


# Comparison of FDG-PET/CT and contrast-enhanced CT for monitoring therapy response in patients with metastatic breast cancer

Christopher C. Riedl<sup>1,2</sup>  · Katja Pinker<sup>1,3</sup> · Gary A. Ulaner<sup>1,2</sup> · Leonard T. Ong<sup>1</sup> · Pascal Baltzer<sup>3</sup> · Maxine S. Jochelson<sup>1,2</sup> · Heather L. McArthur<sup>4</sup> · Mithat Gönen<sup>5</sup> · Maura Dickler<sup>6</sup> · Wolfgang A. Weber<sup>1,2</sup>

Received: 19 December 2016 / Accepted: 9 April 2017 / Published online: 1 May 2017  
© Springer-Verlag Berlin Heidelberg 2017

## Abstract

**Purpose** The aim of this study was to compare fluorodeoxyglucose positron emission tomography/computed tomography (FDG PET/CT) and contrast-enhanced computed tomography (CE-CT) for the prediction of progression-free survival (PFS) and disease-specific survival (DSS) in patients with stage IV breast cancer undergoing systemic therapy.

**Methods** Sixty-five patients with metastatic breast cancer treated with first- or second-line systemic therapy in prospective clinical trials were included. Response to treatment was evaluated by Response Evaluation Criteria in Solid Tumors

(RECIST) 1.1 for CE-CT and by PET Response Criteria in Solid Tumors (PERCIST), respectively.

**Results** All responders by RECIST ( $n = 22$ ) were also responders by PERCIST, but 40% (17/43) of non-responders by RECIST were responders by PERCIST. Responses according to RECIST and PERCIST both correlated with PFS, but PERCIST showed a significantly higher predictive accuracy (concordance index for PFS: 0.70 vs. 0.60). One-year PFS for responders vs. non-responders by RECIST was 59% vs. 27%, compared to 63% vs. 0% by PERCIST. Four-year DSS of responders and non-responders by RECIST was 50% and 38%, respectively ( $p = 0.2$ , concordance index: 0.55) as compared to 58% vs. 18% for PERCIST ( $p < 0.001$ , concordance index: 0.65). Response on PET/CT was also a significantly better predictor for DSS than disease control on CE-CT.

**Conclusions** In patients with metastatic breast cancer, tumor response on PET/CT appears to be a superior predictor of PFS and DSS than response on CE-CT. Monitoring tumor response by PET/CT may increase the power of clinical trials using tumor response as an endpoint, and may improve patient management in clinical routine.

This work has been previously presented at:

- ASCO, May 29, 2015 (abs. 1051)
- SNMMI, June 6, 2015 (abs. 1320)

**Electronic supplementary material** The online version of this article (doi:10.1007/s00259-017-3703-7) contains supplementary material, which is available to authorized users.

✉ Christopher C. Riedl  
riedlc@mskcc.org

<sup>1</sup> Department of Radiology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

<sup>2</sup> Department of Radiology, Weill Cornell Medical College, New York, NY 10065, USA

<sup>3</sup> Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna, 1090 Vienna, Austria

<sup>4</sup> Department of Medicine, Breast Oncology, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

<sup>5</sup> Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

<sup>6</sup> Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

**Keywords** FDG · PET/CT · Breast cancer · Treatment response · Survival

## Introduction

Metastatic breast cancer is an incurable disease that can be palliated by systemic therapy. Imaging plays a pivotal role in determining whether to continue, change, or stop treatment. In drug development, tumor response and progression-free survival (PFS) are increasingly used as endpoints of clinical

studies, because the impact of a drug on overall survival can be confounded by the effects of second- and third-line therapies.

Change in tumor size based on anatomic imaging is the current standard for monitoring tumor response and progression in breast cancer. Size-based response criteria are commonly used as endpoints of clinical trials [1]. Guidelines to standardize assessment of changes in tumor size have been developed and continuously refined over the last 30 years [2–4]. Nonetheless, some inherent limitations cannot be overcome when size criteria are used. Distinguishing viable from nonviable residual tumor tissue is often difficult and osseous metastases are, in general, non-measurable. Meta-analyses of clinical trials have shown only weak correlations between tumor response and overall survival in breast cancer and other malignancies [5].

<sup>18</sup>F-Fluorodeoxyglucose (FDG) positron emission tomography/computed tomography (PET/CT) is an imaging modality that combines morphologic and functional imaging. Measurements of glucose metabolism by FDG PET/CT have a high degree of repeatability [6] and show less inter-observer variability than measurements of tumor size [7]. PET/CT can differentiate an active tumor from post-therapeutic changes and assess metabolic activity in osseous metastases [8]. Changes in FDG uptake have demonstrated a high accuracy for predicting histopathologic response of breast cancer in the neoadjuvant setting [9–11] and appear superior when compared to size measurements alone [12]. Tumor response on PET/CT has also shown promise for assessing treatment response for metastatic breast cancer in the palliative setting [13].

A key limitation of most of these studies has been the lack of a standardized approach for assessing tumor response on PET/CT. PET response evaluation criteria in solid tumors (PERCIST) address this limitation and are increasingly used to monitor tumor response [14]. However, very few studies have compared response assessment by PERCIST with the current standard Response Evaluation Criteria in Solid Tumors (RECIST) in breast cancer or other malignancies. Specifically, it is unknown whether response assessment by PET/CT is better correlated with progression-free survival (PFS) and disease-specific survival (DSS) than response assessment by contrast-enhanced computed tomography (CE-CT).

Thus, the goal of this study was to compare CE-CT to PET/CT for prediction of PFS and DSS in patients with stage IV breast cancer undergoing systemic therapy.

## Materials and methods

This retrospective study was approved by the institutional review board and informed consent was waived.

