

---

---

# Repeatability of Radiomic Features of Brown Adipose Tissue

Aria Nazeri, John P. Crandall, Tyler J. Fraum, and Richard L. Wahl

*Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, Missouri*

---

The aim of this study was to assess the repeatability of activated brown adipose tissue (BAT) radiomic features. To decipher radiomic features that may provide useful information on BAT, the impact of reconstruction methods and imaging modality choice was also evaluated. **Methods:** Twenty-seven healthy adults enrolled in this study. After a cooling procedure to activate BAT, volunteers underwent  $^{18}\text{F}$ -FDG imaging. Participants underwent repeat imaging using the same imaging protocols and a similar  $^{18}\text{F}$ -FDG dose within 14 d. Active BAT was segmented using the BARCIST 1.0 methods. Radiomic features were extracted from each region of interest on high-definition PET (HD PET), non-HD PET, and CT images. Lin's concordance correlation coefficient was used to estimate the repeatability of the extracted radiomic features. To determine whether BAT radiomic feature repeatability correlates with BAT  $\text{SUV}_{\text{max}}$  repeatability, participants were stratified based on the relative difference in  $\text{SUV}_{\text{max}}$  between sessions. Non-HD PET repeatable features were clustered together using hierarchical clustering, and the normalized dynamic range of each feature was computed to identify the most informative feature within each cluster. **Results:** Eighteen of the 27 volunteers had sufficient BAT activity for radiomic analysis. Sixty-six HD PET, 66 non-HD PET, and 6 CT features showed high repeatability (concordance correlation coefficient  $\geq 0.80$ ). Feature repeatability was significantly higher for PET than for CT, but there was no statistically significant difference between HD and non-HD PET in radiomic feature repeatability. The repeatability of radiomic features extracted from each modality and reconstruction method type followed the trend in  $\text{SUV}_{\text{max}}$ , as participants with lower relative differences in  $\text{SUV}_{\text{max}}$  between initial and repeated imaging sessions had higher radiomic feature repeatability. Hierarchical clustering of the high-repeatability PET features resulted in 10 highly correlated clusters ( $R^2 \geq 0.95$ ). Seven features, including  $\text{SUV}_{\text{max}}$ , did not cluster with any other features. **Conclusion:** Several clusters of highly repeatable BAT radiomic features derived from  $^{18}\text{F}$ -FDG PET/CT appear to provide information regarding BAT activity distinct from  $\text{SUV}_{\text{max}}$ . These features might be explored as quantitative imaging biomarkers of BAT activity in future studies.

**Key Words:** brown adipose tissue; radiomics; texture analysis; PET/CT; repeatability

**J Nucl Med 2021; 62:700–706**  
DOI: 10.2967/jnumed.120.248674

---

**B**rown adipose tissue (BAT) is a metabolically active tissue that is the primary facilitator of shivering-independent thermogenesis

Received Apr. 30, 2020; revision accepted Sep. 9, 2020.

For correspondence or reprints contact: Richard L. Wahl, Mallinckrodt Institute of Radiology, Washington University School of Medicine, 510 S. Kingshighway Blvd., Campus Box 8131, St. Louis, MO 63110.

E-mail: [rwahl@wustl.edu](mailto:rwahl@wustl.edu)

Published online Oct. 9, 2020.

COPYRIGHT © 2021 by the Society of Nuclear Medicine and Molecular Imaging.

via uncoupling protein-1 and glucose/lipid homeostasis (1–3). Cold exposure,  $\beta_3$ -adrenergic receptor agonists, and catecholamines have been shown to activate BAT in adult humans. BAT activation has been postulated to raise total body energy consumption, enhance insulin sensitivity, and potentially prevent or reverse weight gain (1,4,5). Hence, it has become a potential target for treating obesity and decreasing insulin resistance.

The presence of BAT in adult humans was initially recognized on  $^{18}\text{F}$ -FDG PET/CT examinations performed for oncologic indications (6–9).  $^{18}\text{F}$ -FDG PET has since become the most commonly used technique, and gold standard, for the in vivo detection of cold-activated BAT in humans (10,11).

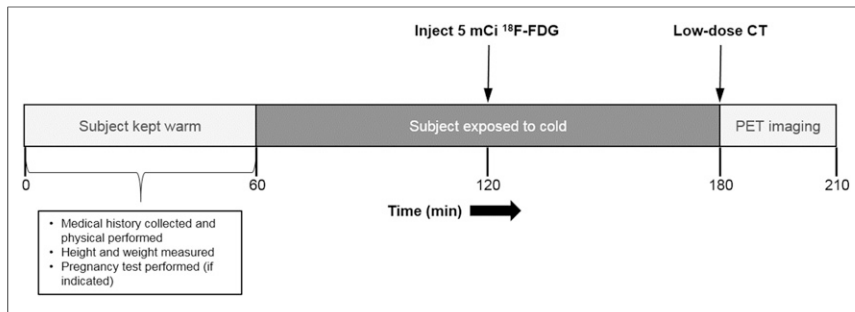
Maximum, peak, and mean standardized uptake values ( $\text{SUV}_{\text{max}}$ ,  $\text{SUV}_{\text{peak}}$  and  $\text{SUV}_{\text{mean}}$ , respectively) are the most commonly studied PET biomarkers (12). However, with the increasing sensitivity and resolution of PET imaging platforms, additional quantitative biomarkers are of interest. The goal of radiomics, which entails the extraction of quantitative features from tomographic images, is the detection of subtle changes in the microenvironment of tumors and other tissues, in addition to the macroscopic features typically reported by radiologists (13,14). Radiomic features have been used as biomarkers for phenotyping tumors, assessing response to treatment, and evaluating prognosis (15–17). Radiomics has the potential to complement our existing knowledge and has therefore emerged as a possible tool for precision medicine (14,18).

