Comparison of EORTC Criteria and PERCIST for PET/CT Response Evaluation of Patients with Metastatic Colorectal Cancer Treated with Irinotecan and Cetuximab

Kristin Skougaard[1], Dorte Nielsen[1], Benny Vittrup Jensen[1], and Helle Westergren Hendel[2]

[1]Department of Oncology, Copenhagen University Hospital Herlev, Herlev, Denmark; and [2]Department of Nuclear Medicine, Copenhagen University Hospital Herlev, Herlev, Denmark

The study aim was to compare European Organization for Research and Treatment of Cancer (EORTC) criteria with PET Response Criteria in Solid Tumors (PERCIST) for response evaluation of patients with metastatic colorectal cancer treated with a combination of the chemotherapeutic drug irinotecan and the monoclonal antibody cetuximab. **Methods:** From 2006 to 2009, patients with metastatic colorectal cancer were prospectively included in a phase II trial evaluating the combination of irinotecan and cetuximab every second week, as third-line treatment. ^{18}F-FDG PET/CT was performed between 1 and 14 d before the first treatment and after every fourth treatment cycle until progression was identified by CT with Response Evaluation Criteria in Solid Tumors (RECIST). Response evaluation with ^{18}F-FDG PET/CT was retrospectively performed according to both EORTC criteria and PERCIST, classifying the patients into 4 response categories: complete metabolic response (CMR), partial metabolic response (PMR), stable metabolic disease (SMD), and progressive metabolic disease (PMD). Individual best overall metabolic response (BOmR) was registered with both sets of criteria, as well as survival within response categories, and compared. **Results:** A total of 61 patients and 203 PET/CT scans were eligible for response evaluation. With EORTC criteria, 38 had PMR, 16 had SMD, and 7 had PMD as their BOmR. With PERCIST, 34 had PMR, 20 had SMD, and 7 had PMD as their BOmR. None of the patients had CMR. There was agreement between EORTC criteria and PERCIST in 87% of the patients, and the corresponding κ-coefficient was 0.76. Disagreements were confined to PMR and SMD. Median overall survival (OS) in months with EORTC criteria was 14.2 in the PMR group and 7.2 in the combined SMD + PMD group. With PERCIST, it was 14.5 in the PMR group and 7.9 in the SMD + PMD group. **Conclusion:** Response evaluation with EORTC criteria and PERCIST gave similar responses and OS outcomes with good agreement on BOmR (κ-coefficient, 0.76) and similar significant differences in median OS between response groups. Compared with EORTC criteria, we find PERCIST unambiguous because of clear definitions and therefore more straightforward to use.

Key Words: EORTC criteria; PERCIST; PET/CT; response evaluation; survival

PET/CT performed with ^{18}F-FDG has become an established imaging modality in oncology (1,2). For several malignancies, including colorectal cancer, ^{18}F-FDG PET/CT examination is today recommended for preoperative staging and detection of recurrence (3–6), and there is growing research attention to expanding the use of ^{18}F-FDG PET/CT in evaluating the response of metastatic disease (7–10). To create the reproducibility that is needed for comparison of response rates between trials, the use of ^{18}F-FDG PET/CT in the setting of metastatic disease requires a fundamental standardization and consensus on response quantification methodology (11–13). Otherwise, the potential benefit for patients and for anticancer drug development could be lost.

PET/CT-based response evaluation has proven to be valuable in chemotherapy and especially in targeted treatment (14–16). Currently, 2 sets of criteria to quantify anticancer treatment response are available: the criteria developed by the European Organization for Research and Treatment of Cancer (EORTC) (17) and PET Response Criteria in Solid Tumors (PERCIST) (18). The EORTC criteria were published in 1999 by Young et al. (17) and are based on baseline-chosen, lesion-specific regions of interest (ROIs) that are followed on each subsequent scan. The chosen lesions should be the most ^{18}F-FDG–avid. PERCIST was published in 2009 by Wahl et al. (18) and operates with a fixed ROI of 1 cm^3 in the most ^{18}F-FDG–avid part of the single most metabolically active tumor in the patient at each PET/CT scan. This region is, on the basis of cancer stem cell theory (19,20), regarded as an indicator of the patient's disease status at the given time point and is not necessarily located in the same lesion at all scans. Apparently, EORTC criteria and PERCIST have quite different approaches to evaluating treatment response; nevertheless, a comparison of response evaluation with the 2 sets of criteria has, to our knowledge, not yet been performed. It is necessary to characterize the potential differences in outcome generated by the 2 sets of criteria in order to elucidate whether the criteria can be used interchangeably or give rise to significantly different results.

