the RECIST1.1 and the mRECIST 1.1. Chi-square test was conducted to compare the ORRs and DCBs between the two criteria. Kappa statistics was used to assess the level of concordance of response categorizations evaluated by these two criteria. Kappa value of $<$0.21 indicated poor concordance, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 good, and 0.81–1.00 excellent. Harrell's concordance index (C-index)[18,19] was calculated to investigate the prognostic performance of treatment response for OS and PFS, when using RECIST1.1 and mRECIST1.1 guidelines. C-index ranges from 0 to 1, with 0.5 denoting random estimation, 0.51–0.70 low accuracy, 0.71–0.90 intermediate accuracy, and 0.91–0.99 high accuracy. We compared median PFS and OS between responders (CR + PR) and non-responders [SD + progressive disease (PD)] using Kaplan–Meier methodology, and the difference was compared using log-rank test. Cox proportional hazard regression analysis was used to examine the association of treatment response and other clinicopathological characteristics with survival outcomes. Statistical analyses were conducted using the statistical software R, version 3.6.1 (R Institute for Statistical Computing, https://www.r-project.org/) and were tested at a two-sided significance level of 0.05.

Results

Patient characteristics

Out of 172 aNSCLC patients screened, 138 were eligible and 87 with at least two measurable lesions in any organ by RECIST1.1 were included in the comparison analysis between mRECIST1.1 and RECIST1.1 (Figure 1). Baseline characteristics of the 138 eligible patients and 87 patients for comparison analysis are summarized in Supplemental Table 2 and Table 1, respectively. Among the 87 patients included in comparison analysis, median age at diagnosis was 55 (range, 28–77) years; 63 (72.4%) were men; 57 (65.5%) had adenocarcinoma and 25 (28.7%) had squamous cell carcinoma; 38 (43.7%) patients had Eastern Cooperative Oncology Group (ECOG) performance status (PS) of 0 and 46 (52.9%) of 1; most patients (82/87, 94.3%) received ICI therapy in second- or later-line setting. Most patients had measurable lesions in lungs (76/87, 87.4%) and LNs (58/87, 66.7%). A total of 54 (62.1%) patients had target lesions in two organs, mostly found in the lungs and LNs. About 15 (17.2%) patients had target lesions only in a single organ, mostly in the lungs. As of data cut-off date, 38 (43.7%) patients were still alive or lost to follow-up: median follow-up was 39.2 [95% confidence interval (CI), 30.5–47.9] months and median OS was 18.2 (95% CI, 12.8–23.7) months.

Number of target lesions

The median (range) number of target lesions was 3 (2–5) by the RECIST1.1 and 2 (1–4) by the mRECIST1.1 (Table 1). The number of target lesions as per the mRECIST1.1 was significantly lower than that assessed per the RECIST1.1 (Wilcoxon p value <0.001). Compared with the RECIST1.1, no patient showed metastatic sites with a newly defined target organ by adopting the mRECIST1.1.

Tumor responses by mRECIST1.1 versus *RECIST1.1*

Of the 138 eligible patients, 28 (20.3%) patients achieved PR, 43 (31.2%) patients had SD, and 67 (48.6%) patients had PD as their best overall response as per the mRECIST1.1 criterion; 27 (19.6%) patients achieved PR, 40 (29.0%) patients had SD, and 71 (51.4%) patients had PD as their best overall response as per the RECIST1.1 criterion. We then focused on conducting comparison analysis between mRECIST1.1 and RECIST1.1 in the 87 patients who had at least two measurable lesions in any organ by RECIST1.1. The waterfall plot in Figure 2 portrays the percentage changes in the sum of the target lesion size according to the RECIST1.1, accompanying by the mRECIST1.1. The differences in the percentage changes of the sum of target lesion size between the RECIST1.1 and mRECIST1.1 were almost within 15%. Differences in change rate between the two criteria exceeded 15% in 9 patients (10.3%) (Supplemental Table 3), among whom 7 (7/9, 77.8%) patients had DR. Comparison of response categorizations between the two criteria is shown in Table 2. Tumor response showed excellent level of concordance when measured using the RECIST1.1 and mRECIST1.1 ($\kappa=0.961$; 95% CI, 0.908–1.000), with only two (2.2%) patients had inconsistent best response. One patient with SD by 15.3% decrease of the sum of tumor measurements according to the RECIST1.1 was recategorized as PR with 50.0% decrease according to the mRECIST1.1. Another patient classified as PD (20.3% increase) according to the RECIST1.1 was recategorized as SD (16.5% increase) according to the mRECIST1.1. We observed no significant difference regarding ORRs ($p=0.853$) and DCBs ($p=1.000$) between the two criteria (20.7% and 29.9% according to the RECIST1.1 *versus* 21.8% and 29.9% according to the mRECIST1.1).

