





RESEARCH ARTICLE

Synergistic impact of motion and acquisition/reconstruction parameters on ^{18}F -FDG PET radiomic features in non-small cell lung cancer: Phantom and clinical studies

Seyyed Ali Hosseini^{1,2}  | Isaac Shiri³  | Ghasem Hajianfar⁴  |
Bahador Bahadorzadeh⁵  | Pardis Ghafarian^{6,7} | Habib Zaidi^{3,8,9,10} |
Mohammad Reza Ay^{1,2}

¹Department of Medical Physics and Biomedical engineering, Tehran University of Medical Sciences, Tehran, Iran

²Research Center for Molecular and Cellular Imaging, Tehran University of Medical Sciences, Tehran, Iran

³Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, Geneva, Switzerland

⁴Rajaie Cardiovascular Medical and Research Center, Iran University of Medical Science, Tehran, Iran

⁵Department of Nuclear Engineering, Shiraz University, Shiraz, Iran

⁶Chronic Respiratory Diseases Research Center, National Research Institute of Tuberculosis and Lung Diseases (NRITLD), Shahid Beheshti University of Medical Sciences, Tehran, Iran

⁷PET/CT and Cyclotron Center, Masih Daneshvari Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran

⁸Geneva University Neurocenter, Geneva University, Geneva, Switzerland

⁹Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, Groningen, Netherlands

¹⁰Department of Nuclear Medicine, University of Southern Denmark, Odense, Denmark

Correspondence

Habib Zaidi, Geneva University Hospital, Division of Nuclear Medicine and Molecular Imaging, CH-1211 Geneva, Switzerland.

Email: habib.zaidi@hcuge.ch

Mohammad Reza Ay, Department of Medical Physics and Biomedical Engineering, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran.

Email: mohammadreza_ay@sina.tums.ac.ir

Funding information

Tehran University of Medical Sciences, Grant/Award Number: 40235; Swiss National Science Foundation under Grant/Award Number: SNSF 320030_176052

Abstract

Objectives: This study is aimed at examining the synergistic impact of motion and acquisition/reconstruction parameters on ^{18}F -FDG PET image radiomic features in non-small cell lung cancer (NSCLC) patients, and investigating the robustness of features performance in differentiating NSCLC histopathology subtypes.

Methods: An in-house developed thoracic phantom incorporating lesions with different sizes was used with different reconstruction settings, including various reconstruction algorithms, number of subsets and iterations, full-width at half-maximum of post-reconstruction smoothing filter and acquisition parameters, including injected activity and test–retest with and without motion simulation. To simulate motion, a special motor was manufactured to simulate respiratory motion based on a normal patient in two directions. The lesions were delineated semi-automatically to extract 174 radiomic features. All radiomic features were categorized according to the coefficient of variation (COV) to select robust features. A cohort consisting of 40 NSCLC patients with adenocarcinoma ($n = 20$) and squamous cell carcinoma ($n = 20$) was retrospectively analyzed.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Medical Physics* published by Wiley Periodicals LLC on behalf of American Association of Physicists in Medicine.

Statistical analysis was performed to discriminate robust features in differentiating histopathology subtypes of NSCLC lesions.

Results: Overall, 29% of radiomic features showed a $\text{COV} \leq 5\%$ against motion. Forty-five percent and 76% of the features showed a $\text{COV} \leq 5\%$ against the test–retest with and without motion in large lesions, respectively. Thirty-three percent and 45% of the features showed a $\text{COV} \leq 5\%$ against different reconstruction parameters with and without motion, respectively. For NSCLC histopathological subtype differentiation, statistical analysis showed that 31 features were significant ($p\text{-value} < 0.05$). Two out of the 31 significant features, namely, the joint entropy of GLCM ($\text{AUC} = 0.71$, $\text{COV} = 0.019$) and median absolute deviation of intensity histogram ($\text{AUC} = 0.7$, $\text{COV} = 0.046$), were robust against the motion (same reconstruction setting).

Conclusions: Motion, acquisition, and reconstruction parameters significantly impact radiomic features, just as their synergies. Radiomic features with high predictive performance (statistically significant) in differentiating histopathological subtype of NSCLC may be eliminated due to non-reproducibility.

KEYWORDS

non-small cell lung cancer, PET/CT, quantitative analysis, radiomics, robustness

1 | INTRODUCTION

Lung cancer is one of the main causes of cancer deaths in men and women in the world.¹ There are two main types of lung cancer, including small cell lung cancer and non-small cell lung (NSCLC). NSCLC is the most prevalent lung cancer type,² which can be divided into three main subtypes: squamous cell carcinoma, adenocarcinoma, and large cell carcinoma. Other subtypes of NSCLC, such as sarcomatoid carcinoma and adenosquamous carcinoma, are much less common.³ The gold standard for differentiating NSCLC is pathology—an invasive procedure presenting risks to the patient.

Positron emission tomography/computed tomography (PET/CT) plays a crucial role in clinical oncology for a variety of indications.⁴ Previous studies established the potential of [^{18}F]-fluorodeoxyglucose (^{18}F -FDG) PET/CT in the assessment of treatment response in lung cancer.^{5,6} PET images may not provide accurate information about the position and size of lesions located in the thoraco-abdominal region.

The potential of PET quantification using the standardized uptake value (SUV) as a semi-quantitative metric in predicting response to therapy has been reported in previous studies.^{7,8} In a recent study, Kumar et al.⁹ showed that SUV changes in test–retest analysis depend on the acquisition/reconstruction parameters. In their study, the intraclass correlation coefficient (ICC) calculated regarding SUV_{mean} and SUV_{max} was 0.92 and 0.93, respectively. They reported that the clinical applications of ^{18}F -FDG PET/CT SUV variation for a single-center was more significant than anticipated. Another multicenter study reported that the SUV varied from 10% to 25% discarding the impact of

reconstruction or other parameters.¹⁰ Because of the large variability of SUV when using different acquisition and reconstruction parameters and the simplified approximations associated with the region of interest-based quantification that ignores the heterogeneity of tracer uptake, there is a need for more sophisticated and reproducible features that are more powerful than the SUV.

Radiomics, an emerging field enabling to extract features from images to support clinical decision making,^{11,12} proved promising in quantitative medical imaging. Radiomics is expected to be commonly integrated in imaging-based oncological studies, to provide a link to personalized medicine through the extraction of quantitative and in-depth features from medical images to help the decision making process in clinical setting.¹³

PET/CT radiomic features have been widely used to determine the malignancy of tumors,^{14,15} predict responses to therapy, patient prognosis,^{16,17} and identification of tumor phenotypes.^{18–20} The number of extracted radiomic features is pretty large,²¹ and hence ensuring their reproducibility is desirable to provide robust imaging biomarkers. Recent studies evaluated PET radiomic feature reproducibility in different scenarios, including respiratory motion,^{22,23} image acquisition,²⁴ test–retest,²⁵ pre-processing,²⁶ image reconstruction,^{27,28} and segmentation²⁹ techniques.

