Nucl Med Mol Imaging (2016) 50:228–239
DOI 10.1007/s13139-016-0403-3

ORIGINAL ARTICLE

ISSN (print) 1869-3482
ISSN (online) 1869-3474

Conspicuity of FDG-Avid Osseous Lesions on PET/MRI Versus PET/CT: a Quantitative and Visual Analysis

Tyler J. Fraum[1] · **Kathryn J. Fowler**[1] · **Jonathan McConathy**[1]

Received: 13 October 2015 / Revised: 14 January 2016 / Accepted: 2 February 2016 / Published online: 19 February 2016

Abstract
Background Because standard MRI-based attenuation correction (AC) does not account for the attenuation of photons by cortical bone, PET/MRI may have reduced sensitivity for FDG-avid focal bone lesions (FFBLs). This study evaluates whether MRI-based AC compromises detection of FFBLs, by comparing their conspicuity both quantitatively and qualitatively on PET/MRI versus PET/CT.
Methods One hundred ninety general oncology patients underwent whole-body PET/CT followed by whole-body PET/MRI, utilizing the same FDG dose. Thirteen patients with a total of 50 FFBLs were identified. Using automated contouring software, a volumetric contour was generated for each FFBL. Adjacent regions of normal background bone (BB) were selected manually. For each contour, SUV-max and SUV-mean were determined. Lesion-to-background SUV ratios served as quantitative metrics of conspicuity. Additionally, two blinded readers evaluated the relative conspicuity of FFBLs on PET images derived from MRI-based AC versus CT-based AC. Visibility of an anatomic correlate for FFBLs on the corresponding CT and MR images was also assessed.
Results SUV-mean was lower on PET/MRI for both FFBLs (-6.5 %, $p = 0.009$) and BB (-20.5 %, $p < 0.001$). SUV-max was lower on PET/MRI for BB (-14.2 %, $p = 0.002$) but not for FFBLs (-6.2 %, $p = 0.068$). The ratio of FFBL SUV-mean to BB SUV-mean was higher for PET/MRI (+29.5 %, $p < 0.001$). Forty of 50 lesions (80 %) were visually deemed to be of equal or greater conspicuity on PET images derived from PET/MRI. Thirty-five of 50 FFBLs (70 %) had CT correlates, while 40/50 FFBLs (80 %) had a correlate on at least one MRI sequence. The mean interval from tracer administration to imaging was longer ($p < 0.001$) for PET/MRI (127 v. 62 min).
Conclusions Both FFBLs and BB had lower mean SUVs on PET/MRI than PET/CT. This finding was likely in part due to differences in the handling of cortical bone by MRI-based AC versus CT-based AC. Despite this systematic bias, FFBLs had greater conspicuity on PET/MRI, both qualitatively and quantitatively. This difference was likely due to the longer tracer uptake times for PET/MRI, which allowed for more tracer accumulation by FFBLs and more tracer washout from BB. Our results suggest that whole-body PET/MRI and PET/CT provide comparable sensitivity for detection of FDG-avid focal bone lesions.

Keywords Clinical oncology · Magnetic resonance imaging · Multimodal imaging · Positron emission tomography · Whole-body imaging

Introduction

Positron emission tomography (PET)/computed tomography (CT) utilizing the glucose analogue 2-deoxy-2-[^{18}F]fluoro-D-glucose (FDG) has become the standard of care for the initial staging of many different malignancies, as well as an important tool for monitoring treatment response. More recently, simultaneous PET/magnetic resonance imaging (MRI) has emerged as a powerful alternative to PET/CT in certain

✉ Tyler J. Fraum
fraumt@mir.wustl.edu

[1] Mallinckrodt Institute of Radiology, Washington University School of Medicine, Campus Box 8131, 510 S. Kingshighway Blvd., Saint Louis, MO 63110, USA

clinical scenarios, as the superb soft tissue contrast of MRI has proven advantageous for local tumor staging and for assessing tumor involvement of organs with significant physiologic tracer uptake, such as the brain and liver [1–4]. The skeletal system is another common location of metastatic disease, and the detection of osseous lesions can have a significant impact on prognosis and treatment [5, 6]. Relative to FDG-PET/CT, the superiority of MRI alone for detecting bone metastasis has already been reported [7]. Consequently, the diagnosis and characterization of malignant osseous lesions represents another potential advantage of PET/MRI over PET/CT.

The segmentation-based approach to PET photon attenuation correction (AC) employed by most clinical PET/MRI systems has important implications for the detection and evaluation of osseous lesions with increased PET tracer uptake. Segmentation-based AC relies on the Dixon method to classify voxels as soft tissue, fat, lung, or air [8]. In contrast to CT-based AC, segmentation-based AC does not account for the attenuation effects of cortical bone. Cortical bone contains relatively few protons that contribute to the MR signal. Hence, standardized uptake values (SUVs) for osseous lesions have been reported to be underestimated by an average of 11 % and as much as 16 % for sclerotic spine lesions [9]. Despite also finding significantly lower SUVs for both FDG-avid focal bone lesions (FFBLs) and normal background bone (BB), Eiber et al. have reported no significant qualitative differences in the visual conspicuity of FFBLs on PET/CT versus PET/MRI [10]. Beiderwellen et al. similarly addressed the issue of visual conspicuity of FFBLs by rating the qualitative visibility on a scale of 0–5 and also recording a maximum SUV (SUV-max) for each lesion. The authors found PET/MRI to be significantly better than PET/CT [11]. While these initial results provide some reassurance that PET/MRI will perform adequately despite potentially imperfect AC, the actual quantitative impact of the AC method on lesion conspicuity (i.e., lesion tracer uptake relative to that of background bone) remains unclear. Our study expands on the above-described visual analysis results by employing lesion-to-background SUV ratios as a quantitative metric of FFBL conspicuity. Additionally, to affirm the reports of prior authors, we compare the conspicuity of 50 FFBLs on PET/CT versus PET/MRI using visual assessment. To our knowledge, our study is the first to address the question of bone lesion conspicuity on PET/MRI compared to PET/CT in this quantitative fashion.

Materials and Methods

Patient Characteristics

As part of an Institutional Review Board-approved PET/MRI optimization protocol, 190 general oncology patients underwent whole-body FDG-PET/CT immediately followed by PET/MRI, utilizing the same FDG dose. Common indications for PET acquisition included initial staging, restaging, or evaluation for suspected recurrence. To identify suitable FDG-avid osseous lesions, the PET/CT was reviewed for each patient. Of the 190 patients imaged, 26 patients with abnormal osseous FDG uptake on PET/CT were identified. Of these 26 patients, five were excluded due to diffuse marrow FDG uptake that precluded separation of discrete FFBLs from BB; eight were excluded due to a lack of PET/MRI acquisitions evaluating the anatomic region where the FFBL had been identified on the PET/CT images. The group of 13 remaining patients included seven men and six women with various oncologic diagnoses (Table 1)

Image Acquisition

Patients fasted for at least 4 hours prior to tracer injection. Blood glucose levels were checked immediately prior to FDG administration; a blood glucose <150 mg/dl was required. FDG dosing was weight-based, with a mean dose of 14.4 mCi (532.8 MBq).

Standard-of-care PET/CT examinations were performed on one of several Biograph PET/CT scanners (Siemens Healthcare; Erlangen, Germany). The CT component was acquired without intravenous contrast for attenuation correction and anatomic correlation according to the following parameters: 120 kV, 110 mAs (effective; utilizing real-time tube current modulation), 5 mm slice thickness, 4 mm slice spacing. For the standard skull base to mid-thighs protocol, PET data were acquired using 5–6 stations, depending on patient height, with emission times of 2–3 min per station. A three-dimensional PET acquisition protocol

Table 1 Distribution of oncologic diagnoses by patient and lesion

Malignancy	Patients with FFBLs suitable for analysis (n = 13)	Number of FFBLs included in analysis (n = 50)
Cervical/uterine cancer	1	1
Lymphoma	0	0
Head/neck SCC	0	0
Multiple myeloma	7	27
Lung cancer	1	5
Vaginal/vulvar cancer	0	0
Thyroid cancer	1	5
Melanoma	0	0
Sarcoma	3	12
Other	0	0

Abbreviations: FFBL – FDG-avid focal bone lesion; SCC – squamous cell carcinoma

was utilized. PET images were then reconstructed via an iterative algorithm (TrueX; 168 x 168 matrix; 2 mm slice thickness). Some patients also underwent imaging of the head/neck region or lower extremities, which required additional stations. The mean interval from tracer administration to imaging for PET/CT was 62 min (standard deviation 15 min; range 50–97 min).