Prognostic value of treatment response by two criteria

The prognostic performance of the RECIST1.1 and the mRECIST1.1 was first compared using C-index. C-index by these two criteria was similar in terms of OS [0.649 (95% CI, 0.586–0.712) *versus* 0.652 (95% CI, 0.589–0.715)] and PFS [0.784 (95% CI, 0.751–0.817) *versus* 0.785 (95% CI, 0.752–0.818)]. The Kaplan–Meier analysis revealed that responders by the two criteria had significantly longer OS and PFS *versus* non-responders; median OS: not reached (NR) (95% CI, NR-NR) months *versus* 14.8 (95% CI, 10.5–19.2) months both for the two criteria [Figure 3(a) and (c)]; median PFS: 16.6 (95% CI, 7.7–25.5) *versus* 2.1 (95% CI, 2.0–2.2) months for the RECIST1.1 and 16.6 (95% CI, 7.6–25.6) *versus* 2.1 (95% CI, 2.0–2.2) months for the mRECIST1.1 criterion [Figure 3(b) and (d)]. Detailed univariate Cox analyses for radiologic responses and other clinicopathological characteristics are shown in Supplemental Table 4. Considering

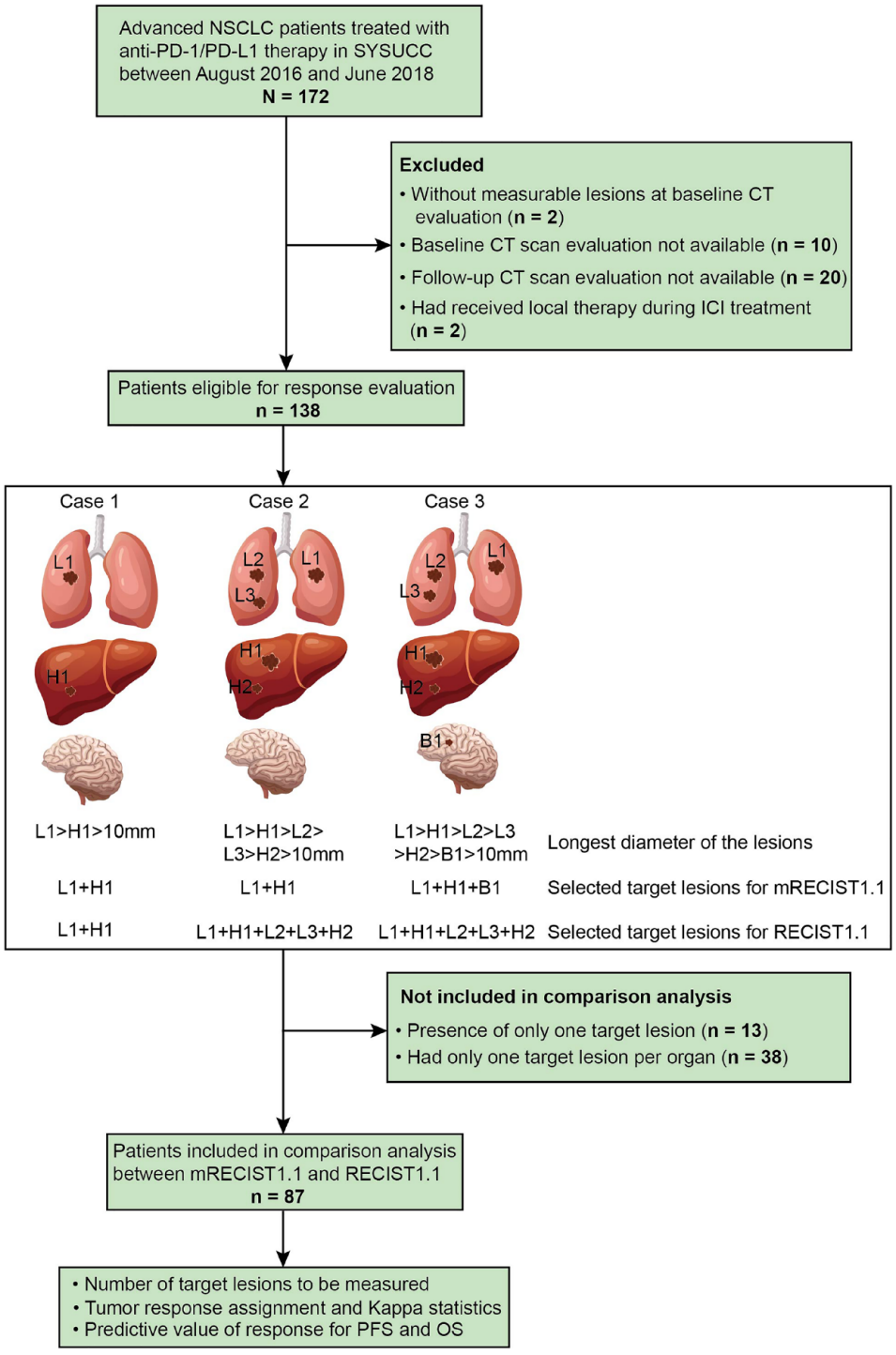

Figure 1. Flow diagram delineating patient screening procedure and comparison of mRECIST1.1 and RECIST1.1. The middle panel of the flow diagram shows how to perform measurement of mRECIST1.1 and RECIST1.1.

B, brain; CT, computed tomography; H, hepar; ICI, immune checkpoint inhibitor; L, lung; mRECIST1.1, modified RECIST1.1; NSCLC, non-small cell lung cancer; PD-(L)1, programed cell death protein (ligand) 1; PFS, progression-free survival; OS, overall survival; RECIST1.1, Response Evaluation Criteria in Solid Tumors version 1.1; SYSUCC, Sun Yat-sen University Cancer Center.

