Title: Repeatability of Radiomic Features of Brown Adipose Tissue

Running Title: BAT Radiomic Repeatability

Aria Nazeri¹, John P Crandall¹, Tyler J Fraum¹, Richard L Wahl^{1,*}

¹Mallinckrodt Institute of Radiology, Washington University School of Medicine, 510 S.

Kingshighway Blvd, Campus Box 8131, St. Louis, MO 63110

Disclaimer: none

Corresponding Author:

Richard L. Wahl

Mallinckrodt Institute of Radiology

Washington University School of Medicine

510 S. Kingshighway Blvd

Campus Box 8131

St. Louis, MO 63110

Phone: (314) 362 – 7100

Fax: (314) 361 – 5428

Email: rwahl@wustl.edu

First Author:

Aria Nazeri

Mallinckrodt Institute of Radiology

Washington University School of Medicine

510 S. Kingshighway Blvd

Campus Box 8131

St. Louis, MO 63110

Phone: (314) 747 – 5561

Email: aria.nazeri@wustl.edu

Financial Support: None. No potential conflicts of interest relevant to this article exist.

Word Count: 4927

ABSTRACT

The aim of this study was to assess the repeatability of activated brown adipose tissue (BAT) radiomic features. In order 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.

Following a cooling procedure to activate BAT, volunteers underwent ^{18}F -fluorodeoxyglucose positron emission / computed tomography (FDG PET/CT) imaging. Participants underwent repeat imaging using the same imaging protocols and a similar FDG dose within 14 days. Active BAT was segmented using the BARCIST 1.0 methods. Radiomic features were extracted from each region of interest (ROI) on high definition PET (HD PET), non-HD PET, and CT images. Lin's concordance correlation coefficient (CCC) was used to estimate the repeatability of the extracted radiomic features. To determine whether BAT radiomic feature repeatability correlates with BAT maximum standardized uptake value (SUV_{max}) repeatability, participants were stratified based on the relative difference in SUV_{max} between sessions. Non-HD PET repeatable features were clustered together using hierarchical clustering, and the normalized dynamic range (NDR) 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 ($\text{CCC} \geq 0.80$). 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 SUV_{max} , as

participants with lower relative differences in SUV_{max} between initial and repeat 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 SUV_{max} , did not cluster with any other features. **Conclusion:** Several clusters of highly repeatable BAT radiomic features derived from FDG PET/CT appear to provide information regarding BAT activity distinct from SUV_{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

INTRODUCTION

Brown adipose tissue (BAT) is a metabolically active tissue that is the primary facilitator of shivering-independent thermogenesis via uncoupling protein-1 (UCP-1) and glucose/lipid homeostasis (1-3). Cold exposure, β 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}F -fluorodeoxyglucose (FDG) positron emission tomography / computed tomography (PET/CT) examinations performed for oncologic indications (6-9). 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).

Mean, peak, and maximum SUV 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 utilized 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 if repeatable radiomic features provide additional information beyond conventional metrics, such as standardized uptake values (19).

MATERIALS AND METHODS

Subjects

This study was approved by the Washington University IRB, 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 (BMI) have been shown to be negatively correlated with BAT activity (5,20). Therefore, healthy adult volunteers ages 18 to 35 with a BMI under 25 were eligible for this study. Pertinent exclusion criteria included: (i) active infections; (ii)

current or past usage of nicotine products; (iii) a history of cold-related injury; (iv) insulin-dependent diabetes; (v) current or past usage of medications that may interfere with BAT activity, such as beta-blockers; (vi) 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 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 weeks.

BAT Activation Protocol

BAT activation was achieved using a previously reported protocol (19). Figure 1 shows a timeline of the imaging sessions. Upon arrival, participants changed into a hospital gown (clo value of 0.44) and were wrapped in warmed blankets for approximately one hour before the start of the cooling procedure. The cooling procedure was initiated about 60 min before the planned time of FDG injection. Participants were outfitted in a cooling suit (CureWrap, MTRE, Advanced Technologies, Yavne, Israel) that circulates chilled water. At the start of the BAT activation procedure, the cooling system was set at 10° 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° C every 60 seconds until shivering ceased. This cooling protocol continued from 60

min prior to FDG administration until approximately 10 min before FDG PET/CT imaging.