The PET/MRI examination, acquired using an integrated whole-body PET/MRI system (Biograph mMR; Siemens Healthcare; Erlangen, Germany), was performed according to either a whole-body protocol or a locally focused protocol, depending on the patient's specific pathology. Consequently, the specific MR protocols varied depending on the clinical scenario. For the purposes of this study, the sequences of interest were as follows: T2-weighted (T2W) half-Fourier acquisition single-shot turbo spin echo (HASTE), T1-weighted (T1W) two-point Dixon, three-dimensional (3D) volumetric interpolated breath-hold examination (VIBE), diffusion-weighted imaging (DWI), and T1W turbo spin echo (TSE) (Table 2).

The T1W Dixon VIBE images were acquired using two TEs in order to generate in/opposed-phase, water-only, and fat-only images for attenuation correction. The other sequences were acquired for anatomic correlation of PET findings. Notably, T1-weighted TSE images were available for only 17 of 50 FFBLs, while the other sequences were available for all 50 FFBLs. For the standard skull base to mid-thighs protocol, PET data were acquired using 4–5 stations, depending on patient height, with emission times of 3–4 min per station. A three-dimensional PET acquisition protocol was again utilized. PET images were then reconstructed via an iterative algorithm (3D-OSEM; 172 x 172 matrix; 2 mm slice thickness). Some patients also underwent lower extremity imaging, which required additional stations. Other patients were scanned according to a locally focused protocol, with imaging performed only for an anatomic region of interest; fewer stations were required for these limited acquisitions. The mean interval from tracer administration to imaging for PET/MRI was 127 min (standard deviation 16 min; range 102–157 min).

Image Analysis

Image analysis occurred in a retrospective fashion, after all patients had enrolled in the above-described protocol and completed the requisite imaging. From the 13 patients included in our study, 50 FFBLs were selected for analysis. The number of FFBLs chosen from a single patient was limited to five to prevent the specific tumor biology of an individual patient from biasing the results (Table 1). In patients (eight of 13) with more than five FFBLs, the FFBLs chosen were those most amenable to analysis as discrete lesions that also had an adjacent area of normal-appearing marrow (within the same bone if possible) to provide a BB contour. For example, confluent FFBLs and areas of ill-defined FDG-avidity were not selected, as such lesions generally cannot be contoured in a reproducible fashion. When no suitable BB could be found in the bone containing the FFBL, the closest adjacent bone containing a region of normal bone marrow of similar size to the FFBL was instead selected for BB delineation.

Software Analysis

All PET/CT and PET/MRI datasets for the 13 patients were transferred into MIM (version 6.4, MIM Software Inc., Cleveland, OH) for analysis. For each of the 50 selected lesions, an automated contouring tool was used to define FFBL boundaries on consecutive transaxial PET images from both the PET/CT and PET/MRI datasets. When judged to be necessary for accuracy, adjustments were made to the computer-generated contours with a two-dimensional manual contour selection (2-D contouring) tool. An adjacent area of normal

Table 2 MR imaging parameters for sequences included in FFBL analysis

Parameter	T1W Dixon 3D VIBE (two-point)	T2W HASTE	T1W TSE	DWI
TR (ms)	3.6	3000	500	10000
TE (ms)	1st – 1.23 2nd – 2.46	116	10	85
ST (mm)	3	5	5	5
FA (degrees)	10	160	150	–
FS	No	No	No	Yes (SPAIR)
FOV (mm)	500	500	450	450
B-values (s/mm^2)	–	–	–	General: 50/500/1000 Abdomen: 50/400/800

Abbreviations: TR – time to repetition; TE – time to echo; ST – slice thickness; FA – flip angle; FS – fat suppression; FOV – field of view; SPAIR – spectral attenuated inversion recovery; T1W – T1-weighted; T2W – T2-weighted; TSE – turbo spin echo; DWI – diffusion-weighted imaging

BB was also defined using the same 2-D contouring tool on consecutive transaxial PET images to achieve a BB volume comparable to that of the relevant FFBL. The BB contours were drawn carefully to avoid cortical bone, thereby including only normal marrow. Similarly, for each of the 13 patients, the 2-D contouring tool was also used to delineate two volumes of normal liver (NL) parenchyma, for a total of 26 liver regions of interest. These liver contours were all drawn in similar regions of the inferior right hemiliver to minimize effects of respiratory motion.

For every FFBL, the MIM software was used to generate 3D isocontours, which isolated the voxels within the original lesion contour with SUVs equal to or greater than 40 % (40 %-iso), 50 % (50 %-iso), and 60 % (60 %-iso) of that lesion's SUV-max. This approach was employed to address any potential effects arising from differences in delineation of lesion borders between the PET/CT and PET/MRI datasets; notably, at lesion borders, voxel SUVs can be significantly lower than those within the lesion as a whole, due to volume averaging with surrounding normal tissue and/or lower densities of tumor cells at the lesion's periphery. For each entire FFBL (i.e., no isocontours), SUV-max, mean SUV (SUV-mean), total glycolysis (TG), and total volume were recorded. For each FFBL isocontour, SUV-mean was determined. For BB, SUV-max, SUV-mean, and total volume were measured. For NL, SUV-max and SUV-mean were measured. Various ratios of these parameters, comparing FFBLs to BB, were also calculated.

Visual Analysis

Two board-certified radiologists, one specializing in nuclear medicine and one specializing in body MRI, reviewed each of the 50 FFBLs on the PET images derived from the PET/CT and PET/MRI datasets in a side-by-side fashion. The readers did not have access to any PET/MRI or PET/CT fusion images. The readers were blinded to the source of the PET images, and the right-left arrangement of the side-by-side PET images for each lesion was randomized. For tracking purposes, the set of PET images on the left side of the screen was always labeled as dataset 1, while the set of PET images on the right side of the screen was always labeled as dataset 2. Each reader independently categorized the conspicuity of each FFBL relative to BB as 'greater on dataset 1' (Gr-1), 'greater on dataset 2' (Gr-2) or 'no difference' (ND). Additionally, the visibility of the 50 FFBLs on the anatomic CT and MR images was assessed by a single reader in a binary 'visible' or 'not visible' fashion.

Statistical Analysis

Statistical analysis was performed using the SPSS Statistics package, version 22.0 (IBM; Armonk, NY, USA). All *p*-values less than 0.05 were considered statistically significant. For each parameter, a Q-Q plot was generated to assess data normality. To detect outliers, z-scores were calculated for each observation within a given parameter; values ≥ 2.5 or ≤ 2.5 were used for defining potential outliers. This analysis identified one outlier lesion and prompted reanalysis of the imaging and the selected FFBL and BB contours (Fig. 1). After the culprit imaging artifact was corrected, this lesion no longer met the above-described outlier criteria and was included in the analysis. To compare FFBL, BB, and NL parameters derived from PET/CT versus PET/MRI, the non-parametric Wilcoxen signed rank test was utilized. For the visual analysis, inter-observer variability was assessed with conventional and weighted kappa test statistics.