Table 1. Baseline characteristics for patients included in comparison analysis ($n = 87$).

Patient characteristics	No. (%)
Age, years	
Median (range)	55 (28–77)
<55	43 (49.4)
⩾55	44 (50.6)
Gender	
Male	63 (72.4)
Female	24 (27.6)
ECOG PS	
0	38 (43.7)
1	46 (52.9)
2	3 (3.4)
Stage	
IIIB	2 (2.3)
IV	85 (97.7)
Smoking status	
Never smoker	53 (60.9)
Current or former smoker	34 (39.1)
Histology	
Squamous cell carcinoma	25 (28.7)
Adenocarcinoma	57 (65.5)
Other types	5 (5.8)
No. of prior lines of therapy	
0	5 (5.8)
1	47 (54.0)
2	16 (18.4)
⩾3	19 (21.8)
No. of metastatic sites	
1	19 (21.8)
2	31 (35.6)
3	23 (26.5)
⩾4	14 (16.1)
EGFR mutation status	
Positive	9 (10.3)
Negative	63 (72.4)
Not available	15 (17.3)
ALK translocation	
Positive	3 (3.4)
Negative	61 (70.1)
Not available	23 (26.5)
Measurable target lesions	
Lungs	76 (87.4)
Lymph nodes	58 (66.7)
Liver	19 (21.8)
Pleura	11 (12.6)
Brains	9 (10.3)
Adrenal glands	7 (8.0)
Kidneys	2 (2.3)
Spleen	1 (1.1)
No. target lesions by mRECIST1.1	
Median (range)	2 (1–4)
1	15 (17.2)
2	54 (62.1)
3	9 (10.3)
4	9 (10.3)
No. target lesions by RECIST1.1	
Median (range)	3 (2–5)
2	15 (17.2)
3	43 (49.5)
4	15 (17.2)
5	14 (16.1)

ALK, anaplastic lymphoma kinase; ECOG PS, Eastern Cooperative Oncology Group performance status; EGFR, epidermal growth factor receptor; mRECIST1.1, modified Response Evaluation Criteria in Solid Tumors version 1.1; RECIST1.1, Response Evaluation Criteria in Solid Tumors version 1.1.