The utility of radiomic features in activated BAT is unknown; however, by capturing differences in BAT activity patterns, radiomic features may add useful information regarding BAT physiology. In addition to conventional metrics, radiomic features might be used as biomarkers to further study BAT and its impact on obesity and metabolic disorders. Repeatability is one of the primary elements of biomarker usefulness, as more highly repeatable biomarkers can potentially characterize a tissue in a more dependable fashion. The primary aim of this study (as part of a broader study of BAT imaging repeatability) was to assess the repeatability of BAT radiomic features and to evaluate whether repeatable radiomic features provide additional information beyond conventional metrics, such as SUVs (19).

## MATERIALS AND METHODS

### Subjects

This study was approved by the Washington University institutional review board, and all participants provided written informed consent. Volunteers were recruited using flyers placed in various locations on the Washington University in St. Louis medical campus. Subjects received modest financial compensation for their investment of time in the study. Age and body mass index have been shown to be negatively correlated with BAT activity (5,20). Therefore, healthy adult volunteers aged 18 to 35 with a body mass index under 25 were eligible for this study. Pertinent exclusion criteria included active infections; current or past usage of nicotine products; a history of cold-related injury; insulin-dependent diabetes; current or past usage of medications that



**FIGURE 1.** BAT activation and imaging process. After a 1-h warming period, volunteers were cooled using a water-perfusion suit for approximately 2 h to activate BAT. After 1 h of cooling,  $^{18}\text{F}$ -FDG was administered and PET/CT imaging was performed immediately after the cooling procedure.

may interfere with BAT activity, such as  $\beta$ -blockers; and nursing or pregnancy.

The study consisted of 2 separate imaging sessions. During each session, participants underwent an individualized cooling procedure intended to activate BAT, followed by  $^{18}\text{F}$ -FDG PET/CT imaging. During the repeat session, participants underwent cooling and imaging using the same protocol. Each imaging session took place at the same time of day, and both sessions were completed within 2 wk.

### BAT Activation Protocol

BAT activation was achieved using a previously reported protocol (19). Figure 1 shows a timeline of the imaging sessions. On arrival, participants changed into a hospital gown (clo value of 0.44) and were wrapped in warmed blankets for approximately 1 h before the start of the cooling procedure. The cooling procedure was initiated about 60 min before the planned time of  $^{18}\text{F}$ -FDG injection. Participants were outfitted in a cooling suit (CureWrap, MTRE; Advanced Technologies) that circulates chilled water. At the start of the BAT activation procedure, the cooling system was set at  $10^{\circ}\text{C}$  for all volunteers. The goal of the procedure was to cool participants as much as possible without inducing sustained shivering. When shivering was observed by the study team or reported by the participant, the temperature of the water was raised by  $0.5^{\circ}\text{C}$  every 60 s until shivering ceased. This cooling protocol continued from 60 min before  $^{18}\text{F}$ -FDG administration until approximately 10 min before  $^{18}\text{F}$ -FDG PET/CT imaging.

### Imaging

PET/CT imaging was performed on a Siemens Biograph 40 (Siemens Healthineers). PET imaging was performed from the umbilicus to the base of the skull. About 60 min before the scan, each participant was administered approximately 185 MBq (5 mCi) ( $\pm 20\%$ ) of  $^{18}\text{F}$ -FDG via intravenous bolus injection. A low-dose CT was performed immediately before PET imaging, using a tube potential of 120 kV, tube mAs, pitch of 0.8, and rotation time of 0.5 s. PET images were reconstructed current of 80 contiguous, 8-min list-mode PET images consisted of 3 using ordered-subset expectation maximization acquisitions. (OSEM; non-HD PET) and OSEM with point-spread function (HD PET).

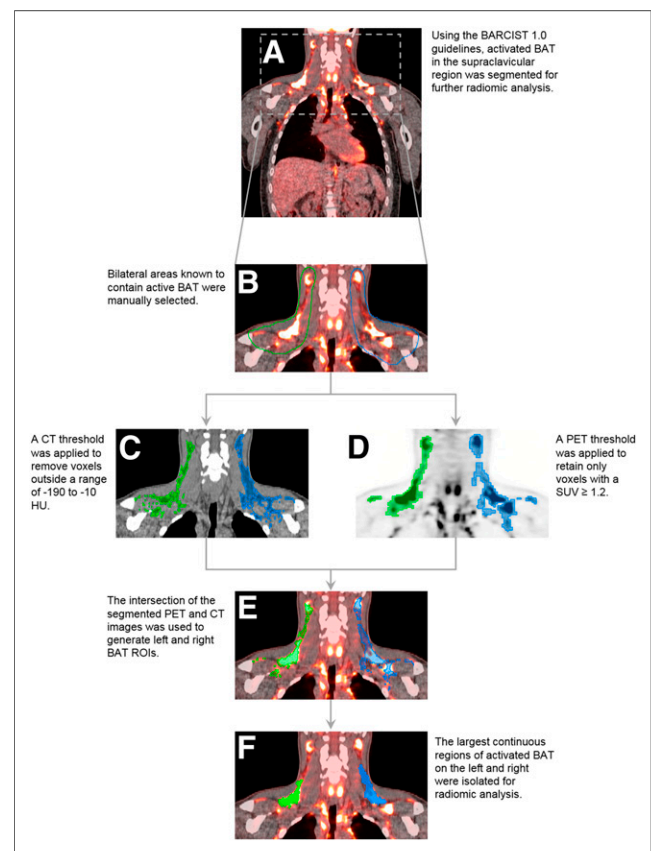
### Image Analysis

Because supraclavicular fossae generally contain the largest BAT depots in adult humans, this region was selected for radiomic analysis (21). The BAT segmentation process is illustrated in Figure 2. BAT segmentation was performed manually using MIM (version 6.7.7; MIM Software, Inc.) and according to a modified BARCIST 1.0 criteria: because all participants in the study were lean, a threshold based on SUV was used in place of SUV adjusted for lean body mass to segregate the activated BAT from the nonactive BAT and white adipose tissue. An