The aim of this study was to compare response evaluation, and corresponding groupwise overall survival (OS), with EORTC criteria to that with PERCIST in patients with metastatic colorectal cancer and to discuss advantages and disadvantages in their clinical applicability.

MATERIALS AND METHODS

Patients

From 2006 to 2009, patients with metastatic colorectal cancer were prospectively included in a Danish phase II multicenter trial. They

were given a combination of the chemotherapeutic drug irinotecan (Fresenius Kabi Oncology), 180 mg/m^2, and the monoclonal antibody cetuximab (Erbitux; Merck), 500 mg/m^2, every second week as a third-line treatment. The protocol was approved by the Danish Regional Research Ethics Committee, the Danish Medicines Agency, and the Data Protection Agency (EudraCT no. 2006-001961-40). Oral and written informed consent was obtained from all patients before inclusion in the trial. Only patients recruited at Copenhagen University Hospital Herlev were included in the current study.

The patients were scanned between 1 and 14 d before the first treatment and after every fourth treatment cycle. Treatment was continued until progression was identified by CT according to Response Evaluation Criteria in Solid Tumors (RECIST) (21). The patients' treatment course was determined solely on the CT-based RECIST response evaluation. PET/CT response evaluation was performed retrospectively. All scans were read twice by the same dedicated specialist, first according to EORTC criteria and then according to PERCIST. The reader was masked to the response outcome of the patients.

PET/CT Examinations

Two different scanners were used: Gemini Dual Slice PET/CT (Philips) and TruFlight 16-Slice PET/CT (Philips). Emission scanning was obtained over 5–6 axial regions (bed positions) of 18 cm each with 50% overlap, at a rate of 2 min per bed position. The delimitations were the base of the skull and mid thigh. Image fusion and standardized uptake value (SUV) calculations were performed with the scanner-specific software. The Tumor Tracking application of Extended Brilliance Workspace nuclear medicine software (version 2.0; Philips) was used to draw ROIs and register maximum SUV (SUVmax), mean SUV, and SD in the recorded ROIs. The intention was to scan each individual patient on the same scanner throughout that patient's treatment course. If patients were scanned on the 2 different scanners in a manner that precluded response evaluation, they were excluded.

^{18}F-FDG was injected intravenously, with a target dose of 370 MBq and an intentional 60 min of uptake time before the start of scanning. Patients fasted for at least 5 h before imaging and were offered water before and after the scan. Blood glucose level was measured immediately before tracer injection using the Ascensia Contour system (Bayer A/S Diabetes Care). Diabetic patients were not excluded, but patients with blood glucose levels of 8 mM or greater were excluded. To avoid potential chemotherapy-induced stunning or flare, the scans were obtained between 10 and 14 d after the last treatment (22–24). The scanning procedure closely resembles the NCI and EANM guidelines (25,26).

The multidetector spiral (2 or 16 slices per rotation) CT scans were standard diagnostic contrast-enhanced examinations covering the region of the thorax, abdomen, and pelvis, performed according to local standard guidelines. Iodinated contrast agent (Omnipaque 350; GE Healthcare) was given orally (20 mL in 500 mL of bottled water [4% solution] half an hour before CT) and intravenously (100 mL with an injection flow of 5 mL/s immediately before the start of scanning). Rotation speed was 0.5 s/rotation, collimation was 5 mm, and the minimal slice thickness was 2.5 mm. The patient stayed in the scanner in the same position during the diagnostic CT scan, the low-dose CT scan, and the PET emission scan. The low-dose CT scan was used for attenuation correction.

Response Evaluation with EORTC Criteria

For SUV normalization, EORTC recommends that body surface area be calculated with the algorithm of Dubois and Dubois. The same algorithm was used in the software. We chose up to 7 of the lesions with the highest ^{18}F-FDG uptake in as many involved organ systems as possible as target lesions at baseline and measured these same lesions on every subsequent follow-up scan. We chose to use SUVmax in preference to mean SUV and therefore did not use isovolumetric measurements of lesion-specific ROIs. SUVmax measurements from all target lesions were summed on each scan, giving ΣSUVmax. At the first follow-up and if ΣSUVmax was decreasing compared with baseline, response was calculated as ΔΣSUVmax between baseline and actual follow-up divided by baseline ΣSUVmax × 100%. If SUVmax increased, response was calculated as ΔΣSUVmax between lowest registered and actual follow-up divided by lowest registered ΣSUVmax × 100%.