## Patient selection

Darwin, a database and web-based analytics platform at Memorial Sloan Kettering Cancer Center (MSK), was used to search for patients with metastatic breast cancer who received first- or second-line chemotherapy, targeted therapy, and/or hormone therapy as part of a clinical trial from 2007 to 2012. Patients were included in this study if a) baseline CE-CT and FDG PET/CT were acquired within 28 days prior to initiation of therapy and within 90 days after initiation of therapy; and b) if disease was present at baseline as defined by RECIST 1.1 [4] or PERCIST [14]. Patients' medical records were reviewed and the following were documented: age at treatment initiation, start date and type of systemic treatment, histologic tumor type, tumor receptor status, and date and cause of death (whether disease-specific or not) or date of last documented visit.

## Imaging techniques

### CE-CT

Multi-slice CE-CTs were acquired using a variety of 16- or 64-detector row CT scanners from the thoracic inlet to the pelvic floor with 120-kVp tube voltage and 120–200-mA tube current time product after intravenous injection of 150 mL of non-ionic contrast agent with a time delay typical of portal venous phase imaging.

All studies included standard transverse and reformatted coronal and sagittal images and were interpreted on picture archiving and communication system (PACS; Centricity, GE Healthcare) workstations.

### FDG PET/CT

Prior to radiotracer <sup>18</sup>F-FDG injection, patients fasted for at least 4 h. If plasma glucose was less than 200 mg/dL, patients were injected with 12–15 mCi (444–555 MBq) of <sup>18</sup>F-FDG intravenously. After a 60- to 90-min uptake period, scans were acquired from the skull base to the mid-thigh using a variety of PET/CT scanners, followed by a CT scan for attenuation correction. Analysis of the PET/CT images was performed with PET-VCAR AW Suite 2.2, Advantage Workstation, GE Healthcare.

## Response evaluation

For both modalities, responses were categorized into four groups: complete response (CR), partial response (PR), stable disease (SD), or progressive disease (PD).

CE-CT response was determined by one board-certified radiologist (KPD) according to the RECIST 1.1 criteria as summarized in Table 1 [4]. The reader was aware of the

**Table 1** Criteria for evaluation of response by CE-CT vs. PET/CT

Response category	Response evaluation criteria in solid tumors (RECIST 1.1) <sup>a</sup> [4]	PET response criteria in solid tumors (PERCIST) <sup>b</sup> [14]
Complete response (CR)	Disappearance of all target lesions	Disappearance of all lesions on PET images to background blood-pool levels, regardless of % change from baseline and anatomic size
Partial response (PR)	Reduction in short axis of target lymph nodes to <10 mm ≥ 30% decrease in sum of target lesion diameter sum	≥ 30% decrease in sum of target lesion SUL and 1 SUL unit absolute change
Progressive disease (PD)	≥ 20% increase in sum of target lesion diameter and minimum of 5 mm total increase or new lesion or unequivocal progression of non-target lesions <sup>c</sup>	30% increase in sum of target lesion SUL and 1 SUL unit absolute change or new FDG avid lesion or unequivocal progression of non-target lesion (e.g., ≥ 30% increase) or unequivocal progression by RECIST.
Stable disease (SD)	Does not meet other criteria	Does not meet other criteria

SUL Standardized uptake values normalized to lean body mass

<sup>a</sup> Measurements are based on the sum of the unidimensional measurement of the greatest diameter of a maximum of five lesions. The longest diameter with a minimum size of 1.0 cm was measured for a maximum of five target lesions (maximum of two lesions per organ). Lymph nodes were considered target lesions if ≥15 mm in short axis and only the short axis contributed to the baseline sum. Lesions <10 mm or pathological lymph nodes ≥10 to <15 mm short axis were considered non-measurable lesions. Other non-measurable lesions included leptomeningeal disease, ascites, pleural or pericardial effusion, inflammatory breast cancer, lymphangitic involvement of skin or lung, and lesions in a previously irradiated area. Bone lesions were not considered measurable unless a lytic lesion with an identifiable soft tissue component that met the aforementioned measurability requirements was present

<sup>b</sup> The sum of all target lesions on baseline and follow-up scan was calculated and the percent change of the sum was used to determine response as defined by PERCIST criteria. Up to five target lesions (maximum of two lesions per organ) were determined based on the highest metabolic activity. An SULpeak below  $1.5 \times \text{liver SUL} + 2$  standard deviations of liver SUL was required to qualify as a target lesion. All lesions other than the target lesion were considered non-target lesions without an SULpeak cut-off. If patients only had lesions below the SULpeak cut-off, they could still be categorized as CR, PR, and PD if lesions completely disappeared, all lesions dropped by ≥30%, or at least one lesion increased by ≥30%. If the change was less than 30%, treatment response was assessed by RECIST 1.1

<sup>c</sup> Defined as a substantial increase in overall disease burden even in the setting of SD or PR of target lesions. If the scan only showed non-measurable lesions, the criteria for progression would be defined as a substantial increase in disease burden

clinical diagnosis but blinded to the results of the FDG PET/CT and the clinical follow-up information.

FDG PET/CT response was determined by a different physician certified in both radiology and nuclear medicine (CCR). The reader was aware of the clinical diagnosis but blinded to clinical follow-up information. For all target lesions identified by the reader, peak standardized uptake values normalized to lean body mass (SULpeak) were determined [14] (PET-VCAR AW Suite 2.2, Advantage Workstation, GE Healthcare). In a separate analysis, we evaluated response classifications based on the clinically more commonly used maximum standardized uptake values (SUVmax) and compared these to response classifications based on SULpeak. Selection of lesions and response categorization was based on PERCIST as summarized in Table 1 [14].

For assessment of PFS and DSS, patients underwent clinical follow-up and CE-CT scans at 3-month intervals until progression, followed by routine follow-up until death. At the discretion of the treating physician, some patients were also followed with PET/CT scans.

