

# Prediction of therapy response in bone-predominant metastatic breast cancer: comparison of [<sup>18</sup>F] fluorodeoxyglucose and [<sup>18</sup>F]-fluoride PET/CT with whole-body MRI with diffusion-weighted imaging

Gurdip K. Azad<sup>1</sup> · Benjamin P. Taylor<sup>2</sup> · Adrian Green<sup>1</sup> · Ines Sandri<sup>2</sup> · Angela Swampillai<sup>2</sup> · Mark Harries<sup>2</sup> · Hartmut Kristeleit<sup>2</sup> · Janine Mansi<sup>2</sup> · Vicky Goh<sup>1</sup> · Gary J. R. Cook<sup>1,3</sup> 

Received: 10 August 2018 / Accepted: 19 November 2018 / Published online: 1 December 2018  
© The Author(s) 2018

## Abstract

**Purpose** To compare [<sup>18</sup>F]-fluorodeoxyglucose (FDG) and [<sup>18</sup>F]-sodium fluoride (NaF) positron emission tomography/computed tomography (PET/CT) with whole-body magnetic resonance with diffusion-weighted imaging (WB-MRI), for endocrine therapy response prediction at 8 weeks in bone-predominant metastatic breast cancer.

**Patients and methods** Thirty-one patients scheduled for endocrine therapy had up to five bone metastases measured [FDG, NaF PET/CT: maximum standardized uptake value ( $SUV_{max}$ ); WB-MRI: median apparent diffusion coefficient ( $ADC_{med}$ )] at baseline and 8 weeks. To detect the flare phenomenon, a 12-week NaF PET/CT was also performed if 8-week  $SUV_{max}$  increased. A 25% parameter change differentiated imaging progressive disease (PD) from non-PD and was compared to a 24-week clinical reference standard and progression-free survival (PFS).

**Results** Twenty-two patients (median age, 58.6 years, range, 40–79 years) completing baseline and 8-week imaging were included in the final analysis.

Per-patient % change in NaF  $SUV_{max}$  predicted 24-week clinical PD with sensitivity, specificity and accuracy of 60, 73.3, and 70%, respectively. For FDG  $SUV_{max}$  the results were 0, 100, and 76.2% and for  $ADC_{med}$ , 0, 100 and 72.2%, respectively.

PFS < 24 weeks was associated with % change in  $SUV_{max}$  (NaF: 41.7 vs. 0.7%,  $p = 0.039$ ; FDG: −4.8 vs. −28.6%,  $p = 0.005$ ) but not  $ADC_{med}$  (−0.5 vs. 10.1%,  $p = 0.098$ ). Interlesional response heterogeneity occurred in all modalities and NaF flare occurred in seven patients.

**Conclusions** FDG PET/CT and WB-MRI best predicted clinical non-PD and both FDG and NaF PET/CT predicted PFS < 24 weeks. Lesional response heterogeneity occurs with all modalities and flare is common with NaF PET/CT.

**Keywords** Whole-body MRI · Diffusion-weighted MRI · [<sup>18</sup>F]-fluorodeoxyglucose · [<sup>18</sup>F]-sodium fluoride · Positron emission tomography/computed tomography · Bone metastases

## Introduction

Bone metastases in patients with advanced breast cancer are common, occurring in at least 70% of patients with advanced

disease, and cause significant morbidity [1]. Patients with breast cancer and bone metastases have a relatively long survival compared to other cancers and coupled with the associated morbidity, there are significant implications for healthcare costs [2, 3]. Despite improved therapeutics, response rates are generally less than 50% and so accurate and timely treatment response-assessment methods are essential for optimal management [4]. However, it is recognized that there is an unmet clinical need for correct categorization of treatment response versus non-response in skeletal metastases at an early time point as conventional methods, e.g., RECIST 1.1 measurements on computed tomography (CT) or magnetic resonance imaging (MRI), usually classify skeletal metastases as non-measurable disease [5, 6]. Similarly, the isotope bone

✉ Gary J. R. Cook  
gary.cook@kcl.ac.uk

<sup>1</sup> Cancer Imaging Department, School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK

<sup>2</sup> Department of Oncology, Guys and St Thomas' Hospital NHS Foundation Trust, London, UK

<sup>3</sup> King's College London & Guy's and St Thomas' PET Centre, St Thomas' Hospital, London, UK

scan is considered to have poor sensitivity and specificity for detecting early response or non-response [7]. This means that without an objective early measure of non-response, patients with bone-predominant metastatic disease may continue with ineffective treatment longer than necessary, delaying therapeutic transition to second or third-line treatment and exposing them to unnecessary treatment-related side effects.

There is increasing evidence that functional imaging methods may be able to address this need with reported studies evaluating individual modalities. There is greatest supporting evidence for [<sup>18</sup>F]-fluorodeoxyglucose (FDG) positron emission tomography/computed tomography (PET/CT) measuring reduction in glucose metabolism in responding metastases [8–12]. There is relatively little evidence for alternative imaging methods, which nevertheless show promise in breast cancer or other cancers, including [<sup>18</sup>F]-sodium fluoride (NaF) PET/CT [13–17] and whole-body MRI including diffusion-weighted sequences (WB-MRI) [18–20]. NaF uptake is dependent on altered blood flow and mineralization in the metastatic bone microenvironment [21] and WB-MRI measures changes in the diffusivity of water molecules within tumors [20]. There is no good comparative evidence of significant superiority of any of these methods in measuring treatment response and no clear guidance on the preferred imaging technique in clinical practice [6].

Our hypothesis was that the three functional imaging methods, WB-MRI, FDG, and NaF PET/CT, can detect functional and metabolic changes in breast cancer skeletal metastases as early as 8 weeks after commencing endocrine-based therapy. The aims were to measure baseline parameters and endocrine therapy related changes with each method, to determine the accuracy of each method to predict progressive disease (PD), or non-PD compared to a clinical reference standard, and to determine if the magnitude of change in any method after 8 weeks of treatment was associated with progression-free survival (PFS).