Response was classified on each scan according to the 4 categories defined in the criteria. Complete metabolic response (CMR) was complete resolution of ^{18}F-FDG uptake within all lesions, making them indistinguishable from surrounding tissue. Partial metabolic response (PMR) was a reduction in ΣSUVmax of at least 25% after more than 1 treatment cycle. Progressive metabolic disease (PMD) was an increase of at least 25% in ΣSUVmax or a new ^{18}F-FDG–avid lesion. Stable metabolic disease (SMD) was a response between PMR and PMD. The best achieved response (CMR, PMR, SMD, or PMD) during a patient's treatment course was assessed from consecutive scans and registered as the best overall metabolic response (BOmR). Metabolic response rate (the rate of patients with CMR and PMR) was calculated from the patients' BOmR.

Response Evaluation with PERCIST

PERCIST recommends the use of lean body mass for SUV normalization, with no particular algorithm stated. SUV normalized to lean body mass is termed SUL. The background area was drawn as a 3-cm-diameter spheric ROI in the right lobe of the liver as defined in the criteria. In patients with liver involvement, the background area was drawn in the descending thoracic aorta. With the available software, it was not possible to extend the ROI from 1 to 2 cm in the z-axis as described in the criteria and it was drawn as a spheric 1-cm-diameter ROI. The SD of the mean SUV of all liver and aorta background ROIs was registered.

The lesion with highest SUL was identified, and a 1.2-cm-diameter spheric ROI was drawn in the hottest part of that lesion. The ROI was placed in the area of the tumor where it resulted in the highest possible mean SUL (SULmean). SULmean of this ROI was SULpeak. Baseline SULpeak had to exceed 1.5 × liver SULmean + 2 × SD of liver SULmean or 2 × aorta SULmean + 2 × SD of aorta SULmean for the tumor to qualify as a target lesion. It was checked that no other lesion could give a higher SULpeak. On subsequent scans, SULpeak could be located in a different lesion from the one measured at baseline, as long as the lesion had been present since baseline. If SULpeak at baseline did not exceed the background value, the patient was not eligible for response evaluation with PERCIST. At the first follow-up and if SULpeak was decreasing, response was calculated as ΔSULpeak between baseline and actual follow-up divided by baseline SULpeak × 100%. If SULpeak increased, response was calculated as ΔSULpeak between lowest registered and actual follow-up divided by lowest registered SULpeak × 100%.

Response was classified on each scan according to the 4 categories defined in the criteria set. CMR was complete resolution of ^{18}F-FDG uptake within all lesions to a level less than or equal to that of mean liver activity and indistinguishable from background blood-pool levels. PMR was a reduction of at least 30% in SULpeak and an absolute drop of 0.8 SULpeak units. PMD was an increase of at least 30% in SULpeak and an absolute increase of 0.8 SULpeak units, or a new ^{18}F-FDG–avid lesion. SMD was between PMR and PMD. Response rate was calculated from the patients' BOmR.

The κ-statistic was used for agreement analysis. The fraction of patients with a shift in the hottest lesion during treatment was calcu-

lated as a percentage. The Kaplan–Meier method was used for OS analysis, with log-rank testing for P value calculation. OS was defined as the time from trial registration of a patient until death from any course.

RESULTS

Among 150 included patients, 131 were examined with PET/CT throughout their treatment course. Of these, 13 were screen failures (never scanned and never treated) and 37 were excluded because of the patient's own wish, anaphylactic reactions to the first treatment infusion, or clinical progression before the first follow-up. One patient with a blood glucose level above the exclusion criterion, 1 patient with no measurable disease on PET, and 4 patients with unavailable PET images were excluded. Furthermore, 2 patients who did not have a target lesion according to PERCIST and 12 patients who were scanned on 2 different scanners in a manner that prevented response evaluation were excluded. A flow diagram is given in Figure 1. Ultimately, 61 patients were eligible for PET/CT response evaluation. Disease involvement was seen in the liver, abdominal lymph nodes, lungs, peritoneum (carcinomatosis), rectum, bones, and spleen. Patient characteristics are outlined in Table 1.

The total number of scans was 230. Of these, 27 scans (12%) were obtained on a scanner other than the one on which the individual patient was scanned at baseline and were therefore excluded, leading to a total of 203 evaluable PET/CT scans. Fifty-six patients were scanned on the Gemini Dual Slice PET/CT scanner and 5 patients on the Gemini TruFlight 16-slice PET/CT scanner. The mean ^{18}F-FDG dose ($\pm$SD) was 371 $\pm$ 25 MBq, and the mean uptake time was 67 $\pm$ 10 min.