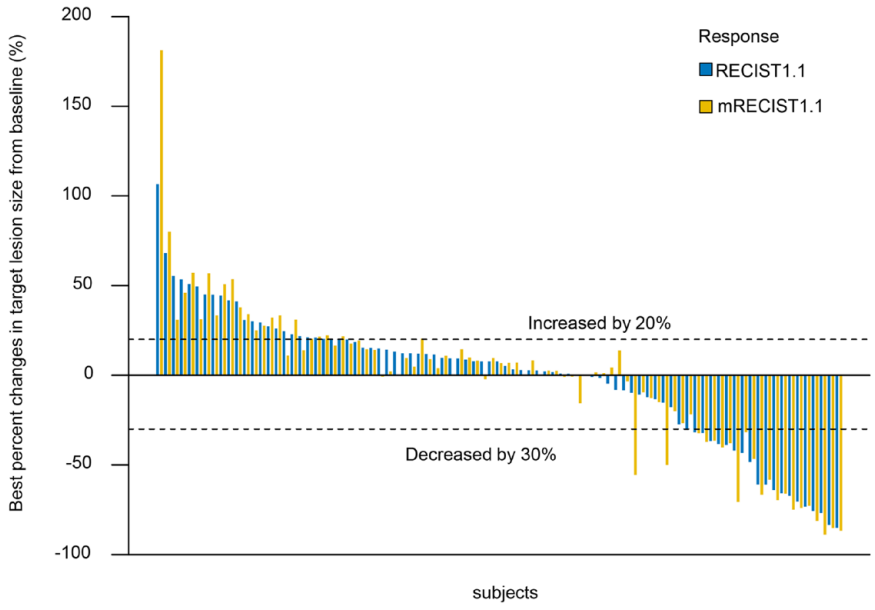

Figure 2. Waterfall plot shows the percent changes in the sum of the target lesion size according to the RECIST1.1 accompanying by the mRECIST1.1.
mRECIST1.1, modified Response Evaluation Criteria in Solid Tumors version 1.1; RECIST1.1, Response Evaluation Criteria in Solid Tumors version 1.1.

Table 2. Best overall response by the RECIST1.1 *versus* the mRECIST1.1.

Tumor response by mRECIST1.1	Tumor response by RECIST1.1				Kappa (95% CI)
	PR	SD	PD	Total	
PR	18	1	0	19	
SD	0	19	1	20	0.961 (0.908–1.000)
PD	0	0	48	48	
Total	18	20	49	87	

mRECIST1.1, modified Response Evaluation Criteria in Solid Tumors version 1.1; PD, progressive disease; PR, partial response; RECIST1.1, Response Evaluation Criteria in Solid Tumors version 1.1; SD, stable disease; CI, confidence interval.

clinical relevance, all factors analyzed in univariate analyses were entered into a multivariate Cox regression model along with one of radiologic parameters assessed by RECIST1.1 and mRECIST1.1. Radiologic response remained significant in predicting OS and PFS, whichever criterion adopted. Detailed adjusted hazard ratios (HRs) for radiologic response from multivariate analysis are presented in Table 3.

Discussion

An accurate and reproducible evaluation of anticancer efficacy is essential for routine patient management as well as clinical trials designed for approval of new drugs. Changes in target lesion size remains a key backbone of response assessment by the RECIST criteria. It would be ideal to measure all lesions in a given patient, with an attempt to accurately capture antitumor efficacy, but this is laborious and time-consuming in clinical practice. Therefore, it is critical to select and follow the most appropriate number of target lesions that can sufficiently reflect the whole tumor burden change. The present study demonstrated that measuring the single largest lesion per organ with five in total (termed mRECIST1.1) showed a concordant response categorization in 97.7% of evaluations when compared with measuring up to two target lesions per organ according