Results

A total of 50 FFBLs occurring in 13 patients with various oncologic diagnoses were identified on PET images derived from separate PET/CT and PET/MRI acquisitions. At the time of interpretation of the standard-of - care PET/CT portion, the majority of the FFBLs (43/50; 86 %) were called metastatic disease; four of 50 (8 %) were felt to represent metastases with superimposed pathologic fractures; and three of 50 (6 %) were considered indeterminate. The 50 FFBLs were distributed throughout the axial and appendicular skeleton as follows: spine (17 lesions), rib (ten lesions), femur (eight lesions), pelvis/sacrum (seven lesions), clavicle (three lesions), scapula (three lesions), sternum (one lesion), and humerus (one lesion). The interval between FDG injection and image acquisition was longer for PET/MRI than PET/CT for all lesions, with means of 127 min and 62 min, respectively ($p < 0.001$).

Quantitative Analysis

Parameters of interest for FFBLs and BB on both PET/CT and PET/MRI are presented in terms of their absolute values (Table 3). To facilitate comparison, the parameter values for PET/MRI are also reported as a percentage of that parameter's corresponding value on PET/CT. For both FFBLs and BB, there was no significant difference between the contour volumes selected for analysis on PET/MRI versus PET/CT. Likewise, FFBL TG was similar for both modalities. For BB, SUV-mean and SUV-max were lower on PET/MRI than on PET/CT (SUV-max: -14.2 %, $p = 0.002$; SUV-mean: -20.5 %, $p < 0.001$). Similarly, for NL, SUV-mean and SUV-max were lower on PET/MRI than on PET/CT (SUV-mean: −27.6 %, $p < 0.001$; SUV-max: −23.4 %, $p < 0.001$). For FFBLs, SUV-mean was lower on PET/MRI than on PET/CT (SUV-mean: −6.5 %, $p = 0.009$). SUV-mean was also significantly lower on PET/MRI than on PET/CT for the FFBL 40 %-iso (−7.0 %, $p = 0.013$) and 50 %-iso (−6.3 %,

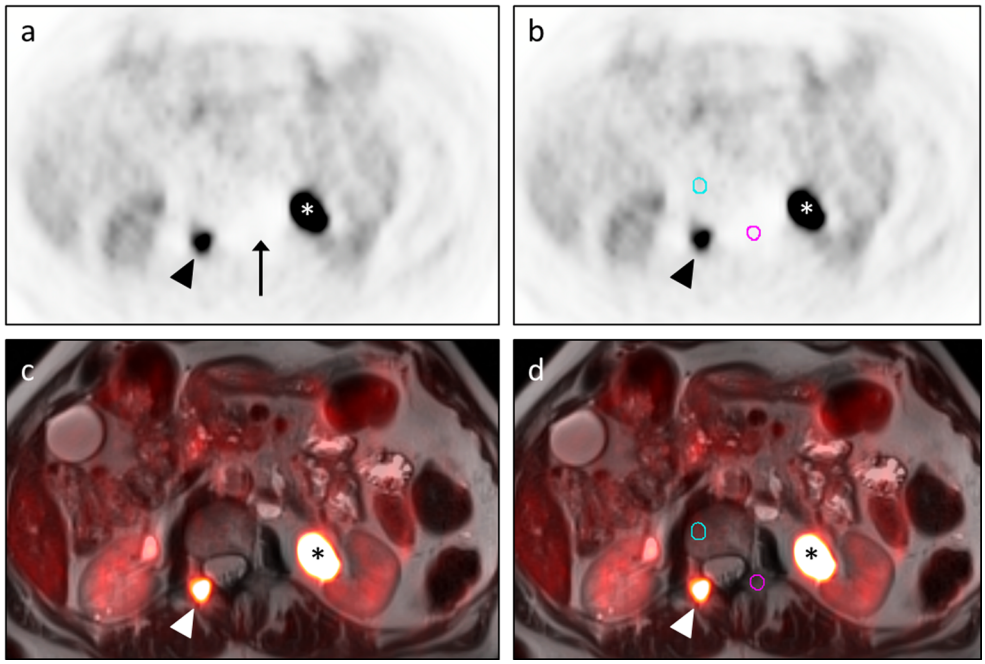

Fig. 1 Outlier FFBL resulting from scatter correction artifact. A 69-year-old man with metastatic papillary thyroid carcinoma status-post multiple radioiodine treatments presented for restaging due to a rising thyroglobulin. Transaxial PET images (**a**, **b**) and transaxial T2W HASTE images with PET fusion (**c**, **d**) demonstrate an FFBL within the right posterior elements of the L1 vertebra (arrowhead) as well as significant tracer activity within a dilated left renal collecting system (asterisk). The initial BB contour (pink circle in **b**, **d**) for this FFBL was drawn in the left posterior elements of the L1 vertebra. Statistical analysis showed this lesion to be an outlier in terms of the ratio of FFBL SUV-mean to BB SUV-mean for both the PET/CT and PET/MRI datasets. SUV-mean for this initial BB contour was 0.39. While not readily apparent on the fusion images, inspection of the PET images (**a**, **b**) revealed an area of photopenia (arrow in **a**) in which this contour (pink circle in **b**) had been placed. This photopenia was consistent with scatter correction artifact from the excreted FDG in the adjacent left renal collecting system (asterisk). Moving the BB contour out of this photopenic region to the right L1 vertebral body (blue circle in **b**, **d**) for both the PET/CT (not shown) and PET/MRI datasets resulted in a new BB SUV-mean of 1.1, similar to values obtained at other vertebral levels. Repeat statistical analysis showed no outliers

$p = 0.028$) but not for the 60 %-iso (−6.1 %, $p = 0.056$). Finally, in contrast to the results for BB, the difference between SUV-max for FFBLs on PET/MRI versus PET/CT did not reach statistical significance ($p = 0.068$), though values were on average 6.2 % lower for PET/MRI.

To assess lesion conspicuity on PET/MRI relative to PET/CT, new parameters were derived from the primary data by dividing FFBL SUV-max and SUV-mean (including isocontours) by the SUV-mean for BB, on a lesion-by-lesion basis. On average, the ratio of FFBL SUV-mean to BB SUV-mean was higher for PET/MRI (+26.7 %, $p < 0.001$) than for PET/CT. This result persisted when evaluating the FFBL isocontours, with higher conspicuity ratios on PET/MRI for the 40 %-iso (+26.1 %, $p < 0.001$), 50 %-iso (+26.6 %, $p < 0.001$), and 60 %-iso (+26.6 %, $p = 0.001$). In contrast, there was no significant difference between PET/MRI and PET/CT for the ratio of FFBL SUV-max to BB SUV-mean (+12.4 %, $p = 0.332$).

Visual Analysis

For PET/CT, 35 of 50 (70 %) FFBLs had correlates on the anatomic CT images; 33 of 35 were lytic, while two of 35 were sclerotic. For PET/MRI, 40 of 50 (80 %) FFBLs had correlates on at least one MR sequence. More specifically, 29 of 50 (58 %) were visible as hyperintense lesions on the T2W HASTE sequence; 32 of 50 (64 %) were visible as hyperintense lesions on DWI; and nine of 17 (53 %) were visible as hypointense lesions on the T1W TSE sequence. Notably, 33 of 50 lesions were not evaluated with T1W TSE.

For the visual conspicuity analysis, two blinded readers compared the PET images of a given FFBL from the PET/MRI and PET/CT datasets in a side-by-side fashion, rating each FFBL relative to BB as Gr-1, Gr-2, or ND. Overall, the two readers agreed on only 24 of 50 FFBLs, resulting in a kappa of 0.23 ($p = 0.003$). However, there were no FFBLs for which one reader chose Gr-1 while the other reader chose Gr-2, a scenario considered 'total disagreement' (TD). All disagreements resulted from one reader choosing ND while the other reader chose either Gr-1 or Gr-2, a scenario considered 'partial disagreement' (PD). By assigning weighting factors to TD and PD, one can account for the relative significance of these disagreements and evaluate the effect on the kappa statistic [12]. Of course, the selection of a specific weighting ratio is arbitrary, so multiple ratios should be considered. Using TD:PD weighting ratios of 2:1, 5:1, and 10:1 resulted in weighted kappa values of 0.32, 0.48, and 0.63, respectively.

