Radiomic Analysis for Predicting Prognosis of Colorectal Cancer From Preoperative ^{18}F-FDG PET/CT

Lilang Lv
Fudan University Shanghai Cancer Center
Bowen Xin
The university of sydney
Yichao Hao
The University of Sydney
Ziyi Yang
Fudan University Shanghai Cancer Center
Junyan Xu
fudan unversity shanghai cancer center
Lisheng Wang
Shanghai Jiao Tong university
Xiuying Wang
The university of Sydney
Shaoli Song (shaoli-song@163.com)
Fudan University Shanghai Cancer Center
Xiaomao Guo
fudan university shanghai cancer center

Research Article

Keywords: stage III Colorectal cancer, 18F-FDG PET/CT, Radiomics, Prognosis, Prediction

Posted Date: November 12th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1036014/v1

Version of Record: A version of this preprint was published at Journal of Translational Medicine on February 2nd, 2022. See the published version at https://doi.org/10.1186/s12967-022-03262-5.

Abstract

Background To develop and validate a survival model with clinico-biological features and ^{18}F- FDG PET/CT radiomic features via machine learning, and for predicting the prognosis from the primary tumor of colorectal cancer.

Methods A total of 196 pathologically confirmed colorectal cancer patients (stage I to stage IV) were included. Preoperative clinical factors, serum tumor markers, and PET/CT radiomic features were included for the recurrence-free survival analysis. For the modeling and validation, patients were randomly divided into the training (n=137) and validation (n=59) set, while the 78 stage III patients [training (n=55), and validation (n=23)] was divided for the further experiment. After selecting features by the log-rank test and variable-hunting methods, random survival forest (RSF) models were built on the training set to analyze the prognostic value of selected features. The performance of models was measured by C-index and was tested on the validation set with bootstrapping. Feature importance and the Pearson correlation were also analyzed.

Results Radiomics signature with four PET/CT features and four clinical factors achieved the best result for prognostic prediction of 196 patients (C-index 0.780, 95% CI 0.634 - 0.877). Moreover, four features (including two clinical features and two radiomics features) were selected in the 78 stage III patients (C-index was 0.820, 95% CI 0.676-0.900). K-M curves of both models significantly stratified low-risk and high-risk groups ($P < 0.0001$). Pearson correlation analysis demonstrated that selected radiomics features were correlated with tumor metabolic factors, such as SUVmean, SUVmax.

Conclusion This study presents integrated clinico-biological-radiological models that can accurately predict the prognosis from the preoperative ^{18}F-FDG PET/CT radiomics in colorectal cancer. It is of potential value in assisting the management and decision making for precision treatment in colorectal cancer.

Trial registration The retrospectively registered study was approved by the Ethics Committee of Fudan University Shanghai Cancer Center (No. 1909207-14-1910) and the data were analyzed anonymously.

Introduction

Colorectal cancer (CRC) is one of the most commonly diagnosed cancers all over the world, though its epidemiology is different in various regions [1]. It is also one of the leading causes of cancer-related mortality despite the advancement in treatment strategies [2, 3]. The prognosis of CRC is one of the essential factors in patient management and selection of treatment strategies [4]. Tumor Node Metastasis (TNM) staging classification system plays an important role in colorectal cancer prognostication [4–6]. But TNM staging system cannot accurately differentiate the prognosis of patients with stage II and III colon cancer. In addition, disease characteristics known to affect the survival of colon cancer, including age, gender, location of primary disease, tumor grade, the number of positive lymph nodes (LNs), the number of LNs examined, lymphatic vessel and peripheral nerve infiltration, intestinal

obstruction or perforation, were not directly included in the TNM staging system. Blood and stool protein markers have also been investigated to identify patients with favorable and poor prognosis [7, 8]. Several studies were dedicated to other prognostic factors in patients with MSI status and chromosome 18q loss of heterozygosity in the coding place [9, 10]. Several studies have attempted to provide clinical assistance in the management strategies of colorectal cancer by utilizing important imaging prognostic features, such as depth of tumor spread, presence of malignant lymph nodes, tumor deposits, extramural vascular invasion, and differentiation of mucinous from nonmucinous tumors [11].

For staging primary colon cancer, contrast-enhanced computed tomography (CT) scans achieved accuracies ranging from 60–80% [12–15]. MRI features are useful in diagnosing locally advanced rectal tumors and also are helpful to assess regional nodal involvement and treatment response [16, 17]. However, the above imaging is based on morphology, cannot provide the metabolic characteristics of the tumor lesion. ^{18}F-fluoro-2-deoxy-D-glucose Positron emission tomography/computed tomography (^{18}F-FDG PET/CT) can sensitively provide the molecular and functional information of not only the primary tumor, but also distant metastasis lesion and the recurrent disease by one-time imaging [18].