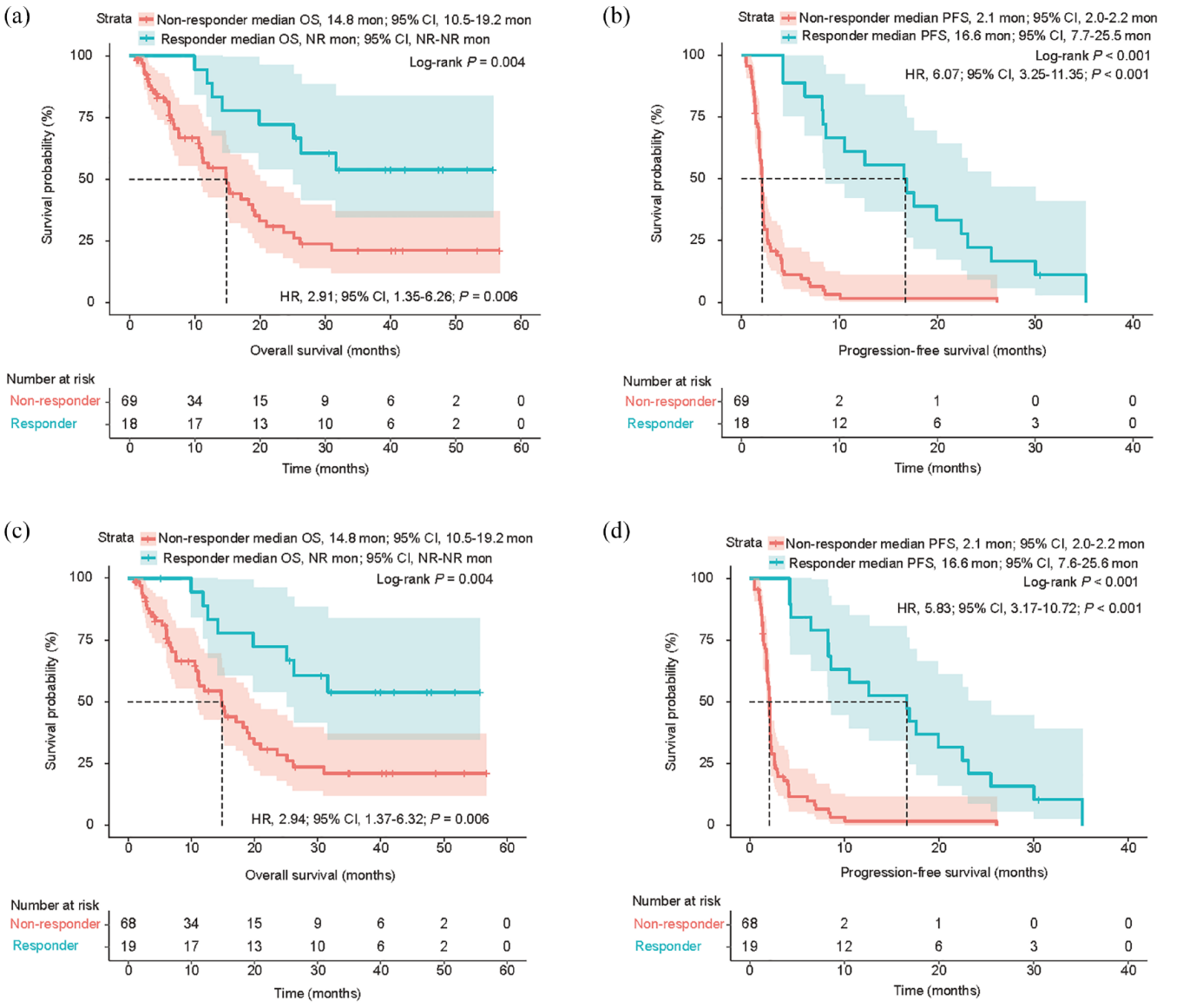

Figure 3. Kaplan–Meier analysis of overall survival (OS) and progression-free survival (PFS). OS (a) and PFS (b) according to radiologic response measured by the RECIST1.1. OS (c) and PFS (d) according to radiologic response measured by the mRECIST1.1.
mRECIST1.1, modified Response Evaluation Criteria in Solid Tumors version 1.1; RECIST1.1, Response Evaluation Criteria in Solid Tumors version 1.1.

to the RECIST1.1. The mRECIST1.1 criteria, meanwhile, significantly reduced the number of target lesions to be measured for response assessment in aNSCLC patients who received single-agent PD1/PD-L1 inhibitors.

The RECIST 1.0 guideline recommends up to five lesions per organ and ten in total to be measured,[7] which was considered to be an arbitrary provision and lacked objective supporting evidence. Along with multiple considerations raised by this provision in practice,[20] the RECIST Working Group published the revised RECIST version 1.1, based on an expanded data warehouse, which was developed to test and validate modifications to the RECIST criteria thus supported the five-target-based measurement in RECIST1.1.[9] Of note, the criterion of up to two lesions per organ in the RECIST1.1 was still an arbitrary decision, without confirmatory ground truth to rely on. When patients had three or more metastatic sites, especially for cases with dissociated response (also termed mixed response, or heterogeneous response), response assessment by the RECIST1.1 may not accurately reflect all involved organs due to limited total number of

Table 3. Detailed median survival time and adjusted HRs from multivariate analysis.