Imaging

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

Image Analysis

Since 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., Cleveland, OH) and according to a modified BARCIST 1.0 criteria: since all participants in the study were lean, a threshold based on SUV was used in place of SUV adjusted for lean body mass (SUL) in order to segregate the activated BAT from the non-active BAT and white adipose tissue (WAT). A SUV_{max} threshold of 1.2 or greater was applied to the supraclavicular fossae in order 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 kg body weight (SUV [g/ml]) and CT images were left as Hounsfield units. 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 (RTSTRUCT) to NRRD format using 3D Slicer v4.8 (<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 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 Co-occurrence Matrix (GLCM), 14 Gray Level Dependence Matrix (GLDM), 16 Gray Level Run Length Matrix (GLRLM), 16 Gray Level Size Zone Matrix (GLSZM), and 5 Neighbouring Gray Tone Difference Matrix (NGTDM) 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 three 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 three 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 to the left side (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 concordance correlation coefficient (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 in order to avoid the additional impact of PSF 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, ρ is the Pearson correlation coefficient, σ is the radiomic feature standard deviation from the indicated visit, and μ is the radiomic feature mean from the indicated visit.

Since 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). In order 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 repeatability, a similar analysis was performed by stratifying subjects into a high BMV relative difference (BMV_{rd}) group and a low BMV_{rd} group using a median BMV_{rd} cutoff. The CCC of each radiomic feature was then recalculated for both SUV_{maxrd} groups and both BMV_{rd} 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 1 SUV_{max} and the visit 2 SUV_{max} .

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).

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

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 v3.4.1 (<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 FDG PET/CT in 18 volunteers. Females

comprised 13 of 18 subjects (72.2%). Subjects had a mean age of 22.9 years (range: 19.3-28.9 years), a mean BMI of 22.6 kg/m² (range: 19.4-26.3 kg/m²), a mean height of 169 cm (range: 152-191 cm), and a mean weight of 65 kg (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. Using a CCC threshold of 0.80, 66/851 (7%) HD PET features and 66/851 (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 ± 0.28) than for CT (mean $\pm$ SD of CCC values, 0.10 ± 0.29). Non-HD PET radiomic feature (mean $\pm$ SD of CCC values, 0.38 ± 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 ($P = 0.521$).

Using features extracted from non-HD PET images, 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 ± 0.31) and z axes (mean $\pm$ SD of CCC values, 0.55 ± 0.25) was significantly higher ($P < 0.001$) than after high pass filter decomposition in y (mean $\pm$ SD of CCC values, 0.26 ± 0.30) and z axes (mean $\pm$ SD of CCC values, 0.13 ± 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 ± 0.4) was significantly ($P < 0.001$) higher than the high SUV_{maxrd} group (mean $\pm$ SD of CCC values, 0.11 ± 0.24). HD PET radiomic feature repeatability of the group with lower SUV_{maxrd} (mean $\pm$ SD of CCC values, 0.54 ± 0.37) was not significantly higher than the group with higher SUV_{maxrd} (mean $\pm$ SD of CCC values, 0.28 ± 0.19). Non-HD PET radiomic feature repeatability of the group with lower SUV_{maxrd} (mean $\pm$ SD of CCC values, 0.42 ± 0.43) was significantly ($P < 0.001$)

higher than the group with higher SUV_{maxrd} (mean $\pm$ SD of CCC values, 0.28 ± 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 ≥ 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 (Figure 6). CT radiomic feature repeatability of the low BMV_{rd} group (mean $\pm$ SD of CCC values, 0.1 ± 0.4) was not significantly ($P < 0.001$) different than the high BMV_{rd} group (mean $\pm$ SD of CCC values, 0.1 ± 0.24). HD PET radiomic feature repeatability of the group with lower BMV_{rd} (mean $\pm$ SD of CCC values, 0.45 ± 0.35) was not significantly different ($P = 1$) from the group with higher BMV_{rd} (mean $\pm$ SD of CCC values, 0.5 ± 0.29) using sign test. Non-HD PET radiomic feature repeatability of the group with lower BMV_{rd} (mean $\pm$ SD of CCC values, 0.41 ± 0.32) was significantly ($P < 0.001$) higher than the group with higher BMV_{rd} (mean $\pm$ SD of CCC values, 0.37 ± 0.4). The stratification of subjects on the basis of SUV_{maxrd} revealed marked differences between subgroups with respect to the number of robust (CCC ≥ 0.8) radiomic features.