Radiomics is a promising translational research field that can provide quantified tumor heterogeneity information from medical images in a non-invasive manner. Studies have shown that radiomic features based on CT or MRI images are related to the prognosis of colorectal cancers [19–21]. Some studies reported the metabolic phenotype could predict genetic alterations of colorectal cancer by ^{18}F-FDG PET/CT radiomics [22]. However, the prognostic value of PET/CT radiomics has not been reported yet, especially in patients with stage III, which accounts for the largest proportion of CRC. In this study, we investigated the prognostic value of ^{18}F-FDG PET/CT-based radiomics features using machine learning for CRC patients of all stages and then applied the same method on CRC patients with stage III to analyze the differences.

Materials And Methods

Patients collection

The study was approved by Fudan University Shanghai Cancer Centre Ethics Committee and Institutional Review Board for clinical investigation. In our study, 196 patients diagnosed with colorectal cancer between January 2010 and July 2018 were retrospectively identified from an electronic database in Fudan University Shanghai Cancer Centre. Patients were followed up until July 2020. All patients who met the following criteria were enrolled: 1) Patients received surgery at the primary colorectal lesion and the final pathology was colorectal adenocarcinoma or mucinous adenocarcinoma; 2) Immunohistochemical results also was received; 3) Patients received no preoperative treatment and underwent preoperative ^{18}F-FDG PET/CT. Thus, the difference in tumor metabolism after adjuvant therapy was avoided; 4) Patients did not receive any chemotherapy, radiation therapy, or molecular

targeted therapy before ^{18}F-FDG PET/CT scans yet; 5) Patients were not lost to follow-up. We reviewed 243 patients diagnosed with colorectal cancer in total and finally enrolled 196 patients for this study.

^{18}F-FDG PET/CT protocol and imaging interpretation

^{18}F-FDG PET/CT scans were performed using a PET/CT scanner (Siemens Medical Systems, Biograph 16 HR). All patients fasted for at least 6 h before ^{18}F-FDG administration and glucose levels in the peripheral blood were confirmed to be 10 mmol/L or less before the ^{18}F-FDG injection (7.4 MBq/kg (0.2 mCi/kg) of body weight) in this study. The scanning included the area from the upper thigh to the skull. Data acquisition started approximately 1h after the injection and the low-dose CT scans were obtained with the following parameters: 40-60 mA, 120 kV, 0.6-s tube rotation, and 3.75-mm section thickness. The spatial resolution of PET images was 168 x 168 x 172 with voxel size 4.06 x 4.06 x 5 mm^3 , while the resolution of CT images was 512 x 512 x 172 with voxel size 1.37 x 1.37 x 5 mm^3. For quantitative analysis, ^{18}F-FDG accumulation on a workstation was assessed by two experienced nuclear medicine physicians through calculating the standardized uptake value (SUV), metabolic tumor volume (MTV) and total lesion glycolysis (TLG) in the regions of interest placed over the suspected lesions and the normal liver. SUV was calculated in a pixel as (radioactivity) / (injected dose/body weight). TLG was calculated as (mean SUV)×(MTV), in which MTV was measured with setting a margin threshold as SUV of 2.5. All values of SUVmax, MTV, and mean SUV were automatically measured by analysis software for each lesion. For evaluating metastatic CRC, the highest SUV in a metastatic tumor was taken as SUVmax and the mean SUV was taken as SUVmean.

Medical Image Delineation

The Volume of Interest (VOIs) in the tumor was segmented slice by slice by two attending nuclear medicine physicians respectively. The open-source software ITK-Snap [23] was used for segmentation. If the two opinions were different, they discussed and made the final decision together. The physicians segmented tumors only on the basis of imaging findings and did not consider pathological findings. Since the PET/CT images were co-registered, only the VOIs of PET images were manually segmented, and then resampled to CT images through coordinate transformation and interpolation. The resulting VOIs for CT images were validated by a radiologist.

Radiomics Feature Extraction

Radiomics Workflow is illustrated in Fig. 1 including three main modules: Feature Engineering, Random Survival Forest (RSF) Models, Statistical analysis. In feature extraction, we applied different settings for PET and CT images to adapt to different image characteristics of these two modalities, as illustrated in Supplementary Fig. S1. For PET images, we firstly applied SUV normalization based on patients' body weight and injection doze. Then, we used a fixed bin-size of 0.25 SUV in intensity discretization to reduce the effect of the image noise [24]. The common parameter for bin size [25, 26] was used to ensure the reproducibility of our model. On the other hand, for CT images, we firstly shifted +1000 HU on image

values to prevent the pixel value from being negative when squared, as the minimum value of HU was -1000. For CT image discretization, we used a fixed bin size of 25 HU, as suggested in previous reports [27–29].

1246 radiomic features were extracted from ROIs delineated by clinicians on PET and CT images respectively, resulting in 2492 radiomic features per patient. Radiomic features include three major types: first-order features, shape features and texture features. First-order features describe the intensity distribution of voxels. Shape features describe the tumor shape characteristics such as volumes and surface areas. Texture features describe the second-order intensity distribution of voxels via Gray Level Co-occurrence Matrix (GLCM), Gray Level Size Zone Matrix (GLSZM), Gray Level Run Length Matrix (GLRLM), and Gray Level Dependence Matrix (GLDM). Wavelet features and Laplacian of Gaussian (LoG) features are texture features extracted from filtered images using wavelet filters and LoG filters. The radiomic feature extraction was implemented with open-source PyRadiomics library [30] (https://github.com/Radiomics/pyradiomics), which is in compliant with Imaging Biomarker Standardization Initiative [31].