## Patients and methods

This prospective study received research ethics committee approval and all patients gave signed informed consent. Thirty-one patients over the age of 18 with histologically confirmed breast cancer, with either de-novo or progressive bone-predominant metastatic disease scheduled for new endocrine therapy, were recruited. Patients who were also scheduled for radiotherapy or colony-stimulating factors were excluded due to potential effects on functional imaging parameters.

All patients underwent standard follow-up with clinical assessments including a pain inventory [22], blood tests, including serum alkaline phosphatase and tumor marker CA15–

3, and standard imaging, including bone scintigraphy and/or diagnostic CT. The reference standard for clinical PD or non-PD was determined by two oncologists (IS and JM, with 10 and 27 years of specialist oncology experience) in consensus using all the listed clinical assessments up to 24 weeks, or earlier if there was clinical PD. Patients were categorized as either having clinical PD or non-PD (stable disease or partial response) as this is the most relevant dichotomization for clinical management, i.e., continue treatment if non-PD without treatment toxicity or change treatment if PD [23].

Prior to commencing endocrine therapy, patients underwent baseline WB-MRI, FDG, and NaF PET/CT, which were repeated using the same imaging protocol after 8 weeks of therapy. When an increase in maximum standardized uptake value ( $SUV_{max}$ ) was measured in any bone lesions, a further 12-week NaF PET/CT scan was performed when possible to help determine if an early increase in activity was due to the flare phenomenon, which has been reported with this tracer [24]. As a flare is not a recognized phenomenon with WB-MRI or FDG PET/CT, this was not performed with these modalities. As RECIST 1.1 precludes using CT for measuring response in bone metastases unless there is a measurable soft tissue component, we did not include stand-alone CT analysis in our protocol.

## WB-MRI

T1-weighted (T1-W), T2-weighted (T2-W), and diffusion-weighted (DWI, b-values: 50, 900 s/mm<sup>2</sup>) axial sequences were acquired for multiple bed positions from the base of the skull to upper thighs on a 1.5-T MRI scanner (Siemens Aera, Erlangen, Germany). Reformatted axial T1-W, T2-W, DWI b900, and apparent diffusion coefficient (ADC) map images (5-mm slice thickness and 1.2-mm in-plane pixel size) were produced for viewing, lesion identification, and analysis.

## FDG PET/CT

Scans were acquired 60 min after intravenous injection of FDG (mean  $348 \pm 18$  MBq) and all patients had blood glucose measurements of < 10 mmol/l. Images were acquired from skull base to upper thighs with 3 min per bed position using a GE Discovery 710 PET/CT scanner (GE Healthcare, Chicago, IL, USA). A low-dose CT scan (140 kV, 10 mA, 0.5 s rotation time, and 40-mm collimation) was performed at the start of imaging to provide attenuation correction and an anatomical reference. PET images were reconstructed with a time-of-flight ordered subset expectation maximization algorithm (2 iterations, 24 subsets) with a reconstructed slice thickness of 3.27 mm and pixel size 4.7 mm.

## NaF PET/CT

Scans were acquired 60 min after injection of NaF (mean 228 ± 15 MBq). All other acquisition and reconstruction parameters were as for FDG PET/CT.

## Scan analysis

Up to five of the largest bone metastases as assessed on the NaF PET scans were analyzed in each patient by the same reader (GC), with 25 years of radiology and PET experience, using the identical lesions in each of the three scan types. For WB-MRI, lesions were identified on T1-W and T2-W sequences and regions drawn on DWI b900, which were automatically mapped to the accompanying ADC images for measurement of the median ADC in mm<sup>2</sup>/s (ADC<sub>med</sub>). A reduction in ADC<sub>med</sub> of > 25% was used to differentiate imaging PD from imaging non-PD [25, 26]. For FDG and NaF PET/CT, the same lesions were selected and regions of interest (ROIs) outlined semi-automatically using a 40% of maximum activity threshold. The maximum standardized uptake value (SUV<sub>max</sub>) was measured from these regions and a > 25% increase used to differentiate imaging PD from imaging non-PD [27–29]. Individual lesion ADC<sub>med</sub> and SUV<sub>max</sub> measurements were recorded for per-lesion analysis and mean values for each patient recorded for per-patient analysis. Both analyses were performed to obtain clinically relevant results on a per-patient basis that can be used in management decisions and also to report on intra-patient response heterogeneity, a topic of interest in oncology and an observation that may impact on treatment decisions. Intra-patient inter-lesional heterogeneity of response was defined when a lesion showed a > 25% change that was discordant with the clinical reference standard for that patient. In NaF PET/CT scans, a flare was defined in each lesion that showed an initial increase in SUV<sub>max</sub> at 8 weeks that then declined on the 12-week scan.

## Statistical methods

It was calculated that 20 patients with baseline and 8-week scans would give 80% power to predict clinical PD from non-PD 80%

of the time (deemed a clinically useful level) for each modality. Differences between parachutes in patients with clinical PD (PFS < 24 weeks) versus non-PD were tested for normality and compared using Student's *t* test or Mann–Whitney *U* test as appropriate. PFS was defined as the time from the first study scan until the time to clinical progression. Patients who had not progressed clinically at the end of the study were censored. Relationships between scans and PFS were tested by comparing scan metrics in patients with PFS < 24 weeks with those > 24 weeks. Statistical analyses were conducted using IBM SPSS statistics software (version 24). A *P* value of < 0.05 was used for statistical significance.

## Results

### Patients

Twenty-two patients (median age 58.6, range, 40–79 years) completed at least one set of both baseline and 8-week imaging (18 all three modalities, two FDG alone, one FDG and NaF and one NaF alone) and hence WB-MRI: 18 patients, 76 lesions; FDG PET/CT: 21 patients, 90 lesions; NaF PET/CT: 20 patients, 85 lesions. Nine patients did not undergo 8-week imaging (eight due to patient choice and one required radiotherapy for incipient cord compression). Six patients had de novo metastatic disease and 16 had progressive disease prior to recruitment and apart from two patients who had small volume lung and liver metastases; all patients had only skeletal metastases. Endocrine therapy consisted of letrozole (*n* = 12), exemestane with everolimus (*n* = 6), tamoxifen (*n* = 3), and famotidine (*n* = 1). Bisphosphonates (zoledronic acid, *n* = 11 or ibandronate, *n* = 4) or denosumab (*n* = 7), were used as adjunctive therapy. By the clinical reference standard up to 24 weeks, five patients had PD and 17 patients had non-PD. Median PFS was 10.3 months (2.6–47.5 months) with 15 patients alive at the end of the study when censored.

