



^{18}F -FDG-PET-based radiomics features to distinguish primary central nervous system lymphoma from glioblastoma



Ziren Kong^{a,1}, Chendan Jiang^{a,1}, Ruizhe Zhu^{b,1}, Shi Feng^b, Yaning Wang^a, Jiatong Li^{a,c}, Wenlin Chen^a, Penghao Liu^a, Dachun Zhao^d, Wenbin Ma^a, Yu Wang^a, Xin Cheng^{b,*}

^a Department of Neurosurgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, No.1 Shuaifuyuan Wangfujing Dongcheng District, Beijing, China

^b Department of Nuclear Medicine, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, No.1 Shuaifuyuan Wangfujing Dongcheng District, Beijing, China

^c School of Medicine, Tsinghua University, Haidian District, Beijing, China

^d Department of Pathology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, No.1 Shuaifuyuan Wangfujing Dongcheng District, Beijing, China

ARTICLE INFO

Keywords:

^{18}F -FDG
Positron emission tomography
Glioblastoma
Lymphoma
Radiomics

ABSTRACT

The differential diagnosis of primary central nervous system lymphoma from glioblastoma multiforme (GBM) is essential due to the difference in treatment strategies. This study retrospectively reviewed 77 patients (24 with lymphoma and 53 with GBM) to identify the stable and distinguishable characteristics of lymphoma and GBM in ^{18}F -fluorodeoxyglucose (FDG) positron emission tomography (PET) images using a radiomics approach. Three groups of maps, namely, a standardized uptake value (SUV) map, an SUV map calibrated with the normal contralateral cortex (ncc) activity (SUV/ncc map), and an SUV map calibrated with the normal brain mean (nbm) activity (SUV/nbm map), were generated, and a total of 107 radiomics features were extracted from each SUV map. The margins of the ROI were adjusted to assess the stability of the features, and the area under the curve (AUC) of the receiver operating characteristic curve of each feature was compared with the SUVmax to evaluate the distinguishability of the features. Ninety-five radiomics features from the SUV map were significantly different between lymphoma and GBM, 46 features were numeric stable after marginal adjustment, and 31 features displayed better performance than SUVmax. Features extracted from the SUV map demonstrated higher AUCs than features from the further calibrated maps. Tumors with solid metabolic patterns were also separately evaluated and revealed similar results. Thirteen radiomics features that were stable and distinguishable than SUVmax in every circumstance were selected to distinguish lymphoma from glioblastoma, and they suggested that lymphoma has a higher SUV in most interval segments and is more mathematically heterogeneous than GBM. This study suggested that ^{18}F -FDG-PET-based radiomics is a reliable noninvasive method to distinguish lymphoma and GBM.

1. Introduction

Primary central nervous system (CNS) lymphoma is a rare CNS disease in immunocompetent patients, which accounts for approximately 2% of all primary CNS tumors (Ostrom et al., 2017). The definitive diagnosis of CNS lymphoma is essential due to the differences in

strategies for surgery, chemotherapy and radiotherapy between CNS lymphomas and other malignant CNS tumors, such as glioblastoma multiforme (GBM) (Dunleavy, 2017; Lapointe et al., 2018). Although magnetic resonance imaging (MRI) is frequently applied for the initial evaluation of CNS neoplasms, the radiological differentiation of lymphoma from GBM remains difficult (i.e., lymphoma and GBM can both

Abbreviations: AUC, area under the curve; CNS, central nervous system; FDG, fluorodeoxyglucose; GBM, glioblastoma multiforme; GLCM, grey level cooccurrence matrix; GLDM, grey level dependence matrix; GLRLM, grey level run length matrix; ICC, intraclass correlation coefficient; IDH, isocitrate dehydrogenase; MRI, magnetic resonance imaging; MTV, metabolic tumor volume; nbm, normal brain mean; ncc, normal contralateral cortex; PET, positron emission tomography; ROC, receiver operating characteristic; ROI, regions of interest; ROIma, region of interest after marginal adjustment; SUV, standardized uptake value; T/N, tumor-to-normal contralateral cortex activity; TLG, total lesion glycolysis

* Corresponding authors.

E-mail address: pumch_chengxin@126.com (X. Cheng).

¹ These authors contributed equally to this work.

<https://doi.org/10.1016/j.nicl.2019.101912>

Received 7 April 2019; Received in revised form 22 June 2019; Accepted 25 June 2019

Available online 27 June 2019

2213-1582/ © 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

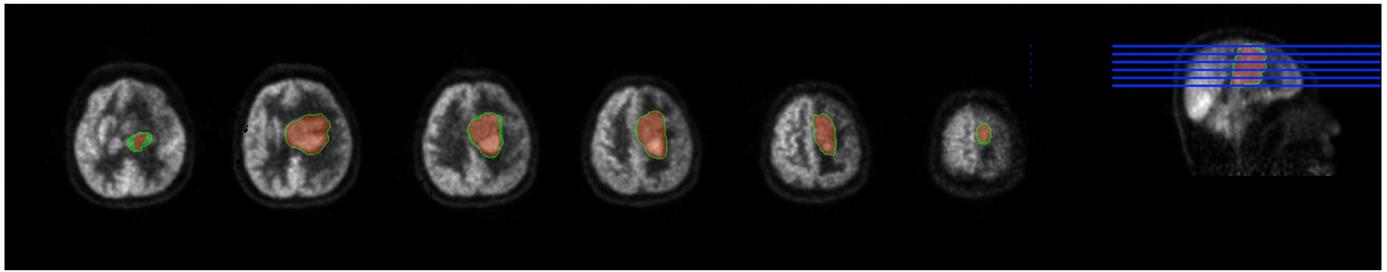


Fig. 1. Example of the three-dimensional tumor segmentation (red area) and marginal expansion (green area). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

present with a solid enhanced pattern without visible necrosis on T1-weighted contrast enhanced images).

^{18}F -fluorodeoxyglucose (FDG) positron emission tomography (PET) is an alternative imaging modality that has been utilized for the assessment of CNS disorders (Kim et al., 2016). CNS lymphomas usually have a relatively high tumor cell density and increased glucose metabolism, resulting in an increased standardized uptake value (SUV) (Kawai et al., 2005; Kasenda et al., 2013; El-Galaly et al., 2018). Although few studies have investigated the role of ^{18}F -FDG-PET for distinguishing CNS lymphoma from GBM (Yamaguchi et al., 2014; Purandare et al., 2017; Zhou et al., 2018), the parameters involved in these studies have been mostly restricted to specific features, such as the maximum SUV (SUV_{max}) and tumor-to-normal contralateral cortex activity (T/N) ratio, which can hardly display the characteristics of the whole tumor.

Radiomics is a recent emerging technique that extracts high-throughput imaging features to quantitatively describe the characteristics of a tumor, and to investigate the relationship between radiomics features and tumor phenotype (Lambin et al., 2017; Lee and Lee, 2018). Radiomics based on multimodal MRI images can detect the heterogeneity within brain tumors and has been widely utilized in the non-invasive prediction of clinical manifestation (Liu et al., 2018a), genetic characteristics (Li et al., 2018) and patient prognosis (Chaddad et al., 2016; Chaddad et al., 2018; Chaddad et al., 2019; Liu et al., 2018b). The ^{18}F -FDG-PET-based radiomics approach extracts conventional and textural features and has been successfully utilized for differential diagnosis, treatment response prediction, mutation detection of lung (Yip et al., 2017), cervical (Tsuji-kawa et al., 2017), and nasopharyngeal cancer (Lv et al., 2018). Although previous studies have demonstrated the differential MRI characteristics of CNS lymphoma and GBM taking advantage of the radiomics approach (Yamasaki et al., 2013; Alcaide-Leon et al., 2017; Kunimatsu et al., 2018; Kang et al., 2018; Suh et al., 2018; Kim et al., 2018), limited studies have focused on differentiation of these two cancers using ^{18}F -FDG-PET radiomics. This study retrospectively investigates the imaging characteristics of CNS lymphoma and GBM on ^{18}F -FDG-PET using a radiomics approach, and selects distinguishable radiomics features for the differential diagnosis of CNS lymphoma from GBM noninvasively.