Feature Selection

Before implementing feature selection, 24 clinico-biological features and 2492 radiomic features were fused to form a feature pool. The feature selection strategy was designed to be outcome-driven, aiming to mine features that capture the prognostic patterns. As illustrated in Fig. 2A, we applied a sequential combination of univariate and multivariate selection on the PET, CT radiomic features and clinico-biological features extracted from training data. In univariate selection, log-rank test was used to select statistically significant features with high prognostic value ($p < 0.05$). Based on the selected features, multivariate selection was deployed to select the final discriminative feature set using RSF-based variable hunting algorithm [32]. To prevent the risk of overfitting, we applied 50 times of five-fold cross-validation in multivariate feature selection to boost the generalizability of selected feature subsets. As the selected features were based on the performance of rotating training sets instead of a single fixed training set, the selected feature subset was more generalizable, thus properly avoiding the risk of overfitting.

Modeling and Validation

The patients were split into training and validation sets (7:3 ratio) using the stratified method. A random survival forest (RSF) model, which captures non-linear effects, was fitted to predict the recurrence-free survival (RFS) on the training set. To select the best performing RSF model with optimized hyperparameter, we used grid search strategy based on the average C-index on the training set with 1000 times of bootstrap. The model performance was evaluated by C-index on the validation set with the 1000 times bootstrap method to reduce model overfitting. Furthermore, the predicted risks of the validation set yielded by the fitted RSF model were dichotomized into low-risk and high-risk groups. Then two groups were compared using the log-rank test to examine whether the model could stratify patients with different RFS.

Statistical Analysis

Statistical analysis was implemented using R package version 3.6.3 (R Foundation for Statistical Computing) and p-value<0.05 was considered statistically significant. The optimal cutoff point for the log-rank test was performed by ‘surv_cutpoint’ function in the ‘survminer’ R package. The random survival forest and variable-hunting algorithm were implemented using the “randomForestSRC” R package.

Results

Demographics of patients

There are 196 patients with colorectal cancer involved in the dataset. Table 1 summarizes the detailed demographics of patients. In the original experiment, 196 patients (ranging from stage I to stage IV) were randomly split into 138 training samples and 58 validation samples with a ratio of 7:3. The dataset used in the primary experiment is denoted as D-1~4. There were 29.6% of patients from D-1~4 who experienced recurrence. In the further experiment, we conducted the prognostic analysis on patients with stage III only, which was split into training and testing sets with a ratio of 7:3. The dataset containing only stage III patients is denoted as D-3. There were 33.3% of patients from D-3 who experienced recurrence.

Table 1
Patients characteristics of the training and validation sets.

Characteristics	Training set(n=138)	Validation set(n=58)	P value*
CEA	18.022±41.604	25.928±51.872	0.090
CA199			0.535
High	36	12	
Normal	102	46	
Lymph nodes	2.326±3.560	2.172±3.320	0.804
Stage			0.486
I	13	3	
II	42	23	
III	55	23	
IV	28	9	
Location			0.331
Ascending	13	4	
Descending	19	8	
Ileocecum	11	1	
Rectum	54	21	
Sigmoid	34	22	
Transverse	7	2	
Follow up time	19	20	0.751

NOTE: Continuous data except follow-up time (which was shown with median) were demonstrated with means ± standard deviations while categorical data were demonstrated with the number of each category and percentage. *p-value was calculated by using χ^2 test for categorical variable and Wilcoxon test for continuous variable.

Result of feature selection

As illustrated in Fig. 2B, feature selection was applied on 2492 radiomics features extracted from PET and CT images and 24 clinico-biological features. For patients with stage I-IV, 12 CT and 48 PET radiomics features and nine clinico-biological features, were selected during univariate selection, while the final feature set composed of two CT, two PET and two clinical features was selected in multivariate selection for model building. For patients with stage III, 27 CT and 22 PET radiomics features and five

clinico-biological features were selected during univariate selection, while one CT, one PET and two clinical features were selected in multivariate selection.

Performance of radiomics signature

The performance of selected radiomic signatures is described in Fig. 3A and 3B for primary experiment D-1~4 and secondary experiment D3, respectively. Fig. 3A and 3B show RSF models built with clinical, CT and PET features outperforms models with solely clinical, PET or CT features, peaking at C-index 0.780 [95% CI 0.634-0.877] and 0.820 [95% CI 0.676-0.900] respectively. The detailed performance of signatures was attached in the supplementary Table S2. K-M curves of radiomics signatures for D-1~4 and D3 are demonstrated in Fig. 3C and 3D respectively ($P < 0.0001$). To evaluate the risk of overfitting, we summarized training and testing C-index during the independent validation in Supplementary Table S1. The table shows the differences between training and testing C-index were less than 0.03 in both experiments, which suggests the risk of overfitting was properly alleviated.