Table 3 Comparison of FFBL conspicuity on PET/MRI versus PET/CT: quantitative analysis

Imaging parameter	PET/MRI (mean ± SD)	PET/CT (mean ± SD)	*p*-value	PET/MRI (as % of PET/CT)
BB SUV-max	1.9 ± 1.0	2.2 ± 0.9	0.002*	−14.2 %*
BB SUV-mean	1.1 ± 0.6	1.4 ± 0.5	<0.001*	−20.5 %*
BB Volume	3.9 ± 4.2	3.8 ± 4.8	0.735	+2.6 %
NL SUV-max	2.4 ± 0.5	3.1 ± 0.6	<0.001*	−23.4 %*
NL SUV-mean	1.8 ± 0.3	2.5 ± 0.5	<0.001*	−27.6 %*
FFBL SUV-max	7.0 ± 4.0	7.5 ± 3.8	0.068	−6.2 %
FFBL SUV-mean	4.0 ± 2.0	4.3 ± 1.8	0.009*	−6.5 %*
FFBL TG (SUV*ml)	31.4 ± 35.5	33.6 ± 41.3	0.469	−6.4 %
FFBL Volume (ml)	9.4 ± 12.2	9.7 ± 13.1	0.757	−2.2 %
FFBL 40 %-iso SUV-mean	4.5 ± 2.5	4.8 ± 2.3	0.013*	−7.0 %*
FFBL 50 %-iso SUV-mean	4.9 ± 2.8	5.2 ± 2.7	0.028*	−6.3 %*
FFBL 60 %-iso SUV-mean	5.4 ± 3.2	5.7 ± 3.0	0.056	−6.1 %
[FFBL SUV-max]/ [BB SUV-mean]	4.2 ± 2.7	3.7 ± 2.0	0.332	+12.4 %
[FFBL SUV-mean]/ [BB SUV-mean]	4.4 ± 2.7	3.4 ± 1.9	<0.001*	+26.7 %*
[FFBL 40 %-iso SUV-mean]/ [BB SUV-mean]	4.9 ± 3.1	3.9 ± 2.2	<0.001*	+26.1 %*
[FFBL 50 %-iso SUV-mean]/ [BB SUV-mean]	5.3 ± 3.4	4.2 ± 2.4	<0.001*	+26.6 %*
[FFBL 60 %-iso SUV-mean]/ [BB SUV-mean]	5.7 ± 3.7	4.6 ± 2.6	0.001*	+26.6 %*

*statistically significant results ($p < 0.05$)

Abbreviations: FFBL – FDG-avid focal bone lesion; BB – background bone; NL – normal liver; SUV – standardized uptake value; iso – isocontour; TG – total glycolysis; SD – standard deviation

The results of the visual analysis are shown in Table 4. Reader A was more likely than reader B to designate FFBLs as ND. For FFBLs not placed in the ND category, both readers felt conspicuity was greater on PET/MRI for the majority of the remaining FFBLs (reader A: 9/13, 69 %; reader B: 26/35, 74 %). For the 24 FFBLs on which the readers agreed, eight were deemed more conspicuous on PET/MRI (Fig. 2); three were felt to be more conspicuous on PET/CT (Fig. 3); and 15 were judged to have similar conspicuity on PET/MRI and PET/CT (Fig. 4). Comparing PET/MRI and PET/CT on the basis of whether either or both readers gave one hybrid modality a conspicuity advantage, 27 FFBLs were deemed more conspicuous on PET/MRI. For only ten FFBLs did one or both readers give PET/CT the conspicuity advantage. In other words, 40 of 50 lesions (80 %) were visually deemed to be of equal (by both readers) or greater (by either or both readers) conspicuity on the PET images derived from PET/MRI.

Table 4 Comparison of FFBL conspicuity on PET/MRI versus PET/CT: visual analysis

Reader	A	B	A and B[a]	A and/or B[b]
PET/MRI	9	26	8[c]	27[d]
No difference	37	15	13[c]	39[d]
PET/CT	4	9	3[c]	10[d]

[a] A and B column – FFBLs for which both readers gave the same conspicuity rating

[b] A and/or B column – FFBLs for which either reader or both readers gave a particular conspicuity rating

[c] Note: numbers in the A and B add to <50 FFBLs, as only lesions for which both readers reached the same assessment are included

[d] Note: numbers in the A and/or B add to >50 FFBLs, as lesions for which readers had 'partial disagreements' are double-counted

Discussion

As anticipated, our study found significantly lower SUVs in bone for PET/MRI than for PET/CT. This result held for both SUV-max (−14.2 %, $p = 0.002$) and SUV-mean (−20.5 %, $p < 0.001$) in the case of BB but only for SUV-mean (−6.5 %, $p = 0.009$) in the case of FFBLs. Notably, the SUV-max for FFBLs was on average 6.2 % lower on PET/MRI, a trend that approached significance ($p = 0.068$); thus, a larger FFBL sample size may have provided sufficient power to detect a statistical difference. For BB SUV-max and SUV-mean, our relative PET/MRI values of −14.2 % and −20.5 %

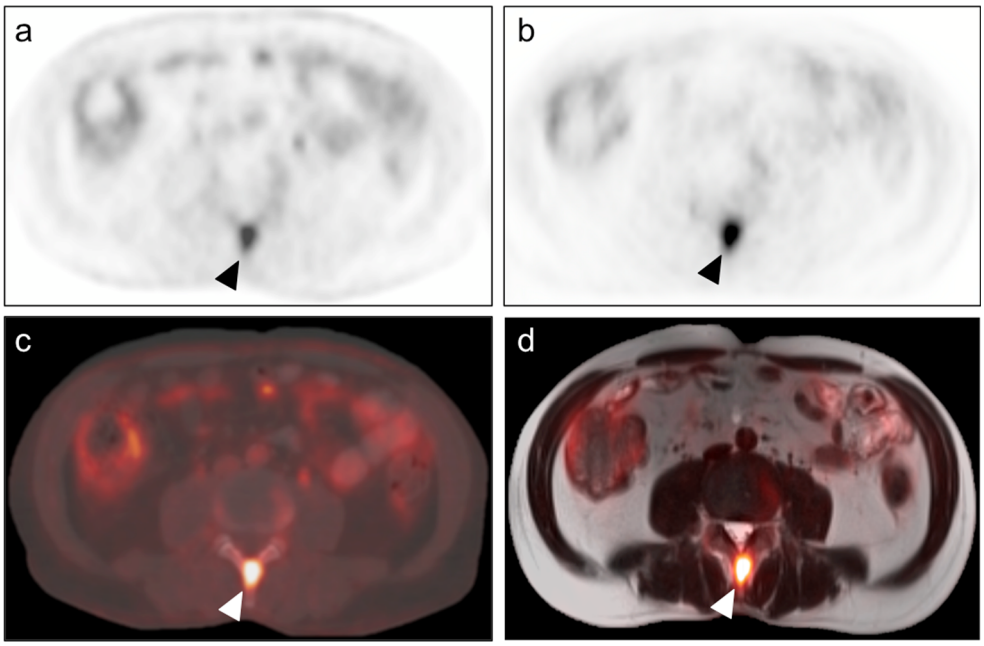

Fig. 2 FFBL deemed more conspicuous on PET images from PET/MRI dataset by both readers. A 53-year-old man with newly diagnosed multiple myeloma presented for initial evaluation of disease extent and underwent both PET/CT (**a** – transaxial PET images; **c** – transaxial CT images with PET fusion) and PET/MRI (**b** – transaxial PET images; **d** – transaxial T2W HASTE images with PET fusion). Among numerous other lesions, an FFBL within the posterior elements of the L3 vertebra (arrowhead) was selected for analysis. Both readers judged this lesion to be more conspicuous on the PET images from PET/MRI (**b**) than on the PET images from PET/CT (**a**)