Guidelines	OS, months (responders *versus* non-responders)	Adjusted HR (95% CI)	*p* value	PFS, months (responders *versus* non-responders)	Adjusted HR (95% CI)	*p* value
RECIST 1.1	NR *versus* 14.8	3.30 (1.51–7.24)	0.003	16.6 *versus* 2.1	8.15 (4.14–16.07)	<0.001
mRECIST 1.1	NR *versus* 14.8	3.32 (1.51–7.27)	0.003	16.6 *versus* 2.1	10.56 (5.14–21.67)	<0.001

CI, confidence interval; HR, hazard ratio; mRECIST 1.1, modified Response Evaluation Criteria in Solid Tumors version 1.1; NR, not reached; OS, overall survival; PFS, progression-free survival; RECIST 1.1, Response Evaluation Criteria in Solid Tumors version 1.1.

targets to be measured. DR was observed in around 5−10% of aNSCLC patients treated with anti-PD1/PD-L1 antibody.[15,16,21,22] We postulated that measuring the single largest lesion per organ could be representative of more metastatic sites in aNSCLC patients and compared response assignment by the RECIST1.1 guideline with that assessed using the mRECIST1.1.

Among 87 aNSCLC patients with two or more measurable lesions in any organ, we identified that 9 (10.3%) patients had target lesions in more than three organs. Independent of our initial expectation, however, we found no patient showed metastatic sites with a newly defined target organ by adopting the mRECIST1.1 rather than the RECIST1.1. Therefore, we were unable to estimate whether the mRECIST1.1 could exert an influence on treatment response assessment in those cases. We thought this may due, in part, to the modest sample size in the study and in-depth studies are warranted in another large-scale cohort. Similar to our observation that the best overall response showed excellent level of concordance between the two criteria ($\kappa = 0.961$), previous studies have demonstrated that measuring the single largest lesion per organ (mRECIST1.1) was comparable to measuring up to two in each organ (RECIST1.1) in treatment response assessment of patients with advanced NSCLC,[11] metastatic colorectal cancer,[12] AGC,[13] and SCLC.[14] Most patients in these series of studies received first-line therapy with cytotoxic drugs, while we recruited patients who underwent anti-PD-1/PD-L1 monotherapy and further pinpointed the association of radiologic response measured by two criteria with survival outcomes at a longitudinal dimension. Prognostic accuracy (calculated as C-index) of radiologic response assessed by two criteria was similar both for PFS and OS. Responders had prolonged PFS and OS compared with non-responders, whichever criterion adopted. Given these observations from our study and previous studies, the mRECIST1.1, with reduced number of lesions to be measured, could produce equivalent response categorization to the conventional RECIST1.1 guideline, without compromising prognostic ability, and thus it was deemed worthwhile to recommend this modified criterion in response assessment of anticancer drugs and to have potential to be included in the next response assessment guideline of NSCLCs. Modifications to different RECIST versions are not discussed in clinical guidelines such as the American Society of Clinical Oncology (ASCO), European Society for Medical Oncology (ESMO), and National Comprehensive Cancer Network (NCCN) guidelines. Currently, detailed changes in RECIST criteria are dissected by the RECIST Working Group and were published separately. We recommend involvement of changes in different RECIST versions in relevant authoritative guidelines.

We noticed that the differences in percent change of the sum of the tumor measurements between the mRECIST1.1 and RECIST1.1 in our study were evident in 9 (10.3%) cases (exceeding 15%), whereas all evaluations were within 10% or 13% in prior studies.[11,14] This might be mainly explained by: (1) different organs in a given patient could demonstrate an apparent heterogeneous sensitivity to PD-1 axis inhibitor, just as noted in a study which showed that lesion-based size change at best response differed significantly across different organs[23]; (2) we did not exclude patients who had shown substantial progression in non-target lesions or the occurrence of new lesions and 3 out of these 9 (33.3%) cases had progressed due to the development of new lesions; however, this screening procedure could, to some extent, make our study less subjected to potential selection bias. Regarding the response categorization, only 2 out of 87 (2.2%) patients showed inconsistent best overall response assessed by the two criteria. One patient classified as PD according to the RECIST1.1 was

recategorized as SD according to the mRECIST1.1. Nonetheless, whether this case could or could not actually benefit from immunotherapy was unevaluable, for treatment was terminated once a progressive disease was assigned at follow-up examination. Another patient had discordant response between SD (RECIST1.1) and PR (mRECIST1.1), in which case treatment regimens would remain the same in clinical practice. It was thus reasonable for us to concluded that the mRECIST1.1 had subtle impact on the shift of treatment decision.