2. Materials and methods

2.1. Patients

This study retrospectively enrolled 77 patients (24 with CNS lymphoma and 53 with GBM) treated at Peking Union Medical College Hospital between January 2010 and October 2018. The inclusion criteria were as follow: 1) age \geq 18 years old; 2) underwent surgical resection or biopsy with pathology confirmed primary CNS lymphoma or GBM; 3) had a preoperative ^{18}F -FDG-PET/CT scan of the brain. The exclusion criteria were as follow: 1) a history of brain tumors; 2) suspected or confirmed peripheral lymphomas; 3) treated with corticosteroids, radiotherapy or chemotherapy before surgery; 4) diabetes

mellitus (blood glucose \geq 10 mmol/L); 5) iatrogenic or disease-related immunosuppression. The study has been approved by the institutional review board, and all patients signed an informed consent form.

2.2. ^{18}F -FDG-PET/CT acquisition and preprocessing

All patients were required to fast for at least 4 h before ^{18}F -FDG-PET/CT scans were performed. The ^{18}F -FDG was produced on-site using RDS-111 Cyclotron (CTI, Knoxville, TN, USA). A dose of 5.55 MBq (0.15 mCi) ^{18}F -FDG per kilogram of body weight was administered intravenously under standardized conditions (in a quiet, dimly lit room with the patient's eyes closed). The scans were obtained 40–60 min after ^{18}F -FDG administration on a Biograph 64 TruePoint TrueV PET/CT system (Siemens Medical Solutions, Erlangen, Germany), which permitted 3-dimensional acquisition with interslice spacing of 3 mm. Images acquired from the PET/CT system were calibrated on the PET/CT workstation and the DICOM images underwent an interpolation with the workstation which doubled the physical resolution.

2.3. Image segmentation

Tumors were manually segmented by two experienced neurosurgeons who were blinded to the patients' information, and three-dimensional regions of interest (ROIs) were delimited on ^{18}F -FDG-PET images using the ITK-SNAP software (<http://www.itksnap.org/pmwiki/pmwiki.php>). A senior nuclear medical scientist subsequently reevaluated the segmented lesions. In cases of discrepancies $<$ 5% between the two neurosurgeons, the final ROI was defined as the overlapping area of their delineations, while in cases of discrepancies $>$ 5% between the two neurosurgeons, the nuclear medical scientist made the final decision. An example of the image segmentation is displayed in Fig. 1.

2.4. Postprocessing and feature extraction

The SUV value of each pixel was calculated with the following formula (NEMA, 2017; Paul Kinahan et al., 2018):

$$\text{SUV} = \frac{\text{Measure Activity}_{(\text{Bq/ml})}}{\frac{\text{Injected Dose}_{(\text{mCi})}}{\text{Decay Factor} \times \text{Body Weight}_{(\text{g})}}}$$

The calculation of SUV was carried out on an in-house software constructed with SimpleITK (<http://www.simpleitk.org>) (Lowekamp et al., 2013) and pydicom (<https://github.com/pydicom/pydicom>). The decay factor was computed with the half-life of the tracer and the time elapsed between the collection and the injection. All of the parameters were integrated in the DICOM metadata. The algorithm was adapted from the suggested vendor neutral pseudocode by the QIBA FDG-PET/CT Standardized Uptake Value (SUV) Technical Subcommittee (NEMA, 2017; Paul Kinahan et al., 2018).

Three groups of maps, namely, an SUV map, an SUV map with each pixel divided by the normal contralateral cortex (ncc) activity (SUV/ncc

map), and an SUV map with each pixel divided by the normal brain mean (nbm) activity (SUV/nbm map), were generated. During the creation of the SUV/nbm map, the skulls were stripped on the CT series and directly resampled with the nearest neighbor algorithm to the PET series. Features of these three groups of maps were extracted separately with identical parameters.

No further normalization of the brightness level was performed on processed maps. No resampling was carried out in order to maintain a higher resolution. Considering the dynamic range of the SUV maps, the bin width for matrices statistics was set to 1. A total of 107 radiomics features were extracted from the ROIs of each SUV map with pyradiomics version 2.1.0 (an open source package covered by the 3-clause BSD License, accessed at <http://www.radiomics.io/>) (van Griethuysen et al., 2017). Radiomics features were defined and calculated according to the formulas from pyradiomics (explanation document available at <http://pyradiomics.readthedocs.io/en/2.1.0/features.html>), and most of which are in accordance with features defined by Imaging Biomarker Standardization Initiative (Zwanenburg et al., 2016). The extracted features are summarized in Supplementary material 1.

2.5. Feature evaluation

Conventional features including SUVmax, T/N ratio, metabolic tumor volume (MTV) and total lesion glycolysis (TLG) were calculated and evaluated their performance in differentiating lymphoma and GBM.

The Mann–Whitney *U* test was performed to test whether a radiomics feature was significantly different between GBM and lymphoma. The discrimination performance of each radiomics feature was evaluated by the area under the curve (AUC) of the receiver operating characteristic (ROC) curve (Robin et al., 2011), whose classifier is a single feature threshold-based simple classifier (a decision tree classifier with one decision node). The AUC of the feature ‘First order_Maximum’ (also known as SUVmax in nonradiomics studies) was used as the threshold for assessing whether a feature was more discriminative than SUVmax.

2.6. Marginal adjustment

To evaluate the robustness of the features and the impact of ROI quality, the edge of the original ROI was expanded with a radius of 2 voxels in all spatial directions (2 mm in X-Y plane and 3 mm in the Z axis) to generate an ROI after marginal adjustment (ROI_{ma}). An example of the marginal adjustment is displayed in Fig. 1. Radiomics features from the SUV map, SUV/ncc map, and SUV/nbm map from the ROI_{ma} were also extracted. Intraclass correlation coefficient (ICC) between the features extracted from the ROI and ROI_{ma} in the SUV map were calculated using the variance components from a one-way ANOVA, and features were considered stable if ICC > 0.50 (Koo and Li, 2016).

2.7. Analysis of solid metabolic tumors

A solid metabolic tumor was defined as a lesion without a significant necrosis or cysts inside the ROI, which is known as atypical glioblastoma among GBM patients (Suh et al., 2018). Solid metabolic tumors for the whole patient group were manually selected, and their radiomics features from the abovementioned 3 radiomics maps (SUV map, SUV/ncc map and SUV/nbm map) with the two ROIs (initial ROI and ROI_{ma}) were and evaluated separately.

2.8. Feature selection and performance evaluation

Radiomics features were selected to differentiate lymphoma from GBM if: 1) significantly different between GBM and lymphoma ($p < 0.05$); 2) relatively stable after marginal adjustment

(ICC > 0.50); 3) performed better than ‘First order_Maximum’ in all 12 situations (in the SUV map, SUV/ncc map and SUV/nbm map with the initial ROI and ROI_{ma}, and in the whole population and in solid metabolic tumors). The threshold values of the selected features were calculated. The accuracy, sensitivity, specificity, precision rate, and recall rate of the selected features in the whole population were also calculated to evaluate the differential performances.

2.9. Feature validation with machine learning approach

AUC values with 5-fold cross validation were calculated for the selected radiomics features under 6 situations with the whole population (in the SUV map, SUV/ncc map and SUV/nbm map, with the initial ROI and ROI_{ma}) to better validate the prediction performances. Randomized grouping of the cross-validation was performed, and a single subgroup is retained as the validation cohort for testing the model while the remaining 4 subgroups are used as training cohort. The grouping was excluded if any of the subgroup contains only one type of tumor. The average AUC value of the 5-rounds of cross-validation (cross-validated AUC) was calculated to reflect the performance of the selected radiomics features.