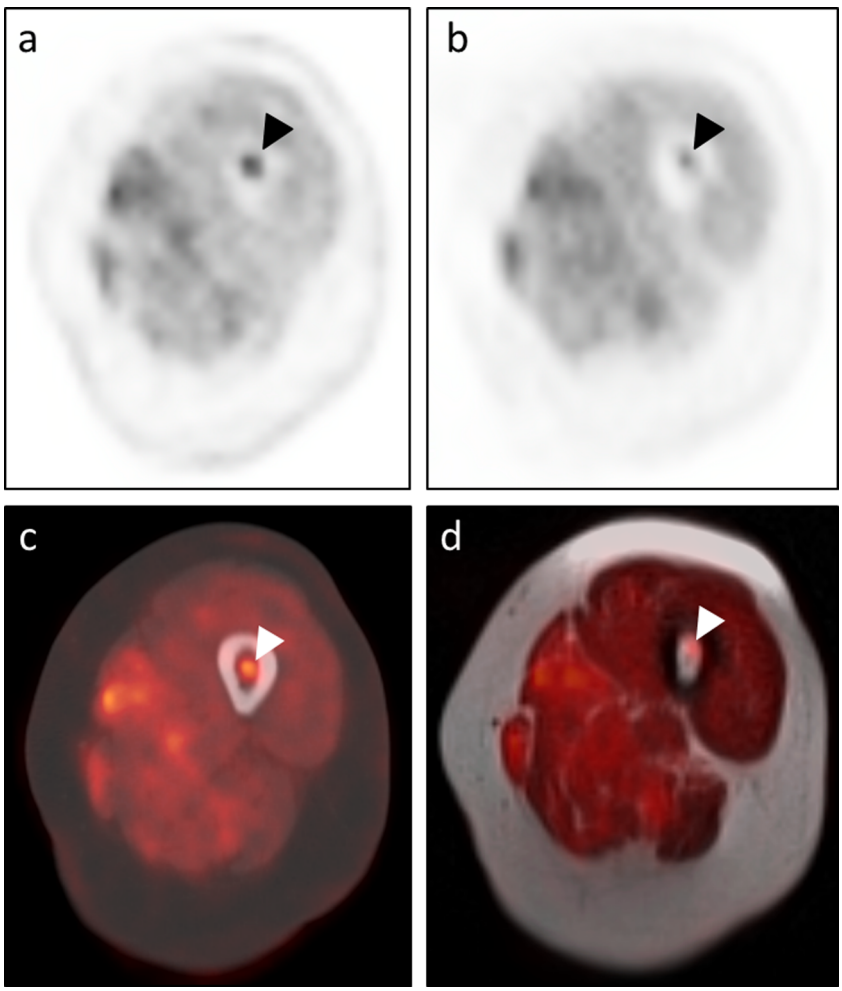

Fig. 3 FFBL deemed more conspicuous on PET images from PET/CT dataset by both readers. An 81-year-old woman with recurrent retroperitoneal liposarcoma presented for restaging and underwent both PET/CT (**a** – transaxial PET images; **c** – transaxial CT images with PET fusion) and PET/MRI (**b** – transaxial PET images; **d** – transaxial T2W HASTE images with PET fusion). Among multiple other lesions, an FFBL within the marrow space of the proximal left femoral diaphysis (arrowhead) was selected for analysis. Both readers judged this lesion to be more conspicuous on the PET images from PET/CT (**a**) than on the PET images from PET/MRI (**b**). Note the dense cortical bone surrounding the lesion on the CT images (**c**). Neglecting cortical bone in the segmentation-based AC algorithm used for the PET images from PET/MRI likely contributed in this substantial difference in FFBL conspicuity (**b**)

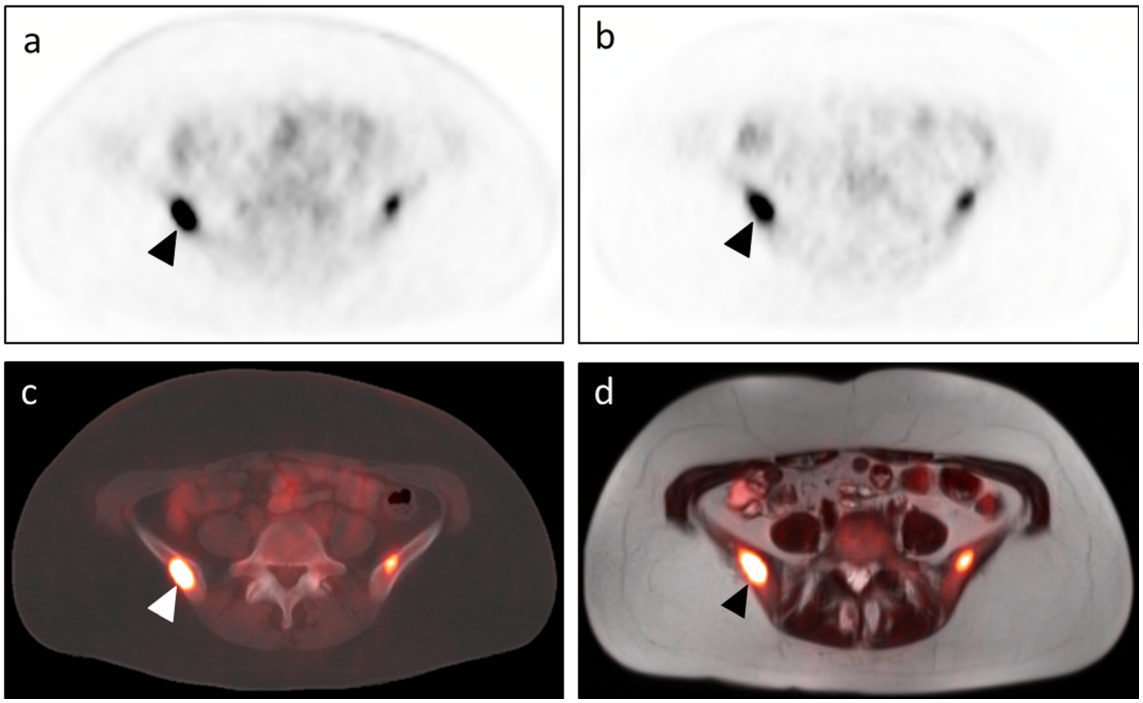

Fig. 4 FFBL deemed to be of equal conspicuity on PET images from the PET/CT and PET/MRI datasets by both readers. A 42-year-old man with epithelioid hemangioendothelioma of the right femur status-post percutaneous cryoablation presented for restaging and underwent both PET/CT (**a** – transaxial PET images; **c** – transaxial CT images with PET fusion) and PET/MRI (**b** – transaxial PET images; **d** – transaxial T2W HASTE images with PET fusion). The patient was noted to have extensive osseous metastatic disease. An FFBL within the right iliac bone (arrowhead) was selected for analysis. Both readers judged this lesion to be of equal conspicuity on the PET images from the PET/CT (**a**) and PET/MRI (**b**) datasets

are comparable to what has been reported by other authors employing similar segmentation strategies (−17 % to −21 % for SUV-max; −14 % to −33 % for SUV-mean) [10, 13, 14]. Likewise, for FFBL SUV-max and SUV-mean, our relative PET/MRI values of −6.2 % and −6.5 % resemble those previously reported by studies with similar segmentation strategies (−11 % for SUV-max; −10 % to −12 % for SUV-mean) [10, 15]. FFBL SUV-mean was also significantly lower on PET/MRI than PET/CT when the 40 % and 50 % isocontours were utilized, suggesting that differences in lesion contours between the two PET datasets did not substantially bias the results. The lack of a significant difference ($p = 0.056$) for the FFBL SUV-mean 60 % isocontour likely reflects insufficient statistical power rather than a substantial difference in lesion behavior when lower SUV voxels are excluded.

The lower SUVs observed in both BB and FFBLs when assessed by PET/MRI are likely in part due to the handling of cortical bone by standard segmentation-derived AC algorithms, which categorize voxels as soft tissue, fat, lung, or air. Our findings corroborate the notion that the high electron density of cortical bone substantially attenuates PET photons arising within the marrow space, resulting in lower SUVs for bone lesions when MRI-based AC is used in place of CT-based AC. Novel approaches to AC for PET/MRI, including atlas-based methods, CT co-registration, and segmentation algorithms employing ultrashort echo times to detect MR signal from cortical bone, may soon be widely available to correct this problem [16–19].

