Journal of Nuclear Medicine, published on October 2, 2020 as doi:10.2967/jnumed.120.247999

Title: Repeatability of ^{18}F-FDG PET Radiomic Features in Cervical Cancer

Running Title: Repeatability of PET Radiomic Features

Authors: John P Crandall[1], Tyler J Fraum[1], MinYoung Lee[1], Linda Jiang[1], Perry Grigsby[2], Richard L Wahl[1, 2]

[1]Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, MO

[2]Department of Radiation Oncology, Washington University in Saint Louis, St Louis, MO

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:

John P. Crandall

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: jcrandall@wustl.edu

Financial Support: This project was supported by NIH grants U01CA140204, 5P30CA006973, and 1R01CA190299. No potential conflicts of interest relevant to this article exist.

Word Count: 5236

ABSTRACT

Knowledge of the intrinsic variability of radiomic features is essential to the proper interpretation of changes in these features over time. The primary aim of this study was to assess the test-retest repeatability of radiomic features extracted from ^{18}F-Fluorodeoxyglucose (FDG) positron emission tomography (PET) images of cervical tumors. The impact of different image pre-processing methods was also explored. **Methods:** Patients with cervical cancer underwent baseline and repeat FDG PET/CT imaging within 7 days. PET images were reconstructed using 2 methods: ordered subset expectation maximization (PET_{OSEM}) or OSEM with point-spread function (PET_{PSF}). Tumors were segmented to produce whole-tumor volumes of interest (VOI_{WT}) and 40% isocontours (VOI_{40}). Voxels were either left at the default size or resampled to 3 mm isotropic voxels. SUV was discretized to a fixed number of bins (32, 64, or 128). Radiomic features were extracted from both VOIs and repeatability was then assessed using Lin's concordance correlation coefficient (CCC). **Results:** Eleven patients were enrolled and completed the test-retest PET/CT imaging protocol. Shape, neighborhood gray-level difference matrix (NGLDM), and gray-level cooccurence matrix (GLCM) features were repeatable with mean CCC values of 0.81. Radiomic features extracted from PET_{OSEM} images showed significantly better repeatability than features extracted from PET_{PSF} images ($P < 0.001$). Radiomic features extracted from VOI_{40} were more repeatable than features extracted from VOI_{WT} ($P < 0.001$). For most features (78.4%), a change in bin number or voxel size resulted in less than 10% change in feature value. All gray-level emphasis and gray-level run emphasis features showed poor repeatability (CCC values < 0.52) when extracted from VOI_{WT}, but were highly repeatable (mean

CCC values > 0.96) when extracted from VOI_{40}. **Conclusion:** Shape, GLCM, and NGLDM radiomic features were consistently repeatable while gray-level run length matrix (GLRLM) and gray-level zone length matrix (GLZLM) features were highly variable. Radiomic features extracted from 40% isocontours were more repeatable than features extracted from whole-tumor contours. Changes in voxel size or SUV discretization parameters typically resulted in relatively small differences in feature value, though several features were highly sensitive to these changes.

Key Words: FDG, PET, radiomics, repeatability

INTRODUCTION

The role of positron emission tomography (PET) / computed tomography (CT) with ^{18}F-fluorodeoxyglucose (FDG) in oncologic diagnosis, staging, and treatment monitoring is well established (*1,2*). With the increasing use of functional imaging modalities, such as PET/CT, interest in the quantification of image data has grown. Despite having established quantitative frameworks for response assessment, such as the Response Evaluation Criteria in Solid Tumors (RECIST) and PET Response Criteria in Solid Tumors (PERCIST), in clinical practice most interpretations are still based primarily on subjective visual interpretations (*3,4*).

Tumor metabolism is commonly quantified with standardized uptake value (SUV) metrics including the voxel of maximum SUV intensity (SUV_{MAX}), mean tumor SUV (SUV_{MEAN}), and the sphere of highest mean intensity within the tumor (SUV_{PEAK}) (*5*). Radiomic features are complex quantitative imaging biomarkers purported to provide additional information beyond intensity-based SUV metrics (*6*). Extraction of radiomic features typically incorporates image preprocessing, segmentation, and feature calculation. Each of these steps has been shown to affect the outcome of radiomic analyses (*7*).

FDG PET is now part of the standard-of-care for the evaluation of cervical cancer in the United States (*8*). In cervical cancer patients, radiomic features extracted from FDG PET images have been shown to predict both survival and disease recurrence (*9,10*). As PET is also useful for monitoring response to cervical cancer treatment, there is

interest in utilizing radiomic methods to further this purpose (*11-13*). However, in order to use these novel features effectively, their intrinsic variability must be formally quantified. Unfortunately, studies focusing on the repeatability of radiomic features are limited and have shown conflicting results (*14,15*).

The primary aim of this study was to assess the repeatability of radiomic features extracted from PET images of cervical cancer patients. The effects of reconstruction, segmentation, voxel resampling, and SUV discretization methods on repeatability were also explored. These images were collected as part of a repeatability study of several commonly used PET/MRI and PET/CT quantitative imaging metrics (*16*).

MATERIALS AND METHODS

Subjects

This prospective study was approved by the Washington University Institutional Review Board and all volunteers provided written, informed consent (ClinicalTrials.gov identifier: NCT02717572). From June 2016 through May 2017, 17 patients with histologically-proven malignancy were enrolled. Two patients who failed to complete the second imaging session were excluded and another was excluded due to a lack of tumor FDG uptake. For the current study focusing on cervical cancer, 3 patients without cervical cancer were also excluded.

Each patient underwent double baseline FDG PET/CT imaging separated by at least 24 h and at most 7 d. Unless otherwise noted, the imaging procedures used in this study

conformed to the EANM tumor imaging guidelines (*2*). The same PET/CT scanner and approximate FDG dose were used for both imaging sessions. Patients were instructed to fast and to avoid liquids other than water for at least 6 h before the planned FDG administration time. Blood glucose levels were measured immediately prior to the FDG injection.

Imaging

All images were aquired using a Siemens Biograph 40 (Siemens AG; Erlangen, Germany). Static PET data were collected for 15 min over a single station (tumor centered within a 21.6 cm field-of-view), starting 60 – 70 min after intravenous injection of 370 MBq FDG. A CT was performed immediately before PET imaging, utilizing a tube potential of 120 kV, maximum tube current of 80 mAs (CareDose^{TM} tube current modulation was used), pitch of 0.8, and rotation time of 0.5 s. PET images were reconstructed using ordered subset expectation maximization (PET_{OSEM}) or OSEM with point-spread function (PET_{PSF}). Image reconstruction parameters, as well as additional acquisition and post-acquisition details, can be found in Supplemental Table 1.

Image Analysis

MIM version 6.9.3 (MIM Software; Cleveland, OH) was used for image segmentation (Figure 1). For each PET/CT session, the lesion volume of interest (VOI) was manually delineated by one expert nuclear medicine reader to generate a whole-tumor contour (VOI_{WT}). Based on previous work involving segmentation of cervical tumors, a 40% isocontour (VOI_{40}) was also generated, containing all voxels with an SUV ≥ 40% of the

SUV_{MAX} of the whole-tumor contour (*17*). Prior to radiomic feature calculation, PET image intensities were normalized to decay-corrected injected activity per kg body weight (SUV [g/ml]). The effects of SUV discretization and spatial resampling were explored by discretizing to a fixed number of bins (32, 64 or 128 bins), over the full SUV range in each image, and either no spatial resampling (NSR; 4.07 mm x 4.07 mm x 5.00 mm) or resampling to 3 mm isotropic voxels (ISR). Radiomic features were then extracted using LIFEx version 5.28 (https://www.lifexsoft.org), which complies with the imaging biomarker standardization initiative (IBSI) recommendations (*18*). Intensity, shape, and textural features (Table 1) were calculated from each VOI. Detailed descriptions, formulas, and computation parameters for each feature can be found on the LIFEx website.

Statistical Methods

Repeatability was assessed using Lin's concordance correlation coefficient (CCC), which provides a quantification of agreement between two repeated measures (*19*). In this study, a CCC value less than 0.80 was considered unrepeatable, greater than or equal to 0.80 repeatable, and greater than or equal to 0.95 highly repeatable. Including the repeatability calculations for all reconstruction, segmentation, voxel resampling, and SUV discretization methods, 24 CCC values were generated for each radiomic feature.

Paired *t* tests or Wilcoxon signed-rank tests were used to assess groupwise repeatability differences. Group normality was evaluated using D'Agostino-Pearson tests. Data analyses were performed using R version 3.6.2 (http://cran.r-project.org/*)*

and Excel 2016 (Microsoft Corporation). A *P*-value less than 0.05 was considered significant, unless otherwise indicated. A Bonferroni correction for mulitple comparisons was applied when necessary to control for type I errors.

RESULTS

A total of 17 patients were enrolled and test-retest images from 11 female patients (Table 2) were eligible for radiomic analysis. The median time between imaging sessions was 2 d with a range of 1 – 7 d. Blood glucose levels (mean ± SD) were 92.2 ± 8.1 mg/dl and 93.2 ± 20.9 mg/dl prior to FDG administration during visits 1 and 2, respecitvely. Mean administered FDG dose was 367.7 ± 20.2 MBq during visit 1 and 371.0 ± 16.8 MBq during visit 2. Mean visit 1 FDG uptake time was 60.4 ± 1.4 min and mean visit 2 FDG uptake time was 61.4 ± 3.1 min.

Radiomic Feature Repeatability

When assessed as groups, standard intensity, shape, NGLDM, and GLCM features showed consistent repeatability with mean CCC values greater than 0.80 (Figure 2). Nonstandard intensity, GLZLM, and GLRLM features were mostly mostly unrepeatable (56.7% of CCC values less than 0.80) when extracted from PET_{PSF} images (Figure 3). Nonstandard intensity, GLZLM, and GLRLM features were also mostly unrepeatable (55.8% of CCC values less than 0.80) when extracted from VOI_{WT} segmentations.