For the BMV_{rd} high repeatability subgroup, there were 75, 170, and 168 robust features for CT, HD PET, and non-HD PET, respectively. For the BMV_{rd} 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_{max} and SUV_{mean} , revealed 10 highly correlated ($R^2 > 0.95$) clusters (Figure 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_{max} , 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 physiological 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 non-redundant, 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, FDG dose, imaging

parameters, etc.), the majority of 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 which have shown the repeatability of standard SUV metrics in activated BAT can vary substantially between test and retest studies, since 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 axis. High frequency wavelet decomposition results in enhancement of the most high-contrast (heterogeneous) areas of the ROI, such as the edge (31,32). Additionally, the low-contrast regions become noisier. Since the boundary regions of test-retest BAT ROIs 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_{max} 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_{max} . Hence, repeatable non-HD

PET radiomic features can potentially provide unique (i.e. separate from SUV_{max}) 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_{max} . Textural radiomic features might also be more sensitive for assessing the drug-induced changes in BAT metabolism than activated BAT volume and SUV_{max} , as these two 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 biological 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 biological significance.

CONCLUSION

Several repeatable FDG PET BAT radiomic features were identified, which may provide information regarding BAT metabolic activation distinct from SUV_{max} and these features potentially can be explored as quantitative imaging biomarkers of BAT in future studies.

KEY POINTS

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

Pertinent Findings: Clusters of repeatable radiomic features were identified which 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 brown adipose tissue.

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:E444-452.
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 18F-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*. 2018;596:363-378.
12. Lodge MA. Repeatability of SUV in Oncologic (18)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. Reuze S, Orlhac F, Chargari C, et al. Prediction of cervical cancer recurrence using textural features extracted from 18F-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 (18)F-Fluorodeoxyglucose in Humans. *Cell Metab*. 2019;30:212-224.e214.
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 18F-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. *Magnetic Resonance 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. 2016.

- 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. *British Journal of Radiology*. 2017;90.
- 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. *The Annals of Statistics*. 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. *Journal of Medical Imaging*. 2018;5.
- 31.** Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 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.** Pfaehler 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. *Medical Physics*. 2019;46:665-678.

FIGURES

FIGURE 1. BAT activation and imaging process. Following a 1 hr warming period, volunteers were cooled using a water-perfusion suit for approximately 2 hr to activate BAT. After 1 hr of cooling, FDG was administered and PET/CT imaging was performed immediately following the cooling procedure.

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 the repeatability of first-order and textural (with and without wavelet decompositions) radiomic features extracted from CT, HD PET, and non-HD PET. The horizontal red line corresponds to a CCC threshold of 0.80.

FIGURE 4. Repeatability comparison of radiomic features extracted following 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.

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

FIGURE 6. The repeatability of radiomic features extracted from PET and CT images, with and without stratification based on BMV_{rd} , at different CCC thresholds. The vertical red line corresponds to the 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 the heat map row and column colors (A).

TABLES

TABLE 1. Pairwise comparison of non-HD PET radiomic feature repeatability using post-hoc Nemenyi test.