2.10. Radiomics map generation and differentiation presentation

The radiomics maps were produced with a 21*21 square mask moving through each pixel in a selected 336*336 pixel layer, making the radiomics map to have 316*316 pixels. At each possible position, a set of radiomics features was calculated with the moving mask and reflected regional radiomics characteristics of the image. Radiomics maps were further utilized to present the diversities in lymphoma and GBM patients.

2.11. Statistical analysis

All statistical analyses were performed with R 3.5.1 (<https://www.r-project.org>), scikit-learn 0.20.1 (<https://scikit-learn.org>) and Python 3.6.6 (<https://www.python.org>). The analysis was performed with a normal computer (Intel Core i9-7940 × CPU @ 3.1 GHz base frequency, 64 GB RAM, Windows 10).

3. Results

3.1. Patient characteristics

The baseline characteristics of the 24 CNS lymphoma patients and 53 GBM patients are demonstrated in Table 1. There were significant differences in sex, lesion number, metabolic pattern, SUVmax, T/N ratio and MTV between two groups, while the variation in age and TLG did not reach statistical significance. The pathological information of the included patients was also acquired. All of the 24 lymphoma patients were diagnosed with diffuse large B cell lymphoma, and 2 (3.8%), 48 (90.6%), and 3 (5.6%) of the GBM patients had molecular features of isocitrate dehydrogenase (IDH)-mutant, IDH-wildtype, and IDH-unknown, respectively. There were 21 (87.5%) lymphoma patients and 21 (39.6%) GBM patients with a solid metabolic tumor pattern, and these patients were included in the analysis for solid tumors.

3.2. Differential performance of conventional features

SUVmax, T/N ratio and MTV and TLG were the conventional quantitative features that been evaluated, and they reached the AUC of 0.943, 0.870, 0.707 and 0.608, accuracy of 0.883, 0.844, 0.662 and 0.701 in differentiating CNS lymphoma from GBM, respectively.

Table 1
Characteristics of the lymphoma and glioblastoma patients.

Characteristics	Lymphoma	Glioblastoma	P-value
Sex			0.045
Male	11 (45.8%)	37 (69.8%)	
Female	13 (54.2%)	16 (30.2%)	
Age (mean \pm SD)	58.83 \pm 12.67	53.42 \pm 14.83	0.125
SUVmax (mean \pm SD)	23.76 \pm 8.26	13.50 \pm 4.37	< 0.001
T/N ratio (mean \pm SD)	5.77 \pm 1.90	4.13 \pm 1.37	0.001
MTV (mean \pm SD)	23.06 \pm 24.07	44.50 \pm 39.40	0.004
TLG (mean \pm SD)	565.24 \pm 539.03	385.89 \pm 404.12	0.109
Lesion number			0.007
Single	16 (66.7%)	52 (98.1%)	
Multiple	8 (33.3%)	1 (1.9%)	
Metabolic pattern			< 0.001
Solid	21 (87.5%)	21 (39.6%)	
Cystic	3 (12.5%)	32 (60.4%)	

Abbreviations: SUV, standardized uptake value; T/N, tumor to normal contralateral cortex activity; MTV, metabolic tumor volume; TLG, total lesion glycolysis.

3.3. Radiomics feature evaluation of the whole population

The heat map of the 107 radiomics features from the SUV map is shown in Fig. 2(A). Among the features extracted from each SUV map, 95 (88.8%), 89 (83.2%), and 90 (84.1%) features (for the SUV map, SUV/ncc map, SUV/nbm map, respectively) were significantly different ($p < 0.05$) between CNS lymphoma and GBM patients, with AUCs for differentiating CNS lymphoma and GBM ranging from 0.644 to 0.999. Most radiomics features from the SUV map demonstrated a similar or slightly higher AUC value compared with features from the SUV/nbm map, and features from both of these maps displayed moderately higher AUC values than features from the SUV/ncc map. The ICC of the 107 features extracted from the SUV map ranged from 0.01 to 0.84 with a mean value of 0.43, and 46 (43.0%) were relatively stable ($ICC > 0.50$) despite marginal adjustment. The statistical properties and differentiation performance of the radiomics features in the whole population are detailed in Supplementary material 2.

The radiomics feature ‘First order_Maximum’ showed an AUC of 0.943, 0.870, and 0.934 when differentiating CNS lymphoma and GBM in the SUV map, SUV/ncc map, and SUV/nbm map, respectively. Thirty-one features from the SUV map displayed better AUCs than ‘First order_Maximum’, and there were 33 features after ncc or nbm correction. The heat map of the 31 discriminating features from the SUV map is shown in Fig. 2(B). Features that had better differentiation performance than ‘First order_Maximum’ in all three SUV maps are shown in

Fig. 3.

There was a slight increase or decrease in the performance of each radiomics feature after margin adjustment, and most features exhibited the same trend of variation in all three SUV maps. As expected, the performance of ‘First order_Maximum’ remained basically the same after margin adjustment, and 2 of the 31 distinguishing features from the SUV map performed worse than ‘First order_Maximum’ after margin adjustment, while 6 new better features were identified. The fluctuation of performance in radiomics features was larger in the SUV/ncc and SUV/nbm maps than in SUV map, and there were also greater number of newly identified features with the two calibrated maps. Features that performed better than ‘First order_Maximum’ were also ranked by AUC values from each of the three radiomics maps before and after margin adjustment to demonstrate stability, and the rankings are displayed in Supplementary material 3.

3.4. Analysis for solid metabolic tumors

The differentiation performance of most radiomics features for solid metabolic tumors had minimal variation compared with that for the whole population, although the number of distinguishing features decreased by 6, 6, and 1 from the SUV map, SUV/ncc map, and SUV/nbm map, respectively. In accordance with the findings in the whole population, most features remained stable after margin adjustment. The heat map of the 25 distinguishing features from the SUV map is shown in Fig. 2(C). The statistical properties and differentiation performances for solid metabolic tumors are shown in Supplementary material 4. Features that had better differentiation performance than ‘First order_Maximum’ for solid metabolic tumors are displayed in Fig. 3. The rankings of features before and after margin adjustment are detailed in Supplementary material 5.

3.5. Feature selection and performance evaluation

Among the 107 involved radiomics features, 95 were significantly different between GBM and lymphoma ($p < 0.05$), 46 were relatively stable after marginal adjustment ($ICC > 0.50$), and 15 outperformed ‘First order_Maximum’ in all of the 12 abovementioned circumstances. Finally, 13 features that met the selection criteria were selected for the noninvasive differential diagnosis of lymphoma and GBM. In the SUV map, the selected radiomics features showed AUCs ranged from 0.971 to 0.998 in distinguish lymphoma and GBM in the whole population, and all of them demonstrated higher accuracy, sensitivity, specificity, precision rate and recall rate than ‘First order_Maximum’. The discrimination performance of the selected radiomics features are