## Statistics

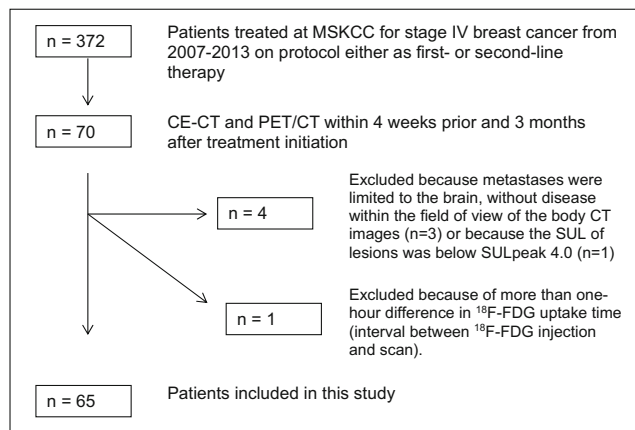
Statistical analysis was performed with Prism 6 (GraphPad, CA, USA) and STATA SE 12.1 (StataCorp, USA). The Pearson correlation coefficient between changes in SUVmax and changes in SULpeak was calculated and differences in response categories were assessed. Concordance between FDG PET/CT- and CE-CT-based response assessment was calculated using kappa statistics. To compare CE-CT and FDG PET/CT for prediction of outcome, patients were dichotomized as follows: In the first analysis, patients with CR and PR were considered responders and patients with SD and PD were considered non-responders. In clinical practice, however, therapy is typically continued as long as the patient shows no progression on anatomic imaging. Furthermore, clinical trials have also used disease control rate (patients with CR, PR, and SD) as an endpoint instead of or in addition to the objective response rate. Therefore, disease control rate on CE-CT was also determined and compared with response on FDG PET/CT. Association of response to treatment with PFS and DSS was assessed using univariate and multivariate Cox hazard regression models. The predictive accuracy of all Cox regression models was assessed and compared by calculating Gonen and Heller's K concordance statistic [15]. The resulting index is a measure of the Cox regression model's discriminative power to predict survival. Its values range between 1, implying perfect prediction, and 0.5, which indicates random guessing. Kaplan–Meier analysis was performed to compare responders with non-responders for each modality as well as disease control for CE-CT vs. response by PET/CT. Additionally, Kaplan–Meier analysis was performed to compare responders with non-responders by both modalities and patients with divergent results. The log-rank (Mantel–Cox)

test was used to evaluate the differences between Kaplan–Meier curves. *P* values <0.05 were considered significant.

**Results**

Sixty-five patients aged 29–85 years (mean age 54 years) were eligible for this study (Fig. 1). Of these 65, 2 patients were enrolled in 2007, 0 in 2008, 2 in 2009, and the remaining 61 from 2010 to 2012. Details about treatment, imaging times, patient follow-up, and cancer baseline characteristics are summarized in Table S1 and Table 2.

Changes in SULpeak and SUVmax were very closely correlated (*r* = 0.998). Consequently, the response classification (CR, PR, SD, or PD) was the same when SUVmax was used instead of SULpeak. Response categorizations as assessed by CE-CT and FDG PET/CT are summarized in Table 3. Overall, there was only fair to moderate agreement between response assessment by CE-CT and FDG PET/CT (kappa = 0.36, weighted kappa = 0.45). However, all 22 patients categorized as CR or PR by CE-CT were also categorized as CR or PR by PET/CT. In contrast, marked differences in response assessment were observed for the 43 patients categorized as SD or PD by CE-CT. Seventeen (40%) of these patients were classified as PR or CR by PET/CT (Table 3, Fig. 2). Of the 15 patients classified as PD by CE-CT, 3 were classified as CR by FDG PET/CT, 1 as SD, and 11 (73%) concordantly as PD. All three patients with discordant CR on PET/CT had new or progressive bone lesions on CE-CT. The 28 patients categorized as SD by CE-CT were classified by FDG PET/CT as CR in 6 cases, as PR in 8 cases, as PD in 6 cases (Fig. 3) and, concordantly, as SD in only 8 cases (29%). Of the 24 patients without osseous involvement, 9 (38%) had SD on CT and 2 (8%) had SD on PET/CT. Only two of the nine patients without osseous involvement and SD by CE-CT were also classified as SD by FDG PET/CT (one was categorized as CR, two were categorized as PR and four as PD; see Table 4 for details).



**Fig. 1** Patient cohort

**Table 2** Imaging intervals and patient follow-up

	Mean	Range
Interval from treatment start to follow-up imaging <sup>a</sup>	55 D	10–90 D
Follow-up time (all patients)	28.0 Mo	4.0–59.3 Mo
Progression-free survival (all patients)	14.0 Mo	0.9–59.3 Mo
Progression-free survival of patients that progressed during the follow-up period (n = 50, 77% of population)	7.1 Mo	0.9–34.7 Mo
Time to disease-specific death (n = 34, 52% of total population)	19.0 Mo	4.4–43.1 Mo

Deaths other than breast cancer, n = 3

Patients actively withdrawing from the study, n = 3

Patients alive at the time of database closure, n = 23

<sup>a</sup> Times given for FDG PET/CTs. CE-CTs were usually performed on the same day or in close proximity to the PET/CT

Differences in response assessment between CE-CT and FDG PET/CT were most commonly observed in patients with osseous metastases (Table 4). However, PET/CT showed substantially fewer cases with SD in all subgroups of histology, receptor status, metastatic pattern, and treatment type (Table 4). This included 23 patients with only soft tissue metastases (9 cases of SD by CE-CT vs. 2 with PET/CT) and 7 patients with invasive lobular cancers (5 cases of SD by CE-CT vs. 1 with PET/CT).