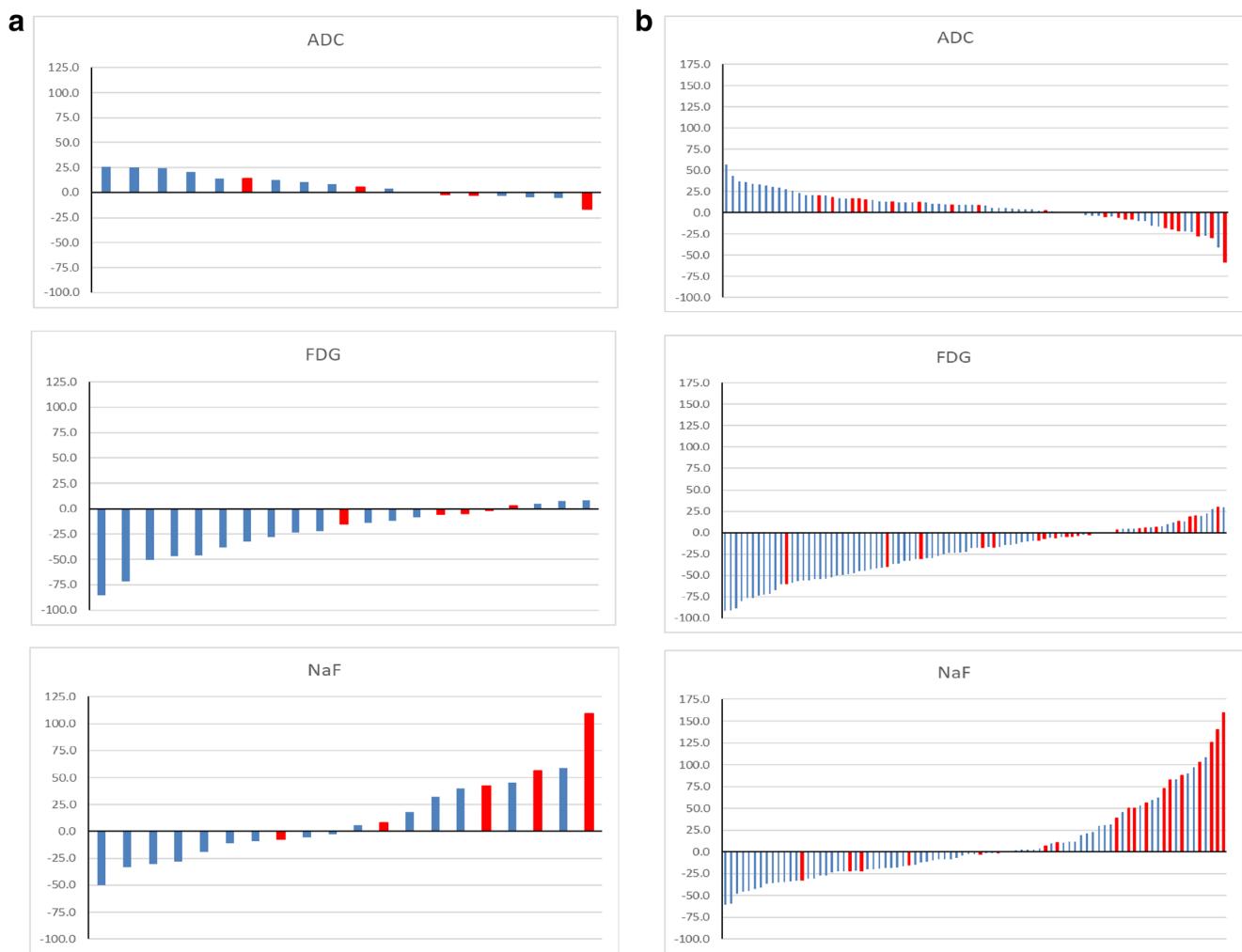
## NaF PET/CT

There was a significant difference in % change in SUV<sub>max</sub> between patients with clinical PD (PFS < 24 weeks) and non-

**Table 1** Performance of % changes in WB-MRI, FDG PET/CT and NaF PET/CT parameters in predicting clinical progressive disease up to 24 weeks on a per-patient and per-lesion basis

	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)
WB-MRI per patient	0	100	0	72.2	72.2
WB-MRI per lesion	15	96.4	60	76.1	75
FDG PET/CT per patient	0	100	0	76.2	76.2
FDG PET/CT per lesion	5	97.1	33.3	78.2	76.7
NaF PET/CT per patient	60	73.3	42.9	84.6	70.0
NaF PET/CT per lesion	55	83.1	50	85.7	76.5

WB-MRI whole-body magnetic resonance with diffusion-weighted imaging, FDG [<sup>18</sup>F]-fluorodeoxyglucose, NaF [<sup>18</sup>F]-sodium fluoride, PET/CT positron emission tomography/computed tomography



**Fig. 1** Waterfall plots showing % change in ADC<sub>med</sub> and SUV<sub>max</sub> in **a** each patient and in **b** each lesion for (top) WB-MRI, (middle) FDG PET/CT, and (bottom) NaF PET/CT [clinical non-progressors (blue) and

progressors (red)]. ADC, apparent diffusion coefficient; FDG, [<sup>18</sup>F]-fluorodeoxyglucose; NaF, [<sup>18</sup>F]-sodium fluoride

PD on a per-patient analysis (41.7 vs. 0.7%,  $p = 0.039$ ) and on a per-lesion analysis (44.4 vs. – 2.6%,  $p = 0.001$ ). Of the 20 patients, 11 showed a less than 25% increase in SUV<sub>max</sub> and were concordant with the clinical reference standard of non-PD. Three out of the five patients with clinical PD showed > 25% increase in SUV<sub>max</sub> (Table 1, Fig. 1a). On analysis of the 85 lesions, 54 were concordant with the clinical reference standard of non-PD and 11 of 20 lesions showed a > 25% increase in SUV<sub>max</sub> in patients with clinical PD (Table 1, Fig. 1b). Baseline SUV<sub>max</sub> was not associated with clinical PD ( $p = 0.6$ ).