More recently, the impact of reconstruction and delineation, post-reconstruction smoothing FWHM filter size, matrix size, time per bed positions, test–retest, and segmentations were studied.^{27,29–31} Lung movement is undeniable in PET imaging. It may impact radiomic features to a large extent. However, the impact of respiratory motion did not receive much attention in most of these studies.^{29,32} These studies reported that the majority of

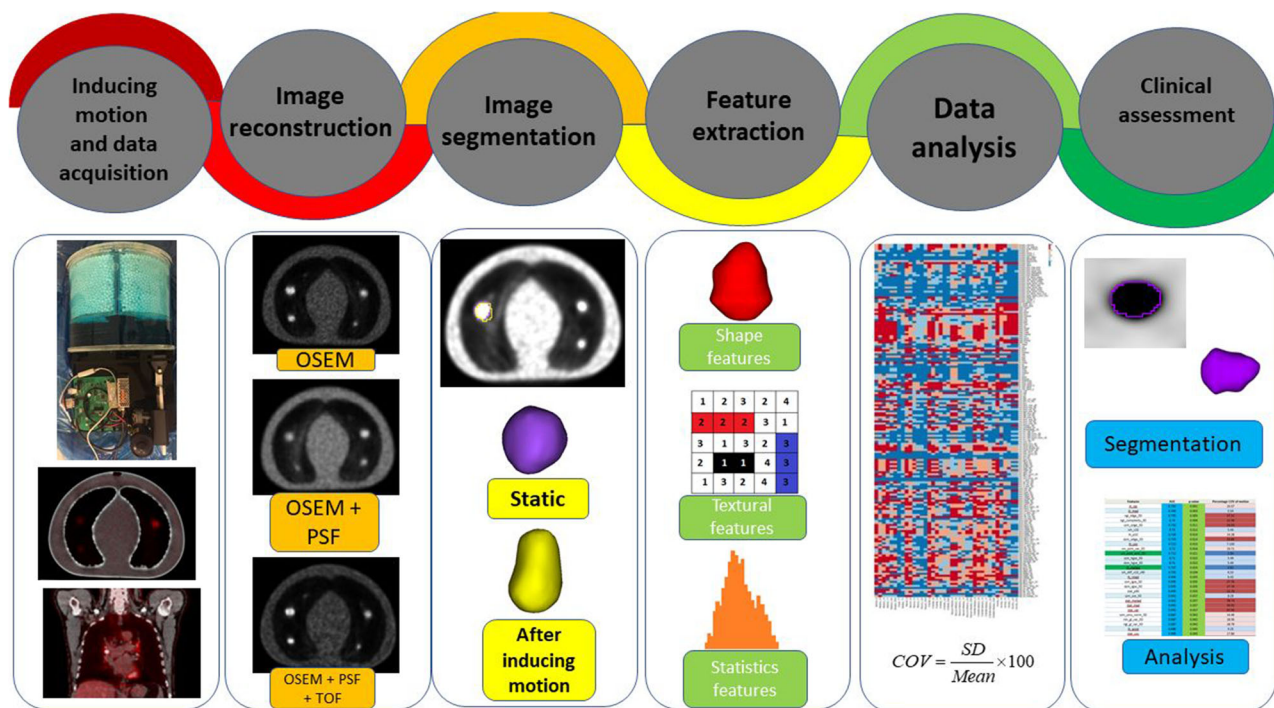


FIGURE 1 Framework adopted in the current study, beginning with inducing motion, and followed by image processing and reconstruction steps and culminating with data analysis

features has 10% to 200% variability.³¹ Oliver et al.²² examined the variability of radiomic features over lung motion between static and 4D gated images and demonstrated that most features had more than 5% variability.

In a recent study,³³ the effect of respiratory motion on radiomic features was investigated. To date, experimental studies evaluating the synergistic impact of motion and acquisition or reconstruction parameters on moving targets from PET images are lacking. This is an important area that has been overlooked in recent studies,³⁴ which requires additional research and development efforts since it is relevant in many thoraco-abdominal PET studies. The scanning time might cover several respiratory periods. As a consequence, motion impacts the images both qualitatively and quantitatively. Furthermore, the thoraco-abdominal motion might impact, along with other influential factors, PET images. This includes multicenter clinical trials, where protocols change among the involved sites, which may impact other parameters.

This study investigated the synergistic impact of motion and acquisition/reconstruction parameters on PET image radiomic features. We evaluated the impact of motion, a wide range of reconstruction algorithms, the number of iterations and subsets, different filter sizes, various lesions sizes, and performed test–retest analysis on PET radiomic features. We selected robust features against those parameters individually and simultaneously. We also assessed the predictive power of

robust features for NSCLC histopathological subtype differentiation.

2 | MATERIALS AND METHODS

2.1 | Phantom data acquisition

In the current study, we selected robust features against multiple acquisition protocols individually and simultaneously. Moreover, we examined the power of robust features in discriminating NSCLC histopathological subtypes. Figure 1 summarizes the workflow adopted in the current study.

Clinical and phantom studies were performed on a GE Discovery 690 PET/CT scanner (General Electric Healthcare, USA). An in-house developed thoracic phantom was used in this study. This phantom has the following features: 180 mm interior length, 9.6 L capacity, six spherical inserts with inner diameters of 8, 10, 13, 17, and 22 (left and right) mm. We classified 8 and 10 mm spheres as small lesion sizes, 13 and 17 mm as medium lesion sizes, and 22 mm as large lesion sizes. To perform a realistic phantom study, all lesions were wall-less to avoid partial volume effect (PVE).

To stimulate respiratory motion, we manufactured a motor placed under the phantom. This motor induces lung movement (two directions: posterior-anterior and lateral) at the typical respiratory rate of a healthy adult

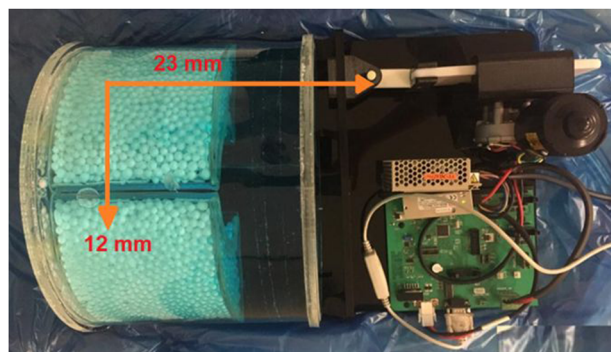


FIGURE 2 Illustration of the thoracic phantom and respiratory motion platform

at rest, that is, 12 breaths per minute.³⁵ The magnitude of thoraco-abdominal lesion's movement due to respiratory motion is between 6 and 23 mm in each direction.³⁶ We induced 23 mm in the posterior-anterior direction and 12 mm in the other direction (Figure 2), 12 times in 1 min. The phantom and lesions were filled with a mixture of ^{18}F -FDG and water with an activity concentration of 5.3 KBq/ml and 2.65 KBq/ml, corresponding to 370 MBq and 185 MBq injected to a 75 kg patient. Lesions to background ratio (LBR) of 4:1 and 2:1 were induced. To examine the effect of post-injection scanning time (one half-life of ^{18}F) on image features, phantom studies were acquired three times. The first scan started 60 min post-injection, which was the baseline of timing. Hence, the three acquisitions refer to 0 min (first time), 30 min, and 110 min from the baseline.