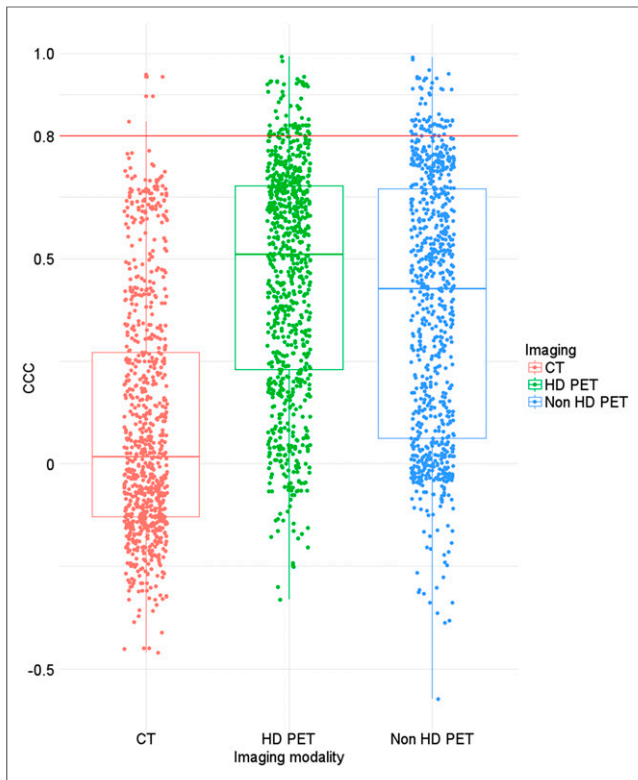
$\text{SUV}_{\text{max}}$  threshold of 1.2 or greater was applied to the supraclavicular fossae to identify potential areas of activated BAT. A CT Hounsfield unit (HU) range was then applied to help ensure the identified areas corresponded to fat.

PET image intensities were normalized to decay-corrected injected activity per kilogram of body weight ( $\text{SUV} [\text{g/mL}]$ ), and CT images were left as HUs. Images were converted from Digital Imaging and Communications in Medicine (DICOM) to Nearly Raw Raster Data (NRRD), and the masks derived from supraclavicular BAT segmentation were converted from DICOM-radiotherapy structure set to NRRD format using 3D Slicer (version 4.8; <https://www.slicer.org/>) (22). Radiomic features were extracted using

PyRadiomics (version 1.3; <http://www.radiomics.io/pyradiomics.html>), which complies with the imaging biomarker standardization initiative (IBSI) recommendations (23,24). Images were discretized to a fixed bin width of 0.1 SUV and 10 HU for PET and CT, respectively, as suggested by Leijenaar et al. (25). Given the sensitivity of textural features to anisotropy, images were resampled to 3-mm isotropic voxels using the Simple Insight Toolkit (ITK) B-Spline algorithm for interpolation (26). The number of analyzed radiomic features and wavelet decompositions per feature class are shown in Supplemental Table 1 (supplemental materials are available at <http://jnm.snmjournals.org>).



**FIGURE 2.** BAT segmentation process. Areas known to contain BAT (A) were manually delineated (B) and then thresholds were applied, using CT HU (C) and SUV (D) cutoffs, to obtain volumes of interest from which radiomic features were extracted (E and F).



**FIGURE 3.** Box plots with overlaid strip plots comparing repeatability of first-order and textural (with and without wavelet decompositions) radiomic features extracted from CT, HD PET, and non-HD PET. Horizontal red line corresponds to a CCC threshold of 0.80.

and additional imaging details (per the IBSI reporting requirements) are provided in Supplemental Table 2. A total of 74 textural features that quantify spatial distribution of voxel intensities were extracted, including 23 gray level cooccurrence matrix, 14 gray level dependence matrix, 16 gray level run length matrix, 16 gray level size zone matrix, and 5 neighboring gray tone difference matrix features. Additionally, we extracted 19 first-order histogram and 14 shape features. A detailed description of each feature can be found on the pyRadiomics website (<https://pyRadiomics.readthedocs.io/en/latest/features.html>). The application

of wavelet filters resulted in 8 decompositions per feature, considering all possible combinations of high-pass (H) or low-pass (L) filters in 3 dimensions (i.e., HHH, HHL, HLH, HLL, LHH, LHL, LLH, LLL). Only first-order and textural features were wavelet-decomposed ( $n = 744$ ). These steps resulted in 851 unique radiomic features for each of the 3 image sets (i.e., HD PET, non-HD PET, and CT). As the volume of the segmented BAT in the right supraclavicular region was more consistent between visits compared with the left side (Lin's concordance correlation coefficient [CCC] values of 0.669 and 0.474, respectively), radiomic features were extracted only from BAT depots on the right side.

### Statistical Analysis

Lin's CCC was used to assess radiomic feature repeatability (Eq. 1). The repeatability of radiomic features extracted from different imaging modalities and reconstruction methods was compared using the Wilcoxon signed-rank test. In all appropriate cases, Levene's test was used to assess variance equality between groups, and variance heterogeneity between groups was remedied using reciprocal transformation. Radiomic feature family comparison and all subsequent analyses were performed using only features extracted from non-HD PET to avoid the additional impact of point-spread function filter incorporation. The repeatability of radiomic feature families was tested using the Kruskal-Wallis H test. The repeatability of radiomic features grouped by wavelet decomposition type was evaluated by the Friedman test and the Nemenyi post hoc test. Additionally, the effect of high versus low frequency decomposition on feature repeatability was separately assessed using the Wilcoxon signed-rank test.

$$CCC = \frac{2\rho\sigma_{V1}\sigma_{V2}}{\sigma_{V1}^2 + \sigma_{V2}^2 + (\mu_{V1} - \mu_{V2})^2} \quad \text{Eq. 1}$$