Highly repeatable radiomic features (mean CCC values above 0.95) included MTV, compacity, entropy (GLCM-calculated), GLNU (GLRLM-calculated), RLNU,

Coarseness, LZHGE, GLNU (GLZLM-caluclated), and ZLNU (Supplemental Tables 2 and 3). Nearly all (98.6%) GLCM CCC values were greater than 0.80 and many (61.8%) were greater than 0.95. Overall, the combination of PET_{OSEM} with isotropic resampling, SUV discretized to 64 bins, and tumors segmented using a 40% isocontour resulted in the most stable features (76.7% were repeatable and 39.5% highly repeatable). The fewest repeatable radiomic features resulted from PET_{PSF} images with no spatial resampling, SUV discretized to 128 bins, and tumors segmented manually (39.5% were repeatable and 20.9% highly repeatable).

PET Reconstruction Method

Mean CCC values of radiomic features extracted from PET_{OSEM} and PET_{PSF} images are provided in Table 3. Shape and textural (i.e. GLCM, GLRLM, NGLDM, and GLZLM) features extracted from PET_{OSEM} images had significantly higher CCC values than those extracted from PET_{PSF} images (0.86 and 0.79, respectively; $P < 0.001$; Figure 4A). GLRLM and GLZLM features were impacted by reconstruction method more than other features (Figure 5A). Standard intensity-based features extracted from both PET reconstructions had approximately the same mean CCC value (0.87; $P = 0.494$; Figure 2).

Segmentation Method

Radiomic features extracted from whole-tumor segmentations were less repeatable than those extracted from 40% isocontours (Tables 4 and 5). Using a CCC cutoff of 0.80, 56.2% of PET_{PSF} VOI_{WT} features and 65.5% of VOI_{40} features were repeatable.

When extracted from PET_{OSEM} images, 68.6% of VOI_{WT} and 89.9% of VOI_{40} features were repeatable. Shape, GLCM, GLRLM, and NGLDM feature groups extracted from PET_{PSF} images were found to have significantly (all *P*-values < 0.015) lower CCC values when extracted from VOI_{WT} than when extracted from VOI_{40}. Nonstandard intensity, GLCM, GLRLM, NGLDM, and GLZLM features extracted from PET_{OSEM} images were found to have significantly lower (all *P*-values < 0.023) CCC values when extracted from VOI_{WT} than when extracted from VOI_{40} images.

Spatial Resampling

The repeatability of most radiomic features was robust against spatial resampling changes with 78.5% of features showing less than 5% relative difference in CCC value between resampling methods (Figures 4C and 5C). Features extracted from PET_{PSF} images were more sensitive to spatial resampling changes than those extracted from PET_{OSEM} images (Tables 6 and 7). When extracted from VOI_{WT}, GLRLM features showed greater repeatability following isotropic resampling (*P*-values < 0.045). The opposite was true when extracted from VOI_{40}, with GLRLM features showing greater repeatability before spatial resampling (*P*-values < 0.005). Few radiomic feature groups (21%) were found to have mean CCC values that differed by more than 0.03 before and after spatial resampling.

SUV Discretization

The repeatability of most radiomic features showed little sensitivity to changes in SUV discretization (Figures 4D and 5D). The majority (60.3%) of textural feature CCC values

varied by less than 5% among bin number groups. GLCM feature repeatability was largely insensitive to SUV discretization changes with 80.2% of CCC values varying less than 5% within SUV bin groups. GLZLM features were considerably impacted by changes in the number of SUV bins with 69.3% of CCC values varying by 5% or more (and by as much as 228.6%) across SUV bin groups.

Lesion Volume Analysis

The absolute relative difference between test-retest radiomic values was correlated with mean test-retest MTV in order to assess the influence of lesion volume on feature repeatability. After controlling for multiple comparisons, there were no significant correlations between MTV and absolute relative difference in feature value (Supplemental Tables 4 and 5).

DISCUSSION

The increasing examination of alternative imaging biomarkers, such as radiomic features, requires an understanding of their test-retest repeatability. The calculation of these features involves several steps including image reconstruction, segmentation, and preprocessing, which have all been shown to impact feature stability (*15*). In the current study, patients with cervical cancer underwent double baseline FDG PET/CT studies. PET images were reconstructed using two different methods and tumors were delineated manually and with a semi-automated technique. Radiomic features were then extracted following various image preprocessing methods.

Features calculated using run length and zone length matrices were found to have low repeatability, while shape, GLCM, and NGLDM features were consistently repeatable. The repeatability of most GLRLM and GLZLM features was quite sensitive to changes in any step of the radiomic extraction process. Tixier et al similarly concluded that GLCM features were repeatable and GLZLM features, extracted from images of esophageal tumors, were unrepeatable (*20*). GLZLM features also showed poor repeatability when extracted from PET images of lung cancer using an OSEM reconstruction, 50% isocontour segmentation, and isotropic spatial resampling (*21*). In our study, GLZLM were mostly unrepeatable, but using an OSEM reconstruction and 40% isocontour, 89.4% of GLZLM CCC values were designated as repeatable.

Two feature groups, GLRLM and GLZLM, were not homogenous in terms of repeatability within each group. Gray-level emphasis and gray-level run emphasis features within the GLRLM group showed poor repeatability. However, the remaining GLRLM features (non-uniformity, run percentage, and run-emphasis calculations) were highly repeatable. Likewise, most GLZLM features were unrepeatable and highly sensitive to changes in image reconstruction, segmentation, and preprocessing methods. However, GLZLM non-uniformity and run percentage features were consistently repeatable.

GLCM features were consistently repeatable in this study, which has also been described in other work (*20-22*). Entropy, in particular, has been consistently found to be reproducible and repeatable, as well as a significant predictor of patient response

(*7,23,24*). In our study, entropy was consistently repeatable, though somewhat sensitive to SUV discretization. When SUV was discretized to 32 bins, mean entropy CCC was 0.90, but increased to 0.97 when SUV was discretized to 64 or 128 bins.

The PET reconstruction methods employed here had a substantial impact on repeatability. This impact was intensified by segmentation method, as shown in Figure 4. Features that were mostly unrepeatable otherwise showed mostly high repeatability when an OSEM reconstruction and a 40% isocontour were used. Yan et al found that 5% to 56% of textural features showed a large variation between values (920%) when reconstruction settings were varied (*25*). Reconstruction method was also found by Gallivanone et al to have a strong impact on the stability of radiomic features (*26*). Using an anthropomorphic phantom, test-retest PET images were acquired and reconstructed using OSEM, OSEM with PSF, and OSEM with PSF and time-of-flight. They concluded fewer than 20% of radiomic features were robust against changes in reconstruction method.

Numerous segmentation methods have been combined with radiomic analyses. One study segmented phantom ROIs using an adaptive Bayesian method or a 60% isocontour and found less than 20% of radiomic features were stable between these segmentations (*26*). Using a 40% isocontour significantly improved the repeatability of most radiomic features in our study, especially when used on images reconstructed without PSF. This improvement was particularly dramatic in some features otherwise found to be unrepeatable. The median CCC values of GLRLM and GLZLM features only

rose above the repeatability cutoff when using a 40% isocontour and all features within both groups were repeatable when also extracted from PET_{OSEM} images. Automated or semi-automated segmentation approaches help to prevent the selection of normal tissue near tumor edges, but often underestimate volume and may not capture regions of necrosis. While this study shows a significant improvement in repeatability with the implementation of an SUV threshold, the utility of textural analyses may suffer when using such methods. Though the threshold applied in this study was based on its previous use in this patient population, other segmentation methods (including fixed or adaptive approaches) may further improve radiomic repeatability and should be evaluated.

Interpolation to isotropic voxel sizes has been routinely employed in radiomic analyses (*27,28*). Since many features are sensitive to changes in voxel size, isotropic resampling effects radiomic feature value and therefore, reproducibility (*29,30*). However, the effect of voxel size resampling on radiomic feature repeatability has not been previously studied and a standard approach to spatial resampling in the context of radiomic analysis has not been described. In the current study, voxels were either not resampled or downsampled to 3mm isotropic voxels. This method requires information inference, while upsampling involves information loss. Here, most features showed similar repeatability between resampling methods, though GLRLM and GLZLM features showed low mean CCC values and high CCC value differences between resampling methods.

Resampling image intensity values is a crucial step in radiomic analysis. Downsampling from a nearly unlimited set of intensity values to a discrete number of bins effectively reduces image noise and allows for comparison of radiomic values between image datasets. SUV discretization has been consistently shown to have a significant impact on radiomic feature value (*7,23*). Though the feature value may fluctuate with varying bin numbers or sizes, consistent repeatability of a discretization method may still allow for effective comparison between repeated evaluations. We found most feature groups were consistently repeatable among SUV bin sizes with GLCM features performing particularly well, and GLZLM quite poorly. Based on the current study and other recent radiomic repeatability studies, GLZLM features may not be useful for test-retest assessments (*7,31*). In this study, we explored varying a fixed number of SUV bins. Another approach is to vary the size of the bins and future studies should explore the repeatability of this alternative method.

This study is limited by a small sample size, though the number of patients included was similar to other published radiomic repeatability studies (*20,21*). Future studies are needed to validate if these results represent radiomic features extracted from cervical tumors, generally. When performing this study, we attempted to adhere to the IBSI guidelines, which only recently became available. Performing additional repeatability studies of other cancer types using the methods reported here, and guided by the IBSI recommendations, may be informative. It is important to note that repeatability thresholds used in this study, as in test-retest studies generally, are somewhat arbitrary and only intended to illustrate which features may be sensitive to changes in radiomic

and image analysis methods. Additional sources of radiomic feature variability should also be explored since any of the various factors that impact SUV repeatability could, in principle, affect radiomic feature stability.

CONCLUSION

Certain radiomic feature groups (shape, GLCM, and NGLDM) were repeatable, while others (GLRLM and GLZLM features) were highly variable. Using a fixed-threshold segmentation method increased the repeatability of most features. Most often, changes in voxel size or SUV discretization parameters resulted in relatively small differences in feature value, though several features were highly sensitive to these changes.

KEY POINTS

Question: Are radiomic features extracted from PET images of cervical cancer repeatable?