2.2 | Clinical data acquisition

A cohort consisting of 40 NSCLC patients with adenocarcinoma ($n = 20$) and squamous cell carcinoma ($n = 20$) (mean age = 60 ± 12 years old, 16 men and 14 women) were retrospectively analyzed. This study was approved by the institutional review board (IRB) of Tehran University of Medical Sciences, and written informed consent of patients was waived. Clinical studies were acquired using the following parameters: fasting for at least 8 h before the scan and injection of a mean activity of 309.26 MBq (range 138.90–572.25 MBq) of ^{18}F -FDG with a mean uptake time of 66.58 min (range 23.08–128.90 min). The acquisition time per bed position was 3 min for both phantom and clinical studies. Moreover, low-dose CT images were acquired for anatomical localization and attenuation correction (Table 1). PET data were reconstructed using ordered subset expectation maximization (OSEM) iterative algorithm with two iterations and 18 subsets, with an image matrix of 256×256 (pixel size = 3.906 mm^2). A

TABLE 1 Patients' clinical and pathological characteristics of the current study

Characteristics	
Gender	
Male	22
Female	18
Height (Cm) (mean \pm SD)	167 ± 14
Weight (Kg) (mean \pm SD)	66 ± 12
Cancer stage	
I	6
II	8
III	12
IV	14
Histology	
Adenocarcinoma	20
Squamous cell	20

TABLE 2 Image reconstruction settings for the phantom study

Parameter	Variation
Reconstruction algorithm	OSEM (HD) OSEM + PSF (HDS) OSEM + PSF + TOF (TS)
Subsets	18, 21
Iterations	2, 3
Post-reconstruction filter	0, 4.5, 5.5, 6.4, 7.5 (mm)
Lesion sizes	8, 10, 13, 17, 22 (mm)

TOF + OSEM (TOF is commercially referred to as VUE.FX), OSEM (commercially referred to as VUE.HD), and PSF + OSEM + TOF (PSF + TOF, referred to as Sharp-IR by the manufacturer) algorithms.

Abbreviations: TOF, time-of-flight; PSF, Point Spread Function; OSEM, Ordered Subset Expectation Maximization.

4.5-mm FWHM post-reconstruction Gaussian filter was also applied to the images.

2.3 | Reconstruction parameters

To examine the effect of reconstruction settings on image features, phantom images were reconstructed using three different reconstruction algorithms, including OSEM with time-of-flight (TOF)³⁷ and resolution recovery or point spread function (PSF) reconstruction.³⁸ Besides, different numbers of iterations and subsets and Gaussian filter sizes were used. All phantom images were acquired with zero Gaussian filter size and 4 FWHM Gaussian filter sizes. The different reconstruction algorithms are listed in Table 2.

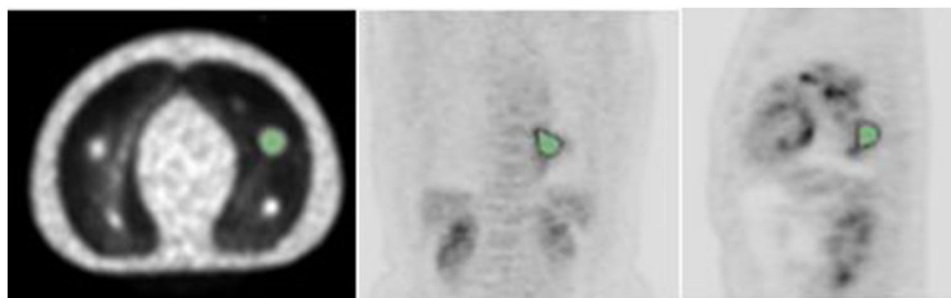


FIGURE 3 Illustration of 3D region-growing-based segmentation of phantom spheres (left) and patient (middle and right) malignant lesions using the Slicer 4.8.0 software

2.4 | Test–retest repeatability

In the test–retest analysis, the time interval for phantom scanning post-injection could be anywhere between a few minutes to hours. Test–retest included measurements taken by the instrument on the same subject under the same conditions. Hence, the same acquisition and reconstruction protocol was used. To minimize the impact of test–retest, size effect, and PVE on the results, we scanned two lesions with 22 mm diameter located in the left and right lung. Test–retest analysis was performed with and without simulating lung motion. Following lesion segmentation, radiomic features were extracted to determine the impact of test–retest on radiomics features with and without motion simulation.

2.5 | Image segmentation

All segmentations were performed using the Slicer 4.8.0 software³⁹ (Figure 3). Malignant lesions were delineated to define volumes of interest (VOIs) using 3D region growing-based segmentation.⁴⁰ One lesion was segmented from each patient's images by a trained nuclear medicine physician with 10 years of experience. The primary tumor was identified and included in the analysis. To minimize the impact of image segmentation on the results, we used a region-growing algorithm for one image and applied it for the different reconstruction settings.

2.6 | ¹⁸F-FDG PET/CT image radiomic features

One hundred and seventy-four 3D radiomic features compliant with IBSI guidelines were extracted using SERA package written in MATLAB.⁴¹ Radiomic features, including morphology ($n = 29$), local intensity ($n = 2$), intensity-based statistics ($n = 18$), intensity histogram ($n = 23$), intensity–volume histogram ($n = 7$), gray level co-occurrence matrix (GLCM) ($n = 25$), gray level run

length matrix (GLRLM) ($n = 16$), gray level size zone matrix (GLSZM) ($n = 16$), gray level distance zone matrix (GLDZM) ($n = 16$), neighborhood gray tone difference matrix (NGTDM) ($n = 5$), neighboring gray level dependence matrix (NGLDM) ($n = 17$) were extracted in the following details: 64 fixed bin width,⁴² combined (for higher-level radiomic features), cubic resampling, using isotropic voxel size $1 \times 1 \times 1 \text{ mm}^3$ without any filter and image discretization. The list of 174 features is provided in Table S1.

2.7 | Data analysis

2.7.1 | Phantom study

The six lesions with 8, 10, 13, 17, and 22 mm diameter (right and left lung) were evaluated for the phantom study. Lesions less than 8 cm^3 were not assessed owing to PVE.³² The impact of PVE was not examined in this work. The inter-setting coefficient of variation (COV) was computed for all 174 radiomic features extracted from the phantom images using the following equation:

$$\text{COV} = \frac{\text{SD}}{\text{Mean}} \times 100 \quad (1)$$

where SD is the standard deviation of image features, and the mean is the average over the different parameters. To classify the differences, 4 categories including extremely low ($\text{COV} \leq 5\%$), low ($5\% < \text{COV} \leq 10\%$), mediocre ($10\% < \text{COV} \leq 20\%$), and high ($\text{COV} > 20\%$) were determined.³² The COV was calculated for each radiomic feature and each specific condition separately. As such, the COV was calculated once for each radiomic feature, based on the number of effective factors examined in this study, including motion, reconstruction, subset and iteration, filter size, and lesion size. All data were analyzed using the R 3.6.3 software. Radiomic features with less than 5% variability over different effective factors in the phantom study were selected as robust features in this study.