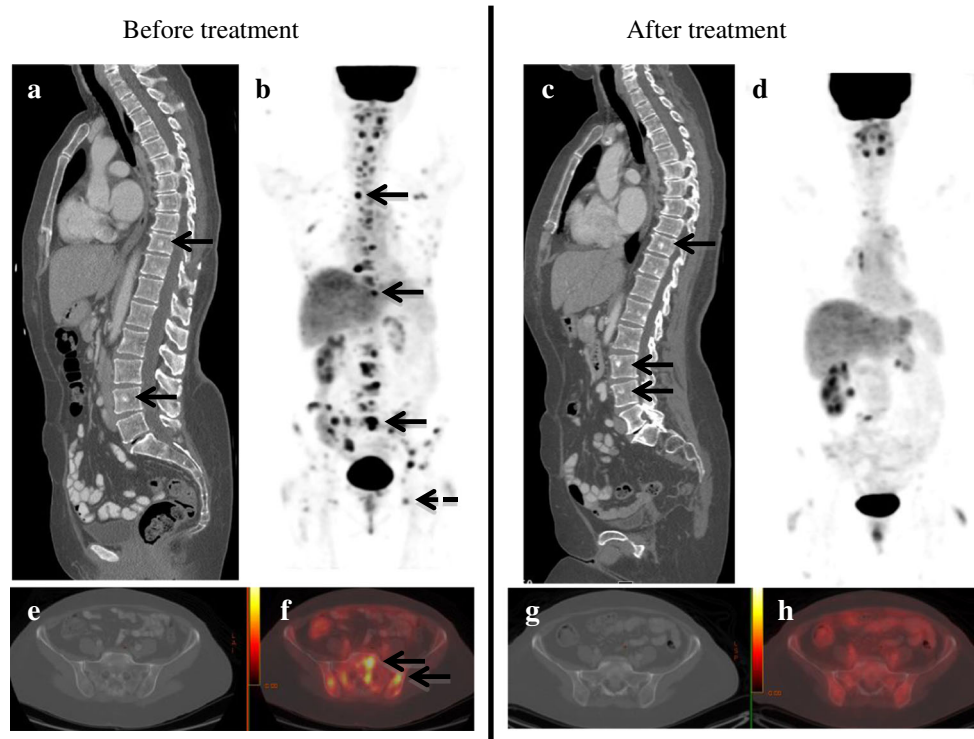
The agreement between response assessment by CE-CT and FDG PET/CT remained moderate when patients were dichotomized as responders and non-responders (kappa = 0.51). Agreement did not improve when disease control on CE-CT was compared with response on FDG PET/CT (kappa = 0.41).

Kaplan–Meier curves for PFS and DSS for FDG PET/CT and CE-CT for responders and non-responders are shown in Fig. 4A and B. One-year PFS for responders vs. non-responders by CE-CT was 59.1% vs. 27.2% (*p* = 0.1954) compared to 63.6% vs. 0% (*p* = 0.0001) by PET/CT. Four-year DSS for responders vs. non-responders by CE-CT was 49.6% vs. 37.5% (*p* = 0.0030) compared to 58.2% vs. 18.0%

**Table 3** Response categorizations as assessed by CE-CT and FDG PET/CT

CE-CT	FDG PET/CT				TOTAL
	CR	PR	SD	PD	
CR	3	0	0	0	3
PR	10	9	0	0	19
SD	6	8	8	6	28
PD	3	0	1	11	15
TOTAL	22	18	8	17	65

CR complete response, PR partial response, SD stable disease, PD progressive disease



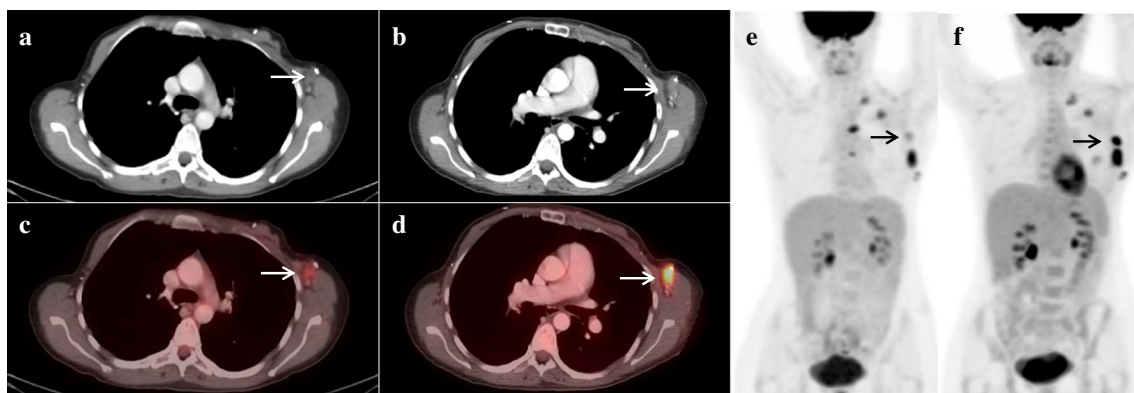
**Fig. 2** Eighty-six-year-old woman with metastatic breast cancer before and after first-line endocrine therapy combined with bevacizumab showing stable disease by CE-CT and complete response by PET/CT. CT images demonstrate several sclerotic foci in the spine, representing biopsy-proven osseous metastases (arrows in **a**), with no discernible change after treatment (**c**). Maximum-intensity projection PET/CT images show multiple FDG-avid metastases including the spine (short

arrows), pelvis (dashed arrow), and ribs (dotted arrow) (**b**), with resolution after treatment initiation (**d**). Axial CT image of the pelvis in the same patient showing no metastases before (**e**) and after therapy (**g**). Corresponding axial PET/CT image showing multiple metastases before (**f**) and resolution of uptake after treatment (**h**). The patient is still alive with a follow-up of 47 months

( $p = 0.0001$ ) by PET/CT. Kaplan–Meier curves for PFS and DSS for responders and non-responders by FDG PET/CT vs. disease control by CE-CT are shown in Fig. 4C and D. When disease control rate on CE-CT was analyzed, 1-year PFS for patients with and without disease control on CE-CT was

43.4% vs. 20.0% ( $p = 0.0118$ ). Four-year DSS for patients with and without disease control on CE-CT was 47.0% vs. 26.7% ( $p = 0.0035$ ).