Pertinent Findings: Utilizing certain image preprocessing techniques, many radiomic features (especially shape, GLCM, and NGLDM features) are stable and repeatable. The impact of PET reconstruction and segmentation methods on radiomic feature repeatability can be substantial.

Implications for Patient Care: Certain radiomic features are repeatable and may be useful in the management of cervical cancer patients.

REFERENCES

1. Hess S, Blomberg BA, Zhu HJ, Hoilund-Carlsen PF, Alavi A. The pivotal role of FDG-PET/CT in modern medicine. *Acad Radiol*. 2014;21:232-249.

2. Boellaard R, Delgado-Bolton R, Oyen WJ, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging*. 2015;42:328-354.

3. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: Evolving Considerations for PET response criteria in solid tumors. *J Nucl Med*. 2009;50 Suppl 1:122S-150S.

4. Schwartz LH, Litiere S, de Vries E, et al. RECIST 1.1-Update and clarification: From the RECIST committee. *Eur J Cancer*. 2016;62:132-137.

5. Lodge MA. Repeatability of SUV in Oncologic (18)F-FDG PET. *J Nucl Med*. 2017;58:523-532.

6. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48:441-446.

7. Altazi BA, Zhang GG, Fernandez DC, et al. Reproducibility of F18-FDG PET radiomic features for different cervical tumor segmentation methods, gray-level discretization, and reconstruction algorithms. *J Appl Clin Med Phys*. 2017;18:32-48.

8. Koh WJ, Abu-Rustum NR, Bean S, et al. Cervical Cancer, Version 3.2019, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw*. 2019;17:64-84.

9. Chen SW, Shen WC, Hsieh TC, et al. Textural features of cervical cancers on FDG-PET/CT associate with survival and local relapse in patients treated with definitive chemoradiotherapy. *Sci Rep*. 2018;8:11859.

10. 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.

11. Lin AJ, Dehdashti F, Grigsby PW. Molecular imaging for radiotherapy planning and response assessment for cervical cancer. *Semin Nucl Med*. 2019;49:493-500.

12. Schwarz JK, Grigsby PW, Dehdashti F, Delbeke D. The role of 18F-FDG PET in assessing therapy response in cancer of the cervix and ovaries. *J Nucl Med*. 2009;50 Suppl 1:64s-73s.

13. Sarabhai T, Tschischka A, Stebner V, et al. Simultaneous multiparametric PET/MRI for the assessment of therapeutic response to chemotherapy or concurrent chemoradiotherapy of cervical cancer patients: Preliminary results. *Clin Imaging*. 2018;49:163-168.

14. Traverso A, Wee L, Dekker A, Gillies R. Repeatability and reproducibility of radiomic features: a systematic review. *Int J Radiat Oncol Biol Phys*. 2018;102:1143-1158.

15. Zwanenburg A. Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis. *Eur J Nucl Med Mol Imaging*. 2019;46:2638–2655.

16. Fraum TJ, Fowler KJ, Crandall JP, et al. Measurement repeatability of (18)F-FDG-PET/CT versus (18)F-FDG-PET/MRI in solid tumors of the pelvis. *J Nucl Med*. 2019;60:1080–1086.

17. Olsen JR, Esthappan J, DeWees T, et al. Tumor volume and subvolume concordance between FDG-PET/CT and diffusion-weighted MRI for squamous cell carcinoma of the cervix. *J Magn Reson Imaging*. 2013;37:431-434.

18. Nioche C, Orlhac F, Boughdad S, et al. LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res*. 2018;78:4786-4789.

19. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989;45:255-268.

20. Tixier F, Hatt M, Le Rest CC, Le Pogam A, Corcos L, Visvikis D. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET. *J Nucl Med*. 2012;53:693-700.

21. Leijenaar RT, Carvalho S, Velazquez ER, et al. Stability of FDG-PET radiomics features: an integrated analysis of test-retest and inter-observer variability. *Acta Oncol*. 2013;52:1391-1397.

22. Desseroit MC, Tixier F, Weber WA, et al. Reliability of PET/CT shape and heterogeneity features in functional and morphologic components of non-small cell lung cancer tumors: a repeatability analysis in a prospective multicenter cohort. *J Nucl Med*. 2017;58:406-411.

23. 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.

24. Tixier F, Le Rest CC, Hatt M, et al. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med*. 2011;52:369-378.

25. Yan J, Chu-Shern JL, Loi HY, et al. Impact of image reconstruction settings on texture features in 18F-FDG PET. *J Nucl Med*. 2015;56:1667-1673.

26. Gallivanone F, Interlenghi M, D'Ambrosio D, Trifiro G, Castiglioni I. Parameters influencing PET imaging features: a phantom study with irregular and heterogeneous synthetic lesions. *Contrast Media Mol Imaging*. 2018;2018:5324517.

27. Zwanenburg A, Leger S, Vallieres M, Lock S, Initiative IBS. Image biomarker standardisation initiative. 2016;arXiv:1612.07003..

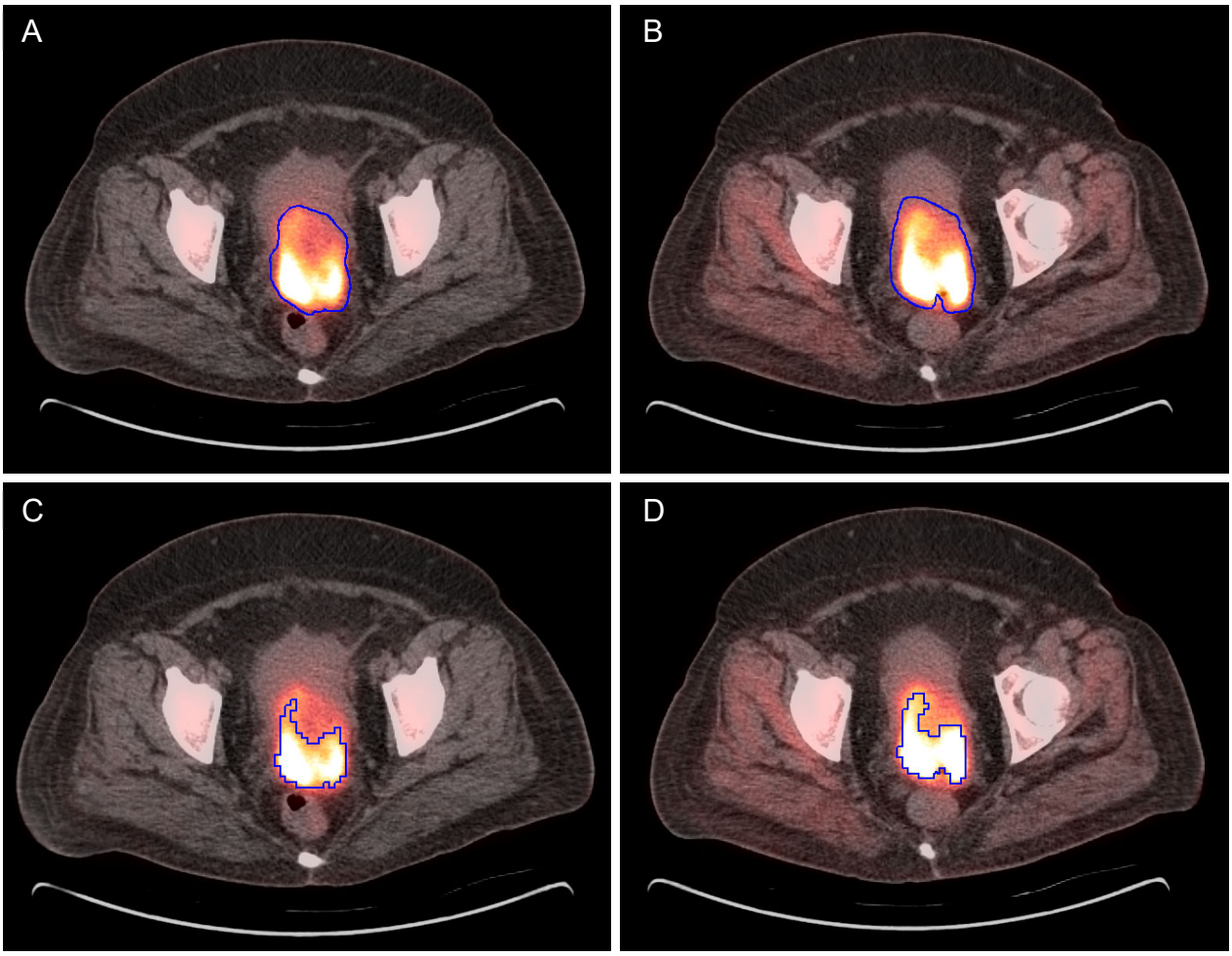
28. Larue RT, Defraene G, De Ruysscher D, Lambin P, van Elmpt W. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. *Br J Radiol*. 2017;90:20160665.

29. Bailly C, Bodet-Milin C, Couespel S, et al. Revisiting the robustness of PET-based textural features in the context of multi-centric trials. *PLoS One*. 2016;11:e0159984.