Despite this systematic underestimation of FFBL SUVs by PET/MRI, our study also generally found greater conspicuity for FFBLs on PET/MRI, when assessed by quantitative methods. The ratio of FFBL SUV-mean to BB SUV-mean was significantly higher (+26.7 %, $p < 0.001$) for PET/MRI relative to PET/CT. Greater quantitative conspicuity on PET/MRI was also observed when the 40 %, 50 %, and 60 % FFBL isocontours were employed, again suggesting that differences in lesion border delineation between the two PET datasets did not considerably affect our results. This trend was not observed ($p = 0.33$) for the ratio of FFBL SUV-max to BB SUV-mean, possibly due to this parameter's susceptibility to the effects of individual outlier voxels. The most likely explanation for the greater quantitative conspicuity of FFBLs on PET/MRI relates to the significantly longer tracer-to-imaging time (127 min v. 62 min, $p < 0.001$) for PET/MRI. Various malignant tumors have been shown to continue accumulating FDG for hours after tracer injection before achieving a maximum intratumoral concentration; in contrast, inflammatory lesions and normal tissues typically reach a steady-state much sooner after injection and then progressively lose tracer as FDG is cleared from the blood stream and excreted from the

body [20, 21]. This phenomenon highlights the potential of delayed and dynamic imaging to increase the sensitivity and specificity of FDG-PET for malignancy. Given their relatively longer acquisition times, many PET/MRI protocols inherently take advantage of these tracer dynamics. Similarly, in our study, the longer tracer-to-imaging time for PET/MRI likely resulted in greater accumulation of FDG by FFBLs and increased FDG washout from BB, producing a higher quantitative conspicuity on the later-acquired PET/MRI than on the earlier-acquired PET/CT images.

In our analysis of normal liver tissue, the relative PET/MRI values of −23.4 % ($p < 0.001$) for SUV-max and −27.6 % ($p < 0.001$) for SUV-mean are comparable to those reported by other authors employing similar segmentation strategies (−32 % for SUV-max; −38 % for SUV-mean) [10]. In contrast to bone marrow, the liver is not surrounded on all sides by dense cortical bone, though some PET photons arising in the liver may have to traverse overlying ribs en route to detectors. Consequently, differences in the handling of cortical bone by MRI-based AC versus CT-based AC likely played only a minimal role in the lower hepatic SUVs observed with PET/MRI. This result indicates that the apparent progressive washout of FDG from bone marrow that occurs with longer PET/MRI tracer uptake times is a more generalized process, also occurring in normal hepatic parenchyma. Furthermore, this finding suggests that the lower SUVs observed in BB on PET/MRI are likely due to the combination of progressive tracer washout and the effects of neglecting cortical bone in MRI-based AC.

In the visual conspicuity analysis, when accounting for the fact that all conflicting assessments were partial disagreements, the *fair* inter-observer agreement indicated by a conventional kappa of 0.23 improved to *moderate* (kappa = 0.48) or *substantial* (kappa = 0.63) inter-observer agreement on a weighted analysis, depending on the weighting ratio selected. In this regard, our chosen method of visual analysis appeared to be reasonably reproducible between different readers. A sizeable percentage of the 50 FFBLs were judged by either or both readers to have greater conspicuity on the PET/MRI dataset than on the PET/CT dataset (27/50 v. 10/50, respectively). This result suggests that the effect of longer PET/MRI tracer-to-imaging times on FFBL tracer dynamics influenced visual conspicuity to a greater extent than the absolute decreases in FFBL SUVs arising from the handling of cortical bone by segmentation-based AC. However, 39 of 50 FFBLs were deemed to have equivalent conspicuity on the two datasets by either or both readers, indicating that any differences in visual conspicuity are likely small. Similarly, using a numerical rating system for visually assessing FFBLs on PET/MRI versus PET/CT, Eiber et al. found no significant difference in visual conspicuity between the two PET datasets [10]. In contrast, employing a similar numerical scale, Beiderwellen et al. did find FFBL visual conspicuity to be significantly better ($p < 0.05$) on PET/MRI than on PET/CT, though these authors noted visual conspicuity to be high for both hybrid modalities with no significant differences between datasets in the readers' diagnostic confidence [11]. Overall, our results and those of these other authors suggest that whole-body PET/MRI and PET/CT provide comparable sensitivity for the detection of FFBLs.

Importantly, there were two rib FFBLs in a patient with metastatic osteosarcoma that reader A felt were not well-visualized on the PET/MRI dataset (Fig. 5). This reader remarked that all other 48 FFBLs were seen well on both datasets, though minor conspicuity differences may have existed between the datasets for some lesions. One of the two rib FFBLs had a sclerotic appearance on the CT images. Interestingly, the other FFBL had no CT correlate at the time of image acquisition, but a sclerotic lesion subsequently developed at that FFBL location on a chest CT obtained several months later. Consequently, it is conceivable that one or both lesions may have represented rib fractures at different stages of healing, explaining why the tracer appeared to exhibit washout from these FFBLs on the later-acquired PET/MRI dataset rather than continuing to accumulate tracer as would be expected for osteosarcoma metastases. For both of these rib lesions as well as for an additional sclerotic lesion included in our analysis, there was no correlate for the FFBL on MRI, highlighting a potential advantage of PET/CT over PET/MRI in bone lesion analysis.

This discussion also highlights the importance of carefully evaluating the corresponding CT or MRI when interpreting the PET images. The presence versus absence of a discrete lesion on the anatomic images can generally influence a reader's decision as to whether a PET finding is likely to represent malignant disease versus a benign inflammatory/infectious process or even normal physiologic tracer uptake. We found that more lesions were visible on the MR images (40/50, 80 %) than on the CT images (35/50, 70 %). Other studies have reported a diagnostic advantage of MRI over CT in evaluating FFBLs, especially when T1W TSE sequences are included [10]. Similarly, we did identify several FFBLs with obvious anatomic correlates on T1W TSE sequences but not on CT images (Fig. 6). As has been suggested by other groups, PET/MRI is likely superior to PET/CT for identifying/delineating confluent marrow infiltration and for detecting FFBLs in the setting of physiologic or pharmacologic marrow stimulation; diffuse marrow FDG uptake can preclude evaluation for malignant osseous disease on the PET images in these scenarios [11].

Our study has several limitations that are important to address. First, the significantly longer tracer uptake time for PET/MRI introduces the confounding factors of continued FFBL tracer accumulation and progressive BB tracer washout on the delayed images. An ideal study design would involve imaging patients with two separate tracer injections so as to