With EORTC criteria, 38 patients had PMR, 16 had SMD, and 7 had PMD as their BOmR. The metabolic response rate was 62% (Table 2). With PERCIST, 34 patients had PMR, 20 had SMD, and 7 had PMD as their BOmR. The response rate was 56% (Table 2). None of the patients had CMR. Twenty patients had the background area drawn in the liver, and 43 patients with liver metastases had it drawn in the aorta. There was agreement on BOmR in
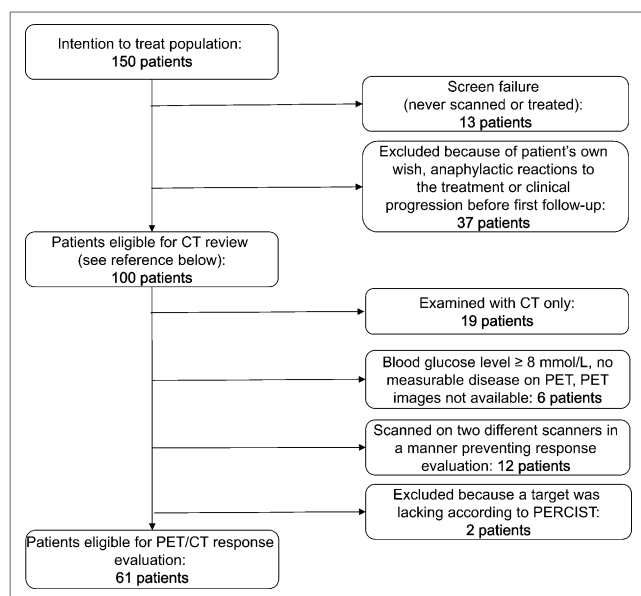
TABLE 1
Characteristics of the 61 Evaluated Patients

Characteristic	Value
Age (y)	
Median	62
Range	36–82
Sex (n)	
Female	24
Male	37
Weight (kg)	
Median	77
Range	49–118
Disease-involved organ systems (n)	
Median	2
Range	1–4
KRAS status	
Wild-type	42 (69%)
Mutated	18 (30%)
Undefined	1 (~1%)
Treatments per patient (n)	
Median	8
Range	4–28
Scans (n)	
Total	230
Median	3
Range	2–8
Scanner (no. of patients)	
Gemini Dual Slice	56
Philips TruFlight 16 Slice	5

87% of the patients, with a corresponding κ-coefficient of 0.76, categorized as good (confidence interval, 0.586–0.900, $P < 0.001$) (Table 2). There was disagreement on BOmR in 13% (8 patients), and the reasons for disagreement are outlined in Table 3. The number of patients with a shift in the hottest lesion during treatment was 28 (46%); the remaining 33 patients (54%) had the same hottest lesion throughout their treatment course; however, 10 of these patients had only 1 lesion (Table 4).

Because of the low number of patients in the PMD group and their diverse length of survival (range, 3.8–39.6 mo), this group was added to the SMD group for the OS plots (Fig. 2). Median OS in months with EORTC criteria was 14.2 in the PMR group, 6.4 in the SMD group, 12.2 in the PMD group, and 7.2 in the combined SMD + PMD group. The difference in median OS between the PMR and the SMD + PMD group was significant, with $P = 0.0001$ (Fig. 2). With PERCIST, the median OS in months was 14.5 in the PMR group, 6.9 in the SMD group, 12.2 in the PMD group, and 7.9 in the combined SMD + PMD group. The difference in median OS between the PMR and the SMD + PMD group was significant, with $P = 0.0008$ (Fig. 2).

DISCUSSION

In lymphomas (27,28) and in several solid cancer types (4,16,29), PET/CT-based response evaluation has been shown to be valuable, especially in visualizing the effect of targeted treatment that induces tumor changes not necessarily followed by tumor shrinkage (14,15,30). Thus, PET/CT is increasingly being tested for response evaluation in clinical cancer trials, and it is therefore important to be familiar with the differences and similarities in the outcome of response evaluation with the existing PET/CT

FIGURE 1. Patient flow diagram of RECIST evaluation interobserver study (39).

TABLE 2
Agreement on BOmR Between EORTC Criteria and PERCIST

Response with EORTC	Response with PERCIST				
	CMR	PMR	SMD	PMD	Total EORTC
CMR	0	0	0	0	0
PMR	0	32	6	0	38
SMD	0	2	14	0	16
PMD	0	0	0	7	7
Total PERCIST	0	34	20	7	61

EORTC criteria and PERCIST agreed on 53 of 61 patients (87%). κ-coefficient was 0.76 and agreement therefore good.

response criteria. In the present study, we compared the results of response evaluation with the 2 currently internationally recognized criteria for PET/CT-based response evaluation—EORTC criteria and PERCIST—in patients with metastatic colorectal cancer.