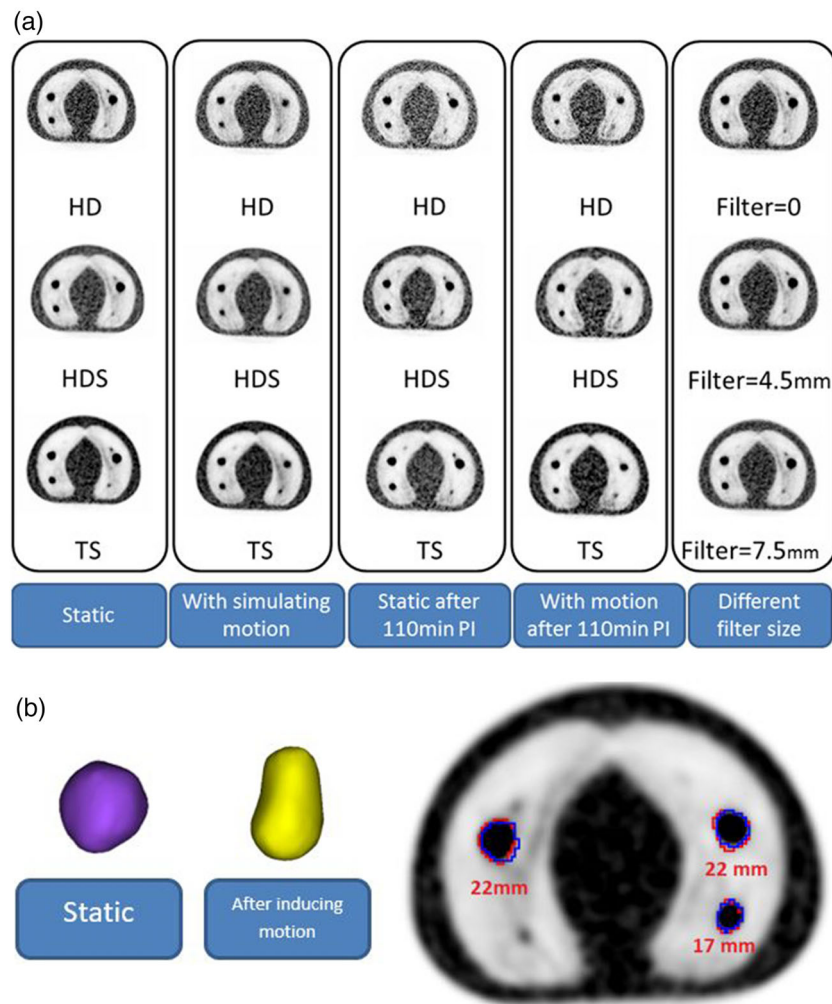


FIGURE 4 (a) The visual impact of different reconstruction algorithms and various subset and iteration numbers, different FWHM filter sizes, and activity after a half-life of ^{18}F -FDG, post-injection (PI) on the phantom images. (b) Display of 2D and 3D representations of how motion affects lesion delineation before (red contour) and after (blue contour) inducing motion for different lesion sizes

2.7.2 | Clinical studies

Statistical analyses were performed using the R software, including the Wilcoxon rank test for p -values,⁴³ false discovery rate correction with Benjamini–Hochberg method (q -value),⁴⁴ and the area under the curve (AUC) univariate analysis.⁴⁵ Our clinical studies were acquired in one center and under a specific condition. Therefore, the only effective factor on radiomic features is breathing motion. As consequence, robust features selected from the phantom study over motion and all 174 IBSI radiomic features were used for statistical analysis to evaluate the impact of using robust features on the accuracy of differentiation. Based on the effective factor in each study, robust features against that factor must be used to increase the repeatability and reproducibility of radiomic features.

3 | RESULTS

Figure 4 shows the qualitative impact of motion, reconstruction, activity, and different FWHM filter sizes on the

resulting PET images. It can be observed that motion, different reconstruction algorithms, and different FWHM filter sizes have a noticeable impact on image quality.

3.1 | Acquisition parameters

3.1.1 | Impact of motion

Four PET acquisitions were performed, in which three were performed with the induced motion to decrease the amount of error in test–retest analysis. We compared the average of three motions for a given scenario. About 53% of the 174 features (92 features) were affected by motion as reflected by a COV $>10\%$. Only 36% of the features (64 features) were robust against motion in large lesions. Our results showed that the robustness of features is highly affected by motion, especially for small lesion size, where 78% of the features have more than 5% variability (Table 4 and Figures 6 and 7). Besides, all features robust against motion are listed based on lesion size in Table S2. All texture features were affected by motion showing a variation of more than 5%. About

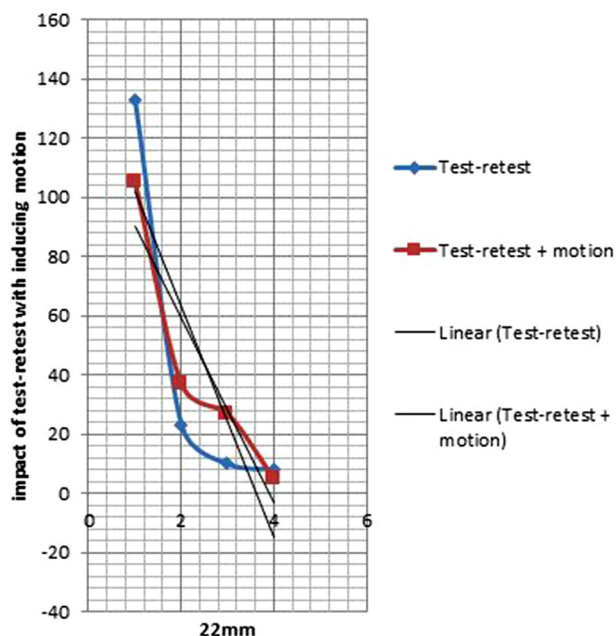


FIGURE 5 Impact of test–retest with motion for the large sphere size from 174 IBSI radiomic features. The X-axis refers to the number of features whereas the Y-axis refers to the number of categories (1 to 4) since the differences were classified into four categories, including 1 = extremely low ($\text{COV} \leq 5\%$), 2 = low ($5\% < \text{COV} \leq 10\%$), 3 = mediocre ($10\% < \text{COV} \leq 20\%$), and 4 = high ($\text{COV} > 20\%$)

56%, 50%, 75%, and 62% of GLCM, GLRLM, GLDZM, and NGLDM features have shown $\text{COV} > 5\%$ against motion in large lesion size, respectively.

3.1.2 | Impact of test–retest with and without motion

We analyzed two spheres with 22 mm diameter (mimicking a large tumor size) with and without motion. To minimize the effect of test–retest in PET/CT imaging, we exploited the availability of two 22 mm spheres in the right and left lung instead of the multiple image acquisitions. The results showed that 81% (142 features) and 89% (156 features) of the 174 features showed a $\text{COV} < 10\%$ against test–retest with and without motion, respectively. The test–retest study showed that PET/CT imaging in static imaging had 24% variability in quantitative analysis. In the test–retest study of static PET/CT imaging, 24% of radiomic features showed high COV ($\text{COV} > 5\%$). This percentage raised to 40% when motion was enabled (Figure 5). Furthermore, robust features against test–retest of the 22 mm lesion are listed in Table S3.

Figure 5 confirms that motion affects the test–retest analysis. Sixteen percent of the 174 features (28 features) are affected by test–retest in the presence of motion. In addition, 76% of the features are robust ($\text{COV} < 5\%$) against test–retest without

motion, but this decreases to 60% when motion is considered.

3.1.3 | Impact of post-injection scanning time

We investigated the effect of post-injection scanning time through scanning the phantom with one half-life time difference (~ 110 min) to assess the effect of activity in the field-of-view on radiomic features. It was shown that 55% of the features (97 features) are robust against post-injection scanning time for large tumor sizes. The downward trend number of robust features from 22 to 8 mm sizes is significant since 47% and 28% of the features are robust for medium and small tumor sizes.

3.2 | Reconstruction parameters

3.2.1 | Impact of different reconstruction settings with and without motion

The impact of reconstruction settings, number of subsets, and iterations are listed in Table 4 and Figures 6 and 7. Moreover, Tables S4 and S6 show robust features against reconstruction parameters based on lesion size with and without the simulation of lung motion. The effect of reconstruction is highly dependent on lesion size, where 66% of the features (116 features) are robust against reconstruction in large lesion size. For the 8 mm sphere, only 18% of the features (33 features) have shown less than 5% variability. Our results showed that synergistic motion and reconstruction highly affect the features, where 53% of the features have more than 5% variation for large lesion size, and about 15% of the features are robust ($\text{COV} < 5\%$) for small lesion size.

3.2.2 | Impact of FWHM filter size with and without motion

Our results demonstrated that 30% of the 174 extracted features (54 features) are robust against different filter sizes. The impact of different post-smoothing FWHM filter size on PET/CT image features is relatively small in the presence of motion, where 80 features were robust against different FWHM filter size, and 73 features were robust when combined with movement for large spheres. The same effect was observed in medium and small lesions size. For instance, in the 8 mm sphere, only 15% of 174 features (27 features) were robust against different filter sizes. The variability of features according to lesion size is shown in Figures 6 and 7. In addition, the robust features against different FWHM filter size are listed based on lesion size in Tables S5 and S7.