Feature analysis and interpretation

There were eight features identified by the feature selection process from D-1~4 including four clinical features (CA199, lymph nodes, stage, and CEA), two PET features (PET-wavelet-LLH-gldm-DV and PET-wavelet-LLL-glcm-imc2) and two CT features (CT-Log-sigma-5.0-3D-glszm-SAE, CT-Log-sigma-4.0-3D-glszm-SALGLE). The RSF model built on these eight features is denoted as M1. There were four features identified for the secondary experiment D-3 including two clinical features (CA199 and lymph nodes) and one PET feature (PET-Wavelet-LLH-glszm-ZV) and one CT feature (CT-Log-sigma-5.0-3D-glszm-SAE). The RSF model built on these four features for the secondary experiment is denoted as M2. Detailed feature explanation was attached in the supplementary Table S3.

We further reveal the contribution of each feature for model M1 and M2 in Fig. 4A and 4B. Bar graphs in Fig. 4A and 4B show the normalized importance of each feature, in which CA199 contributed most in M1 and PET-Wavelet-LLH-glszm-ZV contributed most in M2. Pie graphs in Fig. 4A and 4B illustrate the percentage of contribution of PET, CT and clinical features. PET and CT features contributed 13.3% in M1 while contributed 83.5% in M2. In addition to feature contribution, we compared the features in M1 and M2 then found two common clinical features (CA199 and lymph nodes) and one common CT feature (CT-Log-sigma-5.0-3D-glszm-SAE).

Figure 4C summarizes Pearson correlation between selected radiomic features and clinical features. It shows that radiomic features were significantly correlated to metabolic tumor activity features such as SUVmean, SUVmax, TLG, and 40%MTV.

Case study

We chose two stage III samples from the validation set of the Data-3 to showcase the predictive performance of M2 model built on the radiomics signature. Fig. 5A shows the individualized predictive results for these two samples summarized with survival curves. It shows that the curve of patient 1 was lower than patient 2, which means the recurrence time of patient 1 could be shorter. This was in

accordance with their real recurrence time, eight months for patient 1 and 13 months for patient 2. Values of radiomic features for these two patients were shown in Fig. 5B. The Zone variance (ZV) of PET-Wavelet-LLH image measures the variance in gray level zone size. The larger ZV, the greater heterogeneity. The SAE (Small Area Emphasis) of CT-Log-5.0-3D image measures the distribution of small size gray level zones and the large value indicated fine textures. Detailed clinical information of two patients was included in the supplementary Table S4. The visualization of radiomic features was demonstrated in Fig. 5C. In terms of univariable analysis, patient 1 with CA199 normal tends to have a better prognosis but was in fact associated with shorter recurrence time. In contrast, our model M2 still made the correct prediction dominantly contributed by PET/CT radiomics features (83.5%) while the importance of CA199 is only 14% in M2.

Discussion

Previous studies have shown that ^{18}F-FDG PET/CT radiomics performed well in predicting the prognosis of various malignancies. The newly developed PET/CT radiomic signature was a powerful predictor of gastric cancer survival [33]. Radiomics features of baseline PET/CT images provide complementary prognostic information for nasopharyngeal carcinoma compared with the use of clinical parameters alone [34]. This method was also advantageous to predict the prognosis of lung cancer, breast cancer and other tumors [35–37]. As for colorectal cancer, a few studies demonstrated that FDG PET radiomic held potential towards the improved prediction of clinical outcome in stage IV patients of colorectal cancer and locally-advanced rectal cancer [38, 39]. The explosive researches on the prognostic value of PET/CT-based radiomics methods for the total colorectal cancer were rare, especially for stage III. A study on the National Cancer Data Base (NCDB) showed that CRC patients with stage III accounted for approximately one-third of all stages [40]. Moreover, the 5-year survival rate of this largest proportion of patients was subjected to a large difference in the survival outcomes [41, 42]. Therefore, evaluating the prognosis of stage III colorectal cancer separately and intervening as early as possible according to different individual patients to reduce the risk of recurrence and metastasis is necessary. In this study, we developed an original model to predict the prognosis of CRC patients and further experimented on stage III patients by ^{18}F-FDG PET/CT radiomics.

We originally investigated prognosis models for patients ranging from stage I to IV in the primary experiment. The model M1 trained by a combination of features of all three modalities outperformed other models with a C-index of 0.780 [95% CI 0.634-0.877]. The K-M curves indicated that high-risk and low-risk patient groups could be separated by our model effectively ($P < 0.0001$). For model M1, CA199 was the most important feature. It means that this cancer antigen marker CA199 contributed most to the outcome of the prognostic prediction in model M1. The result is consistent with previous studies [43] that CA199 is a key prognostic biomarker. It should also be noted that the contribution of imaging features was irreplaceable, although they only accounted for 13.3% of the contribution. Both PET and CT features were important and irreplicable in radiomics analysis because they both had positive importance scores, which suggests these features positively contributed to the model accuracy. Experimental results in

Manuscript Fig. 3A verified that the model constructed with multimodalities (C-index 0.780) outperformed the models built with PET (C-index 0.592) or CT (C-index 0.755) alone on D-1~4. Similar trend can be identified on D-3 with Fig. 3B.