We found that EORTC criteria (using ΣSUVmax) and PERCIST performed nearly equally in categorizing the patients into the 4 response groups and that agreement of BOmR on an individual-patient level was good. We also found a significant difference in median OS between patients in the PMR group and patients in the SMD + PMD group. To our knowledge, a comparison of the outcome of EORTC criteria and PERCIST response evaluation has not previously been performed and comparison with literature is therefore not possible. However, Monteil et al. (8) found a high PET/CT response rate of 84%, using EORTC criteria, in patients with metastatic colorectal cancer given chemotherapy alone or in combination with bevacizumab in first-, second-, or third-line treatment. Engels et al. (9) found a response rate of 55%, using modified PERCIST (SUVmax instead of SULpeak), in patients with metastatic colorectal cancer confined to the liver and treated with radiotherapy. These findings are in line with the 62% EORTC criteria response rate and the 56% PERCIST response rate we found in the present study although the studies were performed in slightly different settings with other treatments.

A reasonable explanation for the very similar results is that both sets of criteria confine measurements to the most metabolically active part of the patient's tumor burden, which is regarded as the most viable and aggressive disease fragment, determining for disease development according to cancer stem cell theory (19,20).

Additionally, the similar results reflect the robustness of ΔSUV for assessment of treatment response (3,6). Nevertheless, the 2 sets of criteria differ in several aspects. EORTC criteria focus on the most active tumors, in which ROI volumes and sites with high ^{18}F-FDG uptake as well as whole tumor ^{18}F-FDG uptake should be registered at baseline and sampled on all subsequent scans. Recommendations on the number of target lesions to measure or whether to use SUVmax or mean SUV for response calculation are not given. It is, however, stated that SUVmax, mean SUV, and tumor counts should be registered in all tumors at each scan. The cutoff values of 25% for PMR and PMD are based on a literature review (17). No definitions are stated of a background level of ^{18}F-FDG uptake that a viable tumor should exceed in order to qualify as a target lesion. Normalization of SUVs to body surface area is recommended in order to reduce the influence of body weight on SUVs (31,32). Thus, the EORTC criteria can be applied in several different ways generating different outcomes, and the good agreement found is therefore also noteworthy.

In contrast, PERCIST consider the metabolically most active part of the single most ^{18}F-FDG–avid tumor (or up to 5 of the most ^{18}F-FDG–avid tumors) and regard this part as representative of the activity of the cancer (20). Although maximum values are more resistant to partial-volume effect than mean values (18), SULmean has better test–retest variability (8%–10%) than SULmax (11%–12%) (33), is statistically less susceptible to variance, and is therefore recommended (34). The cutoff value of 30% for PMR and PMD is based on the correlation found between a drop in SUV of more than 30%–35% and good outcome (18) and furthermore exceeds the test–retest variability sufficiently. Because of less test–retest variance, a liver background area is recommended over a mediastinal blood pool background area (35), and clear definitions of target lesion ^{18}F-FDG uptake in proportion to background uptake are given. SUV should be normalized to lean body mass to avoid falsely high organ SUVs in obese patients, as fatty tissue has a much lower ^{18}F-FDG uptake than organ tissue, making SUL more consistent between patients with different body weights (31,36).

EORTC criteria and PERCIST disagreed on the BOmR of 8 patients with either PMR or SMD. The discrepancies were explained by the differences in approach (multiple lesions or single lesion) and in response cutoff values. From a clinical perspective, these distribution differences would not have had implications for the involved patients, as treatment would have been continued with both outcomes.

The EORTC- and PERCIST-defined response groups had significantly different median OSs. Although the SMD and PMD groups were joined, this finding supports the clinical applicability

TABLE 3
Reasons for Disagreement on BOmR

No. of patients	BOmR EORTC	BOmR PERCIST	Reason for disagreement
2	SMD	PMD	Reduction in single-lesion SULpeak was greater than reduction in summed SUVmax, body surface area, for multiple lesions
4	PMR	SMD	Reduction in single-lesion SULpeak was less than reduction in summed SUVmax, body surface area, for multiple lesions
2	PMR	SMD	Decrease in SULpeak and in summed SUVmax, body surface area, was between 25% and 30%, resulting in PMR according to EORTC and in SMD according to PERCIST

TABLE 4
Number of Different Hottest Lesions per Patient According to PERCIST Response Evaluation

Number of lesions	Number of patients	Fraction of patients
1	33	54%
2	22	
3	5	46%
4	1	

Of 33 patients with same hottest lesion during treatment, 10 had only 1 tumor.

of PET/CT-based response evaluation. With a higher number of patients, we assume that the median OS of the PMD group would have been more reliable.