Univariate Cox proportional hazards regression demonstrated that response on FDG PET/CT and CE-CT were both



**Fig. 3** Thirty-eight-year-old woman with metastatic breast cancer 10 days before (**a**, **c**, and **e**) and 80 days after (**b**, **d**, and **f**) initiation of desatinib and paclitaxel as first-line combination therapy showing stable disease by contrast-enhanced CT and progression by FDG PET/CT. Axial CT images before (**a**) and after (**b**) treatment initiation show stable axillary nodal metastases (arrow). Axial FDG PET/CT before (**c**) and after (**d**) treatment initiation show a marked increase in

FDG avidity from an SULpeak of 2.1 to an SULpeak of 6.5. Mixed changes are seen on maximum-intensity projection PET images before (**e**) and after (**f**) treatment initiation. All nodes were stable on CT. The patient stayed on protocol as she did not meet criteria of progression by RECIST 1.1. Two months later, a new left axillary lymph node became palpable and the patient was switched to another treatment regimen

**Table 4** Disease baseline characteristics and types of treatment with subgroup analysis of the number of discrepant cases and number of cases with SD

			Total		Discrepant CE-CT vs. PET/CT results				Cases with SD			
					CR/PR/SD/PD		R vs. NR		CE-CT		PET/CT	
			n	% <sup>a</sup>	n	% <sup>b</sup>	n	% <sup>b</sup>	n	% <sup>b</sup>	n	% <sup>b</sup>
Patients			65	100	35	54	17	26	28	43	8	12
Histology	Invasive ductal		54	83	23	43	10	19	20	37	6	11
	Mixed ductal and lobular		4	6	3	75	0	0	3	75	1	25
	Invasive lobular		7	11	5	71	4	57	5	71	1	14
Receptor status	ER + Her2+		10	15	5	50	3	30	2	20	1	10
	ER + Her2-		39	60	23	59	10	26	21	54	6	15
	ER-Her2+		5	8	2	40	1	20	1	20	0	0
	ER-Her2-		11	17	4	36	2	18	4	36	1	9
Metastatic pattern	Osseous + other metastases		30	46	14	47	5	17	10	33	5	17
	Only osseous metastases		12	18	10	83	8	67	9	75	1	8
	Only other than osseous metastases		23	35	12	52	5	22	9	39	2	9
Type of treatment protocol	Cytotoxic		15	23	8	53	4	27	9	60	3	20
	Immunotherapy + cytotoxic		19	29	10	53	3	16	4	21	1	5
	Immunotherapy		14	22	5	36	4	29	7	50	3	21
	Immunotherapy + hormone therapy		13	20	10	77	6	46	7	54	1	8
	Hormone therapy (anti-androgen)		4	6	1	25	0	0	1	25	0	0

CR complete response, PR partial response, SD stable disease, PD progressive disease, R response, NR non-response, ER + Her2+ estrogen receptor and Her2 receptor-positive, ER + Her2- estrogen receptor-positive and Her2 receptor-negative, ER + Her2- estrogen receptor-negative and Her2 receptor-positive, ER-Her2- estrogen receptor-negative and Her2 receptor-negative

<sup>a</sup> percent of total number of patients

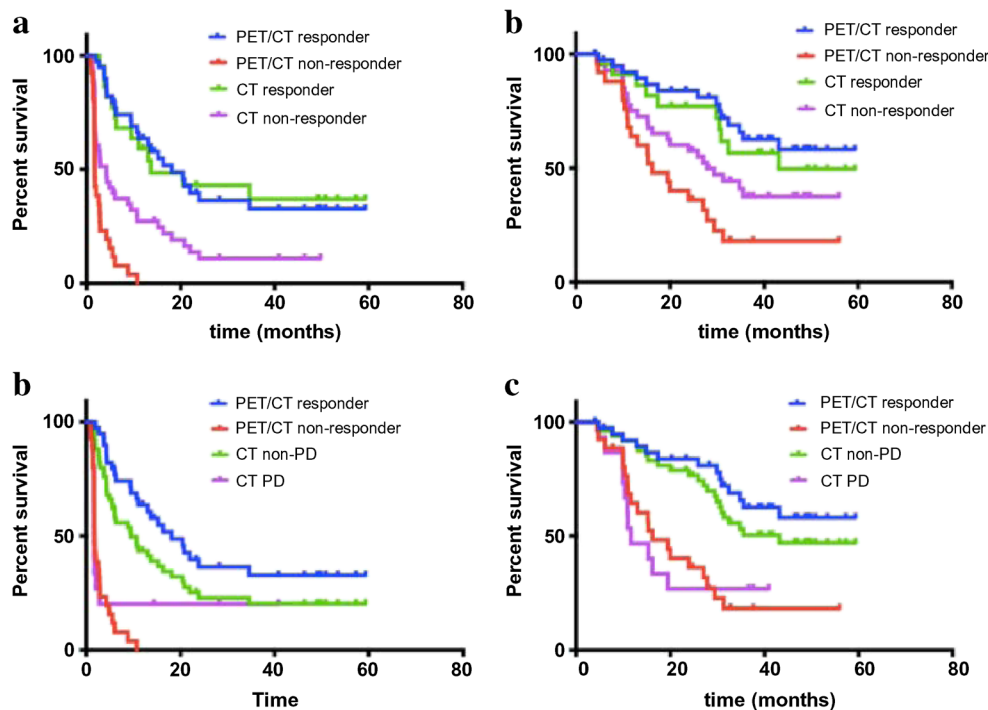
<sup>b</sup> percent of total number in subgroup

significantly associated with PFS. However, the HR for a non-response on FDG PET/CT was four times higher than CE-CT (Table 5). The multivariate model retained only response on PET/CT as a significant and independent predictor of PFS. As a result, the C index of the multivariate model equaled that of the univariate FDG PET/CT model. No incremental value of CE-CT to FDG PET/CT was identified (Table 5).