We also focused on analyzing models for patients with stage III, because the 5-year survival rate is unsatisfactory, though radical surgery and adjuvant chemotherapy were routinely performed. The prediction of prognosis is valuable for supporting individualized treatment. The C-index of M2 was 0.820 [95% CI 0.676-0.900], which means it holds a great potential value of prognostic prediction in colorectal cancer. Its performance was also superior to single modality or double modality models. K-M curves of M2 illustrate the model could significantly separate high-risk and low-risk patient group. For model M2, PET-Wavelet-LLH-glszm-ZV was the most important feature in the predictive model, which means the texture information quantified by this PET feature successfully captured the heterogeneity of colorectal tumour towards prognostic prediction. It was because PET images could provide information not only about the metabolism of the tumor, but also about the total load of the tumor. For further interpretation of this PET feature, we conducted correlation analysis, and found that this PET features positively correlated with 40% MTV and TLG ($p<0.05$). CA199 which contributes most in M1 only made up 14% of all feature contributions.

Moreover, CA199, lymph nodes and CT-Log-sigma-5.0-3D-glszm-SAE were three features identified both in M1 and M2. The feature importance analysis showed that clinical features played the most vital roles in the prognosis of CRC patients of all stages while radiomics features made more contribution when predicting the prognosis of CRC patients with stage III. The case study also demonstrated that features with greater contribution could help the model to overcome the negative impact caused by single features then rectify the prediction. Thereby, it is reasonable to believe that the combination of clinical characteristics and imaging characteristics of ^{18}F-FDG metabolism is more convincing than any single modality model.

We reduced the risk of overfitting through reducing the number of features and employed cross-validation in feature selection. Firstly, we reduced the risk of overfitting by strictly controlling the number of features, as the reduced number of features leaded to the decrease of the number of required parameters inside machine learning models, thus minimizing the risk of overfitting [43]. According to the guideline for radiomics studies [44], we reduced the number of features to less than 1/10 of sample sizes. Secondly, 50 times five-fold cross-validation was deployed during the feature selection on the training dataset to reduce the risk of overfitting [43]. By selecting features on the rotating training instead of a fixed training set, we effectively minimize the risk of overfitting on a fixed proportion of data. Thirdly, we evaluated the risk of overfitting by comparing the performance of the model on training and testing datasets in independent validation. Supplementary Table S1 shows the difference between training and testing C-index was less than 0.03 in both experiments, which suggests the risk of overfitting was properly handled.

This study was partly limited by its retrospective design and relatively modest sample sizes. We will continue to collect more patients who meet the criteria and attempt to conduct prospective studies to

further validate our models. We look forward to further randomized controlled trials in the future on the significance and importance of ^{18}F-FDG PET/CT imaging omics in the diagnosis and treatment of colorectal cancer. We will investigate the effect of spatial resolution of PET/CT images on the parameters of radiomic feature extraction.

Conclusion

We designed and investigate the prognostic value of ^{18}F-FDG PET/CT radiomics for colorectal cancer in this study. ^{18}F-FDG PET/CT radiomics combined with clinical features could be instructive in the predictive prognosis of colorectal cancer, especially in stage III. The proper application of ^{18}F-FDG PET/CT radiomics could optimize the individual treatment strategy by avoiding ineffective or excessive management.

List Of Abbreviations

CRC Colorectal Cancer

^{18}F-FDG ^{18}F-fluoro-2-deoxy-D-glucose

PET/CT Positron emission tomography/computed tomography

SUV Standardized Uptake Value

MTV Metabolic Tumor Volume

TLG Total Lesion Glycolysis

VOI the Volume of Interest

RSF Random Survival Forest

Declarations

Ethics approval and consent to participate

The study was approved by the Ethics Committee of Fudan University Shanghai Cancer Center (No. 1909207-14-1910) and the data were analyzed anonymously. The requirement of written informed consent was waived.

Consent for publication

Written informed consent for the analysis of anonymized clinical and imaging data was obtained from all patients.

Availability of data and materials

The datasets generated and/or analysed during the current study are not publicly available due patient privacy and copyright issues but are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was funded by National Natural Science Foundation of China (Grant number 81771861 and 81971648), Shanghai Scientific and Technological Innovation Program (Grant number 18410711200 and 19142202100).

Authors' contributions

Conceptualization, L.L., S.S. and X.G.; methodology, L.L., B.X. and Y.H.; software, Z.Y. and J.X.; validation, L.L., B.X. and Y.H.; formal analysis, B.X. and Y.H.; investigation, Z.Y. and J.X.; resources, L.W. and X.W.; data curation, L.L., B.X. and Y.H.; writing—original draft preparation, L.L., B.X. and Y.H.; writing—review and editing, all authors; visualization, all authors; supervision, S.S. and X.G.; project administration, S.S. and X.G.; funding acquisition, S.S.