In the present study, we chose to use SUVmax for the EORTC calculations instead of mean SUV. This is a limitation, as calculation of results with both values would have been informative and relevant. However, registration and calculation of mean SUV in tumor-specific ROIs of up to 7 lesions per patient per scan would have been a highly time-consuming process and therefore practically challenging. With future development of automatic software algorithms, this might be possible. Yet, the study is strengthened by the relatively high number of eligible patients and the homogeneity of the patients, who all received the same third-line palliative treatment.

The fact that the parameters we choose for the EORTC evaluation (sum of SUVmax of up to 7 lesions per patient) gave response and OS results similar to those from PERCIST evaluation suggests that, for calculation of treatment response, the change in SULpeak from a single lesion is just as representative as the change in the sum of the SUVmax from multiple lesions. Far more aspects of measurement and evaluation are defined in PERCIST than in EORTC criteria, making PERCIST less complicated and easier to apply in daily work as it is much easier to perform

uniform measurements with guidelines that are unequivocal in all aspects of the evaluation procedure. In the EORTC criteria, these parameters are outlined more as guidelines with options rather than clear definitions. This implies that the observer defines what the size of the ROI should be, whether maximum or mean values should be chosen, how many and which target lesions should be registered, whether the SUVs should be summed or response should be calculated per tumor, and whether there should be a minimum SUV limit that a tumor should exceed in order to qualify as a target lesion. The degree of observer-dependent aspects is considerable and modes of application numerous, rendering the EORTC criteria more susceptible to interobserver differences than PERCIST. Possibly, these interobserver-sensitive elements will be clarified in the next revised version of the EORTC criteria.

The EORTC criteria and PERCIST were published 10 y apart, and the research performed in this time span has allowed for a change in approach from extensive toward concise. Nevertheless, both sets of criteria consider the whole patient, although from different angles: the EORTC criteria embrace quantification of both metabolism and size of the tumor burden, whereas the philosophy of PERCIST is that the patient's metabolically most active tumor part defines the whole patient's disease status. Still, to refine the approach, it remains important to explore and test new methods of assessing response with ^{18}F-FDG PET/CT, but it is equally important to, at the same time, use and correctly apply the established criteria in numerous trials in a broad variety of cancer types to gather and formulate the experiences that are needed for further improvement of the criteria (11,13). Efforts are beginning to be made to stratify anticancer treatment, and this task is strongly dependent on distinct treatment response evaluation (37). Further development and stratification of anticancer treatment is dependent on corresponding optimal response evaluation because decisions in patient care and drug development are based on response rates from clinical trials. Without the discipline to use the same criteria internationally, consensus will not be achieved and advances in cancer treatment will be hampered (11,13,38). In our opinion, PERCIST holds the potential to create this consensus and comparability whereas the EORTC criteria, in their current form, can be applied in too many different ways to fulfill this task.

CONCLUSION

We found response evaluation with EORTC criteria and PERCIST to give similar responses and OS outcomes, with agreement on BOmR in 87% of the patients (κ-coefficient, 0.76) and similar significant differences in median OS between the response groups. Yet, validation of the results is needed. We find PERCIST more uncomplicated to apply than EORTC criteria because far more aspects of application are defined in PERCIST. The detailed and unambiguous definitions make response evaluation with PERCIST theoretically less prone to interobserver variability than response evaluation with EORTC criteria and therefore also theoretically more reproducible. Consensus on quantification of PET/CT response evaluation is needed to develop the possibilities of PET/CT to improve anticancer treatment.
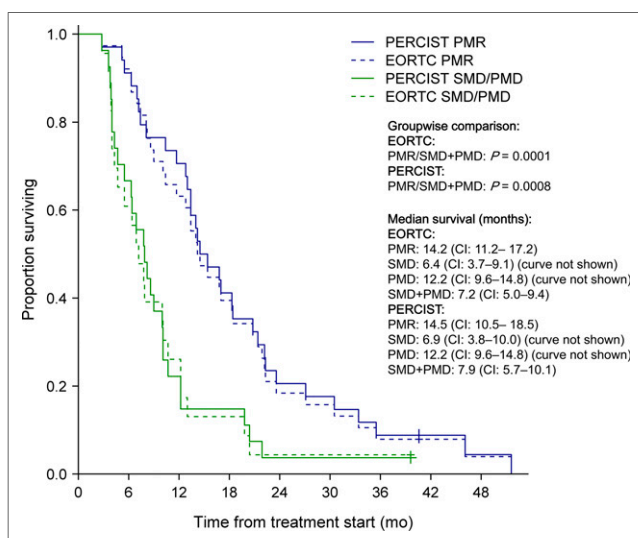
FIGURE 2. OS according to response groups.