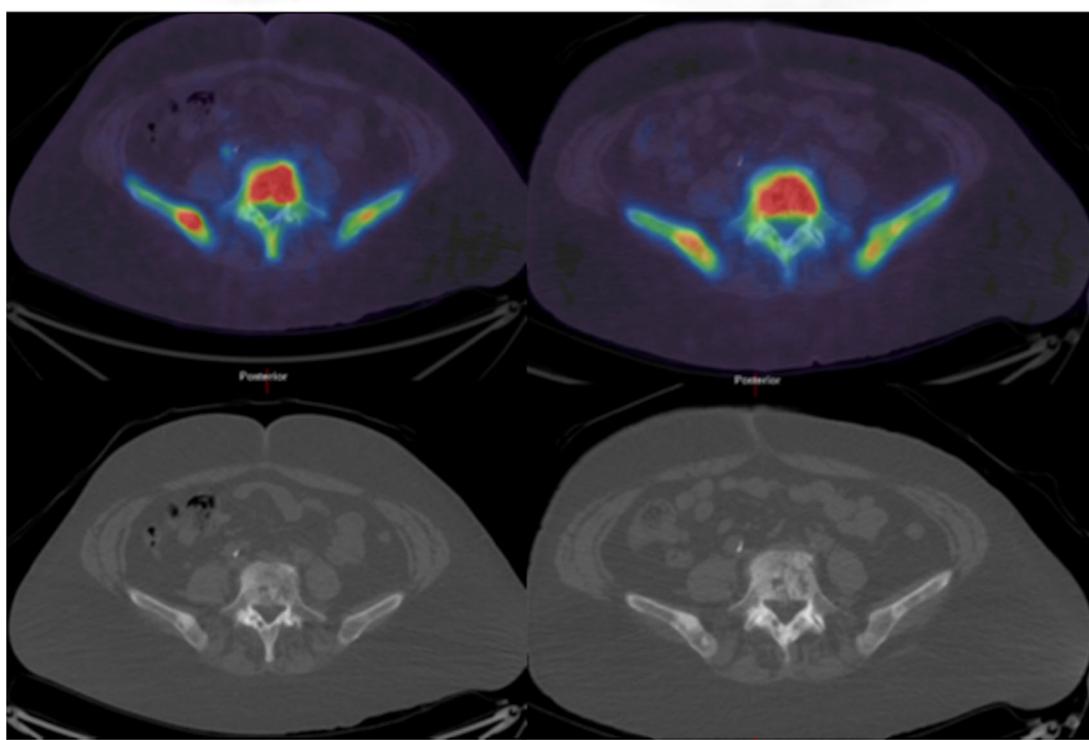
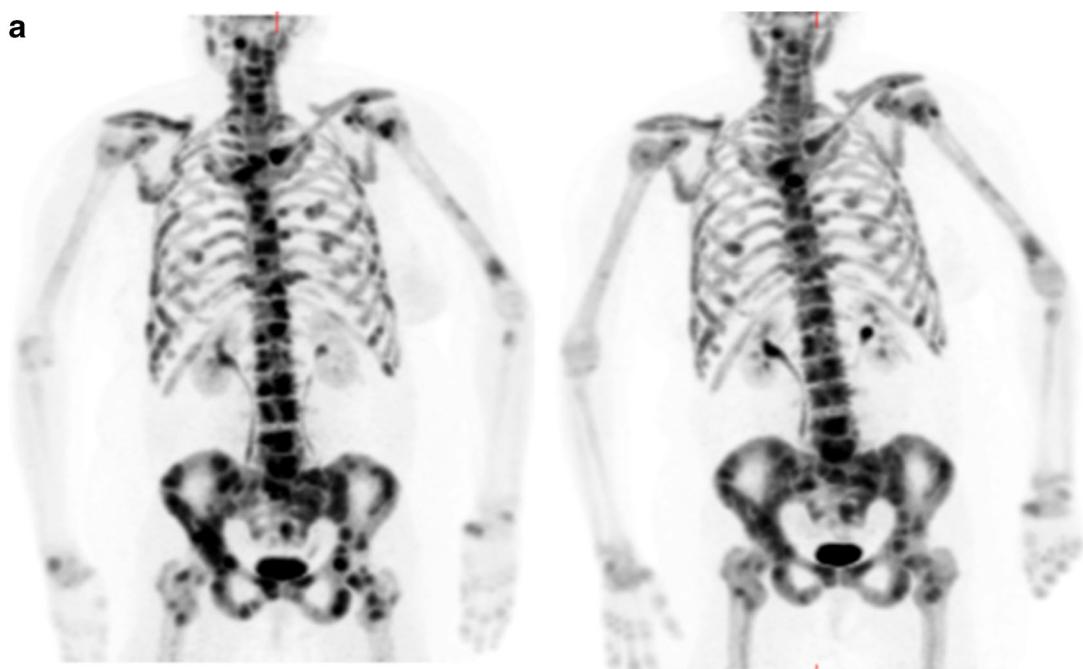
Twelve of the 85 lesions (14.1%) in seven patients showed discordant changes to the clinical reference standard and were categorized as showing an inter-lesional heterogeneous imaging response.

Eighteen lesions (21.2%) in seven patients showed an increase in NaF SUV<sub>max</sub> at 8 weeks followed by a subsequent decrease at 12 weeks, and were therefore categorized as a flare. Four of these patients had clinical non-PD at 24 weeks.