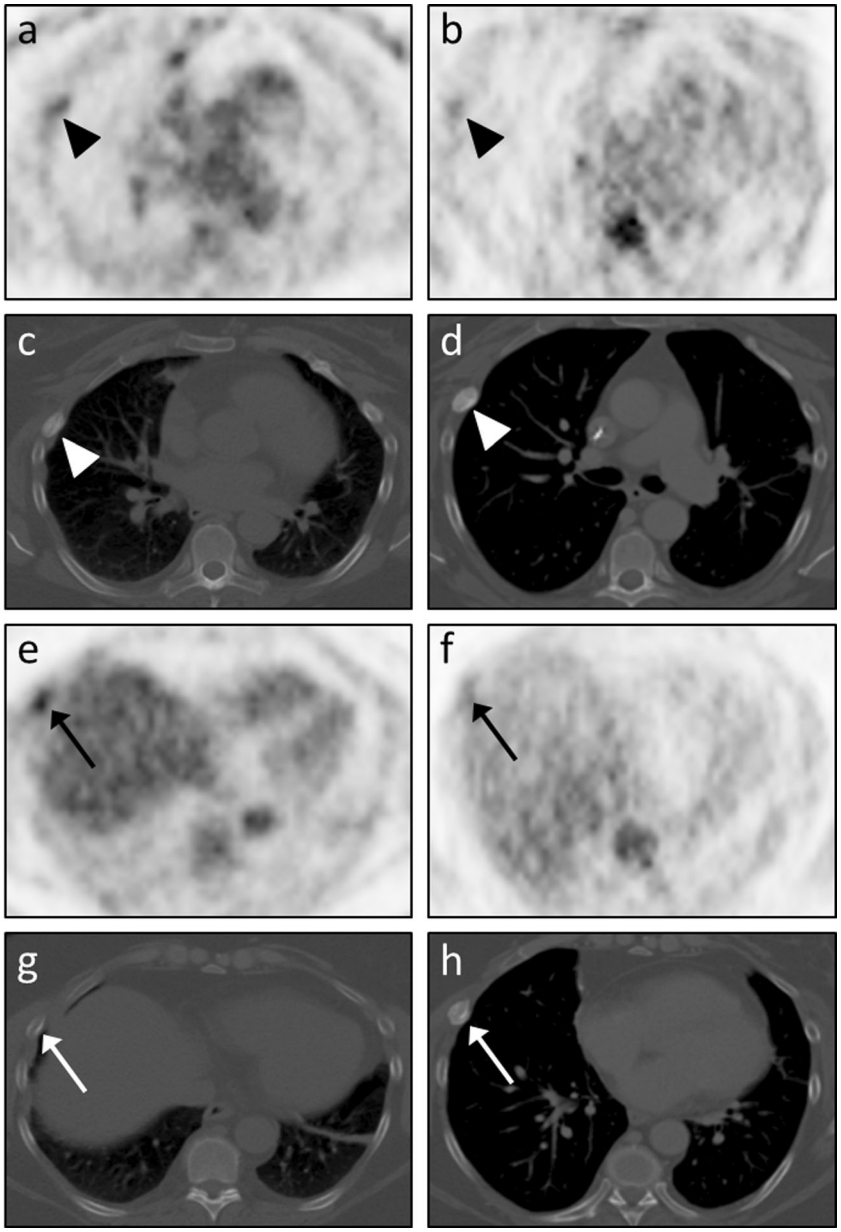

Fig. 5 Tracer washout from indeterminate rib FFBLs. Reader A remarked that only 2 of the 50 FFBLs were not well-visualized on the PET images from the PET/MRI dataset. Both lesions were from the same patient, a 53-year-old woman with metastatic osteosarcoma, and were found in the right 4th (arrowheads in **a** – **d**) and right 6th (arrows in **e** – **h**) ribs. Both readers judged visual conspicuity of the right 4th rib FFBL to be greater on PET images from the PET/CT dataset (**a**) than on PET images from the PET/MRI dataset (**b**). Transaxial CT images from the PET/CT dataset (**c**) showed a sclerotic osseous focus with somewhat ill-defined margins corresponding to this FFBL. A chest CT performed 3 months later (**d**) again showed an osseous focus with increasing sclerosis and more well-defined margins. Similarly, both readers judged visual conspicuity of the right 6th rib FFBL to be greater on PET images from the PET/CT dataset (**e**) than on PET images from the PET/MRI dataset (**f**). Transaxial CT images from the PET/CT dataset (**c**) showed no definite osseous correlate for this FFBL. However, the chest CT performed 3 months later (**d**) showed interval development of a sclerotic osseous focus in the right 6th rib corresponding to the previously seen FFBL. This tracer washout on the relatively delayed PET images (**b**, **f**) is more consistent with an infectious/inflammatory process than metastatic osteosarcoma, suggesting that one or both of these FFBLs may have actually represented rib fractures at different stages of healing. However, no further PET imaging or tissue sampling was performed to evaluate for malignancy at these sites, so these FFBLs are indeterminate. Note the differences in lung volumes between **c** and **d** and between **g** and **h**, causing differential craniocaudal positioning of the heart and liver relative to the rib FFBLs

achieve similar tracer uptake times for PET/MRI and PET/CT. The downside to such a study, of course, would be increased radiation exposure and inconvenience for participants. Furthermore, more than half of FFBLs occurred in patients with multiple myeloma (27/50), and very few FFBLs were sclerotic on CT (two of 50). These sampling biases diminish

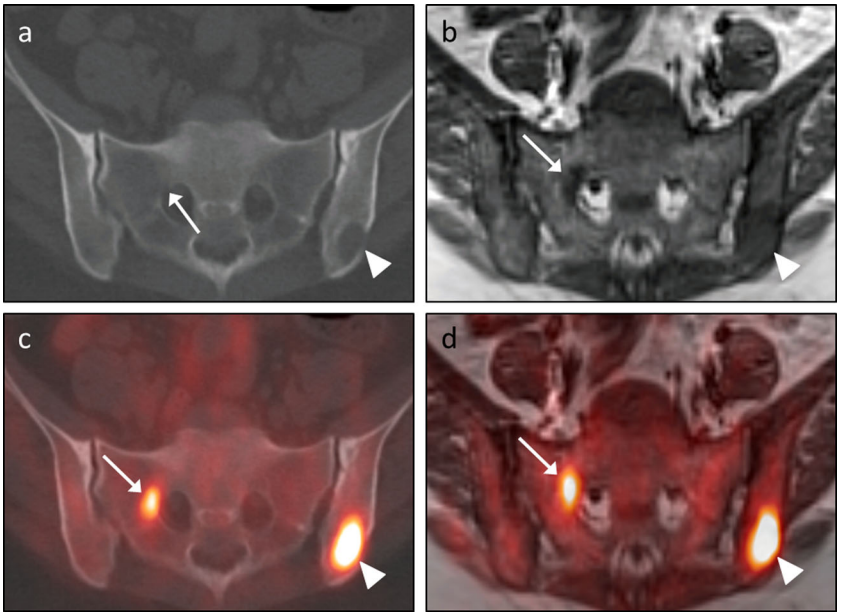

Fig. 6 Advantage of MRI in depicting subtle marrow-replacing lesions. A 42-year-old man with epithelioid hemangioendothelioma of the right femur status-post percutaneous cryoablation presented for restaging and underwent both PET/CT (**a** – transaxial CT images; **c** – transaxial CT images with PET fusion) and PET/MRI (**b** – transaxial T1W TSE images; **d** – transaxial T1W TSE images with PET fusion). An FFBL within the left iliac bone (arrowheads in **c** and **d**) corresponded to an osteolytic lesion on the CT images (arrowhead in **a**) and an area of hypointense marrow replacement on the T1W TSE images (arrowhead in **b**). An additional FFBL along the lateral aspect of the right S1 neural foramen (arrows in **c** and **d**) clearly correlated with an area of hypointense marrow replacement on the T1 TSE images (arrow in **b**). However, there was no corresponding lytic or sclerotic focus on the CT images; the only suggestion of a lesion at this location on CT was subtle cortical loss along the right anterolateral aspect of the right S1 neural foramen (arrow in **b**)

the generalizability of our results to other types of bone metastases. The lack of T1W TSE images for all lesions also hinders our ability to assess the sensitivity of MRI relative to CT for detecting FFBL anatomic correlates, especially given the advantages of T1W TSEs shown by other authors [10]. Finally, our study is limited by the lack of consistent follow-up imaging or pathology to serve as a gold standard for whether the 50 selected FFBLs were malignant. While the intention of our study was to assess FFBLs on both PET/MRI and PET/CT regardless of underlying etiology, a benign versus malignant subgroup analysis would have been beneficial in controlling for potential differences in tracer accumulation and washout. Despite these limitations, our results provide a direct comparison of PET/MRI and PET/CT, in both visual and quantitative fashions, with regard to the important question of whether MRI-based AC reduces the conspicuity of focal FDG-avid bone lesions.