Our study has several limitations needed to be considered. First, the study was retrospectively conducted at a single institution with a moderate sample size, which may limit the statistical power and have challenges of bias. However, the study cohort is representative of the whole advanced NSCLC patient population who received single-agent inhibitor at our institution and inclusion of patients from both clinical trials and real-world clinical practice makes the findings generalizable across various therapeutic settings. Second, our analysis was only conducted in a single malignancy, indicating that whether our observations could be generalized to other types of solid tumors remains to be determined. Therefore, the value of mRECIST1.1 in response assessment requires further investigation in larger prospective cohorts and across various types of malignancies. Another concern is that the largest target lesion may not always be the optimal one to reflect antitumor activity, since large tumors are more likely to undergo necrosis or cavitation in partial region within the tumor and may not sufficiently represent the whole tumor burden change; however, the RECIST1.1 criteria also face challenges when measuring these tumors.[24] 18F-fluorodeoxyglucose (18F-FDG) positron emission tomography (PET), and an alternative measurement that excludes the area of cavitation have been proposed to addressed these issues, but without a definitive guideline to standardize routine application of these methods in clinical practice and clinical trials.[24]

Conclusion

In conclusion, the modified RECIST1.1, with a reduction of number of lesions to be measured, demonstrated a high level of concordance with the conventional RECIST1.1 guideline in response assessment for aNSCLC patients treated with single-agent PD-1 axis inhibitor. Measuring the single largest lesion per organ may be suffice and more convenient to capture the global treatment response of the patient. However, further confirmatory studies with larger sample size and across various solid tumors are warranted before introducing this criterion into clinical practice.

Declarations

Ethics approval and consent to participate

The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Institutional Review Board of Sun Yat-sen University Cancer Center (approved number: SL-B2020-402-02) and individual consent for this retrospective analysis was waived. For protection of the patient's personal data, only anonymized data have been used for the analyses of this study.

Consent for publication

Not applicable.

Author contributions

Li-Na He: Data curation; Formal analysis; Investigation; Resources; Writing – original draft.

Tao Chen: Data curation; Formal analysis; Investigation; Resources; Writing – original draft.

Sha Fu: Data curation; Formal analysis; Investigation; Resources; Writing – original draft.

Yongluo Jiang: Data curation; Formal analysis; Investigation; Resources; Writing – original draft.

Xuanye Zhang: Data curation; Formal analysis; Investigation; Resources.

Chen Chen: Data curation; Formal analysis; Investigation; Resources.

Wei Du: Data curation; Visualization.

Linfeng Luo: Data curation; Visualization.

Anlin Li: Data curation; Visualization.

Yixing Wang: Software; Validation.

Hui Yu: Software; Validation.

Yixin Zhou: Writing – review & editing.

Yuhong Wang: Software; Validation.

Yunpeng Yang: Data curation; Visualization.

Yan Huang: Software; Validation.

Hongyun Zhao: Data curation; Visualization.

Wenfeng Fang: Software; Validation.

Li Zhang: Conceptualization; Data curation; Funding acquisition; Methodology; Project administration; Supervision.

Shaodong Hong: Conceptualization; Data curation; Funding acquisition; Methodology; Project administration; Supervision; Writing – review & editing.

Acknowledgements

We would like to sincerely appreciate all the patients and their family who were included in this retrospective study.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by National Natural Science Foundation of China (Grant numbers: 81972898, 82172713); Guangdong Basic and Applied Basic Research Foundation (Grant number: 2023B1515020008), and the Fundamental Research Funds for the Central Universities, Sun Yat-sen University (Grant number: 22ykqb15).

Competing interests

The authors declare that there is no conflict of interest.

Availability of data and materials

All data and material relevant to the study are included within the article and its additional file. Raw data of this study are available from the corresponding authors upon reasonable request.

Role of the funding source

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

ORCID iDs

Li-Na He https://orcid.org/0000-0002-5416-7290

Yongluo Jiang https://orcid.org/0000-0002-3456-0570

Chen Chen https://orcid.org/0000-0002-8303-2862

Hui Yu https://orcid.org/0000-0003-2288-4671

Supplemental material

Supplemental material for this article is available online.