## FDG PET/CT

There was a significant difference in % change in SUV<sub>max</sub> between patients with clinical PD (PFS < 24 weeks) and non-PD on a per-patient analysis (– 4.8 vs. – 28.6%,  $p = 0.005$ ) and on a per-lesion analysis (– 5.0 vs. – 29.7%,  $p = 0.001$ ). Of the 21 patients, 16 showed < 25% increase in SUV<sub>max</sub> and were concordant with the clinical reference standard of non-PD. None of the five patients with clinical PD showed an increase in SUV<sub>max</sub> > 25% (Table 1, Fig. 1a). On analysis of the 90 individual lesions, 68 were concordant with clinical non-PD but

**Fig. 2** **a** NaF PET maximum intensity projection (MIP) and **b** corresponding transaxial images [PET (top), fused PET/CT (middle), and CT (bottom)] of a lesion in L5 at baseline (left) and 8 weeks (right) after commencing endocrine therapy in a patient who had a response by the clinical reference standard. SUV<sub>max</sub> of the L5 vertebral lesion decreased from 42.7 to 32.5



only one of 20 lesions showed a > 25% increase in  $SUV_{max}$  in patients with clinical PD (Table 1, Fig. 1b). Baseline  $SUV_{max}$  was not associated with clinical PD ( $p = 0.65$ ).

Five of the 90 lesions (5.6%) in four patients showed discordant changes to the clinical reference standard and were categorized as showing an inter-lesional heterogeneous imaging response.

## WB-MRI

There was no significant difference in % change in  $ADC_{med}$  between patients with clinical PD (PFS < 24 weeks) and non-PD on a per-patient analysis (− 0.5 vs. 10.1%,  $p = 0.098$ ) but there was on a per-lesion analysis (− 3.2 vs. 9.2%,  $p = 0.012$ ). Of the 18 patients, 13 showed less than a 25% decrease in  $ADC_{med}$  and were concordant with the clinical reference standard of non-PD. None of the five patients with clinical PD showed a > 25% decrease in  $ADC_{med}$  (Table 1, Fig. 1a). On analysis of the 76 individual lesions, 54 were concordant with clinical non-PD but only three of 20 lesions in patients with clinical PD showed a > 25% decrease in  $ADC_{med}$  (Table 1, Fig. 1b). There was no difference in baseline  $ADC_{med}$  ( $p = 0.46$ ) between patients with PFS < 24 weeks compared to PFS > 24 weeks.

Two of the 76 lesions (2.6%) in two patients showed discordant changes to the clinical reference standard and were categorized as showing an inter-lesional heterogeneous imaging response. Representative NaF, FDG PET, and WB-MRI images from a patient who showed a response by the clinical reference standard are illustrated in Figs. 2, 3, and 4, respectively.

## Discussion

Recognizing the limitations of conventional imaging in predicting treatment response in skeletal metastases and the increasing adoption of novel functional imaging into oncologic practice, it is timely to directly compare three contending functional imaging methods in this role. We have shown in this cohort of bone-predominant metastatic breast cancer patients treated with endocrine therapy that changes in parameters that reflect tumor cellularity (DWI), tumor glucose metabolism (FDG PET/CT), and the bone microenvironment (NaF PET/CT), can be detected and quantified. All three modalities showed a similar overall accuracy in predicting PD/non-PD as determined by a clinical reference standard that used conventional clinical, blood, and imaging methods up to 24 weeks. In addition, significant differences were seen in the magnitude of parameter change in those with PFS < 24 weeks compared to those with longer PFS, for NaF, and FDG PET/CT.

While FDG PET/CT and WB-MRI performed well in predicting non-PD, which would allow patients to continue

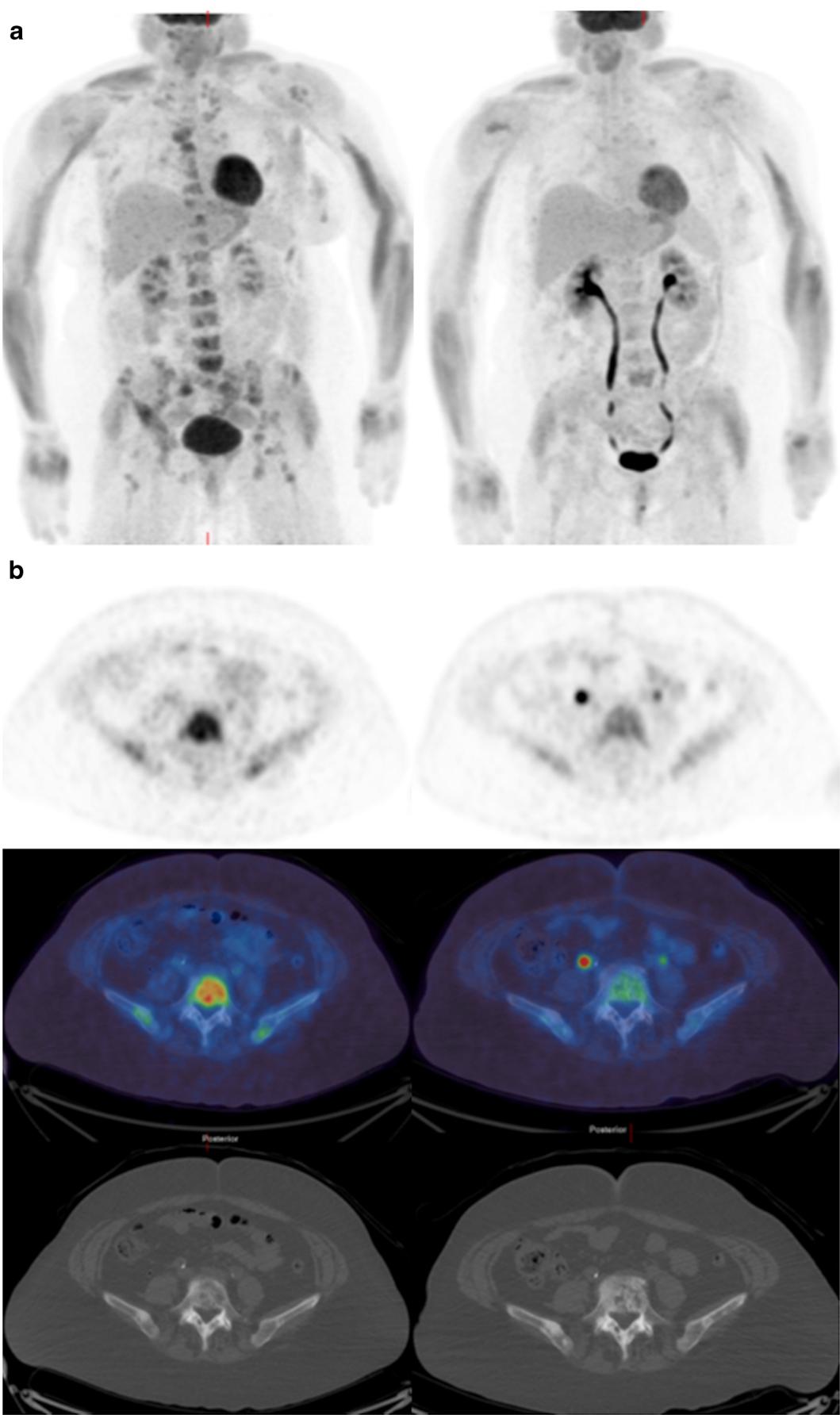
**Fig. 3** **a** FDG PET MIP and **b** corresponding transaxial images [PET (top), fused PET/CT (middle), and CT (bottom)] of a lesion in L5 at baseline (left) and 8 weeks (right) in the same patient as Figs. 2 and 4.  $SUV_{max}$  of the L5 vertebral lesion decreased from 9.9 to 4.9