Conclusions

Neglecting cortical bone in the standard segmentation-based algorithms used for attenuation correction in PET/MRI results in lower mean SUVs for both FDG-avid bone lesions and normal background bone. Despite this systematic underestimation of bone lesion SUVs by PET/MRI, our quantitative analysis, which to our knowledge is the first of its kind in this setting, found superior bone lesion conspicuity on PET/MRI relative to PET/CT. Given the longer tracer uptake time for PET/MRI, this greater conspicuity was likely due to the combination of continued tracer accumulation by FDG-avid bone lesions and progressive tracer washout from normal background bone. The lower SUVs in normal liver on PET/MRI than on PET/CT indicate that this progressive tracer washout occurs in both normal bone marrow and hepatic parenchyma. Our visual analysis demonstrated good lesion conspicuity on both PET/CT and PET/MRI, though a greater proportion of lesions was deemed more conspicuous on PET/MRI by one or both readers. Overall, the results of our study suggest that whole-body PET/MRI and whole-body PET/CT provide comparable sensitivity for the detection of FDG-avid bone metastases. In this regard, whole-body PET/MRI should be considered a viable alternative to PET/CT for the assessment of osseous metastatic disease.

Abbreviations *40 %-iso* 40 % isocontour, *50 %-iso* 50 % isocontour, *60 %-iso* 60 % isocontour, *AC* attenuation correction or attenuation-corrected (depending on context), *BB* background bone, *CT* computed tomography, *DWI* diffusion-weighted imaging, *FA* flip angle, *FDG* 2-deoxy-2-[^{18}F]fluoro-D-glucose, *FFBL* FDG-avid focal bone lesion, *FOV* field of view, *FS* fat saturation, *Gr-1* greater conspicuity on dataset 1, *Gr-2* greater conspicuity on dataset 2, *HASTE* half-Fourier

acquisition single-shot turbo spin echo, *MRI* magnetic resonance imaging, *ND* no difference in conspicuity between datasets, *NL* normal liver, *PD* partial disagreement, *PET* positron emission tomography, *SD* standard deviation, *SPAIR* spectral attenuated inversion recovery, *ST* slice thickness, *SUV* standardized uptake value, *SUV-max* maximum SUV, *SUV-mean* mean SUV, *T1W* T1-weighted, *T2W* T2-weighted, *TD* total disagreement, *TE* time to echo, *TG* total lesional glycolysis, *TR* time to repetition, *TSE* turbo spin echo, *VIBE* volumetric interpolated breath-hold examination

Acknowledgments The authors would like to thank Karishma Furtado, MPH, of the Washington University School of Public Health for her assistance with the statistical analysis.

Compliance with Ethical Standards

Conflict of Interest TJF – None; KJF – Research Support, Bracco Group; JM – Research Support, Eli Lilly & Co.; Research Consultant, General Electric Healthcare; Research Consultant, Blue Earth Diagnostics Ltd.; Research Consultant, Siemens AG

Ethical Statement The study was approved by an institutional review board or equivalent and has been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All subjects in the study gave written informed consent or the institutional review board waived the need to obtain informed consent.

References

1. Kanda T, Kitajima K, Suenaga Y, Konishi J, Sasaki R, Morimoto K, et al. Value of retrospective image fusion of ^{18}F-FDG PET and MRI for preoperative staging of head and neck cancer: comparison with PET/CT and contrast-enhanced neck MRI. Eur J Radiol. 2013;82:2005–10.
2. Plathow C, Aschoff P, Lichy MP, Eschmann S, Hehr T, Brink I, et al. Positron emission tomography/computed tomography and whole-body magnetic resonance imaging in staging of advanced nonsmall cell lung cancer–initial results. Investig Radiol. 2008;43:290–7.
3. Kitajima K, Suenaga Y, Ueno Y, Kanda T, Maeda T, Deguchi M, et al. Fusion of PET and MRI for staging of uterine cervical cancer: comparison with contrast-enhanced (18)F-FDG PET/CT and pelvic MRI. Clin Imaging. 2014;38:464–9.
4. Kitajima K, Suenaga Y, Ueno Y, Kanda T, Maeda T, Takahashi S, et al. Value of fusion of PET and MRI for staging of endometrial cancer: comparison with ^{18}F-FDG contrast-enhanced PET/CT and dynamic contrast-enhanced pelvic MRI. Eur J Radiol. 2013;82:1672–6.
5. Mundy GR. Metastasis to bone: causes, consequences and therapeutic opportunities. Nat Rev Cancer. 2002;2:584–93.
6. Rosenthal DI. Radiologic diagnosis of bone metastases. Cancer. 1997;80:1595–607.
7. Schmidt GP, Schoenberg SO, Schmid R, Stahl R, Tiling R, Becker CR, et al. Screening for bone metastases: whole-body MRI using a 32-channel system versus dual-modality PET-CT. Eur Radiol. 2007;17:939–49.
8. Hofmann M, Bezrukov I, Mantlik F, Aschoff P, Steinke F, Beyer T, et al. MRI-based attenuation correction for whole-body PET/MRI: quantitative evaluation of segmentation- and atlas-based methods. J Nucl Med. 2011;52:1392–9.
9. Samarin A, Burger C, Wollenweber SD, Crook DW, Burger IA, Schmid DT, et al. PET/MR imaging of bone lesions–implications for PET quantification from imperfect attenuation correction. Eur J Nucl Med Mol Imaging. 2012;39:1154–60.
10. Eiber M, Takei T, Souvatzoglou M, Mayerhoefer ME, Fürst S, Gaertner FC, et al. Performance of whole-body integrated 18F-FDG PET/MR in comparison to PET/CT for evaluation of malignant bone lesions. J Nucl Med. 2014;55:191–7.
11. Beiderwellen K, Huebner M, Heusch P, Grueneisen J, Ruhlmann V, Nensa F, et al. Whole-body [^{18}F]FDG PET/MRI vs. PET/CT in the assessment of bone lesions in oncological patients: initial results. Eur Radiol. 2014;24:2023–30.
12. Chmura Kraemer H, Periyakoil VS, Noda A. Kappa coefficients in medical research. Stat Med. 2002;21:2109–29.
13. Heusch P, Buchbender C, Beiderwellen K, Nensa F, Hartung-Knemeyer V, Lauenstein TC, et al. Standardized uptake values for [^{18}F] FDG in normal organ tissues: comparison of whole-body PET/CT and PET/MRI. Eur J Radiol. 2013;82:870–6.
14. Drzezga A, Souvatzoglou M, Eiber M, Beer AJ, Fürst S, Martinez-Möller A, et al. First clinical experience with integrated whole-body PET/MR: comparison to PET/CT in patients with oncologic diagnoses. J Nucl Med. 2012;53:845–55.
15. Aznar MC, Sersar R, Saabye J, Ladefoged CN, Andersen FL, Rasmussen JH, et al. Whole-body PET/MRI: the effect of bone attenuation during MR-based attenuation correction in oncology imaging. Eur J Radiol. 2014;83:1177–83.
16. Hofmann M, Steinke F, Scheel V, Charpiat G, Farquhar J, Aschoff P, et al. MRI-based attenuation correction for PET/MRI: a novel approach combining pattern recognition and atlas registration. J Nucl Med. 2008;49:1875–83.
17. Berker Y, Franke J, Salomon A, Palmowski M, Donker HCW, Temur Y, et al. MRI-based attenuation correction for hybrid PET/MRI systems: a 4-class tissue segmentation technique using a combined ultrashort-echo-time/Dixon MRI sequence. J Nucl Med. 2012;53:796–804.
18. Roy S, Wang W-T, Carass A, Prince JL, Butman JA, Pham DL. PET attenuation correction using synthetic CT from ultrashort echo-time MR imaging. J Nucl Med. 2014;55:2071–7.
19. Marshall HR, Patrick J, Laidley D, Prato FS, Butler J, Théberge J, et al. Description and assessment of a registration-based approach to include bones for attenuation correction of whole-body PET/MRI. Med Phys. 2013;40:082509.
20. Kumar R, Loving VA, Chauhan A, Zhuang H, Mitchell S, Alavi A. Potential of dual-time-point imaging to improve breast cancer diagnosis with (18)F-FDG PET. J Nucl Med. 2005;46:1819–24.
21. Schillaci O. Use of dual-point fluorodeoxyglucose imaging to enhance sensitivity and specificity. Semin Nucl Med. 2012;42:267–80.