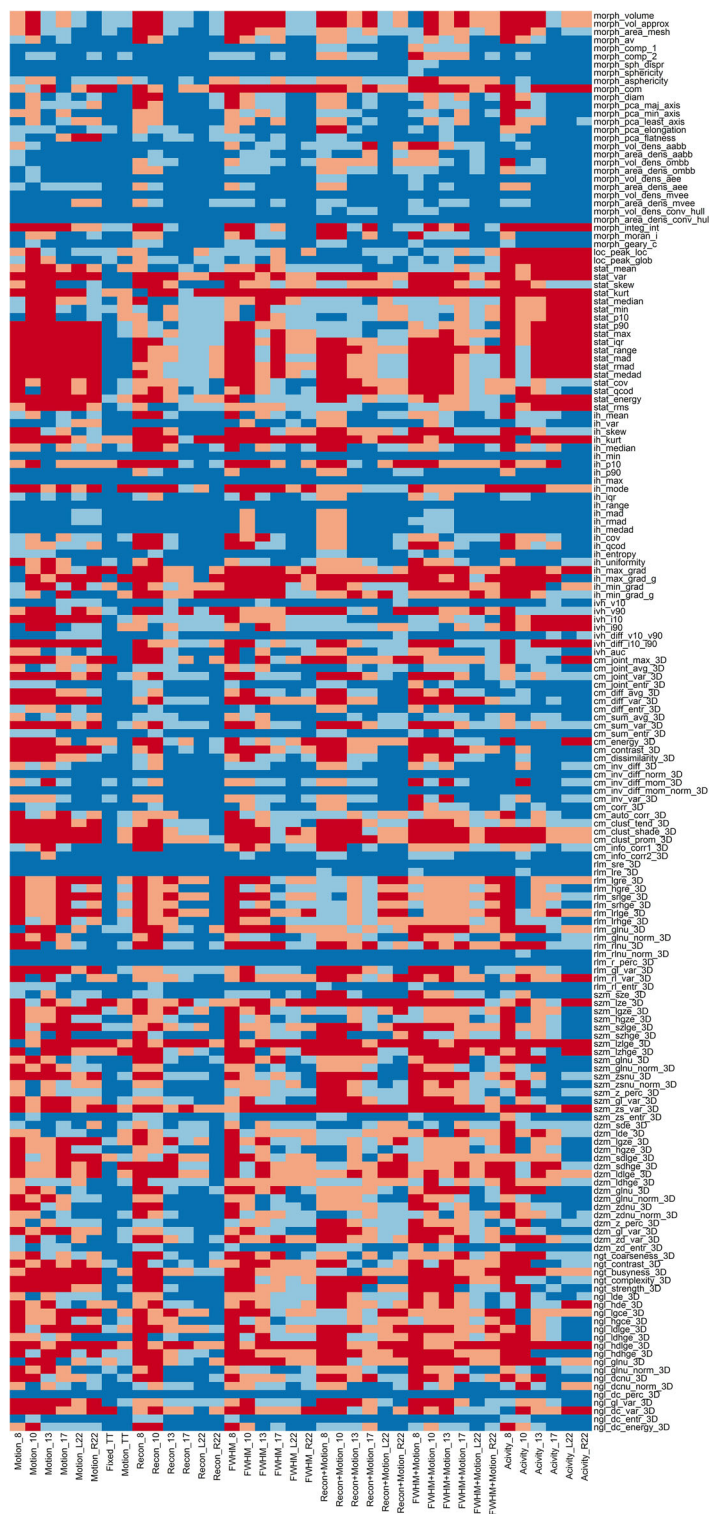


FIGURE 6 Heat map of the variation of 174 radiomic features in the different scenarios

3.3 | Robust and non-robust PET/CT radiomic features in lung lesions

Nineteen features have shown less than 10% variation against all parameters, including motion, different reconstruction algorithms, various numbers of subsets and iterations, activity, and different filter sizes (Table 3). The statistical analysis performed in the current study

showed that the differences between these features are not statistically significant in terms of differentiating the histopathological subtype of tumors. Short runs emphasis from GLRLM was highly robust against motion ($COV = 0.04\%$) but is not significant for differentiating the histopathological subtypes of tumors (p -value = 0.13). Some features showed more than 10% variability against most situations. We consider them as

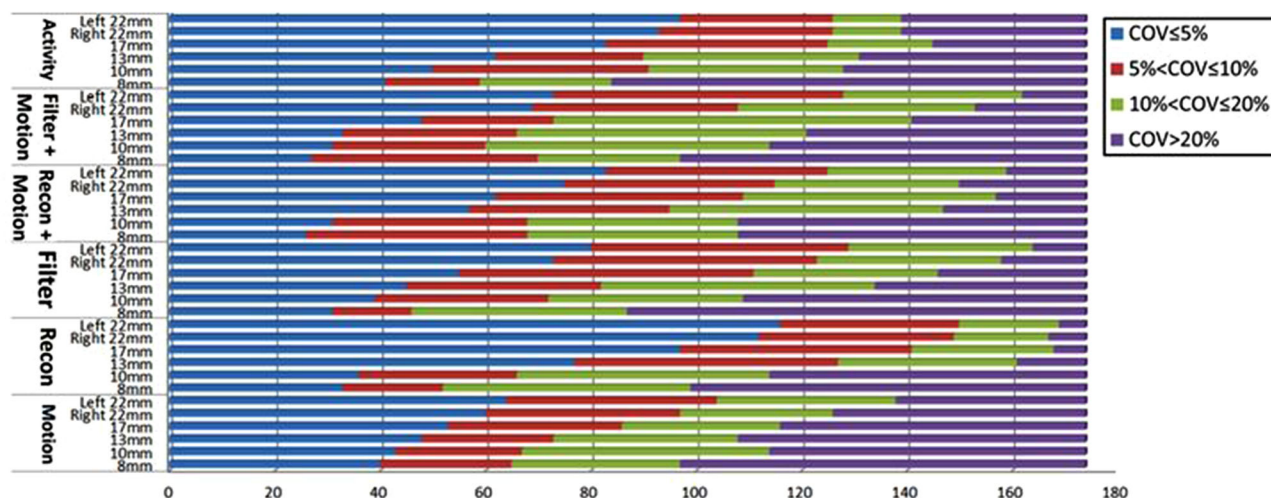


FIGURE 7 Variation of the number of features based on the COV over the different scenarios

TABLE 3 Robust features against various conditions and features with the lowest reproducibility

#	Family	Biomarker
Robust features against various conditions	Morph	Compactness 1, Spherical disproportion, Sphericity, Volume density (AEE), Volume density (MVEE), Volume density (convex hull), Area density (convex hull)
	IH	Minimum, Maximum, Range
	IVH	Volume fraction at 10% intensity, Volume fraction diff between 10% and 90% intensity
	GLCM	Inverse difference normalized, Inverse difference moment normalized,
	GLRLM	Short runs emphasis, Run length non-uniformity normalized, Run percentage, Long runs emphasis
	NGLDM	Dependence count percentage, Dependence count entropy
Features with lowest reproducibility of PET images	Morph	Centre of mass shift
	Stat	Variance, (Excess) kurtosis
	IH	Kurtosis, 10th percentile, Mode, Maximum gradient gray level, Minimum histogram gradient
	IVH	Volume fraction at 90% intensity
	GLCM	Cluster tendency, Cluster shade, Cluster prominence
	GLSZM	Large zone low gray level emphasis, Zone size variance
	GLDSZM	Large distance low gray level emphasis
	NGLDM	High dependence low gray level emphasis, Dependence count variance

features with the lowest reproducibility in PET/CT imaging according to the reproducibility criterion. A feature showing high variability may be significant for differentiation. One such example is the variance from the statistical family that showed high variability over the different scenarios, including motion (COV of motion = 67%, AUC = 0.69, p -value < 0.05).