Groupwise comparison:
EORTC:
PMR/SMD+PMD: $P = 0.0001$
PERCIST:
PMR/SMD+PMD: $P = 0.0008$

Median survival (months):
EORTC:
PMR: 14.2 (CI: 11.2– 17.2)
SMD: 6.4 (CI: 3.7–9.1) (curve not shown)
PMD: 12.2 (CI: 9.6–14.8) (curve not shown)
SMD+PMD: 7.2 (CI: 5.0–9.4)
PERCIST:
PMR: 14.5 (CI: 10.5– 18.5)
SMD: 6.9 (CI: 3.8–10.0) (curve not shown)
PMD: 12.2 (CI: 9.6–14.8) (curve not shown)
SMD+PMD: 7.9 (CI: 5.7–10.1)

DISCLOSURE

fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734. No potential conflict of interest relevant to this article was reported.

REFERENCES

1. Allen-Auerbach M, Weber WA. Measuring response with FDG-PET: methodological aspects. *Oncologist.* 2009;14:369–377.
2. De Maeseneer DJ, Lambert B, Surmont V, Geboes K, Rottey SW. 18-Fluorodeoxyglucose positron emission tomography as a tool for response prediction in solid tumours. *Acta Clin Belg.* 2010;65:291–299.
3. Herbertson RA, Scarsbrook AF, Lee ST, Tebbutt N, Scott AM. Established, emerging and future roles of PET/CT in the management of colorectal cancer. *Clin Radiol.* 2009;64:225–237.
4. Fletcher JW, Djulbegovic B, Soares HP, et al. Recommendations on the use of ^{18}F-FDG PET in oncology. *J Nucl Med.* 2008;49:480–508.
5. Weber WA, Wieder H. Monitoring chemotherapy and radiotherapy of solid tumors. *Eur J Nucl Med Mol Imaging.* 2006;33(suppl 1):27–37.
6. Figueiras RG, Goh V, Padhani AR, Naveira AB, Caamano AG, Martin CV. The role of functional imaging in colorectal cancer. *AJR.* 2010;195:54–66.
7. de Geus-Oei LF, Vriens D, van Laarhoven HW, van der Graaf WT, Oyen WJ. Monitoring and predicting response to therapy with ^{18}F-FDG PET in colorectal cancer: a systematic review. *J Nucl Med.* 2009;50(suppl 1):43S–54S.
8. Monteil J, Mahmoudi N, Leobon S, et al. Chemotherapy response evaluation in metastatic colorectal cancer with FDG PET/CT and CT scans. *Anticancer Res.* 2009;29:2563–2568.
9. Engels B, Everaert H, Gevaert T, et al. Phase II study of helical tomotherapy for oligometastatic colorectal cancer. *Ann Oncol.* 2011;22:362–368.
10. de Geus-Oei LF, van Laarhoven HW, Visser EP, et al. Chemotherapy response evaluation with FDG-PET in patients with colorectal cancer. *Ann Oncol.* 2008;19:348–352.
11. Weber WA. Assessing tumor response to therapy. *J Nucl Med.* 2009;50 (suppl 1):1S–10S.
12. Buckler AJ, Boellaard R. Standardization of quantitative imaging: the time is right, and ^{18}F-FDG PET/CT is a good place to start. *J Nucl Med.* 2011;52:171–172.
13. Michalski MH, Chen X. Molecular imaging in cancer treatment. *Eur J Nucl Med Mol Imaging.* 2011;38:358–377.
14. Desar IM, van Herpen CM, van Laarhoven HW, Barentsz JO, Oyen WJ, van der Graaf WT. Beyond RECIST: molecular and functional imaging techniques for evaluation of response to targeted therapy. *Cancer Treat Rev.* 2009;35:309–321.
15. Milano A, Perri F, Ciarmiello A, Caponigro F. Targeted-therapy and imaging response: a new paradigm for clinical evaluation? *Rev Recent Clin Trials.* 2011;6:259–265.
16. Van den Abbeele AD. The lessons of GIST–PET and PET/CT: a new paradigm for imaging. *Oncologist.* 2008;13(suppl 2):8–13.
17. Young H, Baum R, Cremerius U, et al. Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. European Organization for Research and Treatment of Cancer (EORTC) PET Study Group. *Eur J Cancer.* 1999;35:1773–1782.
18. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med.* 2009;50(suppl 1):122S–150S.
19. Huang EH, Heidt DG, Li CW, Simeone DM. Cancer stem cells: a new paradigm for understanding tumor progression and therapeutic resistance. *Surgery.* 2007;141:415–419.