Univariate Cox proportional hazards regression for prediction of DSS showed a significant association with disease control rate on CT and response on FDG PET/CT, whereas response on CE-CT was not significantly correlated with DSS. Both hazard ratios and predictive accuracy as estimated by the C index were higher using FDG PET/CT response as compared to disease control rate on CE-CT. In a multivariate predictive model, only FDG PET/CT response was an independent predictor of DSS (Table 6).

## Discussion

This first systematic comparison of RECIST and PERCIST for monitoring tumor response in patients with breast cancer indicates that metabolic responses on FDG PET/CT are frequently discordant from morphologic responses on CE-CT. While all patients classified as responders by RECIST were also classified as such by PERCIST, patients with SD or PD by RECIST were frequently classified differently by PERCIST. Because of the differences in these patient groups, response by PERCIST was a significantly stronger predictor for PFS and DSS than response or disease control by RECIST. Our data, therefore, suggest that treatment monitoring with FDG PET/CT could benefit patients with breast cancer in two ways: (i) avoid taking patients off treatment who are considered to have PD by CE-CT but who respond by PET/CT;



**Fig. 4** Progression-free survival (PFS, **a**) and disease-specific survival (DSS, **b**) by tumor response on PET/CT and CT. One-year PFS for responders vs. non-responders by CE-CT was 59.1% vs. 27.2% ( $p = 0.0030$ ), compared to 63.6% vs. 0% ( $p < 0.0001$ ) by PET/CT, respectively. Four-year DSS for responders vs. non-responders by CE-CT was 50% vs. 38% ( $p = 0.2$ ) compared to 58% vs. 18% ( $p < 0.0001$ ) by

PET/CT, respectively. When disease control rate on CE-CT was analyzed and compared with the numbers for response by PET/CT as mentioned above, 1-year PFS (**c**) for patients with and without disease control on CE-CT was 43.4% vs. 20.0%, respectively ( $p = 0.0118$ ). Four-year DSS (**d**) for patients with and without disease control on CE-CT was 47.0% vs. 26.7%, respectively ( $p = 0.0035$ )

and (ii) change therapy in patients considered stable on CE-CT but show no response on FDG PET/CT. According to our data (Table 3), this would alter overall patient management in 25% of the patients.

The marked differences in response classification by FDG PET/CT and CE-CT were often due to changes in the metabolic activity and number of osseous metastases (Fig. 2). Bone metastases can exhibit paradoxical behavior when responding to treatment with an increase in size and number of metastases, or the appearance of new lesions, because CT does not evaluate the bone marrow, but only the osteoblastic reaction in healing bone [16]. For this reason, RECIST 1.1 criteria specify that bone lesions without soft tissue components are non-measurable [4]. This limitation is problematic in patients with breast cancer because bone is often the dominant site of

metastatic disease [17]. In contrast, FDG avidity reflects tumor viability, and can differentiate between tumor progression and healing bone [18, 19]. Additionally, FDG PET/CT is more sensitive than CT for the detection of osseous metastases [20]. Therefore, disease progression is detected earlier by FDG PET/CT than by CE-CT. While this principle advantage of metabolic imaging is well-known, to our knowledge, only a very limited number of studies have investigated the impact of better assessment of osseous disease on prediction of patient outcome [21].

To overcome the low sensitivity of CE-CT for osseous metastases, many clinical trials require CE-CT plus bone scan to evaluate response. However, the osteoblastic reaction of healing bone is well-known to initially increase radiotracer uptake on bone scans, which leads to false-positive findings,

**Table 5** Cox proportional hazards regression results for PFS

Variable	HR (95% CI)	<i>p</i> value	Gonen and Heller's K
Univariate*			
CE-CT response	2.54 (1.34–4.80)	0.004	0.599
PET/CT response	10.39 (5.06–21.34)	<0.0001	0.701
Multivariate			
CE-CT response	1.04 (0.46–2.33)	0.927	
PET/CT response	10.18 (4.37–23.69)	<0.0001	0.703

**Table 6** Cox proportional hazards regression results for DSS

Variable	HR (95% CI)	<i>p</i> value	Gonen and Heller's K
Univariate			
CE-CT response	1.62 (0.77–3.40)	0.200	0.554
CE-CT disease control	2.86 (1.37–5.97)	0.005	0.587
PET/CT response	4.16 (2.04–8.48)	<0.001	0.649
Multivariate <sup>a</sup>			
CE-CT disease control	1.95 (0.90–4.22)	0.092	
PT/CT response	3.58 (1.71–7.54)	0.001	0.675

<sup>a</sup> Between response and disease control by CE-CT, only the one with the higher hazard ratio was used for the multivariate analysis since both were derived from the same interpretation

and there is no established quantitative analysis of bone scans in breast cancer patients. This renders bone scans less valuable than FDG PET/CT for assessment of tumor response [22–24].