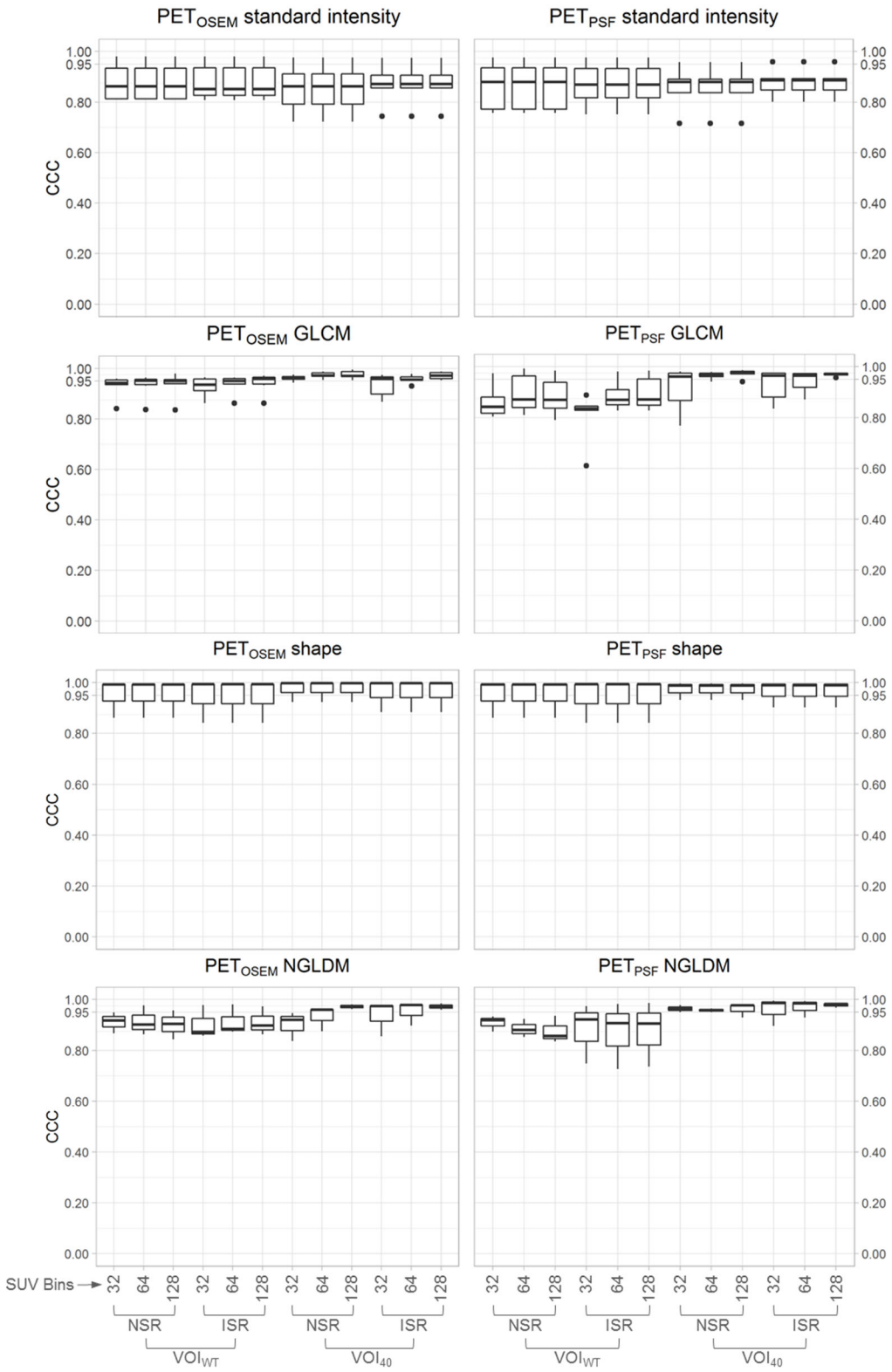
30. Shiri I, Rahmim A, Ghaffarian P, Geramifar P, Abdollahi H, Bitarafan-Rajabi A. The impact of image reconstruction settings on 18F-FDG PET radiomic features: multi-scanner phantom and patient studies. *Eur Radiol*. 2017;27:4498-4509.

31. Papp L, Rausch I, Grahovac M, Hacker M, Beyer T. Optimized feature extraction for radiomics analysis of (18)F-FDG PET imaging. *J Nucl Med*. 2019;60:864-872.

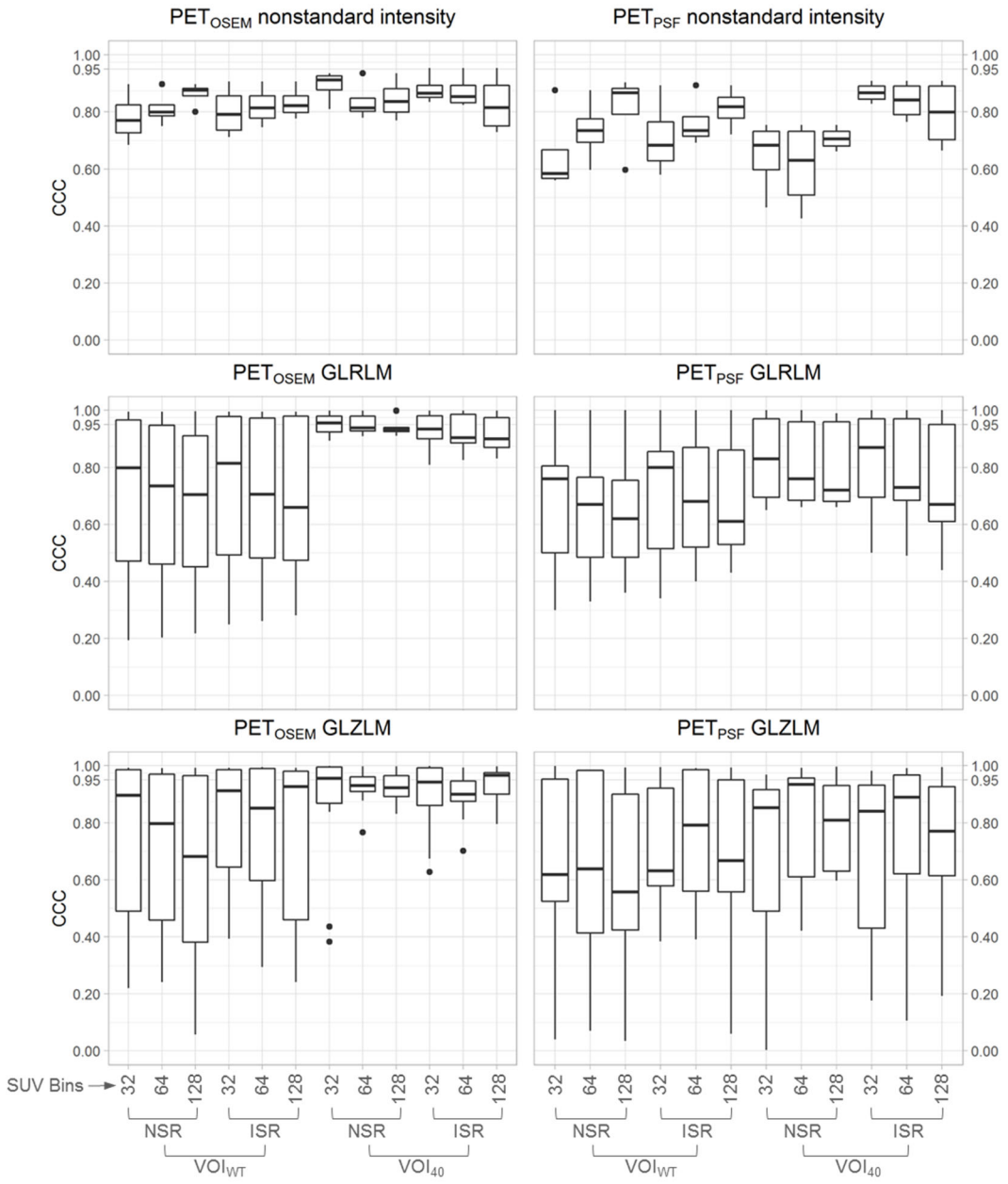
FIGURES



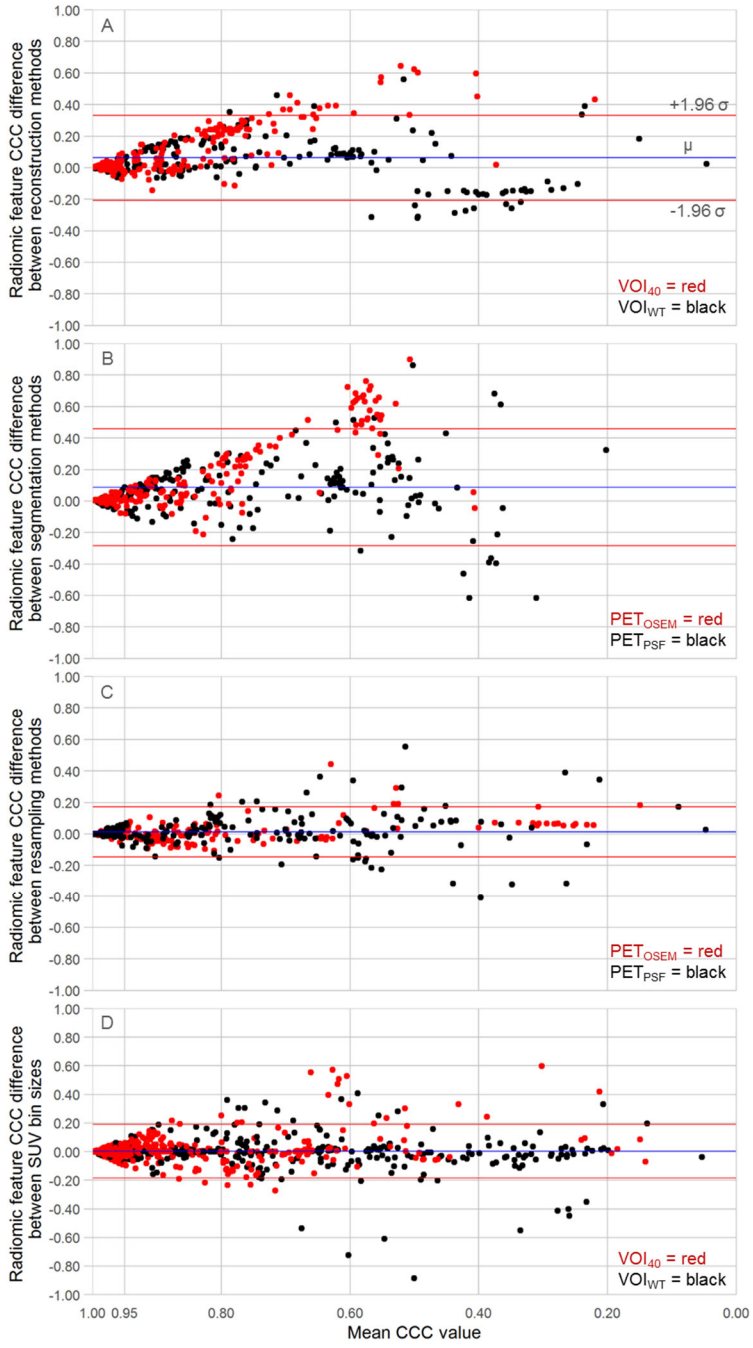
[FIGURE 1. Representative Images. Tumors were manually delineated to generate whole-tumor contours (baseline image, A; repeat image, B). A 40% isocontour was created by removing all voxels with SUVs ≤ 40% of the SUV_{MAX} of the whole-tumor contour (baseline image, C; repeat image, D).]