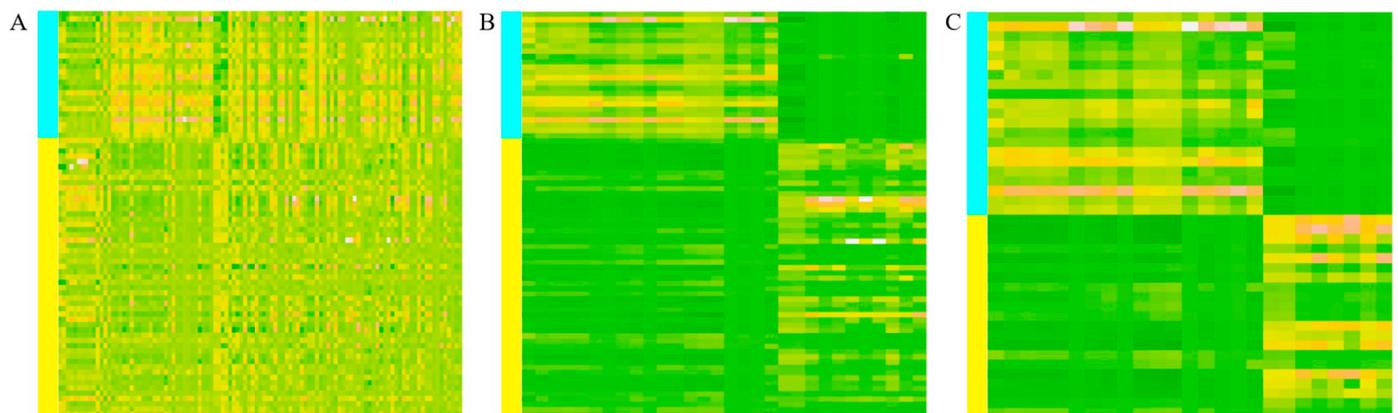


Fig. 2. Heat maps of all radiomics features in the whole population (A), the 31 discriminating features in the whole population (B), and the 25 discriminating features in the solid metabolic tumors (C). Heat maps were clustered by pathological information, with the blue column on the left indicating lymphoma patients and the yellow column indicating glioblastoma patients. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

0.944	0.918	0.95	0.917	0.85	0.92	0.977	0.918	0.984	0.932	0.846	0.957	original_firstorder_10Percentile
0.984	0.921	0.969	0.98	0.918	0.965	0.98	0.923	0.973	0.98	0.921	0.97	original_firstorder_90Percentile
0.897	0.84	0.91	0.94	0.887	0.933	0.907	0.853	0.927	0.943	0.884	0.934	original_firstorder_Entropy
0.9	0.845	0.903	0.94	0.895	0.943	0.925	0.871	0.925	0.94	0.898	0.943	original_firstorder_InterquartileRange
0.98	0.932	0.966	0.974	0.917	0.959	0.986	0.916	0.982	0.977	0.905	0.975	original_firstorder_Mean
0.88	0.831	0.896	0.94	0.885	0.939	0.909	0.857	0.921	0.94	0.889	0.943	original_firstorder_MeanAbsoluteDeviation
0.977	0.926	0.966	0.966	0.902	0.961	0.986	0.914	0.984	0.97	0.896	0.977	original_firstorder_Median
0.925	0.892	0.921	0.883	0.814	0.862	0.918	0.882	0.923	0.789	0.723	0.89	original_firstorder_Minimum
0.895	0.837	0.897	0.94	0.891	0.94	0.918	0.864	0.918	0.939	0.896	0.946	original_firstorder_RobustMeanAbsoluteDeviation
0.984	0.932	0.973	0.977	0.918	0.967	0.984	0.925	0.98	0.973	0.907	0.973	original_firstorder_RootMeanSquared
0.903	0.863	0.888	0.942	0.881	0.924	0.948	0.896	0.896	0.94	0.875	0.916	original_firstorder_Uniformity
0.887	0.822	0.893	0.94	0.877	0.936	0.905	0.846	0.918	0.939	0.884	0.948	original_firstorder_Variance
0.944	0.864	0.948	0.956	0.91	0.964	0.941	0.837	0.948	0.955	0.9	0.97	original_glcmm_Autocorrelation
0.856	0.785	0.873	0.921	0.857	0.928	0.876	0.787	0.88	0.925	0.857	0.936	original_glcmm_ClusterProminence
0.886	0.804	0.881	0.929	0.87	0.921	0.891	0.828	0.907	0.934	0.873	0.926	original_glcmm_ClusterTendency
0.986	0.929	0.962	0.991	0.944	0.958	0.989	0.923	0.966	0.996	0.936	0.962	original_glcmm_Contrast
0.99	0.93	0.962	0.995	0.945	0.958	0.993	0.923	0.966	0.999	0.936	0.962	original_glcmm_DifferenceAverage
0.978	0.925	0.96	0.98	0.94	0.955	0.98	0.918	0.964	0.98	0.936	0.959	original_glcmm_DifferenceEntropy
0.971	0.924	0.961	0.974	0.938	0.955	0.977	0.914	0.964	0.977	0.932	0.959	original_glcmm_DifferenceVariance
0.991	0.931	0.962	0.998	0.945	0.958	0.991	0.923	0.966	0.999	0.936	0.962	original_glcmm_Id
0.991	0.931	0.962	0.997	0.945	0.958	0.991	0.923	0.966	0.998	0.936	0.962	original_glcmm_Idm
0.926	0.895	0.877	0.92	0.888	0.866	0.92	0.923	0.862	0.932	0.909	0.858	original_glcmm_Idn
0.726	0.932	0.962	0.763	0.946	0.958	0.671	0.925	0.966	0.715	0.941	0.962	original_glcmm_InverseVariance
0.946	0.864	0.947	0.957	0.911	0.966	0.943	0.834	0.948	0.959	0.896	0.975	original_glcmm_JointAverage
0.926	0.889	0.934	0.96	0.909	0.943	0.936	0.909	0.936	0.968	0.914	0.943	original_glcmm_JointEnergy
0.932	0.875	0.932	0.962	0.903	0.945	0.936	0.884	0.939	0.962	0.898	0.941	original_glcmm_JointEntropy
0.902	0.88	0.906	0.945	0.882	0.905	0.92	0.909	0.918	0.962	0.88	0.907	original_glcmm_MaximumProbability
0.946	0.864	0.947	0.957	0.911	0.966	0.943	0.834	0.948	0.959	0.896	0.975	original_glcmm_SumAverage
0.885	0.86	0.924	0.936	0.896	0.942	0.891	0.873	0.936	0.94	0.896	0.939	original_glcmm_SumEntropy
0.873	0.829	0.906	0.93	0.877	0.928	0.896	0.857	0.927	0.936	0.882	0.932	original_glcmm_SumSquares
0.725	0.882	0.943	0.764	0.908	0.951	0.76	0.882	0.941	0.794	0.896	0.952	original_gldmm_DependenceEntropy
0.902	0.897	0.969	0.94	0.894	0.965	0.952	0.864	0.968	0.969	0.845	0.968	original_gldmm_DependenceNonUniformityNormalized
0.879	0.821	0.874	0.923	0.797	0.447	0.891	0.594	0.465	0.957	0.773	0.517	original_gldmm_DependenceVariance
0.687	0.835	0.906	0.94	0.881	0.93	0.905	0.859	0.927	0.939	0.882	0.932	original_gldmm_GrayLevelVariance
0.937	0.851	0.948	0.953	0.907	0.968	0.925	0.828	0.948	0.952	0.896	0.97	original_gldmm_HighGrayLevelEmphasis
0.987	0.936	0.97	0.98	0.948	0.965	0.991	0.923	0.973	0.998	0.939	0.966	original_gldmm_LargeDependenceEmphasis
0.688	0.81	0.906	0.673	0.859	0.94	0.732	0.789	0.689	0.742	0.639	0.94	original_gldmm_LargeDependenceHighGrayLevelEmphasis
0.952	0.888	0.96	0.947	0.92	0.971	0.986	0.859	0.964	0.982	0.909	0.989	original_gldmm_LargeDependenceLowGrayLevelEmphasis
0.919	0.84	0.946	0.892	0.877	0.96	0.793	0.785	0.948	0.95	0.846	0.97	original_gldmm_LowGrayLevelEmphasis
0.999	0.94	0.954	0.998	0.924	0.965	0.998	0.923	0.932	1	0.907	0.975	original_gldmm_SmallDependenceEmphasis
0.961	0.875	0.964	0.968	0.903	0.978	0.957	0.839	0.959	0.968	0.878	0.98	original_gldmm_SmallDependenceHighGrayLevelEmphasis
0.602	0.609	0.658	0.94	0.87	0.937	0.607	0.607	0.664	0.939	0.666	0.946	original_glrmm_GrayLevelNonUniformityNormalized
0.894	0.807	0.877	0.932	0.876	0.943	0.9	0.8	0.875	0.936	0.886	0.946	original_glrmm_GrayLevelVariance
0.926	0.827	0.944	0.943	0.888	0.968	0.927	0.796	0.943	0.948	0.88	0.968	original_glrmm_HighGrayLevelRunEmphasis
0.986	0.924	0.961	0.987	0.934	0.962	1	0.912	0.955	1	0.925	0.962	original_glrmm_LongRunEmphasis
0.954	0.906	0.969	0.946	0.928	0.974	0.984	0.889	0.98	0.973	0.923	0.99	original_glrmm_LongRunLowGrayLevelEmphasis
0.921	0.813	0.948	0.899	0.87	0.968	0.93	0.777	0.941	0.948	0.85	0.975	original_glrmm_LowGrayLevelRunEmphasis
0.998	0.942	0.98	1	0.945	0.969	0.998	0.93	0.975	1	0.932	0.966	original_glrmm_RunLengthNonUniformityNormalized
0.991	0.936	0.969	0.996	0.948	0.965	0.996	0.921	0.968	1	0.939	0.966	original_glrmm_RunPercentage
0.974	0.911	0.942	0.979	0.927	0.958	0.989	0.898	0.93	0.995	0.918	0.95	original_glrmm_RunVariance
0.998	0.924	0.946	1	0.924	0.958	1	0.907	0.93	1	0.912	0.95	original_glrmm_ShortRunEmphasis
0.941	0.846	0.958	0.956	0.882	0.972	0.939	0.823	0.95	0.95	0.864	0.966	original_glrmm_ShortRunHighGrayLevelEmphasis
0.962	0.88	0.906	0.946	0.832	0.898	0.943	0.814	0.855	0.916	0.762	0.866	original_glszmm_LargeAreaEmphasis
0.962	0.902	0.944	0.946	0.853	0.943	0.97	0.85	0.927	0.957	0.807	0.943	original_glszmm_LargeAreaLowGrayLevelEmphasis
0.986	0.886	0.902	0.957	0.774	0.66	0.973	0.823	0.864	0.927	0.714	0.68	original_glszmm_ZonePercentage
0.954	0.88	0.878	0.943	0.836	0.856	0.932	0.823	0.794	0.912	0.755	0.742	original_glszmm_ZoneVariance
0.962	0.807	0.807	0.963	0.848	0.855	0.934	0.667	0.719	0.946	0.726	0.78	original_ngtdmm_Busyness
0.913	0.85	0.924	0.926	0.886	0.933	0.902	0.837	0.927	0.923	0.896	0.936	original_ngtdmm_Complexity
0.958	0.811	0.759	0.98	0.816	0.772	0.962	0.889	0.72	0.968	0.853	0.742	original_ngtdmm_Contrast
0.943	0.82	0.84	0.961	0.87	0.888	0.909	0.696	0.789	0.932	0.786	0.79	original_ngtdmm_Strength
0.943	0.87	0.934	0.943	0.868	0.934	0.946	0.878	0.934	0.946	0.875	0.934	original_firstorder_Maximum
T	T_ncc	T_nbm	T_ma	T_manc	T_marbm	T_sm	T_sm_ncc	T_sm_nbm	T_sm_ma	T_sm_manc	T_sm_marbm	