References

1. Blumenthal GM, Zhang L, Zhang H, *et al.* Milestone analyses of immune checkpoint inhibitors, targeted therapy, and conventional therapy in metastatic Non-Small cell lung cancer trials: a meta-analysis. *JAMA Oncol* 2017; 3: e171029.
2. Brahmer J, Reckamp KL, Baas P, *et al.* Nivolumab versus docetaxel in advanced squamous-cell Non–Small-Cell lung cancer. *N Engl J Med* 2015; 373: 123–135.
3. Borghaei H, Paz-Ares L, Horn L, *et al.* Nivolumab versus docetaxel in advanced nonsquamous Non–Small-Cell lung cancer. *N Engl J Med* 2015; 373: 1627–1639.
4. Reck M, Rodríguez-Abreu D, Robinson AG, *et al.* Pembrolizumab versus chemotherapy for PD-L1–Positive Non–Small-Cell lung cancer. *N Engl J Med* 2016; 375: 1823–1833.
5. Rittmeyer A, Barlesi F, Waterkamp D, *et al.* Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial. *J Lancet* 2017; 389: 255–265.
6. Wu YL, Lu S, Cheng Y, *et al.* Nivolumab versus docetaxel in a predominantly Chinese patient population with previously treated advanced NSCLC: checkmate 078 Randomized Phase III clinical trial. *J Thorac Oncol* 2019; 14: 867–875.
7. Therasse P, Arbuck SG and Eisenhauer EA. New guidelines to evaluate the response to treatment in solid tumors. European Organization for research and treatment of cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000; 92: 205–216.
8. 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: 228–247.
9. Bogaerts J, Ford R, Sargent D, *et al.* Individual patient data analysis to assess modifications to the RECIST criteria. *Eur J Cancer* 2009; 45: 248–260.

10. Moskowitz CS, Jia X, Schwartz LH, *et al.* A simulation study to evaluate the impact of the number of lesions measured on response assessment. *Eur J Cancer* 2009; 45: 300–310.

11. Kim HS, Kim JH and Yang I. Tumor response assessment by measuring the single largest lesion per organ in patients with advanced non-small cell lung cancer. *Lung Cancer* 2014; 85: 385–389.

12. Kim HS and Kim JH. Ideal number of target lesions per organ to measure in metastatic colorectal cancer. *Oncol Lett* 2014; 8: 1896–1900.

13. Kim HS, Kim JW, Kim JH, *et al.* Single-lesion measurement per organ for assessing tumor response in advanced gastric cancer. *Oncology* 2015; 88: 69–75.

14. Jung SG, Kim JH, Kim HS, *et al.* Tumor response assessment by the single-lesion measurement per organ in small cell lung cancer. *Chin J Cancer Res* 2016; 28: 161–167.

15. Sato Y, Morimoto T, Hara S, *et al.* Dissociated response and clinical benefit in patients treated with nivolumab monotherapy. *Investig New Drugs* 2021; 39: 1170–1178.

16. Tazdait M, Mezquita L, Lahmar J, *et al.* Patterns of responses in metastatic NSCLC during PD-1 or PDL-1 inhibitor therapy: comparison of RECIST 1.1, irRECIST and iRECIST criteria. *Eur J Cancer* 2018; 88: 38–47.

17. Schwartz LH, Seymour L, Litière S, *et al.* RECIST 1.1 – Standardisation and disease-specific adaptations: perspectives from the RECIST Working Group. *Eur J Cancer* 2016; 62: 138–145.

18. Zhou X-H, McClish DK and Obuchowski NA. *Statistical methods in diagnostic medicine*. New York: John Wiley & Sons, lnc, 2002.

19. Lee FE Jr., Mark KL, DB, Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996; 15: 361–387.

20. Hillman SL, An MW, O'Connell MJ, *et al.* Evaluation of the optimal number of lesions needed for tumor evaluation using the response evaluation criteria in solid tumors: a north central cancer treatment group investigation. *J Clin Oncol* 2009; 27: 3205–3210.

21. Tozuka T, Kitazono S, Sakamoto H, *et al.* Dissociated responses at initial computed tomography evaluation is a good prognostic factor in non-small cell lung cancer patients treated with anti-programmed cell death-1/ligand 1 inhibitors. *BMC Cancer* 2020; 20: 207.

22. Humbert O, Cadour N, Paquet M, *et al.* (18)FDG PET/CT in the early assessment of non-small cell lung cancer response to immunotherapy: frequency and clinical significance of atypical evolutive patterns. *Eur J Nucl Med Mol Imaging* 2020; 47: 1158–1167.

23. Nishino M, Ramaiya NH, Chambers ES, *et al.* Immune-related response assessment during PD-1 inhibitor therapy in advanced non-small-cell lung cancer patients. *J Immunother Cancer* 2016; 4: 84.

24. Nishino M, Jagannathan JP, Ramaiya NH, *et al.* Revised RECIST guideline version 1.1: what oncologists want to know and what radiologists need to know. *AJR Am J Roentgenol* 2010; 195: 281–289.