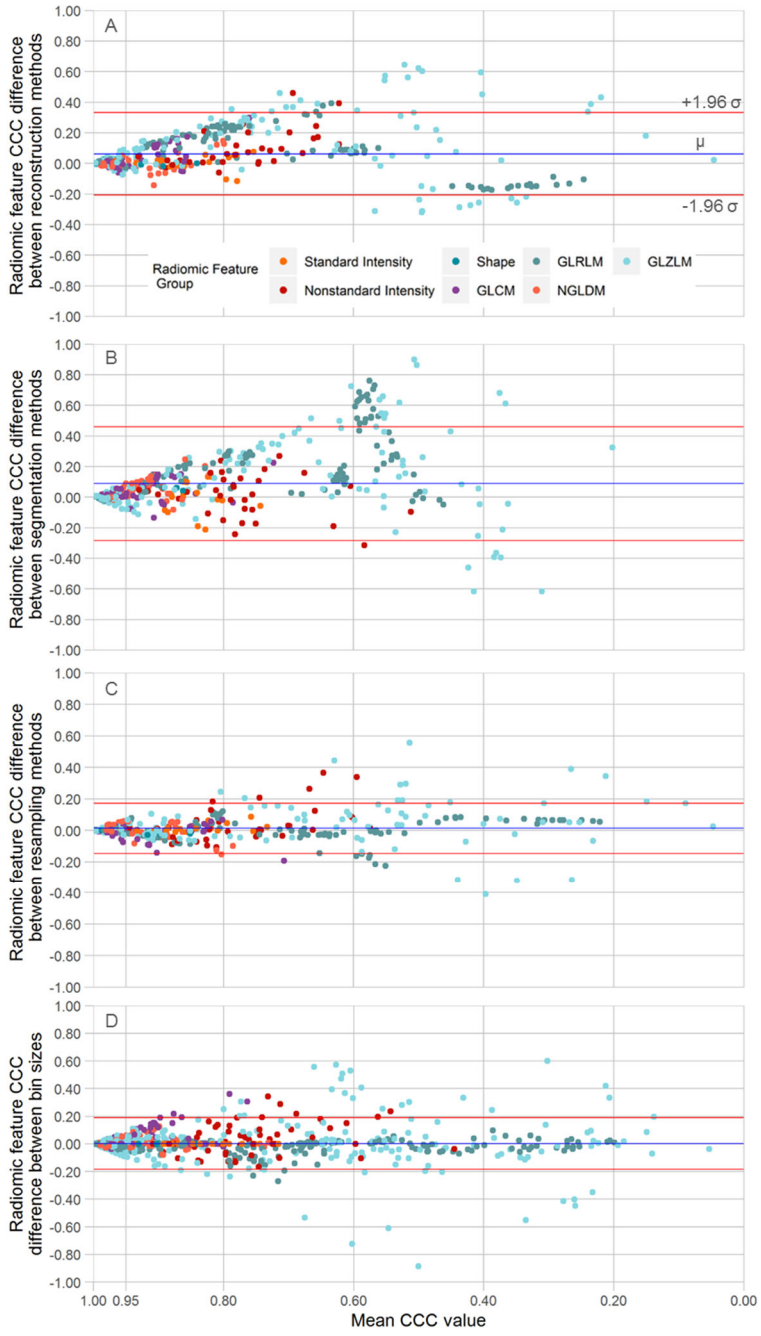
[FIGURE 2. Repeatable radiomic feature groups. Boxplots show repeatability ranges for all features within the specified feature group.]



[FIGURE 3. Unrepeatable radiomic feature groups.]



[FIGURE 4. Bland-Altman plots show the effect of changes in reconstruction (A; PET_{PSF} CCC subtracted from PET_{OSEM} CCC), segmentation (B; VOI_{WT} CCC subtracted from VOI_{40} CCC), spatial resampling (C; NSR CCC subtracted from ISR CCC), and SUV discretization (D; 32 bin CCC subtracted from 64 bin CCC, 32 bin CCC subtracted from 128 bin CCC, and 64 bin CCC subtracted from 128 bin CCC) on radiomic feature repeatability. Plots were generated by calculating the mean of, and difference between, corresponding radiomic feature CCC values of each method.]



[FIGURE 5. Bland-Altman plots show how changes in reconstruction (A), segmentation (B), spatial resampling (C), and SUV discretization (D) affected the repeatability of radiomic feature groups.]

TABLES

[TABLE 1. Radiomic feature groups.]

Radiomic Group	No. of Features	Radiomic Feature Names
Intensity (Standard)	5	maximum standardized uptake value (SUV_{MAX}), mean SUV (SUV_{MEAN}), mean SUV of a sphere of 10 mm diameter (SUV_{PEAK}), SUV standard deviation (SUV_{SD}), total lesion glycolysis (TLG)
Intensity (Nonstandard)	4	skewness, kurtosis, entropy, uniformity
Shape	3	metabolic tumor volume (MTV), sphericity, compacity
Gray-Level Co-occurrence Matrix (GLCM)	6	homogeneity, energy, contrast, correlation, entropy, dissimilarity
Gray-Level Run Length Matrix (GLRLM)	11	short-run emphasis (SRE), long-run emphasis (LRE), low gray-level run emphasis (LGRE), high gray-level run emphasis (HGRE), short-run low gray-level emphasis (SRLGE), short-run high gray-level emphasis (SRHGE), long-run low gray-level emphasis (LRLGE), long-run high gray-level emphasis (LRHGE), gray-level non-uniformity (GLNU), run length non-uniformity (RLNU), run percentage (RP)
Neighborhood Gray-Level Difference Matrix (NGLDM)	3	coarseness, contrast, busyness
Gray-Level Zone Length Matrix (GLZLM)	11	short-zone emphasis (SZE), long-zone emphasis (LZE), low gray-level zone emphasis (LGZE), high gray-level zone emphasis (HGZE), short-zone low gray-level emphasis (SZLGE), short-zone high gray-level emphasis (SZHGE), long-zone low gray-level emphasis (LZLGE), long-zone high gray-level emphasis (LZHGE), gray-level non-uniformity (GLNU), zone length non-uniformity (ZLNU), zone percentage (ZP)

[TABLE 2. Patient characteristics.]

Characteristic	Value
Age (yr)	46.8 ± 11.0*
Height (m)	1.66 ± 0.06*
Weight (kg)	77.8 ± 17.3*
Race/ethnicity	
Caucasian (non-Hispanic)	9 (82)†
Caucasian (Hispanic)	1 (9)†
African-American	1 (9)†

*Reported as mean ± standard deviation

†Reported as frequency (percentage)

[TABLE 3. A comparison of radiomic feature repeatability between PET image reconstruction methods.]

Feature Group	Feature	Mean PET_{PSF} CCC	Mean PET_{OSEM} CCC	*P*
Intensity (Standard)	SUV_{MEAN}	0.82	0.86	< 0.001*
	SUV_{SD}	0.89	0.83	0.066
	SUV_{MAX}	0.88	0.86	< 0.001*
	SUV_{PEAK}	0.78	0.82	< 0.001*
	TLG	0.97	0.98	< 0.001*
Intensity (Nonstandard)	Skewness	0.85	0.92	0.012
	Kurtosis	0.74	0.83	0.003
	Entropy	0.69	0.80	0.040
	Energy	0.70	0.80	0.003
Shape	MTV	0.99	1.00	0.313
	Sphericity	0.88	0.88	0.016
	Compacity	0.99	1.00	0.313
GLCM	Homogeneity	0.89	0.97	< 0.001*
	Energy	0.89	0.94	0.110
	Contrast	0.90	0.95	0.176
	Correlation	0.93	0.91	< 0.001*
	Entropy	0.93	0.96	0.266
	Dissimilarity	0.91	0.96	0.176
GLRLM	SRE	0.89	0.96	0.007
	LRE	0.89	0.97	0.002
	LGRE	0.52	0.58	0.176
	HGRE	0.62	0.78	< 0.001*
	SRLGE	0.52	0.59	0.176
	SRHGE	0.62	0.78	< 0.001*
	LRLGE	0.55	0.58	0.569
	LRHGE	0.71	0.82	0.002
	GLNU	0.99	1.00	< 0.001*
	RLNU	0.99	0.99	0.036
	RP	0.89	0.96	0.009
NGLDM	Coarseness	0.96	0.97	0.493
	Contrast	0.88	0.89	0.522
	Busyness	0.93	0.92	0.092
GLZLM	SZE	0.54	0.74	0.005
	LZE	0.92	0.97	0.036
	LGZE	0.43	0.56	0.176
	HGZE	0.54	0.73	< 0.001*
	SZLGE	0.44	0.53	0.447
	SZHGE	0.59	0.77	< 0.001*
	LZLGE	0.55	0.77	< 0.001*
	LZHGE	0.97	0.98	0.970
	GLNU	0.98	0.99	0.009
	ZLNU	0.94	0.98	0.003
	ZP	0.83	0.94	0.007

*Significant *P*-value, based on a paired Wilcoxon signed-rank test. After correcting for multiple comparisons, a *P*-value less than 0.001 was considered significant.

[TABLE 4. Groupwise comparison of the repeatability of radiomic features extracted from PET_{OSEM} images using 2 different segmentation methods.]

Feature Group	Mean VOI_{WT} CCC*	Mean VOI_{40} CCC	*P*
Intensity (Standard)	0.88	0.86	0.057
Intensity (Nonstandard)	0.82	0.86	0.010
Shape	0.95	0.97	0.140
GLCM	0.93	0.96	< 0.001†
GLRLM	0.70	0.94	< 0.001†
NGLDM	0.91	0.94	0.023
GLZLM	0.73	0.90	< 0.001†

*Mean value calculation includes CCC values from both resampling methods and each SUV bin level (i.e. 6 CCC values per feature per segmentation).