	Mean ± SD	None	HHH*	HHL	HLH	HLL	LHH	LHL	LLH
None	0.68 ± 0.14	-							
HHH	0.05 ± 0.18	p < 0.001	-						
HHL	0.45 ± 0.18	p < 0.001	p < 0.001	-					
HLH	0.29 ± 0.23	p < 0.001	p < 0.001	p = 1	-				
HLL	0.61 ± 0.29	p = 1	p < 0.001	p = 0.77	p < 0.001	-			
LHH	0.01 ± 0.17	p < 0.001	p = 1	p < 0.001	p < 0.001	p < 0.001	-		
LHL	0.54 ± 0.24	p = 1	p < 0.001	p = 0.28	p < 0.001	p = 1	p < 0.001	-	
LLH	0.18 ± 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 ± 0.24	p = 0.16	p < 0.001	p = 1	p = 0.003	p = 1	p < 0.001	p = 1	p < 0.001

*Abbreviations: H, High-pass; L, Low-pass

Supplemental Figure 1. Box plots overlaid with strip plots comparing the repeatability of non-HD PET first-order and textural (with and without wavelet decompositions) radiomic features. The horizontal red line corresponds to a CCC threshold of 0.8.

Feature Type	Feature Count	Wavelet Decompositions	Total
First-order	19	152	171
Shape	14	0	14
Texture	74	592	666
GLCM	23	184	207
GLDM	14	112	126
GLRLM	16	128	144
GLSZM	16	128	144
NGTDM	5	40	45
Total	107	744	851

Supplemental Table 1. First-order, shape, and textural features were extracted from HD PET, non-HD PET, and CT images. In total, 851 features (including features extracted from each wavelet-decomposed image) were extracted from both PET reconstruction and the CT images.

Supplemental Table 2. Imaging Details.

Item	Description
Region of interest	Right Supraclavicular BAT Depot
Imaging modality	PET/CT
Scanner	Siemens Biograph 40
Radiotracer	18 F-Fluorodeoxyglucose
Administered method	Intravenous injection
Scan type	Static imaging
Contrast agents	None
PET matrix size	168 X 168
PET slice thickness(mm)	3mm
PET reconstruction method 1(Non-HD)	OSEM
Iterations	4
Subsets	8
Smoothing filter	5 mm Gaussian
PET reconstruction method 2(HD)	OSEM with PSF
Iterations	2
Subsets	21
Smoothing Filter	2 mm Gaussian
Attenuation correction	CT based
Randoms correction	Delayed event subtraction
Interporlation method	Trilinear
PET voxel size (mm ³)	4.07 x 4.07 x 3.00 mm ³
CT voxel size(mm ³)	0.98 x 0.98 x 3.00 mm ³
Resampled voxel size for radiomic analysis	3 x 3 x 3 mm ³
Interpolation method for resampling	B-Spline algorithm

Imaging Modality	Feature Family	Original	HHH	HHL	HLH	HLL	LHH	LHL	LLH	LLL	Total
Non-HD PET	First order	2	0	0	0	9	0	1	0	2	14
	GLCM	5	0	0	0	8	0	0	0	3	16
	GLDM	3	0	0	0	4	0	1	0	3	11
	GLSZM	3	0	1	0	3	0	2	0	3	12
	GLRLM	3	0	0	0	4	0	0	0	3	10
	NGTDM	0	0	0	0	2	0	0	0	0	2
	Shape	1	-	-	-	-	-	-	-	-	1
HD PET	First order	2	0	0	0	13	0	0	0	2	17
	GLCM	4	0	0	0	9	0	0	0	5	18
	GLDM	2	0	0	0	4	0	0	0	3	9
	GLSZM	2	0	0	0	4	0	0	0	3	9
	GLRLM	3	0	0	0	4	0	0	0	3	10
	NGTDM	0	0	0	0	2	0	0	0	0	2
	Shape	1	-	-	-	-	-	-	-	-	1
CT	First order	5	0	0	0	0	0	0	0	0	5
	GLCM	0	0	0	0	0	0	0	0	0	0
	GLDM	0	0	0	0	0	0	0	0	0	0
	GLSZM	0	0	0	0	0	0	0	0	0	0
	GLRLM	0	0	0	0	0	0	0	0	0	0
	NGTDM	0	0	0	0	0	0	0	0	0	0
	Shape	1	-	-	-	-	-	-	-	-	1

Supplemental Table 3. Number of radiomic features extracted from non-HD PET, HD PET, and CT with a minimum CCC ≥ 0.80 before and after wavelet decomposition.