Acknowledgments

This work was funded by National Natural Science Foundation of China (Grant number 81771861 and 81971648), Shanghai Scientific and Technological Innovation Program (Grant number 18410711200 and 19142202100).

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin.* 2019; 69:7-34.
2. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ, He J. Cancer statistics in China, 2015. *CA Cancer J Clin.* 2016; 66:115-132.
3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin.* 2016; 66:7-30.
4. Gunderson LL, Jessup JM, Sargent DJ, Greene FL, Stewart A. Revised tumor and node categorization for rectal cancer based on surveillance, epidemiology, and end results and rectal pooled analysis outcomes. *J Clin Oncol.* 2010; 28:256-263.
5. Gunderson LL, Sargent DJ, Tepper JE, Wolmark N, O'Connell MJ, Begovic M, Allmer C, Colangelo L, Smalley SR, Haller DG, et al. Impact of T and N stage and treatment on survival and relapse in adjuvant rectal cancer: a pooled analysis. *J Clin Oncol.* 2004; 22:1785-1796.

6. Gunderson LL, Jessup JM, Sargent DJ, Greene FL, Stewart AK. Revised TN categorization for colon cancer based on national survival outcomes data. *J Clin Oncol.* 2010; 28:264-271.
7. Mala T, Bøhler G, Mathisen Ø, Bergan A, Søreide O. Hepatic resection for colorectal metastases: can preoperative scoring predict patient outcome? *World J Surg.* 2002; 26:1348-1353.
8. Lech G, Słotwiński R, Słodkowski M, Krasnodębski IW. Colorectal cancer tumour markers and biomarkers: Recent therapeutic advances. *World J Gastroenterol.* 2016; 22:1745-1755.
9. Sinicrope FA, Sargent DJ. Clinical implications of microsatellite instability in sporadic colon cancers. *Curr Opin Oncol.* 2009; 21:369-373.
10. Sarli L, Bottarelli L, Bader G, Iusco D, Pizzi S, Costi R, D'Adda T, Bertolani M, Roncoroni L, Bordi C. Association between recurrence of sporadic colorectal cancer, high level of microsatellite instability, and loss of heterozygosity at chromosome 18q. *Dis Colon Rectum.* 2004; 47:1467-1482.
11. García-Figueiras R, Baleato-González S, Padhani AR, Luna-Alcalá A, Marhuenda A, Vilanova JC, Osorio-Vázquez I, Martínez-de-Alegría A, Gómez-Caamaño A. Advanced Imaging Techniques in Evaluation of Colorectal Cancer. *Radiographics.* 2018; 38:740-765.
12. Smith NJ, Bees N, Barbachano Y, Norman AR, Swift RI, Brown G. Preoperative computed tomography staging of nonmetastatic colon cancer predicts outcome: implications for clinical trials. *Br J Cancer.* 2007; 96:1030-1036.
13. Hundt W, Braunschweig R, Reiser M. Evaluation of spiral CT in staging of colon and rectum carcinoma. *Eur Radiol.* 1999; 9:78-84.
14. Engelmann BE, Loft A, Kjær A, Nielsen HJ, Berthelsen AK, Binderup T, Brinch K, Brünner N, Gerds TA, Høyer-Hansen G, et al. Positron emission tomography/computed tomography for optimized colon cancer staging and follow up. *Scand J Gastroenterol.* 2014; 49:191-201.
15. Nerad E, Lahaye MJ, Maas M, Nelemans P, Bakers FC, Beets GL, Beets-Tan RG. Diagnostic Accuracy of CT for Local Staging of Colon Cancer: A Systematic Review and Meta-Analysis. *AJR Am J Roentgenol.* 2016; 207:984-995.
16. Horvat N, Carlos Tavares Rocha C, Clemente Oliveira B, Petkovska I, Gollub MJ. MRI of Rectal Cancer: Tumor Staging, Imaging Techniques, and Management. *Radiographics.* 2019; 39:367-387.
17. Liu LH, Lv H, Wang ZC, Rao SX, Zeng MS. Performance comparison between MRI and CT for local staging of sigmoid and descending colon cancer. *Eur J Radiol.* 2019; 121:108741.
18. Evans J, Patel U, Brown G. Rectal cancer: primary staging and assessment after chemoradiotherapy. *Semin Radiat Oncol.* 2011; 21:169-177.
19. Badic B, Desseroit MC, Hatt M, Visvikis D. Potential Complementary Value of Noncontrast and Contrast Enhanced CT Radiomics in Colorectal Cancers. *Acad Radiol.* 2019; 26:469-479.
20. Dai W, Mo S, Han L, Xiang W, Li M, Wang R, Tong T, Cai G. Prognostic and predictive value of radiomics signatures in stage I-III colon cancer. *Clin Transl Med.* 2020; 10:288-293.
21. Li Y, Liu W, Pei Q, Zhao L, Güngör C, Zhu H, Song X, Li C, Zhou Z, Xu Y, et al. Predicting pathological complete response by comparing MRI-based radiomics pre- and postneoadjuvant radiotherapy for

locally advanced rectal cancer. *Cancer Med.* 2019; 8:7244-7252.