(caption on next page)

Fig. 3. Name and performance of the features that had better differentiation performance than ‘First order_Maximum’ in any of the 12 situations. The performances of ‘First order_Maximum’ are labeled in green at the bottom of the figure, and features that outperformed ‘First order_Maximum’ are labeled in red, with a stronger color indicating higher AUC values. Features that performed worse than ‘First order_Maximum’ are labeled blue. Abbreviations: T, tumor; ncc, calibrated by the normal contralateral cortex activity; nbm, calibrated by the normal brain mean activity; ma, marginal adjustment; sm, solid metabolic. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2

The discrimination performance of the selected radiomics features in comparison with ‘First order_Maximum’.

Selected features	AUC	ACC	SEN	SPE	PRE	REC	ICC
First order_90 percentile	0.984	0.948	0.943	0.958	0.980	0.943	0.793
First order_Mean	0.980	0.948	0.962	0.917	0.962	0.962	0.803
First order_Median	0.977	0.948	0.962	0.917	0.962	0.962	0.792
First order_Root mean squared	0.984	0.961	0.981	0.917	0.963	0.981	0.807
GLCM_Contrast	0.986	0.948	0.962	0.917	0.962	0.962	0.593
GLCM_Difference average	0.990	0.948	0.943	0.958	0.980	0.943	0.765
GLCM_Difference entropy	0.976	0.935	0.962	0.875	0.944	0.962	0.789
GLCM_Difference variance	0.971	0.909	0.906	0.917	0.960	0.906	0.580
GLCM_Inverse difference	0.991	0.935	0.906	1.000	1.000	0.906	0.811
GLCM_Inverse difference moment	0.991	0.935	0.906	1.000	1.000	0.906	0.819
GLRLM_Run length non-uniformity normalized	0.998	0.974	0.962	1.000	1.000	0.962	0.835
GLRLM_Run percentage	0.991	0.974	0.981	0.958	0.981	0.981	0.793
GLDM_Large dependence emphasis	0.987	0.974	0.981	0.958	0.981	0.981	0.724
First order_Maximum (comparison)	0.943	0.883	0.887	0.875	0.940	0.887	0.716

Abbreviations: GLCM, grey-level co-occurrence matrix; GLRLM, grey level run length matrix; GLDM, grey level dependence matrix; AUC, area under curve; ACC, accuracy; SEN, sensitivity; SPE, specificity; PRE, precision rate; REC, recall rate; ICC, intraclass correlation coefficient.

displayed in Table 2.

3.6. Feature validation with machine learning approach

The cross-validated AUC values of the ‘First order_Maximum’ ranged from 0.690 to 0.815 in all of the 6 circumstances with the whole population, which were lower than AUCs without cross-validation. Most of the 13 selected radiomics features displayed a cross-validated AUCs higher than that of the ‘First order_Maximum’ but also lower than their original AUCs. The 13 features displayed cross-validated AUCs ranging from 0.784 to 0.969 in the SUVmap in the differential diagnosis with the whole population. The cross-validated AUCs of the selected radiomics features with the whole population under the 6 circumstances are presented in Table 3.

3.7. Differentiation presentation

One lymphoma and one GBM patient who shared similar metabolic types of tumors and had comparable SUVmax values were selected for discrimination using radiomics maps. The radiomics maps of these two patients are demonstrated in Fig. 4. The selected features,

Table 3

The area under curve values based on cross validation in six circumstances with whole population.

Selected features	T	T.ncc	T.nbm	T.ma	T.ma.ncc	T.ma.nbm
First order_90 percentile	0.850	0.941	0.967	0.929	0.801	0.829
First order_Mean	0.833	0.757	0.878	0.750	0.900	0.857
First order_Median	0.969	0.801	0.944	0.857	0.800	0.900
First order_Root mean squared	0.938	0.740	0.955	0.800	0.821	0.889
GLCM_Contrast	0.893	0.705	1.000	0.864	0.847	0.875
GLCM_Difference average	0.917	0.714	0.969	0.958	0.717	1.000
GLCM_Difference entropy	0.857	0.795	0.967	0.857	0.878	0.964
GLCM_Difference variance	0.784	0.958	0.929	0.875	0.781	0.917
GLCM_Inverse difference	0.955	0.917	0.829	0.929	0.815	0.892
GLCM_Inverse difference moment	0.857	0.702	0.864	0.900	0.750	0.839
GLRLM_Run length non-uniformity normalized	0.964	0.833	1.000	0.969	0.833	0.962
GLRLM_Run percentage	0.955	1.000	0.857	0.964	0.800	0.900
GLDM_Large dependence emphasis	0.929	0.900	0.900	1.000	0.900	0.889
First order_Maximum (comparison)	0.750	0.690	0.815	0.750	0.739	0.762

Abbreviations: T, tumor; ncc, calibrated by the normal contralateral cortex activity; nbm, calibrated by the normal brain mean activity; ma, marginal adjustment; GLCM, grey-level co-occurrence matrix; GLRLM, grey level run length matrix; GLDM, grey level dependence matrix.