3.4 | Clinical studies

The summary of statistical analysis provided in Table 5 showed that 31 features are significant (p -value < 0.05),

107 features are acceptable for differentiation (AUC > 0.6). Two of the 31 significant features including joint entropy of GLCM (AUC = 0.71, p -value = 0.021, COV = 0.019) and median absolute deviation of intensity histogram (AUC = 0.7, p -value = 0.024, COV = 0.046) have a COV < 5% whereas 10 of the 31 features showed 5% < COV ≤ 10% against motion. In addition, 12 of the 31 significant features showed a high COV against motion (COV > 20%). The results of the statistical analysis regarding the differentiation between adenocarcinoma and squamous cells are listed in Table S9. According to Table 5, the results of false discovery rate correction (q -value) indicated that none of

TABLE 4 Summary of parameters affecting the number of robust and non-robust radiomic features (from a total of 174 radiomics features)

Parameter	Size (mm)	COV \leq 5%	5% < COV \leq 10%	10% < COV \leq 20%	COV > 20%
Motion	8	40	25	32	77
	10	43	24	47	60
	13	48	25	35	66
	17	53	33	30	58
	R22	60	37	29	48
	L22	64	40	34	36
Reconstruction	8	33	19	47	75
	10	36	30	48	60
	13	77	50	34	13
	17	97	44	27	5
	R22	112	37	18	7
	L22	116	34	19	4
FWHM filter size	8	31	15	41	87
	10	39	33	37	65
	13	45	37	52	40
	17	55	56	35	28
	R22	73	50	35	16
	L22	80	56	35	28
Reconstruction + Motion	8	26	42	40	66
	10	31	37	40	66
	13	57	38	52	27
	17	62	47	48	17
	R22	75	40	35	24
	L22	83	42	34	15
FWHM filter size + Motion	8	27	43	27	77
	10	31	29	54	60
	13	33	33	55	53
	17	48	25	68	33
	R22	69	39	45	21
	L22	73	55	34	12

the radiomic features was significant for differentiating NSCLC histopathology subtypes.

4 | DISCUSSION

In this study, we aimed to investigate the synergistic impact of motion and parameters affecting PET/CT image features and understand the effectiveness of their use as robust features for differentiating the histopathological subtypes of tumors. The reproducibility and robustness of image features are the main aspects determining the success of radiomic analysis in clinical studies.⁴⁶ We performed a realistic phantom study using wall-less spheres mimicking lesions to avoid the partial volume effect. It must be noticed that the wall in a sphere represents a cold area within a warm area⁴⁷

to provide necrotic regions for a more realistic simulation of tumors. According to Vallières et al.⁴⁸ standardization of radiomic methods is highly recommended. As a consequence, IBSI was set to concatenate the radiomic workflow. Each step from feature extraction, to image processing and image segmentation are covered in the IBSI guidelines. Our study is compliant with IBSI guidelines and the “quality factors in radiomics studies” from radiomic feature extraction to reporting the results.

Previous research examined the reproducibility of texture features, but the impact of motion on PET radiomic features was rarely considered.^{27,29–31} In this study, we did not evaluate the impact of partial volume on the features as Hatt et al.⁴⁹ examined the impact of reconstruction considering the partial volume and the delineation of lesions concluding that the impact of partial volume correction on the features is relatively small.

TABLE 5 PET/CT images radiomic features categorized based on p -values (p -value < 0.05)

Features	AUC	p -value	q -value	COV of motion
ih_var	0.792	0.001	0.189	10.37
ih_mad	0.765	0.003	0.189	5.54
ngl_ldlge_3D	0.755	0.005	0.189	37.31
ngt_complexity_3D	0.75	0.006	0.189	21.56
szm_szlge_3D	0.732	0.011	0.189	34.54
ivh_v10	0.73	0.012	0.189	5.44
ih_p10	0.726	0.014	0.189	14.28
dzm_sdlge_3D	0.725	0.014	0.189	33.85
ih_cov	0.722	0.015	0.189	7.13
cm_joint_var_3D_comb	0.72	0.016	0.189	19.71
cm_joint_entr_3D_comb	0.712	0.021	0.189	1.94
szm_hgze_3D	0.71	0.022	0.189	5.49
dzm_hgze_3D	0.71	0.022	0.189	5.49
ih_medad	0.707	0.024	0.189	4.62
ivh_diff_v10_v90	0.705	0.026	0.189	6.32
ih_rmad	0.695	0.035	0.189	6.43
szm_lgze_3D	0.695	0.035	0.189	27.76
dzm_lgze_3D	0.695	0.035	0.189	27.76
stat_p90	0.695	0.035	0.189	32.76
szm_sze_3D	0.692	0.037	0.189	8.29
stat_medad	0.692	0.037	0.189	38.74
stat_mad	0.692	0.037	0.189	39.55
stat_var	0.692	0.037	0.189	67.93
szm_zsnu_norm_3D	0.687	0.042	0.189	14.46
rlm_gl_var_3D_comb	0.687	0.042	0.189	18.56
ngl_gl_var_3D	0.687	0.042	0.189	18.79
ih_qcod	0.686	0.045	0.189	9.25
stat_cov	0.685	0.045	0.189	17.86
stat_rmad	0.685	0.045	0.189	39.89
rlm_lrhge_3D_comb	0.682	0.049	0.189	8.20
ivh_diff_i10_i90	0.682	0.049	0.189	39.30
morph_area_dens_mvce	0.68	0.052	0.189	10.17

The radiomic features are sorted by p -value and the first 31 radiomic features showing p -values less than 5% are listed. In addition to p -values, the results of false discovery rate correction are shown in q -values column. In the first column, radiomic features with a COV less than 5% are highlighted in green. For a better display of the features showing a COV between 5% and 10% are highlighted in blue. In the last column, a COV less than 5%, $5\% < \text{COV} < 10\%$, $10\% < \text{COV} < 20\%$, and a COV $> 20\%$ are highlighted in dark blue, pale blue, pale red, and dark red, respectively.

Van Velden et al.²⁹ examined the impact of reconstruction and delineation on PET/CT images radiomic features of NSCLC. They showed that 60% of the features have $0.9 \leq$ intraclass correlation coefficient (ICC). Galavis et al.³¹ studied the reproducibility of radiomic features over different reconstructions and various protocols and demonstrated that 90% of texture features

showed a high variability ranging between 10% and 200%. Our results showed that 47% and 66% of the features are robust (COV $< 5\%$) against reconstruction with and without motion, respectively. In another study, Yan et al.³⁰ reported that 5% to 56% of radiomic features are highly affected by different reconstruction settings, including skewness and zone percentage. Compared to our study, these features showed a high variation over different reconstruction and other parameters, including kurtosis of intensity histogram and most of the statistics family, especially when considering motion, such as variance, energy, and mean absolute deviation radiomic features.