†Significant *P*-value, based on a paired *t* test. After correcting for multiple comparisons, a *P*-value less than 0.004 was considered significant.

[TABLE 5. Groupwise comparison of the repeatability of radiomic features extracted from PET_{PSF} images using 2 different segmentation methods.]

Feature Group	Mean VOI_{WT} CCC	Mean VOI_{40} CCC	*P*
Intensity (Standard)	0.87	0.87	0.984
Intensity (Nonstandard)	0.75	0.74	0.784
Shape	0.95	0.97	0.015
GLCM	0.87	0.95	< 0.001*
GLRLM	0.68	0.81	< 0.001*
NGLDM	0.88	0.96	< 0.001*
GLZLM	0.68	0.73	0.139

*Significant *P*-value, based on a paired *t* test. After correcting for multiple comparisons, a *P*-value less than 0.004 was considered significant.

[TABLE 6. Groupwise comparisons of repeatability of radiomic features extracted from PET_{OSEM} images using 2 different spatial resampling methods.]

	Mean VOI_{WT} CCC			**Mean VOI_{40} CCC**		
Resampling	**NSR**	**ISR**	***P***	**NSR**	**ISR**	***P***
Intensity (Standard)	0.88	0.88	0.857	0.85	0.87	0.021
Intensity (Nonstandard)	0.82	0.82	0.970	0.86	0.86	0.890
Shape	0.95	0.94	0.090	0.97	0.96	0.250
GLCM	0.93	0.94	0.333	0.97	0.95	0.068
GLRLM	0.69	0.71	0.005	0.95	0.92	0.001†
NGLDM	0.91	0.91	0.928	0.93	0.95	0.031
GLZLM	0.69	0.76	< 0.001*	0.90	0.90	0.249

*Mean value calculation includes CCC values from each SUV bin level (i.e. 3 CCC values per feature per resampling method.

†Significant *P*-value, based on a paired Wilcoxon signed-rank test. After correcting for multiple comparisons, a *P*-value less than 0.002 was considered significant.

[TABLE 7. Groupwise comparison of repeatability of radiomic features extracted from PET_{PSF} images using 2 different spatial resampling methods.]

	Mean VOI_{WT} CCC			**Mean VOI_{40} CCC**		
Resampling	**NSR**	**ISR**	***P***	**NSR**	**ISR**	***P***
Intensity (Standard)	0.86	0.87	0.670	0.86	0.88	0.376
Intensity (Nonstandard)	0.73	0.76	0.156	0.65	0.83	< 0.001*
Shape	0.95	0.94	0.683	0.97	0.96	0.729
GLCM	0.88	0.86	0.259	0.95	0.95	0.490
GLRLM	0.66	0.70	< 0.001*	0.82	0.79	0.045
NGLDM	0.89	0.88	0.655	0.96	0.97	0.345
GLZLM	0.63	0.72	0.001*	0.75	0.71	0.077

*Significant *P*-value, based on a paired Wilcoxon signed-rank test. After correcting for multiple comparisons, a *P*-value less than 0.002 was considered significant.

[SUPPLEMENTAL TABLE 1. Imaging details (per the IBSI guidelines).]

Item	Description
Region of interest	Cervical tumors
Imaging modality	PET/CT
Scanner	Siemens Biograph 40
Radiotracer	18F-Fluorodeoxyglucose
Administration method	Intravenous injection
Injected activity	370 MBq (fixed dose)
Radiotracer uptake time	60 min
Contrast agents	None
Scan type (dynamic/static)	Static imaging
Scanner calibration schedule	Daily
Scan duration	15 min
Time-of-flight	Not used
PET matrix size (pixels)	168 x 168
PET slice thickness (mm)	5
PET reconstruction method 1	OSEM
Iterations	4
Subsets	8
Smoothing filter	5 mm Gaussian
PET reconstruction method 2	OSEM with PSF
Iterations	2
Subsets	21
Smoothing filter	2 mm Gaussian
Attenuation correction	CT-based
Scatter correction	Model-based
Randoms correction	Delayed event subtraction
Interpolation method	Trilinear

[SUPPLEMENTAL TABLE 2. Test-Retest repeatability of PET_{OSEM} radiomic features (CCC).]

	VOI	Mean VOI_{WT} CCC						Mean VOI_{40} CCC					
	Resampling	NSR			ISR			NSR			ISR		
	SUV Bins	32	64	128	32	64	128	32	64	128	32	64	128
Intensity (Standard)	SUV_{MEAN}	0.81	0.81	0.81	0.81	0.81	0.81	0.91	0.91	0.91	0.91	0.91	0.91
	SUV_{SD}	0.93	0.93	0.93	0.94	0.94	0.94	0.72	0.72	0.72	0.74	0.74	0.74
	SUV_{MAX}	0.86	0.86	0.86	0.85	0.85	0.85	0.86	0.86	0.86	0.87	0.87	0.87
	SUV_{PEAK}	0.81	0.81	0.81	0.83	0.83	0.83	0.79	0.79	0.79	0.86	0.86	0.86
	TLG	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.97	0.97
Intensity (Non-standard)	Skewness	0.90	0.90	0.90	0.91	0.91	0.91	0.94	0.94	0.94	0.95	0.95	0.95
	Kurtosis	0.80	0.80	0.80	0.84	0.84	0.84	0.81	0.81	0.81	0.87	0.87	0.87
	Entropy	0.68	0.75	0.87	0.71	0.75	0.78	0.92	0.82	0.86	0.84	0.82	0.76
	Energy	0.74	0.80	0.88	0.74	0.79	0.81	0.90	0.78	0.77	0.86	0.83	0.73
Shape	MTV	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00
	Sphericity	0.86	0.86	0.86	0.84	0.84	0.84	0.92	0.92	0.92	0.88	0.88	0.88
	Compacity	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00
GLCM	Homogeneity	0.96	0.96	0.95	0.96	0.96	0.97	0.98	0.97	0.97	0.97	0.98	0.98
	Energy	0.93	0.93	0.94	0.91	0.94	0.93	0.94	0.98	0.99	0.88	0.93	0.99
	Contrast	0.95	0.95	0.95	0.95	0.96	0.96	0.95	0.95	0.95	0.96	0.96	0.96
	Correlation	0.84	0.84	0.84	0.86	0.86	0.86	0.96	0.97	0.97	0.96	0.95	0.95
	Entropy	0.94	0.96	0.98	0.92	0.95	0.97	0.96	0.99	1.00	0.87	0.95	0.99
	Dissimilarity	0.96	0.96	0.96	0.96	0.96	0.96	0.97	0.97	0.97	0.97	0.97	0.97
GLRLM	SRE	0.96	0.95	0.90	0.98	0.97	0.98	0.97	0.98	0.94	0.98	0.98	0.97
	LRE	0.97	0.95	0.91	0.98	0.97	0.97	0.98	0.98	0.94	0.98	0.99	0.98
	LGRE	0.28	0.26	0.24	0.35	0.33	0.31	0.91	0.91	0.91	0.83	0.84	0.84
	HGRE	0.66	0.66	0.66	0.62	0.63	0.64	0.93	0.93	0.93	0.89	0.90	0.90
	SRLGE	0.30	0.27	0.25	0.37	0.34	0.31	0.89	0.91	0.91	0.81	0.83	0.83
	SRHGE	0.64	0.65	0.65	0.61	0.63	0.63	0.94	0.94	0.93	0.91	0.90	0.90
	LRLGE	0.19	0.20	0.22	0.25	0.26	0.28	0.96	0.93	0.92	0.93	0.89	0.86
	LRHGE	0.80	0.73	0.70	0.81	0.71	0.66	0.92	0.92	0.93	0.92	0.88	0.88
	GLNU	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	RLNU	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	0.99	1.00	1.00
	RP	0.96	0.94	0.91	0.98	0.97	0.98	0.98	0.98	0.94	0.98	0.99	0.97
NGLDM	Coarseness	0.95	0.98	0.96	0.98	0.98	0.97	0.95	0.96	0.97	0.98	0.98	0.97
	Contrast	0.92	0.90	0.90	0.87	0.87	0.86	0.84	0.87	0.96	0.86	0.90	0.96
	Busyness	0.87	0.86	0.84	0.86	0.88	0.90	0.92	0.96	0.98	0.97	0.98	0.98
GLZLM	SZE	0.43	0.80	0.68	0.62	0.75	0.93	0.38	0.94	0.86	0.67	0.88	0.93
	LZE	0.99	0.96	0.95	0.99	0.99	0.98	1.00	0.93	0.97	1.00	0.91	0.97
	LGZE	0.22	0.24	0.23	0.39	0.29	0.28	0.84	0.88	0.88	0.84	0.81	0.82
	HGZE	0.55	0.51	0.57	0.67	0.54	0.58	0.90	0.91	0.92	0.89	0.89	0.90
	SZLGE	0.38	0.34	0.28	0.42	0.41	0.34	0.44	0.77	0.83	0.63	0.70	0.80
	SZHGE	0.69	0.58	0.48	0.83	0.65	0.64	0.91	0.91	0.90	0.88	0.87	0.90
	LZLGE	0.94	0.41	0.06	0.96	0.85	0.24	1.00	0.92	0.96	1.00	0.90	0.97
	LZHGE	0.99	0.98	0.98	0.99	0.99	0.98	1.00	0.94	0.96	1.00	0.92	0.98
	GLNU	0.98	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	0.99	0.99	1.00
	ZLNU	0.99	0.99	0.99	0.98	0.99	0.99	0.97	0.98	0.98	0.94	0.98	0.98
	ZP	0.90	0.91	0.86	0.91	0.95	0.96	0.96	0.98	0.92	0.97	0.97	0.97

[SUPPLEMENTAL TABLE 3. Test-Retest repeatability of PET_{PSF} radiomic features (CCC).]