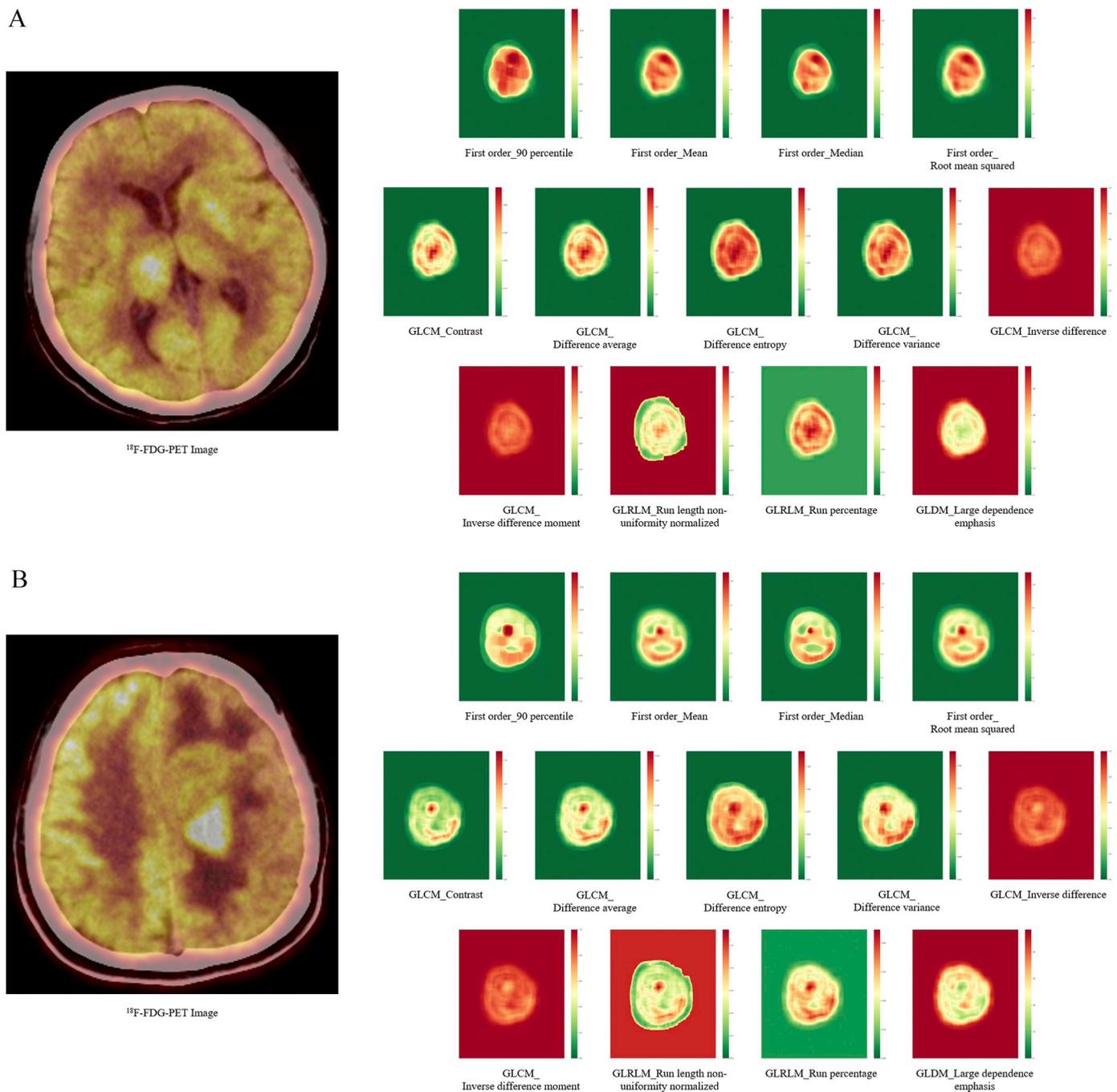


Fig. 4. Radiomics maps of the selected radiomics features in a lymphoma patient (A) and a glioblastoma patient (B). First order features demonstrated a higher SUV in the region of interest and the texture features displayed a more mathematically heterogeneous of lymphomas. Although the radiomics maps presented here are different from what have been calculated in radiomics feature analysis, they provide an intuitive way for visualization.

SUVmax or the T/N ratio and reached an AUC of approximately 0.90 (Yamaguchi et al., 2014; Purandare et al., 2017; Zhou et al., 2018). Our study demonstrated a similar result, with AUC of SUVmax reached 0.943 and AUC of the T/N ratio reached 0.870 using a single feature threshold-based simple classifier (consistent with previous studies). However, these two parameters only reflect limited (although representative) information regarding the whole tumor, and the diagnostic performance can be further improved if the imaging characteristics of the tumor are comprehensively investigated. The current study explored the distinguishing role of shape, first-order and texture features on ^{18}F -FDG-PET using a radiomics approach and selected 13 distinguishing radiomics features. Each single selected radiomics feature displayed excellent discrimination performance, reaching an AUC

of 0.971–0.998 in the whole population and 0.977–1.00 in solid metabolic tumors, and this outstanding performance weakened the needs for establishing a mathematic model that combines the selected features together to achieve better performance than each individual feature. As a comparison, previous studies based on multiparametric MRI radiomics (including conventional MRI sequences and advanced MRI sequences) to differentiate CNS lymphoma from GBM selected radiomics features and built mathematic models, and these models reached AUCs ranging from 0.877–0.956 (not superior to all of the single selected ^{18}F -FDG-PET radiomics features) (Yamasaki et al., 2013; Alcaide-Leon et al., 2017; Kunimatsu et al., 2018; Kang et al., 2018; Suh et al., 2018; Kim et al., 2018), suggesting the dependability of ^{18}F -FDG-PET radiomics for differentiating CNS lymphoma from GBM. In addition to the higher

Table 4
Selected radiomics features, their discrimination performances and threshold value.

Selected features	Threshold	Values of two specific patients		
		Lymphoma	Glioblastoma	Performance
First order_90 percentile	11.72	14.51	10.81	Accurate
First order_Mean	8.380	12.12	7.508	Accurate
First order_Median	8.171	11.98	7.236	Accurate
First order_Root mean squared	9.430	12.24	7.822	Accurate
GLCM_Contrast	1.126	1.167	0.7556	Accurate
GLCM_Difference average	0.747	0.7672	0.5762	Accurate
GLCM_Difference entropy	1.500	1.504	1.313	Accurate
GLCM_Difference variance	0.484	0.5463	0.4126	Accurate
GLCM_Inverse difference	0.686	0.6763	0.7396	Accurate
GLCM_Inverse difference moment	0.670	0.6561	0.7298	Accurate
GLRLM_Run length non-uniformity normalized	0.508	0.5174	0.3889	Accurate
GLRLM_Run percentage	0.651	0.6555	0.5280	Accurate
GLDM_Large dependence emphasis	122.4	121.8	205.2	Accurate

Abbreviations: GLCM, grey-level co-occurrence matrix; GLRLM, grey level run length matrix; GLDM, grey level dependence matrix.