Issues linked with harmonizing the results for reproducing radiomic features in the previous study^{22,32} are considerable owing to differences in data acquisition protocols. Shiri et al.³² studied the impact of different reconstruction algorithms on PET radiomic features and demonstrated that the variability of radiomic features depends on the features. Our results also showed that lesion size and motion have a high impact on the features since a number of features are robust for large lesion size but are not for small lesion size. This observation should be considered in studies when different lesion sizes are involved. The above-referenced study showed that 44 of 100 features are robust (COV $< 5\%$) against different reconstruction settings, including GLCM (Entropy, Correlation), GLRLM (Long and short run emphasis), GLSZM (Zone percentage, small zone emphasis), NGLCM (Homogeneity, Entropy, Dissimilarity), SUV (Entropy, SUV_{mean}, SUV_{peak}). Thirty-six of the 44 robust features selected in this study were also robust in our work over different reconstruction settings without motion. Furthermore, 116 from 174 features were robust in large lesion size. It appears that these 36 features were basically robust against different reconstruction algorithms for different lesion sizes.

Oliver et al.²² investigated the variability of radiomic features between a 3D static acquisition and 4D gated images and reported a significant difference among radiomic features of a static and respiratory-gated image. They found that 26% of the features were robust between static and gated PET images, hence concluding that the variability of the features between two images is primarily due to respiratory motion. From 56 features examined in their study, they reported 14 robust features, of which 12 of 14 were robust in our study. In addition, our results showed that 64 of the 174 features were robust against motion for large lesion size. In this study, we aimed to examine the impact of respiratory motion on PET/CT images using a 3D static protocol with and without inducing movement. Motion and other changes in acquisition/reconstruction parameters were introduced individually and combined to evaluate the synergistic impact. The synergistic impact of motion and other parameters, such as image reconstruction and

lesion size, is severe and should be further considered in future studies, especially in multicenter clinical trials involving studies acquired on scanners from different vendors. Our results demonstrated that the robustness of features varies and that the overall change is not the average of the impact of two or more parameters taking place simultaneously.

The robustness of the features is lesion size-dependent. An increase in lesion size results in an increase in the number of robust features. Motion and lesion size have a major impact on the features. Motion had the highest impact on the features when other parameters were not changed. The combination of motion and other parameters decreased the effect of motion on the features or increased the number of robust features. For example, in the 22 mm sphere, 64, 105, 83, and 73 features were robust against motion and the combination of motion with test-retest, reconstruction, and FWHM filter size, respectively. This illustrates that the synergistic impact of acquisition/reconstruction parameters with motion mitigates the impact of lung movement. Motion had the most negligible impact against other parameters for small lesion size. Motion had the least impact when combined with different filter sizes. Changing the filter size decreases the impact of other parameters because of the averaging effect of the FWHM filter on images.

Sixty-four and 116 features showed less than 5% variability against motion and different reconstruction algorithms (individually) in large lesion size. This result will change if different reconstruction algorithms and motion are combined. It was observed that 83 features were robust against both factors. The same effect occurs in medium and small lesion sizes, where it was demonstrated that 35% and 17% of the features (62 and 31 features) were robust for medium and small tumor sizes. For multicenter studies, where different reconstruction algorithms and magnitude of lesions motion are commonly encountered, a combination of motion and reconstruction should be considered. Ninety-seven, 83, and 50 features were robust against post-injection scanning time in large, medium, and small lesions. These features belonged mostly to second-order texture features inferring from gray-tone spatial dependency matrices created from image intensity values. These features provide a metric linked to the spatial alignment of voxel intensities, and hence intra-lesion dissimilarity.⁴⁶ Overall, features showing a high COV against motion are not robust against the synergistic impact of combined factors. For instance, most of the robust features against the synergistic impact of motion and various reconstruction parameters were initially robust against motion individually, with the addition of some robust features against reconstruction individually.

Reproducibility is an aspect of radiomic analysis that requires careful assessment. Robust features

are not persistent, and changes in image acquisition/processing protocols might induce changes in their behavior. Features that are deemed robust and significant for differentiation include joint entropy of GLCM (AUC = 0.71, p -value < 0.05, COV = 0.019) and median absolute deviation of intensity histogram (AUC = 0.7, p -value < 0.05, COV = 0.046), which appear to be the best candidates for radiomic analysis. The results of the phantom study may be reliable owing to repeated imaging. The main limitation of this study was the low number of clinical studies. The effect of pre-processing on radiomic features was not investigated in this work and will be evaluated in future studies.

5 | CONCLUSION

We assessed the synergistic impact of motion and various reconstruction and acquisitions settings on the robustness of PET image radiomic features. The robustness of PET radiomic features appears to depend on the category they belong to and the lesion size. Low COV features are not guaranteed to be good candidates for differentiating between the histopathological subtypes of tumors in patients with NSCLC. It was observed that some features might have a low COV while lacking the significance of differentiation of the histopathological subtypes NSCLC. In addition, a number of features showed a high variability over different situations, including motion, while being significant with a high differentiation power. The suggested candidates for NSCLC differentiation using radiomic features are those presenting with a low COV while bringing significant differentiation power. The joint entropy of GLCM and median absolute deviation of intensity histogram respond to these criteria.

ACKNOWLEDGMENTS

This work was supported by the Tehran University of Medical Sciences under Grant No. 40235 and the Swiss National Science Foundation under Grant No. SNSF 320030_176052.

Open access funding provided by Universite de Geneve.

CONFLICT OF INTEREST


The authors have no conflicts to disclose.