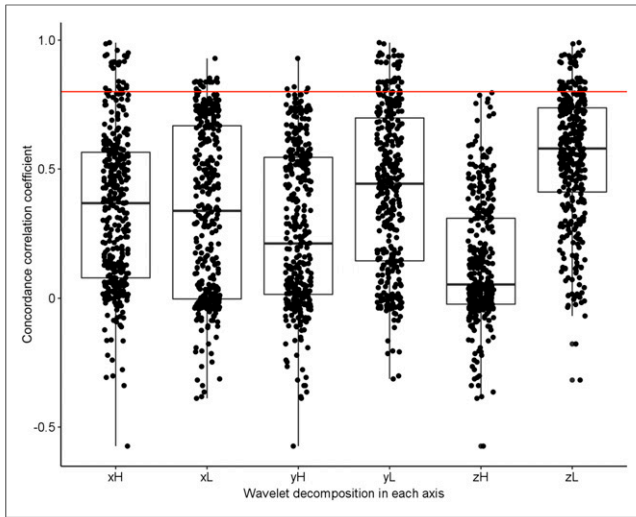
where  $V1$  and  $V2$  indicate the first and second imaging visits, respectively,  $\rho$  is the Pearson correlation coefficient,  $\sigma$  is the radiomic feature SD from the indicated visit, and  $\mu$  is the radiomic feature mean from the indicated visit.

Because  $SUV_{max}$  has been shown to be relatively consistent between test and retest imaging sessions, we hypothesized that radiomic feature repeatability would follow the trend in  $SUV_{max}$  relative difference ( $SUV_{maxrd}$ ) between imaging sessions (Eq. 2) (19,27). To test this, subjects were stratified into a high- $SUV_{maxrd}$  group and a low- $SUV_{maxrd}$  group using a median  $SUV_{maxrd}$  cutoff. In order to assess the possible impact of changes in BAT metabolic volume (BMV) on radiomic feature

**TABLE 1**  
Pairwise Comparison of Non-HD PET Radiomic Feature Repeatability Using Post Hoc Nemenyi Test

Wavelet Decomposition Filter	Mean $\pm$ SD	None	HHH*	HHL	HLH	HLL	LHH	LHL	LLH	LLL
None	0.68 $\pm$ 0.14	—								
HHH	0.05 $\pm$ 0.18	$P < 0.001$	—							
HHL	0.45 $\pm$ 0.18	$P < 0.001$	$P < 0.001$	—						
HLH	0.29 $\pm$ 0.23	$P < 0.001$	$P < 0.001$	$P = 1$	—					
HLL	0.61 $\pm$ 0.29	$P = 1$	$P < 0.001$	$P = 0.77$	$P < 0.001$	—				
LHH	0.01 $\pm$ 0.17	$P < 0.001$	$P = 1$	$P < 0.001$	$P < 0.001$	$P < 0.001$	—			
LHL	0.54 $\pm$ 0.24	$P = 1$	$P < 0.001$	$P = 0.28$	$P < 0.001$	$P = 1$	$P < 0.001$	—		
LLH	0.18 $\pm$ 0.24	$P < 0.001$	$P = 1$	$P < 0.001$	$P = 0.51$	$P < 0.001$	$P = 0.01$	$P < 0.001$	—	
LLL	0.60 $\pm$ 0.24	$P = 0.16$	$P < 0.001$	$P = 1$	$P = 0.003$	$P = 1$	$P < 0.001$	$P = 1$	$P < 0.001$	$P < 0.001$

H = high pass; L = low pass.

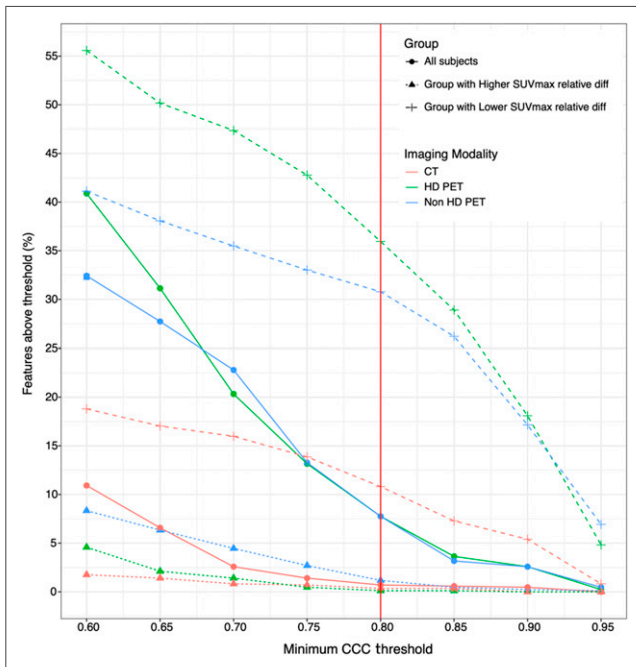


**FIGURE 4.** Repeatability comparison of radiomic features extracted after wavelet decomposition in  $x$ ,  $y$ , and  $z$  axes using high- (H) and low- (L) pass filters. High-pass filters significantly reduced radiomic feature repeatability versus low-pass filters in the  $y$  and  $z$  axes.

repeatability, a similar analysis was performed by stratifying subjects into a high-BMV<sub>rd</sub> relative difference (BMV<sub>rd</sub>) group and a low-BMV<sub>rd</sub> group using a median BMV<sub>rd</sub> cutoff. The CCC of each radiomic feature was then recalculated for both SUV<sub>maxrd</sub> groups and both BMV<sub>rd</sub> groups, and the Wilcoxon signed-rank test was used for groupwise comparison. Levene's test was used to assess variance equality between groups.

$$SUV_{maxrd} = \frac{|SUV_{maxV1} - SUV_{maxV2}|}{\mu_{V1,V2}} \quad \text{Eq. 2}$$

where  $\mu_{V1,V2}$  specifies the mean of the visit  $SUV_{maxV1}$  and the visit  $SUV_{maxV2}$ .



**FIGURE 5.** Repeatability of radiomic features extracted from PET and CT images, with and without stratification based on  $SUV_{maxrd}$ , at different CCC thresholds. Vertical red line corresponds to CCC repeatability threshold of 0.80.