AUC values, the 13 selected features also demonstrated higher performances in alternative evaluating indicators (e.g., accuracy, sensitivity, specificity, precision rate and recall rate), which illustrates the balance and reliability of these features. The selected features were also validated using a machine learning approach, and most of their cross-validated AUCs were higher than that of the ‘First order_Maximum’, indicating an insignificant overfitting of the final features. It is also worth noting that despite the considerable differences in solid metabolic tumors and cystic metabolic tumors recognized by the naked eye, there was a high degree of coincidence in the distinguishable features in the whole population and in solid metabolic tumors, suggesting the capability of the radiomics approach to differentiate CNS lymphoma and GBM regardless of metabolic patterns. Nevertheless, ^{18}F -FDG-PET has a relatively lower resolution than conventional MRI and the Gaussian convolution filter, which is a part of the standard protocol in the reconstruction of PET image, unsharpens the details of the images. Taking the loss of information when resampling into consideration, no resampling was carried out in order to maintain the relatively higher resolution. In addition, features from specific frequency domains in ^{18}F -FDG-PET may be inaccurate, and therefore, wavelet features were excluded in our study.

In general, radiomics features extracted from the SUV map performed the best compared to the features extracted from the SUV/ncc and SUV/nbm maps, although it did least in calibrating differences among patients. Features from the SUV/ncc map displayed acceptable but lower discrimination performance than features derived from the SUV/nbm map, suggesting that normal contralateral cortex activity may not be the most accurate parameter for correction in radiomics studies. An unavoidable shortcoming of the additional calibration of the SUV map is the reduction in the dynamic range of the map. Since the calibration factor (e.g., ncc or nbm in our study) is usually > 1.0 , each pixel would have lower values after calibration, leading to a decrease in matrix resolution and fewer details presented by radiomics features. Since the generation of SUV has experienced a number of corrections, additional calibration may be unnecessary in CNS ^{18}F -FDG-PET radiomics studies.

The accuracy of most radiomics features relies on precise segmentation of the tumor, yet manual segmentations may have minor inaccuracies at the edges of the ROI, and stability of radiomics features were measured through marginal adjustment. Marginal expansion of 2 voxels (2 mm in X-Y plane and 3 mm in the Z axis) was chosen because it was a possible distance with a temperate size considering the resolution of the DICOM data. Features that were distinguishable in all of the 12 circumstances but numeric unstable after marginal adjustment (e.g., ‘GLRLM_Long run low grey level emphasis’) were excluded during the final selection of features since the significant influence of margin to

feature value may affect the settlement of cutoff value and thereby restrict their clinical application, and the mean ICC of the final selected features reached 0.761, demonstrating a good intraclass agreement. Features that were numeric stable but not distinguishable in all circumstances (e.g., ‘First order_10 percentile’) were also excluded due to the inter-reader discrepancy of segmentations. Interestingly, the number of distinguishable features increased after margin adjustment regardless of the calibration method or analyzed population, suggesting that the peri-tumor area may also contain information that can distinguish CNS lymphoma and GBM. Radiomics features in such peri-tumor area as well as the intratumoral habitat regions from multimodal imaging may provide additional information about tumor properties, and the biological processes underlying these imaging features remains to be investigated (Wu et al., 2016; Wu et al., 2017; Cui et al., 2017; Wu et al., 2018).

First-order features refer to imaging characteristics exhibited by voxels alone. In addition to ‘First order_Maximum’, the ‘90 percentile’, ‘10 percentile’, ‘Mean’, ‘Median’, ‘Root mean squared’ of the SUV values in ROIs can also distinguish CNS lymphoma from GBM (both in the whole population and in patients with solid metabolic tumors), suggesting a higher SUV of most interval segments of lymphomas no matter the metabolic patterns. Although lymphoma is typically considered to have a uniform metabolic pattern compared with glioma, this characteristic may have less discriminatory capacity because the first order features that are relevant to distribution deviation (e.g., ‘Skewness’, ‘Kurtosis’, ‘Skewness’, ‘Variance’) did not surpass ‘First order_Maximum’. Texture features refer to the presentation of textures exhibited by special defined grey level matrices, and each type of grey level matrix reflects certain aspects of the image, including the relationships to the adjacent voxel, voxel blocks and linear scales. The grey level cooccurrence matrix (GLCM) expresses the distribution of neighboring voxels, and 6 of all the 13 selected features were derived from the GLCM, with ‘GLCM_Contrast’ measures the local intensity variation, ‘GLCM_Difference average’, ‘GLCM_Entropy’ and ‘GLCM_Variance’ describe the differences in neighborhood intensity, and ‘GLCM_Inverse difference’ and ‘GLCM_Inverse difference moment’ measure the local homogeneity. These 6 GLCM features indicated that GBM had more homogeneity and lower contrast adjacent voxel relationships than lymphoma did (note: the ‘homogeneity’ here is supported by the value of radiomics features calculated directly from the pixels of SUV maps instead of the ratio of pixels to a certain threshold [e.g., SUVmax], which is more favorable for judgment made by the naked eye). Two selected features were derived from the grey level run length matrix (GLRLM), which counts the maximum length of voxels with identical grey levels in a certain direction. ‘GLRLM_Run length non-uniformity normalized’ measures the similarity of run lengths, and

'GLRLM_Run percentage' illustrates the coarseness of textures. The final selected feature was derived from grey level dependence matrix (GLDM), which is constructed with a given distance and calculated based on the center voxel. The differences in these 3 GLRLM and GLDM derived features also suggested that GBM was more uniform and homogeneous than lymphoma. It is also worth mentioning that alternative GLRLM features (e.g., 'GLRLM_Short run emphasis') are also numeric stable and outperformed 'First order_Maximum' in 11 of the 12 situations and may also be included for distinguishing lymphoma from GBM. On the other hand, shape features and other texture features (features from the grey level size zone matrix and neighboring grey tone difference matrix) may not be distinguishing characteristics compared with 'First order_Maximum'.

The current study has a few limitations. First, the patient population included in this study was relatively small, and ^{18}F -FDG-PET examination is not mandatory in patients with CNS diseases, both of which may have introduced selection bias. The limited number of patients also prevented validation with an independent cohort to assess our findings, and further large-scale prospective studies are needed. Second, the study was performed without information regarding post therapy changes due to the variations in therapeutic principles and regimens, and more comprehensive clinical data are needed to discover the prognostic values of the current findings. Third, in addition to GBM, CNS lymphomas may need to differentiate with other CNS malignancies (e.g., high grade astrocytoma or oligodendroglioma with contrast-enhancement on MRI and solid metabolic pattern on ^{18}F -FDG-PET) or alternative CNS disorders (e.g., infection, inflammation or demyelination), and further study for noninvasive differentiation may be carried out. And finally, despite the good discriminating capacities of the selected features, the biological processes that underlie these features remains to be understood.

5. Conclusions

^{18}F -FDG-PET-based radiomics provides a reliable noninvasive method for distinguishing CNS lymphoma from GBM. Several stable and highly discriminative first-order and texture features suggest that lymphoma has a higher SUV in most interval segments and is more mathematically heterogeneous than GBM.

Funding

This study was supported by the Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences [grant numbers 2016-I2M-2-001 and No. 2018-I2M-3-001] and the Fundamental Research Funds for the Central Universities [grant number 3332018029].

Declaration of interests

None.

Acknowledgment

The authors thank American Journal Experts for providing language help.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.nicl.2019.101912>.

References

Alcaide-Leon, P., Dufort, P., Geraldo, A.F., et al., 2017. Differentiation of enhancing glioma and primary central nervous system lymphoma by texture-based machine learning. *Am. J. Neuroradiol.* 38 (6), 1145–1150.