In clinical trials and in clinical practice, therapy is typically continued in patients with controlled disease on CT, which includes patients with CR, PR, or SD. However, a limitation of this approach is that patients with slow-growing tumors may exhibit SD in the absence of a drug effect. In placebo-controlled trials, the SD rate in the placebo arm can be as high as 30%, even for aggressive malignancies such as metastatic non-small cell lung cancer [25]. Consequently, the beneficial effect of a drug can be markedly overestimated if SD is considered as a drug effect. In contrast, PR and CR by RECIST or PERCIST are extremely rare in the absence of active treatment. Therefore, it may be preferable to define the effectiveness of a therapy by response rather than by disease stabilization. This could potentially avoid an overestimation of the effectiveness of a drug in a clinical trial and reduce the side-effects and costs of ineffective therapies in clinical practice.

These theoretical considerations are supported by the findings of the present study, where response on PET/CT was also a better predictor for patient outcome than disease control on CE-CT. The disease control rate (CR + PR + SD) was 77% (50/65) in the current study (Table 3), whereas the response rate on PET/CT was only 62% (40/65). This difference occurred because 46% (13/28) of patients showing SD by CE-CT were non-responders according to FDG PET/CT. DSS of PERCIST non-responders was very similar to patients with RECIST PD, with even a tendency for a worse outcome for the non-responders on FDG PET/CT (Fig. 4D). This suggests that FDG PET/CT correctly identified a significant fraction of patients who were not benefiting from the selected therapies despite showing SD by RECIST (i.e., 50% of the 28 patients showing SD by CE-CT and 20% of the total patient population). There was also a smaller group of patients that progressed by CE-CT, but were responders on PET/CT. These three patients demonstrated new osseous lesions on CT, but showed an excellent clinical outcome.

In patients with metastatic breast cancer, effective second- and third-line therapies can dilute the prognostic value of the

initial response to therapy for OS and DSS. Therefore, it is encouraging that response by PERCIST was significantly correlated with DSS despite these confounding factors.

While these observations are encouraging for the clinical use of FDG PET/CT, the following limitations should be considered. The patients in our study received therapy as part of multiple protocols including cytotoxic, hormone, and targeted therapies, as well as a combination of these therapies. Furthermore, various breast cancer subtypes were included in the study. While this heterogeneity may be considered a limitation, we would argue that the strong prognostic value of tumor response on FDG PET/CT in this heterogeneous patient population supports the robustness of response assessment by PERCIST. This is supported by the data in Table 4, which show no obvious differences in the prognostic value of FDG PET/CT in various patient subgroups. While theoretically superior, it seems impractical to define different response criteria for specific therapies and different histological subtypes of breast cancer. For the same reason, RECIST is used (with very few exceptions) for all solid tumors and all therapies. We cannot make an argument for or against the value of CE-CT and PET/CT in case of hormone therapy only, since the number of cases with such treatments was too small ( $n = 4$ ).

Further limitations of this study include the retrospective analysis and the relatively small number of patients. Thus, the reported findings need to be confirmed in a prospective, multicenter trial. However, we believe that our study provides reliable estimates of the prognostic value of FDG PET/CT to design and power such confirmatory trials, because only protocol patients were included in our analysis and, hence, standardized patient follow-up was performed and all outcome data collected prospectively.

Tumor FDG uptake can be measured in various ways, including SULpeak and SUVmax, raising the question of which parameter is most suitable for response assessment. We have used SUVmax in a preliminary analysis of our data, which showed very similar differences in PFS and DSS between responders and non-responders [26]. In the present study, the response classification was unchanged when SUVmax was



used instead of SULpeak for response assessment. This suggests that response assessment by PERCIST is not significantly affected by differences in SUV quantification, which is encouraging for its broader clinical use.

In conclusion, our study suggests that in patients with metastatic breast cancer, response assessment by FDG PET/CT may be superior to CE-CT, because tumor response on FDG PET/CT correlated significantly better with PFS and DSS than tumor response or tumor control on CE-CT. The differences in response assessment by FDG PET/CT and CE-CT could cause changes in patient management in 25% of the patients. Thus, monitoring tumor response to therapy with FDG PET/CT may not only improve the quality of clinical trials using response as an endpoint, but may also reduce the morbidity and costs of ineffective therapies in clinical practice.

#### Compliance with ethical standards

**Funding** This study was funded by a Susan G. Komen for the Cure research grant (KG110441) and MSKCC Biostatistics Core (P30 CA008748).

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** This HIPAA-compliant, retrospective, single-institution study was performed under institutional review board approval.

Informed consent was waived due to the retrospective nature of the study.