	VOI	VOI_{WT}						VOI_{40}					
	Resampling	NSR			ISR			NSR			ISR		
	SUV Bins	32	64	128	32	64	128	32	64	128	32	64	128
Intensity (Standard)	SUV_{MEAN}	0.76	0.76	0.76	0.75	0.75	0.75	0.89	0.89	0.89	0.89	0.89	0.89
	SUV_{SD}	0.94	0.94	0.94	0.93	0.93	0.93	0.84	0.84	0.84	0.85	0.85	0.85
	SUV_{MAX}	0.88	0.88	0.88	0.87	0.87	0.87	0.88	0.88	0.88	0.89	0.89	0.89
	SUV_{PEAK}	0.77	0.77	0.77	0.82	0.82	0.82	0.72	0.72	0.72	0.80	0.80	0.80
	TLG	0.98	0.98	0.98	0.98	0.98	0.98	0.96	0.96	0.96	0.96	0.96	0.96
Intensity (Non-standard)	Skewness	0.88	0.88	0.88	0.89	0.89	0.89	0.72	0.72	0.72	0.91	0.91	0.91
	Kurtosis	0.60	0.60	0.60	0.72	0.72	0.72	0.75	0.75	0.75	0.88	0.88	0.88
	Entropy	0.56	0.74	0.90	0.64	0.75	0.84	0.46	0.43	0.66	0.83	0.76	0.66
	Energy	0.57	0.73	0.86	0.58	0.69	0.80	0.64	0.54	0.69	0.85	0.80	0.72
Shape	MTV	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	Sphericity	0.86	0.86	0.86	0.84	0.84	0.84	0.93	0.93	0.93	0.90	0.90	0.90
	Compacity	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00
GLCM	Homogeneity	0.81	0.81	0.79	0.84	0.85	0.85	0.95	0.94	0.94	0.97	0.97	0.97
	Energy	0.80	0.99	0.95	0.61	0.92	0.97	0.77	0.96	0.99	0.84	0.87	0.97
	Contrast	0.83	0.83	0.83	0.83	0.83	0.83	0.98	0.98	0.98	0.97	0.97	0.97
	Correlation	0.89	0.89	0.89	0.89	0.89	0.89	0.98	0.98	0.98	0.96	0.96	0.96
	Entropy	0.97	0.99	0.98	0.83	0.98	0.98	0.84	0.96	0.98	0.85	0.90	0.97
	Dissimilarity	0.85	0.85	0.85	0.85	0.85	0.85	0.97	0.97	0.97	0.97	0.97	0.98
GLRLM	SRE	0.77	0.76	0.76	0.85	0.87	0.86	0.97	0.96	0.96	0.97	0.97	0.95
	LRE	0.82	0.77	0.74	0.86	0.87	0.86	0.96	0.95	0.96	0.97	0.96	0.94
	LGRE	0.44	0.40	0.39	0.50	0.48	0.48	0.66	0.68	0.67	0.51	0.51	0.46
	HGRE	0.57	0.57	0.58	0.55	0.56	0.57	0.68	0.68	0.69	0.67	0.68	0.66
	SRLGE	0.47	0.42	0.40	0.52	0.50	0.49	0.65	0.66	0.66	0.50	0.49	0.44
	SRHGE	0.53	0.55	0.57	0.51	0.54	0.57	0.71	0.70	0.69	0.72	0.69	0.67
	LRLGE	0.30	0.33	0.36	0.34	0.40	0.43	0.83	0.76	0.72	0.85	0.73	0.58
	LRHGE	0.76	0.67	0.62	0.80	0.68	0.61	0.82	0.69	0.67	0.87	0.71	0.64
	GLNU	0.99	0.99	0.99	0.99	0.99	0.99	1.00	1.00	0.99	1.00	1.00	1.00
	RLNU	1.00	1.00	1.00	1.00	1.00	1.00	0.97	0.98	0.98	0.97	0.98	0.98
	RP	0.79	0.76	0.75	0.84	0.86	0.86	0.97	0.96	0.96	0.98	0.97	0.95
NGLDM	Coarseness	0.92	0.92	0.94	0.97	0.98	0.99	0.96	0.96	0.98	0.99	0.98	0.97
	Contrast	0.87	0.88	0.84	0.75	0.73	0.74	0.98	0.95	0.98	0.90	0.93	0.98
	Busyness	0.93	0.85	0.86	0.92	0.91	0.91	0.95	0.96	0.93	0.99	0.99	0.98
GLZLM	SZE	0.04	0.24	0.37	0.38	0.79	0.67	0.36	0.67	0.87	0.34	0.87	0.91
	LZE	0.97	0.99	0.83	0.96	0.99	0.91	0.92	0.97	0.81	0.91	0.96	0.77
	LGZE	0.48	0.47	0.44	0.56	0.58	0.54	0.27	0.51	0.60	0.20	0.19	0.28
	HGZE	0.57	0.46	0.51	0.59	0.39	0.55	0.66	0.60	0.62	0.52	0.48	0.56
	SZLGE	0.62	0.65	0.56	0.57	0.72	0.65	0.00	0.42	0.60	0.18	0.11	0.19
	SZHGE	0.57	0.36	0.41	0.63	0.54	0.56	0.62	0.62	0.64	0.64	0.77	0.71
	LZLGE	0.48	0.07	0.03	0.61	0.46	0.06	0.85	0.93	0.72	0.85	0.91	0.67
	LZHGE	1.00	0.98	0.97	1.00	0.99	0.99	0.97	0.99	0.91	0.96	0.99	0.87
	GLNU	0.99	0.99	0.99	0.98	0.99	0.99	0.97	0.99	1.00	0.95	0.97	1.00
	ZLNU	0.93	0.99	0.99	0.88	0.99	0.99	0.89	0.94	0.97	0.84	0.89	0.94
	ZP	0.63	0.64	0.66	0.73	0.79	0.81	0.91	0.94	0.95	0.98	0.98	0.96

[SUPPLEMENTAL TABLE 4. Regression analysis of mean tumor volume and absolute relative difference in PET_{OSEM} radiomic features between test and retest FDG PET studies. After controlling for multiple comparisons, an r value of 0.90 or greater was considered significant.]

	VOI	VOI_{WT} (Pearson's r)						VOI_{40} (Pearson's r)					
	Resampling	NSR			ISR			NSR			ISR		
	SUV Bins	32	64	128	32	64	128	32	64	128	32	64	128
Intensity (Standard)	SUV_{MEAN}	0.41	0.41	0.41	0.40	0.40	0.40	0.37	0.37	0.37	0.32	0.32	0.32
	SUV_{SD}	0.10	0.10	0.10	0.14	0.14	0.14	0.35	0.35	0.35	0.39	0.39	0.39
	SUV_{MAX}	0.40	0.40	0.40	0.45	0.45	0.45	0.40	0.40	0.40	0.49	0.49	0.49
	SUV_{PEAK}	0.45	0.45	0.45	0.49	0.49	0.49	0.39	0.39	0.39	0.43	0.43	0.43
	TLG	0.13	0.13	0.13	0.18	0.18	0.18	0.40	0.40	0.40	0.47	0.47	0.47
Intensity (Non-standard)	Skewness	0.21	0.21	0.21	0.24	0.24	0.24	-0.19	-0.19	-0.19	-0.01	-0.01	-0.01
	Kurtosis	-0.18	-0.18	-0.18	-0.11	-0.11	-0.11	-0.29	-0.29	-0.29	-0.14	-0.14	-0.14
	Entropy	0.28	0.23	0.34	0.39	0.42	0.36	-0.08	-0.19	-0.26	-0.21	-0.15	-0.25
	Energy	0.24	0.22	0.25	0.30	0.39	0.32	-0.28	-0.32	-0.34	-0.35	-0.25	-0.27
Shape	MTV	-0.41	-0.41	-0.41	-0.48	-0.48	-0.48	-0.34	-0.34	-0.34	-0.21	-0.21	-0.21
	Sphericity	-0.50	-0.50	-0.50	-0.44	-0.44	-0.44	-0.71	-0.71	-0.71	-0.63	-0.63	-0.63
	Compacity	-0.38	-0.38	-0.38	-0.37	-0.37	-0.37	-0.38	-0.38	-0.38	-0.08	-0.08	-0.08
GLCM	Homogeneity	-0.23	-0.22	-0.26	-0.21	-0.19	-0.19	-0.22	-0.23	-0.30	-0.49	-0.49	-0.45
	Energy	0.47	-0.22	-0.30	0.30	0.39	-0.14	-0.25	-0.29	-0.26	-0.18	-0.05	-0.04
	Contrast	0.02	0.03	0.03	-0.08	-0.08	-0.07	-0.17	-0.13	-0.17	-0.18	-0.19	-0.21
	Correlation	-0.43	-0.44	-0.44	-0.34	-0.34	-0.34	-0.41	-0.40	-0.40	-0.33	-0.35	-0.34
	Entropy	0.11	-0.31	-0.37	0.49	0.03	-0.31	-0.44	-0.38	-0.42	0.00	-0.05	-0.11
	Dissimilarity	-0.01	-0.01	0.00	-0.08	-0.08	-0.08	-0.19	-0.16	-0.18	-0.24	-0.25	-0.26
GLRLM	SRE	-0.17	-0.26	-0.47	0.35	0.18	0.17	-0.15	-0.12	-0.45	-0.39	-0.27	0.01
	LRE	-0.27	-0.35	-0.48	0.06	0.13	-0.03	-0.24	-0.05	-0.36	-0.31	-0.25	-0.04
	LGRE	0.01	0.00	0.00	0.11	0.11	0.11	-0.06	-0.06	-0.08	-0.34	-0.34	-0.33
	HGRE	-0.23	-0.25	-0.26	-0.19	-0.22	-0.23	0.00	0.01	-0.01	-0.17	-0.17	-0.17
	SRLGE	0.03	0.01	0.01	0.14	0.11	0.11	-0.07	-0.07	-0.09	-0.34	-0.34	-0.33
	SRHGE	-0.22	-0.24	-0.26	-0.18	-0.21	-0.23	0.03	0.04	0.01	-0.15	-0.15	-0.16
	LRLGE	-0.05	-0.03	0.00	0.01	0.08	0.10	0.03	-0.03	-0.01	-0.39	-0.35	-0.34
	LRHGE	-0.27	-0.30	-0.28	-0.22	-0.22	-0.24	-0.18	-0.13	-0.10	-0.34	-0.27	-0.23
	GLNU	-0.46	-0.44	-0.43	-0.55	-0.55	-0.50	-0.50	-0.60	-0.62	-0.42	-0.40	-0.43
	RLNU	-0.39	-0.41	-0.41	-0.41	-0.45	-0.47	-0.18	-0.32	-0.39	-0.21	-0.21	-0.16
	RP	-0.20	-0.31	-0.48	0.19	0.17	0.07	-0.21	-0.07	-0.40	-0.35	-0.27	0.00
NGLDM	Coarseness	-0.48	-0.32	-0.55	-0.38	-0.39	-0.47	-0.39	-0.53	-0.49	-0.42	-0.37	-0.44
	Contrast	-0.39	-0.42	-0.28	-0.48	-0.41	-0.37	-0.36	-0.65	-0.44	-0.31	-0.17	-0.03
	Busyness	-0.45	-0.44	-0.41	-0.63	-0.48	-0.40	-0.42	-0.48	-0.23	-0.45	-0.35	-0.20
GLZLM	SZE	-0.42	-0.25	-0.38	0.14	-0.27	-0.51	-0.42	-0.12	-0.41	-0.23	-0.26	-0.19
	LZE	-0.21	0.01	-0.29	-0.27	0.05	-0.12	-0.45	0.66	-0.03	-0.48	0.28	0.05
	LGZE	0.06	0.05	0.02	0.07	0.03	0.05	0.11	0.03	-0.15	-0.46	-0.35	-0.38
	HGZE	0.32	-0.05	-0.20	0.30	0.06	-0.10	0.23	0.01	-0.01	-0.33	-0.16	-0.13
	SZLGE	-0.29	0.06	0.00	-0.30	-0.07	0.03	-0.23	0.01	-0.42	-0.17	-0.28	-0.54
	SZHGE	0.32	-0.08	-0.18	0.50	-0.06	-0.17	0.01	-0.19	-0.07	-0.18	-0.39	-0.14
	LZLGE	-0.39	-0.21	-0.08	-0.54	-0.26	-0.13	-0.43	0.74	0.43	-0.33	0.32	0.01
	LZHGE	-0.54	0.12	-0.37	-0.35	-0.18	-0.23	-0.51	0.37	-0.17	-0.49	0.21	-0.14
	GLNU	-0.32	-0.62	-0.56	0.14	-0.33	-0.55	-0.15	-0.46	-0.42	-0.37	-0.04	-0.30
	ZLNU	-0.57	-0.19	-0.33	-0.42	-0.32	-0.27	-0.19	-0.15	-0.37	-0.51	-0.24	0.01
	ZP	-0.39	-0.39	-0.40	-0.36	-0.17	-0.10	-0.16	0.13	-0.20	-0.33	0.25	0.05