- Chaddad, A., Desrosiers, C., Hassan, L., Tanougast, C., 2016. A quantitative study of shape descriptors from glioblastoma multiforme phenotypes for predicting survival outcome. *Br. J. Radiol.* 89 (1068), 20160575.
- Chaddad, A., Sabri, S., Niazi, T., Abdulkarim, B., 2018. Prediction of survival with multi-scale radiomic analysis in glioblastoma patients. *Med. Biol. Eng. Comput.* 56 (12), 2287–2300.
- Chaddad, A., Daniel, P., Desrosiers, C., Toews, M., Abdulkarim, B., 2019. Novel Radiomic features based on joint intensity matrices for predicting glioblastoma patient survival time. *IEEE J. Biomed. Health Inform.* 23 (2), 795–804.
- Cui, Y., Ren, S., Tha, K.K., Wu, J., Shirato, H., Li, R., 2017. Volume of high-risk intratumoral subregions at multi-parametric MR imaging predicts overall survival and complements molecular analysis of glioblastoma. *Eur. Radiol.* 27 (9), 3583–3592.
- Dunleavy, K., 2017. Advances in the management of primary central nervous system lymphoma. *Clin. Adv. Hematol. Oncol.* 15 (9), 677–679.
- El-Galaly, T.C., Villa, D., Gormsen, L.C., Baech, J., Lo, A., Cheah, C.Y., 2018. FDG-PET/CT in the management of lymphomas: current status and future directions. *J. Intern. Med.* 284 (4), 358–376.
- Kang, D., Park, J.E., Kim, Y.H., et al., 2018. Diffusion radiomics as a diagnostic model for atypical manifestation of primary central nervous system lymphoma: development and multicenter external validation. *Neuro-Oncology* 20 (9), 1251–1261.
- Kasenda, B., Haug, V., Schorb, E., et al., 2013. ^{18}F -FDG PET is an independent outcome predictor in primary central nervous system lymphoma. *J. Nucl. Med.* 54 (2), 184–191.
- Kawai, N., Nishiyama, Y., Miyake, K., Tamiya, T., Nagao, S., 2005. Evaluation of tumor FDG transport and metabolism in primary central nervous system lymphoma using [^{18}F]fluorodeoxyglucose (FDG) positron emission tomography (PET) kinetic analysis. *Ann. Nucl. Med.* 19 (8), 685–690.
- Kim, M.M., Parolia, A., Dunphy, M.P., Venneti, S., 2016. Non-invasive metabolic imaging of brain tumours in the era of precision medicine. *Nat. Rev. Clin. Oncol.* 13 (12), 725–739.
- Kim, Y., Cho, H.-h., Kim, S.T., Park, H., Nam, D., Kong, D.-S., 2018. Radiomics features to distinguish glioblastoma from primary central nervous system lymphoma on multi-parametric MRI. *Neuroradiology* 60 (12), 1297–1305.
- Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. *J. Chiropractic Med.* 15 (2), 155–163.
- Kunimatsu, A., Kunimatsu, N., Kamiya, K., Watadani, T., Mori, H., Abe, O., 2018. Comparison between glioblastoma and primary central nervous system lymphoma using MR image-based texture analysis. *Magn. Reson. Med.* 17 (1), 50–57.
- Lambin, P., Leijenaar, R.T.H., Deist, T.M., et al., 2017. Radiomics: the bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* 14 (12), 749–762.
- Lapointe, S., Perry, A., Butowski, N.A., 2018. Primary brain tumours in adults. *Lancet* 392 (10145), 432–446.
- Lee, J.W., Lee, S.M., 2018. Radiomics in oncological PET/CT: clinical applications. *Nucl. Med. Mol. Imaging* 52 (3), 170–189.
- Li, Y., Qian, Z., Xu, K., et al., 2018. MRI features predict p53 status in lower-grade gliomas via a machine-learning approach. *Neuroimage Clin.* 17, 306–311.
- Liu, Z., Wang, Y., Liu, X., et al., 2018a. Radiomics analysis allows for precise prediction of epilepsy in patients with low-grade gliomas. *Neuroimage Clin.* 19, 271–278.
- Liu, X., Li, Y., Qian, Z., et al., 2018b. A radiomic signature as a non-invasive predictor of progression-free survival in patients with lower-grade gliomas. *Neuroimage* 20, 1070–1077.
- Loweckamp, B.C., Chen, D.T., Ibanez, L., Blezek, D., 2013. The design of SimpleITK. *Front. Neuroinform.* 7, 45.
- Lv, W., Yuan, Q., Wang, Q., et al., 2018. Robustness versus disease differentiation when varying parameter settings in radiomics features: application to nasopharyngeal PET/CT. *Eur. Radiol.* 28 (8), 3245–3254.
- NEMA, 2017. The DICOM Standard. www.dicomstandard.org/NEMA.
- Ostrom, Q.T., Gittleman, H., Liao, P., et al., 2017. CBRUS Statistical Report: primary brain and other central nervous system tumors diagnosed in the United States in 2010–2014. *Neuro-Oncology* 19 (suppl.5), v1–v88.
- Paul Kinahan, D.C., Boellaard, Ronald, Casey, Michael, Christian, Paul E., Cole, Patricia E., DeBruin, Simon, Hoffman, John M., Kohlmyer, Steve, Maniawski, Piotr, Nelson, Dennis, Perlman, Eric S., Shao, Ling X., Wolodzko, John G., Yap, Jeffrey, Zimmerman, Brian, 2018. Vendor-neutral pseudo-code for SUV calculation. PDF. http://qibawiki.rsna.org/images/8/86/SUV_vendorneutral_pseudocode_20180626_DAC.pdf, Accessed date: February 2018.
- Purandare, N.C., Puranik, A., Shah, S., et al., 2017. Common malignant brain tumors: can ^{18}F -FDG PET/CT aid in differentiation? *Nucl. Med. Commun.* 38 (12), 1109–1116.
- Robin, X., Turck, N., Hainard, A., et al., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* 12 (1), 77.
- Suh, H.B., Choi, Y.S., Bae, S., et al., 2018. Primary central nervous system lymphoma and atypical glioblastoma: differentiation using radiomics approach. *Eur. Radiol.* 28 (9), 3832–3839.
- Tsujikawa, T., Rahman, T., Yamamoto, M., et al., 2017. (^{18}F)-FDG PET radiomics approaches: comparing and clustering features in cervical cancer. *Ann. Nucl. Med.* 31 (9), 678–685.
- van Griethuysen, J.J.M., Fedorov, A., Parmar, C., et al., 2017. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 77 (21), e104–e107.
- Wu, J., Gensheimer, M.F., Dong, X., et al., 2016. Robust intratumor partitioning to identify high-risk subregions in lung cancer: A Pilot Study. *Int. J. Radiat. Oncol. Biol. Phys.* 95 (5), 1504–1512.
- Wu, J., Li, B., Sun, X., et al., 2017. Heterogeneous enhancement patterns of tumor-adjacent parenchyma at MR imaging are associated with dysregulated signaling pathways and poor survival in breast cancer. *Radiology* 285 (2), 401–413.
- Wu, J., Cao, G., Sun, X., et al., 2018. Intratumoral spatial heterogeneity at perfusion MR imaging predicts recurrence-free survival in locally advanced breast cancer treated

- with neoadjuvant chemotherapy. *Radiology* 288 (1), 26–35.
- Yamaguchi, S., Hirata, K., Kobayashi, H., et al., 2014. The diagnostic role of (18)F-FDG PET for primary central nervous system lymphoma. *Ann. Nucl. Med.* 28 (7), 603–609.
- Yamasaki, T., Chen, T., Hirai, T., Murakami, R., 2013. Classification of cerebral lymphomas and glioblastomas featuring luminance distribution analysis. *Comput. Math. Methods Med.* 2013, 619658.
- Yip, S.S., Kim, J., Coroller, T.P., et al., 2017. Associations between somatic mutations and metabolic imaging phenotypes in non-small cell lung cancer. *J. Nucl. Med.* 58 (4), 569–576.
- Zhou, W., Wen, J., Hua, F., et al., 2018. (18)F-FDG PET/CT in immunocompetent patients with primary central nervous system lymphoma: differentiation from glioblastoma and correlation with DWI. *Eur. J. Radiol.* 104, 26–32.
- Zwanenburg, A., Leger, S., Vallières, M., Löck, S., 2016. Image Biomarker Standardisation Initiative.