The most informative BAT radiomic features were identified using cluster analysis. The most stable radiomic features extracted from non-HD PET (defined as those with  $CCC \geq 0.8$ ), as well as  $SUV_{mean}$  and  $SUV_{max}$ , were median centered, then cross-correlated. The coefficient of determination ( $R^2$ ) between the features was computed, and features were then clustered together using hierarchical clustering with complete linkage and a Pearson correlation-based dissimilarity index ( $1-R^2$ ). The cluster tree was cut at the height of 0.05 to designate clusters with highly correlated features (i.e.,  $R^2 > 0.95$ ). This step allowed a dimensionality decrease and the identification of redundant features. The most informative radiomic feature within each cluster was found by calculating the normalized dynamic range (NDR) of each feature (Eq. 3).

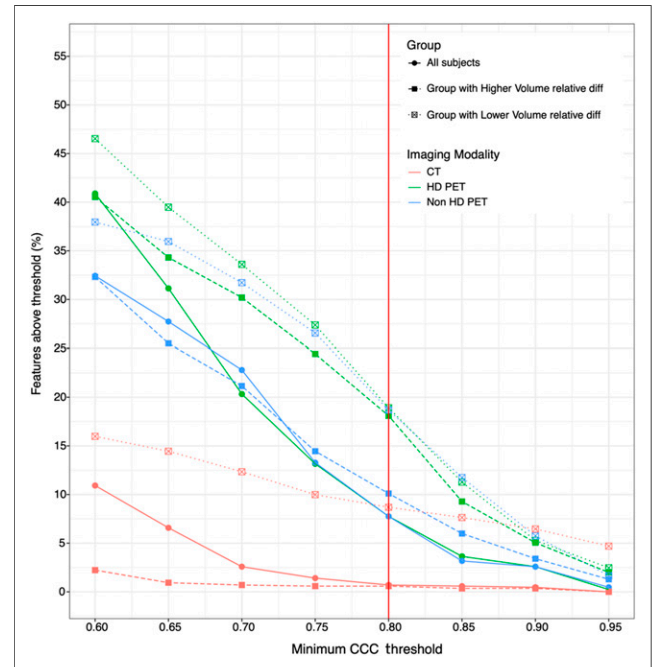
$$NDR = 1 - \frac{1}{n} \sum_{i=1}^n \frac{|R_{V1}(i) - R_{V2}(i)|}{R_{max} - R_{min}} \quad \text{Equation 3}$$

where  $n$  is the number of subjects in the group and  $R$  is the radiomic feature value from the indicated visit for subject  $i$ .  $R_{max}$  and  $R_{min}$  are the maximum and minimum radiomic feature values within the tested group, respectively.

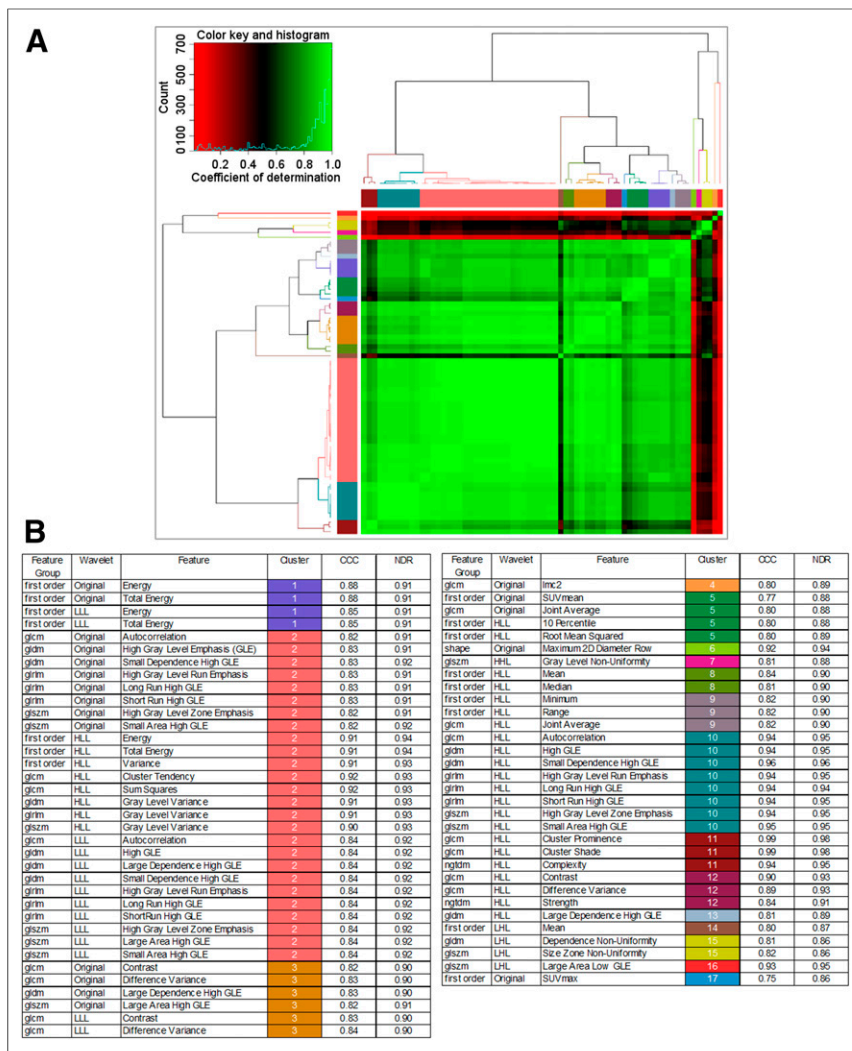
All data analyses were performed in R (version 3.4.1; <http://cran.r-project.org/>).  $P$  values less than 0.01 were considered statistically significant, unless otherwise stated, and the Benjamini-Yekutieli method was used to control for multiple comparisons (28).