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

#### References

1. United States Food and Drug Administration Backgrounder. Cancer therapies: accelerating approval and expanding access. *Ann Pharmacother*. 1996;30(7–8):907.
2. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer*. 1981;47(1):207–14.
3. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, 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(3):205–16.
4. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45(2):228–47. doi:10.1016/j.ejca.2008.10.026.
5. Burzykowski T, Buyse M, Piccart-Gebhart MJ, Sledge G, Carmichael J, Luck HJ, et al. Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. *J Clin Oncol*. 2008;26(12):1987–92. doi:10.1200/JCO.2007.10.8407.
6. Weber WA, Gatsonis CA, Mozley PD, Hanna LG, Shields AF, Aberle DR, et al. Repeatability of 18F-FDG PET/CT in advanced non-small cell lung cancer: prospective assessment in 2 multicenter trials. *J Nucl Med*. 2015;56(8):1137–43. doi:10.2967/jnumed.114.147728.
7. Jacene HA, Leboulloux S, Baba S, Chatzifotiadis D, Goudarzi B, Teytelbaum O, et al. Assessment of interobserver reproducibility in quantitative 18F-FDG PET and CT measurements of tumor response to therapy. *J Nucl Med*. 2009;50(11):1760–9. doi:10.2967/jnumed.109.063321.
8. Hildebrandt MG, Gerke O, Baun C, Falch K, Hansen JA, Farahani ZA, et al. [18F]fluorodeoxyglucose (FDG)-positron emission tomography (PET)/computed tomography (CT) in suspected recurrent breast cancer: a prospective comparative study of dual-time-point FDG-PET/CT, contrast-enhanced CT, and bone scintigraphy. *J Clin Oncol*. 2016;34(16):1889–97. doi:10.1200/JCO.2015.63.5185.
9. Groheux D, Giacchetti S, Delord M, de Roquancourt A, Merlet P, Hamy AS, et al. Prognostic impact of 18F-FDG PET/CT staging and of pathological response to neoadjuvant chemotherapy in triple-negative breast cancer. *Eur J Nucl Med Mol Imaging*. 2015;42(3):377–85. doi:10.1007/s00259-014-2941-1.
10. Mortazavi-Jehanno N, Giraudet AL, Champion L, Lerebours F, Le Stanc E, Edeline V, et al. Assessment of response to endocrine therapy using FDG PET/CT in metastatic breast cancer: a pilot study. *Eur J Nucl Med Mol Imaging*. 2012;39(3):450–60. doi:10.1007/s00259-011-1981-z.
11. Schelling M, Avril N, Nahrig J, Kuhn W, Romer W, Sattler D, et al. Positron emission tomography using [(18F)]fluorodeoxyglucose for monitoring primary chemotherapy in breast cancer. *J Clin Oncol*. 2000;18(8):1689–95.
12. Rousseau C, Devillers A, Sagan C, Ferrer L, Bridji B, Campion L, et al. Monitoring of early response to neoadjuvant chemotherapy in stage II and III breast cancer by [18F]fluorodeoxyglucose positron emission tomography. *J Clin Oncol*. 2006;24(34):5366–72. doi:10.1200/JCO.2006.05.7406.
13. Gebhart G, Lamberts LE, Wimana Z, Garcia C, Emonts P, Ameye L, et al. Molecular imaging as a tool to investigate heterogeneity of advanced HER2-positive breast cancer and to predict patient outcome under trastuzumab emtansine (T-DM1): the ZEPHIR trial. *Ann Oncol*. 2016;27(4):619–24. doi:10.1093/annonc/mdv577.
14. 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–50S. doi:10.2967/jnumed.108.057307.
15. Gonen M, Heller G. Concordance probability and discriminatory power in proportional hazards regression. *Biometrika*. 2005;92(4):965–70. doi:10.1093/biomet/92.4.965.
16. Hendlisz A, Golfinopoulos V, Garcia C, Covas A, Emonts P, Ameye L, et al. Serial FDG-PET/CT for early outcome prediction in patients with metastatic colorectal cancer undergoing chemotherapy. *Ann Oncol*. 2012;23(7):1687–93. doi:10.1093/annonc/mdr554.
17. Coleman RE. Clinical features of metastatic bone disease and risk of skeletal morbidity. *Clin Cancer Res*. 2006;12(20 Pt 2):6243s–9s. doi:10.1158/1078-0432.CCR-06-0931.
18. Specht JM, Tam SL, Kurland BF, Gralow JR, Livingston RB, Linden HM, et al. Serial 2-[18F] fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) to monitor treatment of bone-dominant metastatic breast cancer predicts time to progression (TTP). *Breast Cancer Res Treat*. 2007;105(1):87–94. doi:10.1007/s10549-006-9435-1.
19. Stafford SE, Gralow JR, Schubert EK, Rinn KJ, Dunnwald LK, Livingston RB, et al. Use of serial FDG PET to measure the response of bone-dominant breast cancer to therapy. *Acad Radiol*. 2002;9(8):913–21.
20. Niikura N, Costelloe CM, Madewell JE, Hayashi N, Yu TK, Liu J, et al. FDG-PET/CT compared with conventional imaging in the

- detection of distant metastases of primary breast cancer. *Oncologist*. 2011;16(8):1111–9. doi:10.1634/theoncologist.2011-0089.
21. Tateishi U, Gamez C, Dawood S, Yeung HW, Cristofanilli M, Macapinlac HA. Bone metastases in patients with metastatic breast cancer: morphologic and metabolic monitoring of response to systemic therapy with integrated PET/CT. *Radiology*. 2008;247(1):189–96. doi:10.1148/radiol.2471070567.
  22. Caglar M, Kupik O, Karabulut E, Hoiland-Carlsen PF. Detection of bone metastases in breast cancer patients in the PET/CT era: do we still need the bone scan? *Rev Esp Med Nucl Imagen Mol*. 2016;35(1):3–11. doi:10.1016/j.remnm.2015.08.006.
  23. Glendenning J, Cook G. Imaging breast cancer bone metastases: current status and future directions. *Semin Nucl Med*. 2013;43(4):317–23. doi:10.1053/j.semnuclmed.2013.02.002.
  24. Morris PG, Lynch C, Feeney JN, Patil S, Howard J, Larson SM, et al. Integrated positron emission tomography/computed tomography may render bone scintigraphy unnecessary to investigate suspected metastatic breast cancer. *J Clin Oncol*. 2010;28(19):3154–9. doi:10.1200/JCO.2009.27.5743.
  25. Thatcher N, Chang A, Parikh P, Rodrigues Pereira J, Ciuleanu T, von Pawel J, et al. Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: results from a randomised, placebo-controlled, multicentre study (Iressa survival evaluation in lung cancer). *Lancet*. 2005;366(9496):1527–37. doi:10.1016/S0140-6736(05)67625-8.
  26. Riedl C, Pinker K, Ong L, Jochelson MS, Ulaner G, McArthur HL, et al. FDG-PET/CT versus contrast enhanced CT for prediction of progression-free and disease-specific survival in stage IV breast cancer patients. *J Clin Oncol*. 2015;33(Abstract). <http://meetinglibrary.asco.org/content/153557-156>