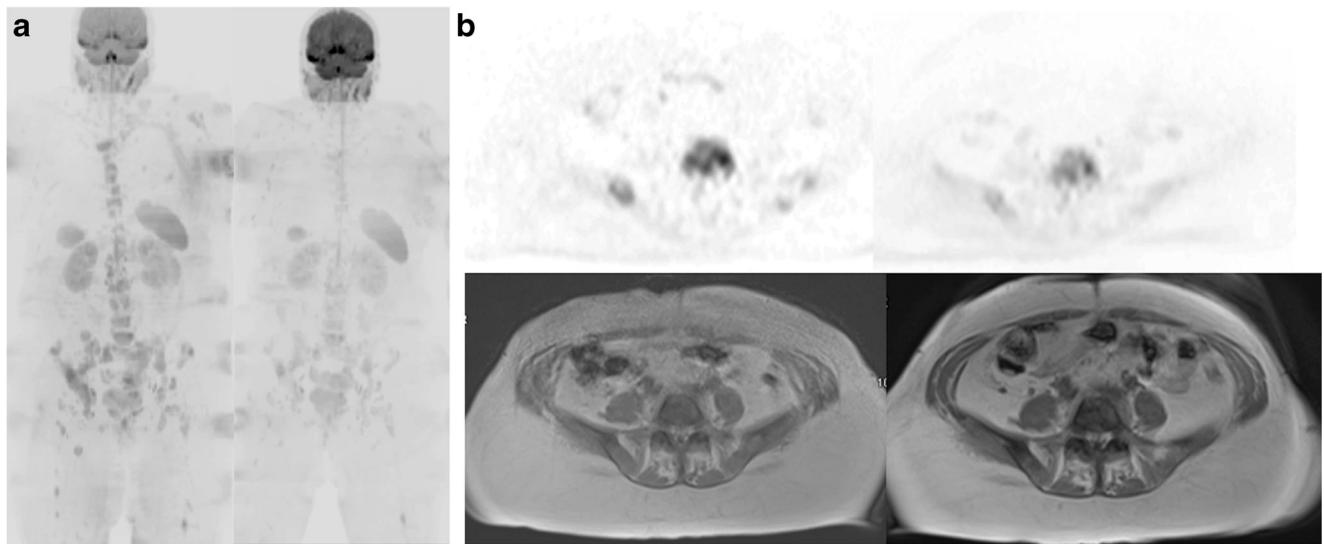
with therapy [26], the magnitude of change (reduction in  $SUV_{max}$  or increase in  $ADC_{med}$ ) on a per-patient or per-lesion basis was greater with FDG (Fig. 1). However, neither WB-MRI nor FDG PET/CT predicted PD at this early 8-week time point. Both FDG PET/CT and WB-MRI primarily reflect tumor cell effects (glucose metabolism [30] and restriction in water molecule motion influenced by cellularity and other tumor-related factors [31], respectively) and while these demonstrated > 25% changes in less than 8 weeks in many responding metastases, the biological changes associated with tumor progression were not of sufficient magnitude to be detected this early at this threshold, implying a non-linear relationship between changes in image parameters and clinical PD.

Nevertheless, a significant difference in % change of FDG  $SUV_{max}$  was seen between patients with PFS < 24 weeks and those with longer PFS but with no significant difference in  $ADC_{med}$ , implying % change in FDG  $SUV_{max}$  may be a better prognostic metric. Our findings augment previous reports where FDG PET/CT  $SUV_{max}$  has previously been shown to be associated with PFS in skeletal, nodal, or visceral metastases from breast cancer [8, 9] or to be associated with changes in tumor markers [10, 11]. To our knowledge, no literature exists for response prediction or assessment using WB-MRI in skeletal metastases from breast cancer in humans. Preclinical data, using a breast cancer model treated with the antiangiogenic agent bevacizumab, rather than endocrine treatment or chemotherapy, found DWI to be insensitive [32]. However, several small series report an increase in ADC in responding prostate cancer bone metastases [19, 20, 33].

While NaF PET/CT feasibility has previously been shown for monitoring treatment response in breast cancer bone metastases [13], to our knowledge definitive results have only been shown in prostate cancer [14–17]. In our series, NaF PET/CT showed modest sensitivity for predicting clinical PD (three of five patients). However, the clinical utility of NaF would be limited, as imaging PD would not be able to be differentiated from a treatment-induced flare, as observed in some of our patients. Despite these observations, the results for NaF PET/CT are of academic interest and suggest that the bone microenvironment changes reflected by this tracer are more rapid and larger in amplitude than the changes we saw with tumor cellular processes demonstrated by FDG PET/CT and WB-MRI.