ORCID

Seyyed Ali Hosseini 

<https://orcid.org/0000-0001-7542-7541>

Isaac Shiri  <https://orcid.org/0000-0002-5735-0736>

Ghasem Hajianfar 

<https://orcid.org/0000-0001-5359-2407>

Bahador Bahadorzadeh 

<https://orcid.org/0000-0002-7244-5215>

REFERENCES

1. Fitzmaurice C, Dicker D, Pain A, et al. The global burden of cancer 2013. *JAMA Oncol.* 2015;1(4):505-527.
2. Grootjans W, de Geus-Oei LF, Troost EG, Visser EP, Oyen WJ, Bussink J. PET in the management of locally advanced and metastatic NSCLC. *Nat Rev Clin Oncol.* 2015;12(7):395-407.
3. Garofano F, Gonzalez-Carmona MA, Skowasch D, et al. Clinical trials with combination of cytokine-induced killer cells and dendritic cells for cancer therapy. *Int J Mol Sci.* 2019;20(17):4307.
4. Czernin J, Allen-Auerbach M, Schelbert HR. Improvements in cancer staging with PET/CT: literature-based evidence as of September 2006. *J Nucl Med.* 2007;48(1):78S-88S.
5. Huang W, Fan M, Liu B, et al. Value of metabolic tumor volume on repeated 18F-FDG PET/CT for early prediction of survival in locally advanced non-small cell lung cancer treated with concurrent chemoradiotherapy. *J Nucl Med.* 2014;55(10):1584-1590.
6. Toma-Dasu I, Uhrdin J, Lazzeroni M, et al. Evaluating tumor response of non-small cell lung cancer patients with 18F-fluorodeoxyglucose positron emission tomography: potential for treatment individualization. *Int J Radiat Oncol Biol Phys.* 2015;91(2):376-384.
7. Chicklore S, Goh V, Siddique M, Roy A, Marsden PK, Cook GJ. Quantifying tumour heterogeneity in 18 F-FDG PET/CT imaging by texture analysis. *Eur J Nucl Med Mol Imaging.* 2013;40(1):133-140.
8. El Naqa I, Grigsby PW, Apte A, et al. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. *Pattern Recognit.* 2009;42(6):1162-1171.
9. Kumar V, Nath K, Berman CG, et al. Variance of standardized uptake values for FDG-PET/CT greater in clinical practice than under ideal study settings. *Clin Nucl Med.* 2013;38(3):175.
10. Fahey FH, Kinahan PE, Doot RK, Kocak M, Thurston H, Poussaint TY. Variability in PET quantitation within a multicenter consortium. *Med Phys.* 2010;37(7 Part1):3660-3666.
11. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014;5(1):1-9.
12. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology.* 2016;278(2):563-577.
13. Lambin P, Leijenaar RT, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol.* 2017;14(12):749-762.
14. Tsujikawa T, Yamamoto M, Shono K, et al. Assessment of intratumor heterogeneity in mesenchymal uterine tumor by an 18 F-FDG PET/CT texture analysis. *Ann Nucl Med.* 2017;31(10):752-757.
15. Xu R, Kido S, Suga K, et al. Texture analysis on 18 F-FDG PET/CT images to differentiate malignant and benign bone and soft-tissue lesions. *Ann Nucl Med.* 2014;28(9):926-935.
16. Nakajo M, Jinguji M, Shinaji T, et al. Texture analysis of 18F-FDG PET/CT for grading thymic epithelial tumours: usefulness of combining SUV and texture parameters. *Br J Radiol.* 2018;91(1083):20170546.
17. Yang F, Thomas MA, Dehdashti F, Grigsby PW. Temporal analysis of intratumoral metabolic heterogeneity characterized by textural features in cervical cancer. *Eur J Nucl Med Mol Imaging.* 2013;40(5):716-727.
18. Antunovic L, Gallivanone F, Sollini M, et al. [18F] FDG PET/CT features for the molecular characterization of primary breast tumors. *Eur J Nucl Med Mol Imaging.* 2017;44(12):1945-1954.
19. Tsujikawa T, Rahman T, Yamamoto M, et al. 18 F-FDG PET radiomics approaches: comparing and clustering features in cervical cancer. *Ann Nucl Med.* 2017;31(9):678-685.
20. Shiri I, Maleki H, Hajianfar G, et al. Next generation radiogenomics sequencing for prediction of EGFR and KRAS mutation status in NSCLC patients using multimodal imaging and machine learning approaches. *ArXiv Prepr ArXiv190702121.* Published online 2019.
21. Aggarwal N, Agrawal RK. First and second order statistics features for classification of magnetic resonance brain images. *J Signal Inf Process.* 2012;3(2):146-153.
22. Oliver JA, Budzevich M, Zhang GG, Dilling TJ, Latifi K, Moros EG. Variability of image features computed from conventional and respiratory-gated PET/CT images of lung cancer. *Transl Oncol.* 2015;8(6):524-534.
23. Cook GJ, Siddique M, Taylor BP, Yip C, Chicklore S, Goh V. Radiomics in PET: principles and applications. *Clin Transl Imaging.* 2014;2(3):269-276.
24. Tan S, Kligerman S, Chen W, et al. Spatial-temporal [18F] FDG-PET features for predicting pathologic response of esophageal cancer to neoadjuvant chemoradiation therapy. *Int J Radiat Oncol Biol Phys.* 2013;85(5):1375-1382.
25. Shiri I, Hajianfar G, Sohrabi A, et al. Repeatability of radiomic features in magnetic resonance imaging of glioblastoma: test-retest and image registration analyses. *Med Phys.* 2020;47(9):4265-4280.
26. Hosseini SA, Shiri I, Hajianfar G, Ghafarian P, Karam MB, Ay MR. The impact of preprocessing on the PET-CT radiomics features in non-small cell lung cancer. *Front Biomed Technol.* 2021;8(4).
27. Gallivanone F, Interlenghi M, D'Ambrosio D, Trifirò G, Castiglioni I. Parameters influencing PET imaging features: a phantom study with irregular and heterogeneous synthetic lesions. *Contrast Media Mol Imaging.* 2018;2018:5324517.
28. Edalat-Javid M, Shiri I, Hajianfar G, et al. Cardiac SPECT radiomic features repeatability and reproducibility: a multi-scanner phantom study. *J Nucl Cardiol.* 2021;28(6):2730-2744.
29. van Velden FH, Kramer GM, Frings V, et al. Repeatability of radiomic features in non-small-cell lung cancer [18F] FDG-PET/CT studies: impact of reconstruction and delineation. *Mol Imaging Biol.* 2016;18(5):788-795.
30. 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(11):1667-1673.
31. Galavis PE, Hollensen C, Jallow N, Paliwal B, Jeraj R. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. *Acta Oncol.* 2010;49(7):1012-1016.
32. 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(11):4498-4509.
33. Xu H, Lv W, Zhang H, Ma J, Zhao P, Lu L. Evaluation and optimization of radiomics features stability to respiratory motion in 18F-FDG 3D PET imaging. *Med Phys.* 2021;48(9):5165-5178.
34. Carles M, Torres-Espallardo I, Alberich-Bayarri A, et al. Evaluation of PET texture features with heterogeneous phantoms: complementarity and effect of motion and segmentation method. *Phys Med Biol.* 2016;62(2):652.
35. Li Y, Liu X, Qian Z, et al. Genotype prediction of ATRX mutation in lower-grade gliomas using an MRI radiomics signature. *Eur Radiol.* 2018;28(7):2960-2968.
36. Xu Q, Yuan K, Ye D. Respiratory motion blur identification and reduction in ungated thoracic PET imaging. *Phys Med Biol.* 2011;56(14):4481.
37. Surti S. Update on time-of-flight PET imaging. *J Nucl Med.* 2015;56(1):98-105.
38. Alessio AM, Rahmim A, Orton CG. Resolution modeling enhances PET imaging. *Med Phys.* 2013;40(12):120601.
39. 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(9):1323-1341.
40. Velazquez ER, Parmar C, Jermoumi M, et al. Volumetric CT-based segmentation of NSCLC using 3D-slicer. *Sci Rep.* 2013;3(1):1-7.
41. Ashrafinia S. Quantitative nuclear medicine imaging using advanced image reconstruction and radiomics. Published online 2019.

42. Orlhac F, Soussan M, Chouahnia K, Martinod E, Buvat I. ^{18}F -FDG PET-derived textural indices reflect tissue-specific uptake pattern in non-small cell lung cancer. *PLoS ONE*. 2015;10(12): e0145063.
43. Coroller TP, Agrawal V, Narayan V, et al. Radiomic phenotype features predict pathological response in non-small cell lung cancer. *Radiother Oncol*. 2016;119(3):480-486.
44. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol*. 1995;57(1):289-300.
45. van Timmeren JE, Carvalho S, Leijenaar RT, et al. Challenges and caveats of a multi-center retrospective radiomics study: an example of early treatment response assessment for NSCLC patients using FDG-PET/CT radiomics. *PLoS ONE*. 2019;14(6): e0217536.
46. Rizzo S, Botta F, Raimondi S, et al. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp*. 2018;2(1):1-8.
47. Wollenweber SD, Alessio AM, Kinahan PE. A phantom design for assessment of detectability in PET imaging. *Med Phys*. 2016;43(9):5051-5062.
48. Vallières M, Zwanenburg A, Badic B, Le Rest CC, Visvikis D, Hatt M. Responsible radiomics research for faster clinical translation. *J Nucl Med*. 2018;59(2):189-193.
49. Hatt M, Tixier F, Le Rest CC, Pradier O, Visvikis D. Robustness of intratumour ^{18}F -FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. *Eur J Nucl Med Mol Imaging*. 2013;40(11):1662-1671.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Hosseini SA, Shiri I, Hajianfar G, et al. Synergistic impact of motion and acquisition/reconstruction parameters on ^{18}F -FDG PET radiomic features in non-small cell lung cancer: Phantom and clinical studies. *Med Phys*. 2022;49:3783–3796.
<https://doi.org/10.1002/mp.15615>