20. Al-Hajj M, Becker MW, Wicha M, Weissman I, Clarke MF. Therapeutic implications of cancer stem cells. *Curr Opin Genet Dev.* 2004;14:43–47.
21. Therasse P, Arbuck SG, Eisenhauer EA, et al. 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.
22. Engles JM, Quarless SA, Mambo E, Ishimori T, Cho SY, Wahl RL. Stunning and its effect on ^{3}H-FDG uptake and key gene expression in breast cancer cells undergoing chemotherapy. *J Nucl Med.* 2006;47:603–608.
23. Akhurst T, Kates TJ, Mazumdar M, et al. Recent chemotherapy reduces the sensitivity of [^{18}F]fluorodeoxyglucose positron emission tomography in the detection of colorectal metastases. *J Clin Oncol.* 2005;23:8713–8716.
24. Bjurberg M, Abedinpour P, Brun E, et al. Early metabolic flare in squamous cell carcinoma after chemotherapy is a marker of treatment sensitivity in vitro. *Nucl Med Mol Imaging.* 2010;44:165–169.
25. Boellaard R, O'Doherty MJ, Weber WA, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging.* 2010;37:181–200.
26. Shankar LK, Hoffman JM, Bacharach S, et al. Consensus recommendations for the use of ^{18}F-FDG PET as an indicator of therapeutic response in patients in National Cancer Institute Trials. *J Nucl Med.* 2006;47:1059–1066.
27. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol.* 2007;25:579–586.
28. Cheson BD. Role of functional imaging in the management of lymphoma. *J Clin Oncol.* 2011;29:1844–1854.
29. Lordick F, Ott K, Krause BJ, et al. PET to assess early metabolic response and to guide treatment of adenocarcinoma of the oesophagogastric junction: the MUNICON phase II trial. *Lancet Oncol.* 2007;8:797–805.
30. Contractor KB, Aboagye EO. Monitoring predominantly cytostatic treatment response with ^{18}F-FDG PET. *J Nucl Med.* 2009;50(suppl 1):97S–105S.
31. Zasadny KR, Wahl RL. Standardized uptake values of normal tissues at PET with 2-[fluorine-18]-fluoro-2-deoxy-D-glucose: variations with body weight and a method for correction. *Radiology.* 1993;189:847–850.
32. Kim CK, Gupta NC, Chandramouli B, Alavi A. Standardized uptake values of FDG: body surface area correction is preferable to body weight correction. *J Nucl Med.* 1994;35:164–167.
33. Nahmias C, Wahl LM. Reproducibility of standardized uptake value measurements determined by ^{18}F-FDG PET in malignant tumors. *J Nucl Med.* 2008;49:1804–1808.
34. Krak NC, Van der Hoeven JJ, Hoekstra OS, Twisk JW, van der Wall E, Lammertsma AA. Measuring [^{18}F]FDG uptake in breast cancer during chemotherapy: comparison of analytical methods. *Eur J Nucl Med Mol Imaging.* 2003;30:674–681.
35. Paquet N, Albert A, Foidart J, Hustinx R. Within-patient variability of ^{18}F-FDG: standardized uptake values in normal tissues. *J Nucl Med.* 2004;45:784–788.
36. Sugawara Y, Zasadny KR, Neuhoff AW, Wahl RL. Reevaluation of the standardized uptake value for FDG: variations with body weight and methods for correction. *Radiology.* 1999;213:521–525.
37. Shankar LK, Van den Abbeele A, Yap J, Benjamin R, Scheutze S, Fitzgerald TJ. Considerations for the use of imaging tools for phase II treatment trials in oncology. *Clin Cancer Res.* 2009;15:1891–1897.
38. Herrmann K, Benz MR, Krause BJ, Pomykala KL, Buck AK, Czernin J. ^{18}F-FDG-PET/CT in evaluating response to therapy in solid tumors: where we are and where we can go. *Q J Nucl Med Mol Imaging.* 2011;55:620–632.
39. Skougaard K, McCullagh MJ, Nielsen D, Hendel HW, Jensen BV, Johannesen HH. Observer variability in a phase II trial: assessing consistency in RECIST application. *Acta Oncol.* 2012;51:774–780.