We observed a heterogeneous response between metastases most frequently with NaF PET/CT, predominantly reflecting





**Fig. 4** **a** b900 DWI MIP and **b** corresponding transaxial b900 DWI (*top*) and T1-W (*bottom*) images at baseline (*left*) and 8 weeks (*right*) of a L5 lesion in the same patient as Figs. 2 and 3.  $\text{ADC}_{\text{med}}$  of the L5 vertebral lesion decreased from 1047 to  $1150 \text{ mm}^2/\text{s}$

the flare phenomenon seen with this tracer. Some inter-lesional response heterogeneity was also observed with FDG PET/CT (5.6%) and WB-MRI (2.6%), suggesting that biological response heterogeneity exists and may reflect tumor resistance to therapy in some clones.

This study has some potential limitations. Additional test-retest scans were not performed for measurement of repeatability as the imaging protocol was already intensive for patients and as good repeatability and inter-observer variation have previously been reported for all three imaging methods employed in this study [19, 25, 26, 28, 29, 34]. Partly due to the intensity of the protocol, nine patients did not complete any 8-week imaging and a small number of patients did not undergo all three imaging tests. This led to a lower number of evaluable patients than preferred but nevertheless, enabled a comparison between all three modalities in most patients and a large number of metastases ( $n = 90$ ) were included in the analysis. We adopted previously published thresholds of 25% change in image metrics to differentiate PD from non-PD but using alternative thresholds would not have significantly improved the ability to differentiate in this series with 25% appearing to be a satisfactory level cross the three modalities. Potential limitations with the clinical reference standard we employed were offset by using consensus from two blinded oncologists and allowing all standard clinical, blood, and imaging to be included while allowing up to 24 weeks for assessment in a method we have previously shown to be robust [35]. While specific treatments differed slightly between patients, they were all endocrine-based regimens without chemotherapy or other non-endocrine treatment-based regimes in order to minimize any heterogeneity due to different classes of treatment as was practically possible. Finally, we have only tested the imaging metrics in breast cancer patients undergoing endocrine-based therapy and we cannot exclude different biological effects from

other therapeutic regimes, e.g., chemotherapy, that could affect imaging parameters differently.

## Conclusions

Changes in tumor cell characteristics and bone microenvironment can be measured with functional imaging methods in bone-predominant metastatic breast cancer at 8 weeks after commencing endocrine-based therapy, although the amplitude of changes did not always reach the threshold for response categorization. Overall accuracy in predicting PD is similar between the three tested modalities but FDG PET/CT and WB-MRI are more reliable than NaF PET/CT in determining non-PD at 8 weeks. Given the larger quantitative percent changes in FDG  $\text{SUV}_{\text{max}}$  compared to  $\text{ADC}_{\text{med}}$  and the fact that larger percent changes in FDG  $\text{SUV}_{\text{max}}$  are associated with PFS, this method has an advantage for determining non-PD and would allow an early decision for patient therapy to continue if there were no limiting side effects. In contrast, none of the three methods were reliable at 8 weeks in predicting subsequent clinical PD, NaF PET/CT performing best with a sensitivity of 60%. However, a flare and inter-lesional heterogeneity is relatively common at 8 weeks with NaF PET/CT and because of these factors, this method would not be sufficiently reliable to change a patient's treatment at 8 weeks.

**Acknowledgements** The authors acknowledge financial support from the King's College London / University College London Comprehensive Cancer Imaging Centres funded by Cancer Research UK and Engineering and Physical Sciences Research Council in association with the Medical Research Council and the Department of Health (C1519/A16463), Breast Cancer Now (2012NovPR013), the Wellcome Trust EPSRC Centre for Medical Engineering at King's College London (WT203148/Z/16/Z), the

Royal College of Radiologists and the support of the National Institute of Health Research Clinical Research Network (NIHR CRN).