22. Chen SW, Shen WC, Chen WT, Hsieh TC, Yen KY, Chang JG, Kao CH. Metabolic Imaging Phenotype Using Radiomics of [(18)F]FDG PET/CT Associated with Genetic Alterations of Colorectal Cancer. *Mol Imaging Biol.* 2019; 21:183-190.
23. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, Gerig G. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage (Orlando, Fla).* 2006; 31:1116-1128.
24. Ha S, Choi H, Paeng JC, Cheon GJ. Radiomics in Oncological PET/CT: a Methodological Overview. *Nucl Med Mol Imaging.* 2019; 53:14-29.
25. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges JP, Corcos L, Visvikis D. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med.* 2011; 52:369-378.
26. Pfaehler E, van Sluis J, Merema BBJ, van Ooijen P, Berendsen RCM, van Velden FHP, Boellaard R. Experimental Multicenter and Multivendor Evaluation of the Performance of PET Radiomic Features Using 3-Dimensionally Printed Phantom Inserts. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine.* 2020; 61:469-476.
27. Welch ML, McIntosh C, Haibe-Kains B, Milosevic MF, Wee L, Dekker A, Huang SH, Purdie TG, O'Sullivan B, Aerts H, Jaffray DA. Vulnerabilities of radiomic signature development: The need for safeguards. *Radiother Oncol.* 2019; 130:2-9.
28. Dou TH, Coroller TP, van Griethuysen JJM, Mak RH, Aerts H. Peritumoral radiomics features predict distant metastasis in locally advanced NSCLC. *PLoS One.* 2018; 13:e0206108.
29. Yuan R, Shi S, Chen J, Cheng G. Radiomics in RayPlus: a Web-Based Tool for Texture Analysis in Medical Images. *J Digit Imaging.* 2019; 32:269-275.
30. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillion-Robin JC, Pieper S, Aerts H. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* 2017; 77:e104-e107.
31. Zwanenburg A, Vallieres M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, Ashrafinia S, Bakas S, Beukinga RJ, Boellaard R, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology.* 2020; 295:328-338.
32. Ishwaran H, Kogalur UB, Chen X, Minn AJ. Random survival forests for high-dimensional data. 2011; 4:115-132.
33. Jiang Y, Yuan Q, Lv W, Xi S, Huang W, Sun Z, Chen H, Zhao L, Liu W, Hu Y, et al. Radiomic signature of (18)F fluorodeoxyglucose PET/CT for prediction of gastric cancer survival and chemotherapeutic benefits. *Theranostics.* 2018; 8:5915-5928.
34. Lv W, Yuan Q, Wang Q, Ma J, Feng Q, Chen W, Rahmim A, Lu L. Radiomics Analysis of PET and CT Components of PET/CT Imaging Integrated with Clinical Parameters: Application to Prognosis for Nasopharyngeal Carcinoma. *Mol Imaging Biol.* 2019; 21:954-964.

35. Oikonomou A, Khalvati F, Tyrrell PN, Haider MA, Tarique U, Jimenez-Juan L, Tjong MC, Poon I, Eilaghi A, Ehrlich L, Cheung P. Radiomics analysis at PET/CT contributes to prognosis of recurrence and survival in lung cancer treated with stereotactic body radiotherapy. *Sci Rep.* 2018; 8:4003.
36. Huang SY, Franc BL, Harnish RJ, Liu G, Mitra D, Copeland TP, Arasu VA, Kornak J, Jones EF, Behr SC, et al. Exploration of PET and MRI radiomic features for decoding breast cancer phenotypes and prognosis. *NPJ Breast Cancer.* 2018; 4:24.
37. Wang H, Zhao S, Li L, Tian R. Development and validation of an (18)F-FDG PET radiomic model for prognosis prediction in patients with nasal-type extranodal natural killer/T cell lymphoma. *Eur Radiol.* 2020; 30:5578-5587.
38. Rahmim A, Bak-Fredslund KP, Ashrafinia S, Lu L, Schmidtlein CR, Subramaniam RM, Morsing A, Keiding S, Horsager J, Munk OL. Prognostic modeling for patients with colorectal liver metastases incorporating FDG PET radiomic features. *Eur J Radiol.* 2019; 113:101-109.
39. Lovinfosse P, Polus M, Van Daele D, Martinive P, Daenen F, Hatt M, Visvikis D, Koopmansch B, Lambert F, Coimbra C, et al. FDG PET/CT radiomics for predicting the outcome of locally advanced rectal cancer. *European Journal of Nuclear Medicine and Molecular Imaging.* 2018; 45:365-375.
40. Chagpar R, Xing Y, Chiang YJ, Feig BW, Chang GJ, You YN, Cormier JN. Adherence to stage-specific treatment guidelines for patients with colon cancer. *J Clin Oncol.* 2012; 30:972-979.
41. Hari DM, Leung AM, Lee JH, Sim MS, Vuong B, Chiu CG, Bilchik AJ. AJCC Cancer Staging Manual 7th edition criteria for colon cancer: do the complex modifications improve prognostic assessment? *J Am Coll Surg.* 2013; 217:181-190.
42. Webber C, Gospodarowicz M, Sobin LH, Wittekind C, Greene FL, Mason MD, Compton C, Brierley J, Groome PA. Improving the TNM classification: findings from a 10-year continuous literature review. *Int J Cancer.* 2014; 135:371-378.
43. Mayerhoefer ME, Materka A, Langs G, Haggstrom I, Szczypinski P, Gibbs P, Cook G. Introduction to Radiomics. *J Nucl Med.* 2020; 61:488-495.
44. Sollini M, Antunovic L, Chiti A, Kirienko M. Towards clinical application of image mining: a systematic review on artificial intelligence and radiomics. *Eur J Nucl Med Mol Imaging.* 2019; 46:2656-2672.