[SUPPLEMENTAL TABLE 5. Regression analysis of mean tumor volume and absolute relative difference in PET_{PSF} radiomic features between test and retest FDG PET studies. After controlling for multiple comparisons, an r value of 0.90 or greater was considered significant.]

	VOI	VOI_{WT} (Pearson's r)						VOI_{40} (Pearson's r)					
	Resampling	NSR			ISR			NSR			ISR		
	SUV Bins	32	64	128	32	64	128	32	64	128	32	64	128
Intensity (Standard)	SUV_{MEAN}	0.36	0.36	0.36	0.35	0.35	0.35	0.25	0.25	0.25	0.29	0.29	0.29
	SUV_{SD}	0.14	0.14	0.14	0.11	0.11	0.11	0.48	0.48	0.48	0.46	0.46	0.46
	SUV_{MAX}	-0.03	-0.03	-0.03	0.12	0.12	0.12	-0.03	-0.03	-0.03	0.22	0.22	0.22
	SUV_{PEAK}	0.41	0.41	0.41	0.29	0.29	0.29	0.37	0.37	0.37	0.35	0.35	0.35
	TLG	0.16	0.16	0.16	0.22	0.22	0.22	0.50	0.50	0.50	0.46	0.46	0.46
Intensity (Non-standard)	Skewness	0.74	0.74	0.74	0.66	0.66	0.66	-0.60	-0.60	-0.60	-0.36	-0.36	-0.36
	Kurtosis	-0.32	-0.32	-0.32	-0.22	-0.22	-0.22	-0.43	-0.43	-0.43	-0.25	-0.25	-0.25
	Entropy	-0.37	-0.35	-0.27	-0.41	-0.40	-0.39	0.08	-0.10	-0.03	0.53	0.51	0.46
	Energy	-0.34	-0.31	-0.23	-0.44	-0.47	-0.48	0.03	-0.12	0.07	0.20	0.21	0.33
Shape	MTV	-0.41	-0.41	-0.41	-0.48	-0.48	-0.48	0.23	0.23	0.23	0.08	0.08	0.08
	Sphericity	-0.50	-0.50	-0.50	-0.44	-0.44	-0.44	-0.36	-0.36	-0.36	-0.46	-0.46	-0.46
	Compacity	-0.38	-0.38	-0.38	-0.37	-0.37	-0.37	0.16	0.16	0.16	0.13	0.13	0.13
GLCM	Homogeneity	-0.55	-0.56	-0.52	-0.52	-0.53	-0.54	-0.20	-0.29	-0.20	0.19	0.26	0.39
	Energy	-0.30	-0.06	-0.54	-0.40	-0.25	-0.26	0.40	0.51	0.62	0.38	0.57	0.56
	Contrast	-0.48	-0.46	-0.46	-0.40	-0.40	-0.40	0.60	0.58	0.59	0.57	0.58	0.58
	Correlation	-0.49	-0.47	-0.47	-0.39	-0.39	-0.39	-0.51	-0.52	-0.51	-0.64	-0.64	-0.63
	Entropy	-0.07	-0.31	-0.54	-0.21	0.22	-0.38	0.48	0.46	0.39	0.53	0.55	0.40
	Dissimilarity	-0.54	-0.53	-0.53	-0.46	-0.47	-0.47	0.52	0.48	0.50	0.51	0.51	0.49
GLRLM	SRE	-0.45	-0.49	-0.35	-0.38	-0.42	-0.45	0.54	0.17	0.36	0.88	0.81	0.83
	LRE	-0.33	-0.43	-0.28	-0.25	-0.28	-0.30	0.57	0.10	0.43	0.89	0.82	0.79
	LGRE	0.03	-0.01	0.00	0.05	0.06	0.07	-0.09	-0.08	-0.08	-0.07	-0.07	-0.09
	HGRE	-0.14	-0.16	-0.17	-0.13	-0.16	-0.16	-0.03	-0.02	-0.03	0.01	0.01	-0.01
	SRLGE	0.05	0.00	0.00	0.07	0.07	0.08	-0.16	-0.09	-0.10	-0.16	-0.13	-0.12
	SRHGE	-0.14	-0.17	-0.17	-0.11	-0.16	-0.16	0.01	-0.01	-0.01	0.11	0.06	0.02
	LRLGE	-0.04	-0.04	-0.01	0.03	0.06	0.08	0.09	-0.04	0.00	0.29	0.15	0.04
	LRHGE	-0.18	-0.17	-0.18	-0.22	-0.19	-0.20	-0.26	-0.07	-0.11	-0.40	-0.23	-0.17
	GLNU	-0.43	-0.42	-0.39	-0.57	-0.56	-0.57	-0.21	-0.27	-0.16	-0.10	0.19	0.09
	RLNU	-0.41	-0.41	-0.42	-0.56	-0.50	-0.49	0.42	0.29	0.30	0.42	0.29	0.17
	RP	-0.38	-0.45	-0.31	-0.29	-0.34	-0.36	0.55	0.13	0.40	0.89	0.82	0.82
NGLDM	Coarseness	-0.54	-0.58	-0.69	-0.34	-0.33	-0.25	-0.48	-0.44	-0.22	-0.01	-0.23	-0.41
	Contrast	-0.09	-0.05	-0.34	-0.41	-0.31	-0.36	0.81	0.56	-0.07	0.34	0.38	0.58
	Busyness	-0.54	-0.64	-0.52	-0.35	-0.47	-0.51	-0.07	-0.46	-0.19	-0.62	-0.67	-0.52
GLZLM	SZE	-0.23	-0.37	-0.36	-0.18	-0.24	-0.40	-0.67	-0.46	-0.44	-0.42	-0.32	-0.09
	LZE	-0.52	-0.51	0.03	-0.45	-0.42	-0.27	0.05	-0.19	0.62	0.83	0.45	0.75
	LGZE	0.05	0.01	0.02	0.01	0.01	0.05	-0.32	-0.28	-0.09	-0.15	-0.11	-0.06
	HGZE	0.12	-0.03	-0.11	0.34	0.15	0.09	-0.19	-0.08	-0.04	-0.14	-0.05	0.00
	SZLGE	0.00	0.09	0.08	-0.11	0.10	0.11	-0.06	0.01	-0.15	-0.26	-0.14	-0.11
	SZHGE	-0.10	-0.18	-0.17	0.15	0.14	0.14	-0.42	-0.30	-0.07	-0.12	0.25	0.27
	LZLGE	-0.54	-0.45	-0.11	-0.44	-0.32	0.02	-0.09	-0.16	0.62	0.80	0.41	0.68
	LZHGE	-0.45	-0.21	-0.42	-0.43	-0.06	-0.42	0.24	0.08	0.26	0.66	0.35	0.53
	GLNU	-0.22	-0.17	-0.44	-0.34	-0.04	-0.58	0.36	0.05	-0.48	0.58	0.78	0.22
	ZLNU	-0.24	-0.50	-0.41	0.17	0.02	-0.38	-0.41	0.01	0.40	0.52	0.77	0.57
	ZP	-0.26	-0.37	-0.29	-0.22	-0.31	-0.37	0.10	0.23	0.25	0.87	0.84	0.77