## RESULTS

Twenty-seven healthy adults were enrolled, and 26 participants completed both imaging sessions. Active BAT was identified using  $^{18}F$ -FDG PET/CT in 18 volunteers. Females comprised 13 of 18 subjects (72.2%). Subjects had a mean age of 22.9 y (range, 19.3–28.9 y), a mean body mass index of 22.6 kg/m<sup>2</sup> (range, 19.4–26.3 kg/m<sup>2</sup>), a mean height of 169 cm (range, 152–191 cm), and a mean weight of 65 kg



**FIGURE 6.** Repeatability of radiomic features extracted from PET and CT images, with and without stratification based on  $BMV_{rd}$ , at different CCC thresholds. Vertical red line corresponds to CCC repeatability threshold of 0.80.



**FIGURE 7.** Hierarchical clustering of repeatable radiomic features (A) extracted from non-HD PET images. Features were median-centered and clustered using complete linkage with a coefficient of determination-based dissimilarity index ( $1-R^2$ ). (B) Repeatable radiomic features, their respective clusters, NDR, and CCC values. Cluster groups are numbered and color-coded to match heat map row and column colors (A).

(range, 52–83 kg). Self-reported races and ethnicities were as follows: 14 White, 3 Asian, and 1 African-American (1 Hispanic; 17 non-Hispanic).

The number of repeatable features extracted from each imaging modality is shown in Figure 3 and Supplemental Table 3. When a CCC threshold of 0.80 was used, 66 of 851 (7%) HD PET features and 66 of 851 (7%) non-HD PET features were repeatable. The CCC values and absolute relative differences between test and retest radiomic feature values for all non-HD, HD, and CT features are provided in Supplemental Tables 4, 5, and 6, respectively. Fifty-five features were repeatable on both HD and non-HD PET. On CT, only 5 first-order histogram and 1 shape-based feature (0.7%) were repeatable. Averaged across all first-order histogram and texture features (with and without wavelet decompositions), there was significantly higher ( $P < 0.001$ ) repeatability for HD PET (mean  $\pm$  SD of CCC values,  $0.45 \pm 0.28$ ) than for CT (mean  $\pm$  SD of CCC values,  $0.10 \pm 0.29$ ). Non-HD PET radiomic feature (mean  $\pm$  SD of CCC values,  $0.38 \pm 0.32$ ) repeatability was also significantly higher ( $P < 0.001$ ) than CT radiomic feature repeatability. There was no

significant difference between HD and non-HD PET radiomic feature repeatability ( $P = 0.521$ ).

When features extracted from non-HD PET images were used, no significant difference in repeatability was found between radiomic feature families ( $P \approx 0.510$ ). A significant repeatability difference was found between features when grouped by wavelet decomposition type ( $P < 0.001$ ). Table 1 summarizes the results of a pairwise, post hoc analysis of the repeatability of features extracted from images with and without wavelet filtration. These radiomic feature CCC differences before and after wavelet filtering are also shown in Supplemental Figure 1. The differences in repeatability between high- and low-pass filter decomposition in each axis are displayed in Figure 4. Repeatability of features after low-pass filter decomposition in  $y$  (mean  $\pm$  SD of CCC values,  $0.42 \pm 0.31$ ) and  $z$  axes (mean  $\pm$  SD of CCC values,  $0.55 \pm 0.25$ ) was significantly higher ( $P < 0.001$ ) than after high-pass filter decomposition in  $y$  (mean  $\pm$  SD of CCC values,  $0.26 \pm 0.30$ ) and  $z$  axes (mean  $\pm$  SD of CCC values,  $0.13 \pm 0.23$ ). However, no significant difference in repeatability was found between high-pass and low-pass filters ( $P \approx 0.021$ ) in the  $x$  axis.

Stratification by  $SUV_{maxrd}$ , extracted from non-HD PET images, revealed  $SUV_{max}$  CCC values of 0.94 and 0.26 for the low- and high- $SUV_{maxrd}$  groups, respectively. The number of repeatable radiomic features at varying CCC thresholds and grouped by  $SUV_{maxrd}$  classification is shown in Figure 5. CT radiomic feature repeatability of the low- $SUV_{maxrd}$  group (mean  $\pm$  SD of CCC values,  $0.13 \pm 0.4$ ) was significantly ( $P < 0.001$ ) higher than the high- $SUV_{maxrd}$  group (mean  $\pm$  SD of CCC values,  $0.11 \pm 0.24$ ). HD PET radiomic feature repeatability of the

group with lower  $SUV_{maxrd}$  (mean  $\pm$  SD of CCC values,  $0.54 \pm 0.37$ ) was not significantly higher than the group with higher  $SUV_{maxrd}$  (mean  $\pm$  SD of CCC values,  $0.28 \pm 0.19$ ). Non-HD PET radiomic feature repeatability of the group with lower  $SUV_{maxrd}$  (mean  $\pm$  SD of CCC values,  $0.42 \pm 0.43$ ) was significantly ( $P < 0.001$ ) higher than the group with higher  $SUV_{maxrd}$  (mean  $\pm$  SD of CCC values,  $0.28 \pm 0.24$ ). The stratification of subjects on the basis of  $SUV_{maxrd}$  revealed marked differences between subgroups with respect to the number of robust (CCC  $\geq 0.8$ ) radiomic features. For the  $SUV_{maxrd}$  high-repeatability subgroup, there were 92, 306, and 262 robust features for CT, HD PET, and non-HD PET, respectively. For the  $SUV_{maxrd}$  low-repeatability subgroup, there were 3, 1, and 10 robust features for CT, HD PET, and non-HD PET, respectively.