## Compliance with ethical standards

**Conflict of interest** The authors declare they have no conflicts of interest.

**Ethical approval** 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.

**Informed consent** Informed consent was obtained from all individual participants included in the study.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Jung S, Rosenzweig M, Sereika S, Linkov F, Brufsky A, Weissfeld JL. Factors associated with mortality after breast cancer metastasis. *Cancer Causes Control*. 2012;23:103–12.
- Plunkett TA, Smith P, Rubens RD. Risk of complications from bone metastases in breast cancer: implications for management. *Eur J Cancer*. 2000;36:476–82.
- Hoefeler H, Duran I, Hechmati G, Garzon Rodriguez C, Luftner D, Ashcroft J, et al. Health resource utilization associated with skeletal-related events in patients with bone metastases: results from a multinational retrospective—prospective observational study—a cohort from 4 European countries. *J Bone Oncol*. 2014;3:40–8.
- Mouridsen H, Gershonovich M, Sun Y, Perez-Carrion R, Boni C, Monnier A, et al. Phase III study of letrozole versus tamoxifen as first-line therapy of advanced breast cancer in postmenopausal women: analysis of survival and update of efficacy from the international Letrozole breast cancer group. *J Clin Oncol*. 2003;21:2101–9.
- 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;45:228–47.
- Lecouvet FE, Talbot JN, Messiou C, Bourguet P, Liu Y, de Souza NM, et al. Monitoring the response of bone metastases to treatment with magnetic resonance imaging and nuclear medicine techniques: a review and position statement by the European Organisation for Research and Treatment of Cancer imaging group. *Eur J Cancer*. 2014;50:2519–31.
- Coombes RC, Dady P, Parsons C, McCready VR, Ford HT, Gazet JC, et al. Assessment of response of bone metastases to systemic treatment in patients with breast cancer. *Cancer*. 1983;52:610–4.
- 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;39:450–60.
- 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;247:189–96.
- Katayama T, Kubota K, Machida Y, Toriihara A, Shibuya H. Evaluation of sequential FDG PET/ CT for monitoring bone metastasis of breast cancer during therapy: correlation between morphological and metabolic changes with tumor markers. *Ann Nucl Med*. 2012;26:426–35.
- 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;9:913–21.
- Specht J, Tam S, Kurland B, 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;105:87–94.
- Doot RK, Muzy M, Peterson LM, Schubert EK, Gralow JR, Specht JM, et al. Kinetic analysis of 18F-fluoride PET images of breast cancer bone metastases. *J Nucl Med*. 2010;51:521–7.
- Hillner BE, Siegel BA, Hanna L, Duan F, Quinn B, Shields AF. 18F-fluoride PET used for treatment monitoring of systemic cancer therapy: results from the National Oncologic PET registry. *J Nucl Med*. 2015;56:222–8.
- Cook GJR, Parker C, Chua S, Johnson B, Aksnes AK, Lewington VJ. 18F-fluoride PET: changes in uptake as a method to assess response in bone metastases from castrate-resistant prostate cancer patients treated with 223Ra-chloride (Alpharadin). *Eur J Nucl Med Mol Imaging Res*. 2011;1:4.
- Yu EY, Duan F, Muzy M, Deng X, Chin BB, Alumkal JJ, et al. Castration-resistant prostate cancer bone metastasis response measured by 18F-fluoride PET after treatment with dasatinib and correlation with progression-free survival: results from American College of Radiology Imaging Network 6687. *J Nucl Med*. 2015;56:354–60.
- Harmon SA, Perk T, Lin C, Eickhoff J, Choyke PL, Dahut WL, et al. Quantitative assessment of early [18F]sodium fluoride positron emission tomography/computed tomography response to treatment in men with metastatic prostate cancer to bone. *J Clin Oncol*. 2017;35:2829–37.
- Padhani AR, Makris A, Gall P, Collins DJ, Tunari N, de Bono JS. Therapy monitoring of skeletal metastases with whole-body diffusion MRI. *J Magn Reson Imaging*. 2014;39:1049–78.
- Perez-Lopez R, Mateo J, Mossop H, Blackledge MD, Collins DJ, Rata M, et al. Diffusion-weighted imaging as a treatment response biomarker for evaluating bone metastases in prostate cancer: a pilot study. *Radiology*. 2017;283:168–77.
- Reischauer C, Froehlich JM, Koh DM, Graf N, Padavit C, John H, et al. Bone metastases from prostate cancer: assessing treatment response by using diffusion-weighted imaging and functional diffusion maps-initial observations. *Radiology*. 2010;257:523–31.
- Blake GM, Park-Holahan SJ, Cook GJ, Fogelman I. Quantitative studies of bone with the use of 18F-fluoride and 99mTc-methylene diphosphonate. *Semin Nucl Med*. 2001;31:28–49.
- Cleeland CS. The measurement of pain from metastatic bone disease: capturing the patient's experience. *Clin Cancer Res*. 2006;12:6236s–42s.
- Rugo HS, Rumble RB, Macrae E, Barton DL, Connolly HK, Dickler MN, et al. Endocrine therapy for hormone receptor-positive metastatic breast cancer: American Society of Clinical Oncology guideline. *J Clin Oncol*. 2016;34:3069–103.
- Wade AA, Scott JA, Kuter I, Fischman AJ. Flare response in 18F-fluoride ion PET bone scanning. *Am J Roentgenol*. 2006;86:1783–6.
- Blackledge MD, Tunari N, Orton MR, Padhani AR, Collins DJ, Leach MO, et al. Inter- and intra-observer repeatability of quantitative whole-body, diffusion-weighted imaging (WBDWI) in metastatic bone disease. *PLoS One*. 2016;11:e0153840.

26. Messiou C, Collins DJ, Morgan VA, Desouza NM. Optimising diffusion weighted MRI for imaging metastatic and myeloma bone disease and assessing reproducibility. *Eur Radiol.* 2011;21:1713–8.
27. Young H, Baum R, Cremerius U, Herholz K, Hoekstra O, Lamertsma AA, et al. Measurement of clinical and subclinical tumour response using [<sup>18</sup>F]-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;35:1773–82.
28. Frost ML, Cook GJ, Blake GM, Marsden PK, Benatar NA, Fogelman I. A prospective study of risedronate on regional bone metabolism and blood flow at the lumbar spine measured by <sup>18</sup>F-fluoride positron emission tomography. *J Bone Miner Res.* 2003;18:2215–22.
29. Lin C, Bradshaw T, Perk T, Hamon S, Eickhoff J, Jallow N, et al. Repeatability of quantitative <sup>18</sup>F-NaF PET: a multicenter study. *J Nucl Med.* 2016;57:1872–9.
30. Koo HR, Park JS, Kang KW, Cho N, Chang JM, Bae MS, et al. <sup>18</sup>F-FDG uptake in breast cancer correlates with immunohistochemically defined subtypes. *Eur Radiol.* 2014;24:610–8.
31. Merz M, Seyler L, Bretschi M, Semmler W, Bauerle T. Diffusion-weighted imaging and dynamic contrast-enhanced MRI of experimental breast cancer bone metastases—a correlation study with histology. *Eur J Radiol.* 2015;84:623–30.
32. Bäuerle T, Bartling S, Berger M, Schmitt-Graff A, Hilbig H, Kauczor HU, et al. Imaging anti-angiogenic treatment response with DCE-VCT, DCE-MRI and DWI in an animal model of breast cancer bone metastasis. *Eur J Radiol.* 2010;73:280–7.
33. Messiou C, Collins DJ, Giles S, de Bono JS, Bianchini D, de Souza NM. Assessing response in bone metastases in prostate cancer with diffusion weighted MRI. *Eur Radiol.* 2011;21:2169–77.
34. Lodge MA. Repeatability of SUV in oncologic (<sup>18</sup>)F-FDG PET. *J Nucl Med.* 2017;58:523–32.
35. Azad GK, Swampillai A, Sandri I, Bassett P, Kristeleit H, Harries M, et al. Measuring validity of response assessment in bone predominant metastatic breast cancer based on clinicoradiological evaluation: an inter-observer study. *J Clin Oncol.* 2018;36:e13055.