Stratification by  $BMV_{rd}$  revealed mean CCC values of 0.96 and 0.32 for the low- and high- $BMV_{rd}$  groups, respectively (Fig. 6). CT radiomic feature repeatability of the low- $BMV_{rd}$  group (mean  $\pm$  SD of CCC values,  $0.1 \pm 0.4$ ) was not significantly ( $P < 0.001$ )

different from the high-BMV<sub>rd</sub> group (mean  $\pm$  SD of CCC values, 0.1  $\pm$  0.24). HD PET radiomic feature repeatability of the group with lower BMV<sub>rd</sub> (mean  $\pm$  SD of CCC values, 0.45  $\pm$  0.35) was not significantly different ( $P = 1$ ) from the group with higher BMV<sub>rd</sub> (mean  $\pm$  SD of CCC values, 0.5  $\pm$  0.29) using the sign test. Non-HD PET radiomic feature repeatability of the group with lower BMV<sub>rd</sub> (mean  $\pm$  SD of CCC values, 0.41  $\pm$  0.32) was significantly ( $P < 0.001$ ) higher than the group with higher BMV<sub>rd</sub> (mean  $\pm$  SD of CCC values, 0.37  $\pm$  0.4). The stratification of subjects on the basis of SUV<sub>maxrd</sub> revealed marked differences between subgroups with respect to the number of robust (CCC  $\geq$  0.8) radiomic features.

For the BMV<sub>rd</sub> high-repeatability subgroup, there were 75, 170, and 168 robust features for CT, HD PET, and non-HD PET, respectively. For the BMV<sub>rd</sub> low-repeatability subgroup, there were 5, 164, and 86 robust features for CT, HD PET, and non-HD PET, respectively.

Hierarchical clustering of robust non-HD PET radiomic features, along with SUV<sub>max</sub> and SUV<sub>mean</sub>, revealed 10 highly correlated ( $R^2 > 0.95$ ) clusters (Fig. 7). Repeatable features and their respective clusters, along with their NDR and CCC values, are provided in Figure 7B. Six robust features (CCC  $\geq$  0.8), as well as SUV<sub>max</sub>, did not cluster with any other feature. As mentioned previously, the features with the highest NDR values can be interpreted as the representative features of their respective clusters. Though several high-performing features with similar NDR indices in each cluster were identified.

## DISCUSSION

Radiomics has been mainly focused on characterizing tumors. However, these methods may be relevant for other physiologic structures and functions. BAT is a complex tissue with a variable and dynamic mixture of lipid and water density. This heterogeneity may not be well captured using simple CT HU or PET SUV tissue measurements.

Identification of nonredundant, repeatable PET/CT radiomic features of BAT may help elucidate a BAT radiomic signature and its activity pattern in healthy adults, which is potentially quite informative. To the best of our knowledge, this is the first study assessing radiomic feature repeatability in BAT. Although we attempted to scan participants under the same conditions (e.g., cooling procedure, <sup>18</sup>F-FDG dose, imaging parameters, etc.), most features did not show strong repeatability. We assume the dynamic response of BAT to cold temperature, and its complex biology, may help to explain the observed low repeatability of radiomic features in BAT. This is supported by recent studies that have shown the repeatability of standard SUV metrics in activated BAT can vary substantially between test and retest studies, because all the factors that impact SUV measurements can hypothetically influence radiomic repeatability (19,27).

In our study, PET radiomic features were more repeatable than CT radiomic features. This poor performance of CT-based features might be attributable to the low-dose technique used for the CTs in our study, as low tube currents result in higher levels of noise that might obscure radiomic features and negatively impact their repeatability (29,30). Our results also indicate that high-pass filter decomposition in the  $y$  and  $z$  axes leads to lower repeatability when compared with low-pass filter decomposition in the same axes. High-frequency wavelet decomposition results in enhancement of the most high-contrast (heterogeneous) areas of the region of interest, such as the edge (31,32). Additionally, the low-contrast regions become noisier. Because the boundary regions of test–retest BAT regions of

interest are highly variable, this may help explain why high-pass filters lowered repeatability.

Our results suggest the repeatability of BAT radiomic features mostly follows the SUV<sub>max</sub> repeatability trend. A similar trend was not consistently identified when subjects were stratified based on differences in BAT metabolic volume. Interestingly, none of the repeatable non-HD radiomic features clustered with SUV<sub>max</sub>. Hence, repeatable non-HD PET radiomic features can potentially provide unique (i.e., separate from SUV<sub>max</sub>) information regarding activated BAT. These features may potentially be used as biomarkers for assessing the effects of pharmacologic BAT activation in conjunction with conventional metrics such as activated BAT volume and SUV<sub>max</sub>. Textural radiomic features might also be more sensitive for assessing the drug-induced changes in BAT metabolism than activated BAT volume and SUV<sub>max</sub>, as these 2 more conventional BAT metrics have been shown to exhibit lower repeatability than some of the radiomic features and they may capture activation pattern rather than single voxel value (19,27). The precise biologic significance of the features we have evaluated will require further study.

Limitations of this study include using a low-dose CT scan, which decreases radiation exposure but has been shown to negatively impact the reproducibility of CT-derived radiomic features (30). Given the dynamic nature of BAT, its contour changes between visits. The delineation of supraclavicular BAT, thus, may not be identical between the visits, which could negatively impact radiomic repeatability (33). In this study, intensity discretization was achieved using a fixed SUV bin width. An alternative approach is to vary the number of SUV bins. Radiomic feature repeatability may vary between these methods, and the impact of a fixed bin number approach on radiomic feature repeatability should be evaluated in future studies. Another limitation is the risk of false discovery of correlation given the multiple radiomic features assessed. Thus, the highly repeatable features should be prospectively evaluated for further validity of biologic significance.

## CONCLUSION

Several repeatable <sup>18</sup>F-FDG PET BAT radiomic features were identified, which may provide information regarding BAT metabolic activation distinct from SUV<sub>max</sub>, and these features potentially can be explored as quantitative imaging biomarkers of BAT in future studies.

## DISCLOSURE

No potential conflict of interest relevant to this article was reported.

## KEY POINTS

**QUESTION:** Are radiomic features extracted from PET images of brown fat repeatable?

**PERTINENT FINDINGS:** Clusters of repeatable radiomic features were identified that may offer information beyond what is provided by standard SUV measures.

**IMPLICATIONS FOR PATIENT CARE:** Certain radiomic features, found here to be repeatable, may be useful for evaluating BAT.

## REFERENCES

1. Cannon B, Nedergaard J. Brown adipose tissue: function and physiological significance. *Physiol Rev*. 2004;84:277–359.

2. Nedergaard J, Golozoubova V, Matthias A, Asadi A, Jacobsson A, Cannon B. UCP1: the only protein able to mediate adaptive non-shivering thermogenesis and metabolic inefficiency. *Biochim Biophys Acta*. 2001;1504:82–106.
3. Scheele C, Nielsen S. Metabolic regulation and the anti-obesity perspectives of human brown fat. *Redox Biol*. 2017;12:770–775.
4. Chondronikola M, Volpi E, Borsheim E, et al. Brown adipose tissue improves whole-body glucose homeostasis and insulin sensitivity in humans. *Diabetes*. 2014;63:4089–4099.
5. Wang Q, Zhang M, Ning G, et al. Brown adipose tissue in humans is activated by elevated plasma catecholamines levels and is inversely related to central obesity. *PLoS One*. 2011;6:e21006.
6. Nedergaard J, Bengtsson T, Cannon B. Unexpected evidence for active brown adipose tissue in adult humans. *Am J Physiol Endocrinol Metab*. 2007;293:E444–E452.
7. van Marken Lichtenbelt WD, Vanhommerig JW, Smulders NM, et al. Cold-activated brown adipose tissue in healthy men. *N Engl J Med*. 2009;360:1500–1508.
8. Cypess AM, Lehman S, Williams G, et al. Identification and importance of brown adipose tissue in adult humans. *N Engl J Med*. 2009;360:1509–1517.
9. Cohade C, Osman M, Pannu HK, Wahl RL. Uptake in supraclavicular area fat (“USA-Fat”): description on <sup>18</sup>F-FDG PET/CT. *J Nucl Med*. 2003;44:170–176.
10. Chen KY, Cypess AM, Laughlin MR, et al. Brown Adipose Reporting Criteria in Imaging Studies (BARCIST 1.0): recommendations for standardized FDG-PET/CT experiments in humans. *Cell Metab*. 2016;24:210–222.
11. Chondronikola M, Beeman SC, Wahl RL. Non-invasive methods for the assessment of brown adipose tissue in humans. *J Physiol (Lond)*. 2018;596:363–378.
12. Lodge MA. Repeatability of SUV in oncologic <sup>18</sup>F-FDG PET. *J Nucl Med*. 2017;58:523–532.
13. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278:563–577.
14. Aerts HJ. The potential of radiomic-based phenotyping in precision medicine: a review. *JAMA Oncol*. 2016;2:1636–1642.
15. Zhao B, Tan Y, Tsai WY, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci Rep*. 2016;6:23428.
16. Balagurunathan Y, Gu Y, Wang H, et al. Reproducibility and prognosis of quantitative features extracted from CT images. *Transl Oncol*. 2014;7:72–87.
17. Reuzé S, Orhac F, Chargari C, et al. Prediction of cervical cancer recurrence using textural features extracted from <sup>18</sup>F-FDG PET images acquired with different scanners. *Oncotarget*. 2017;8:43169–43179.
18. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14:749–762.
19. Fraum TJ, Crandall JP, Ludwig DR, et al. Repeatability of quantitative brown adipose tissue imaging metrics on positron emission tomography with <sup>18</sup>F-fluorodeoxyglucose in humans. *Cell Metab*. 2019;30:212–224.e4.
20. Pfannenberger C, Werner MK, Ripkens S, et al. Impact of age on the relationships of brown adipose tissue with sex and adiposity in humans. *Diabetes*. 2010;59:1789–1793.
21. Ouellet V, Routhier-Labadie A, Bellemare W, et al. Outdoor temperature, age, sex, body mass index, and diabetic status determine the prevalence, mass, and glucose-uptake activity of <sup>18</sup>F-FDG-detected BAT in humans. *J Clin Endocrinol Metab*. 2011;96:192–199.
22. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging*. 2012;30:1323–1341.
23. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res*. 2017;77:e104–e107.
24. Zwanenburg A, Leger S, Vallieres M, Lock S, Initiative IBS. Image biomarker standardisation initiative. arXiv.org website (Cornell University). <https://arxiv.org/abs/1612.07003>. 2016. Accessed March 12, 2021.
25. Leijenaar RT, Nalbantov G, Carvalho S, et al. The effect of SUV discretization in quantitative FDG-PET radiomics: the need for standardized methodology in tumor texture analysis. *Sci Rep*. 2015;5:11075.
26. Larue RTHM, Defraene G, De Ruyscher D, Lambin P, Van Elmpt W. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. *Br J Radiol*. 2017;90:20160665.
27. Crandall JP, Gajwani P, O JH, Mawhinney DD, Sterzer F, Wahl RL. Repeatability of brown adipose tissue measurements on FDG PET/CT following a simple cooling procedure for BAT activation. *PLoS One*. 2019;14:e0214765.
28. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat*. 2001;29:1165–1188.
29. Litmanovich DE, Tack DM, Shahrzad M, Bankier AA. Dose reduction in cardiothoracic CT: review of currently available methods. *Radiographics*. 2014;34:1469–1489.
30. Midya A, Chakraborty J, Gonen M, Do RKG, Simpson AL. Influence of CT acquisition and reconstruction parameters on radiomic feature reproducibility. *J Med Imaging (Bellingham)*. 2018;5:011020.
31. Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans Pattern Anal Mach Intell*. 1989;11:674–693.
32. Soufi M, Arimura H, Nagami N. Identification of optimal mother wavelets in survival prediction of lung cancer patients using wavelet decomposition-based radiomic features. *Med Phys*. 2018;45:5116–5128.
33. Pfähler E, Beukinga RJ, de Jong JR, et al. Repeatability of F-18-FDG PET radiomic features: a phantom study to explore sensitivity to image reconstruction settings, noise, and delineation method. *Med Phys*. 2019;46:665–678.