Figures

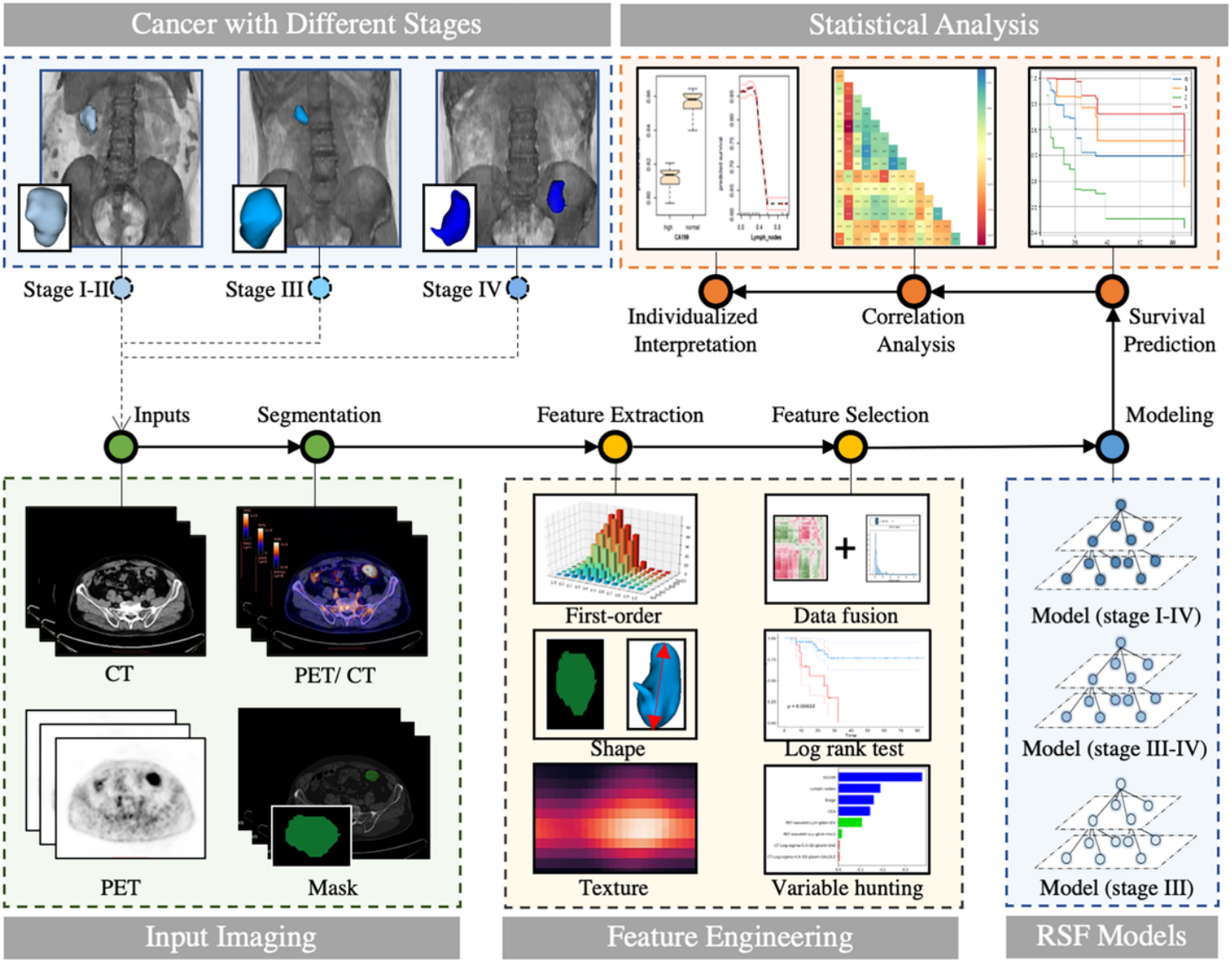

Figure 1

Radiomics workflow.

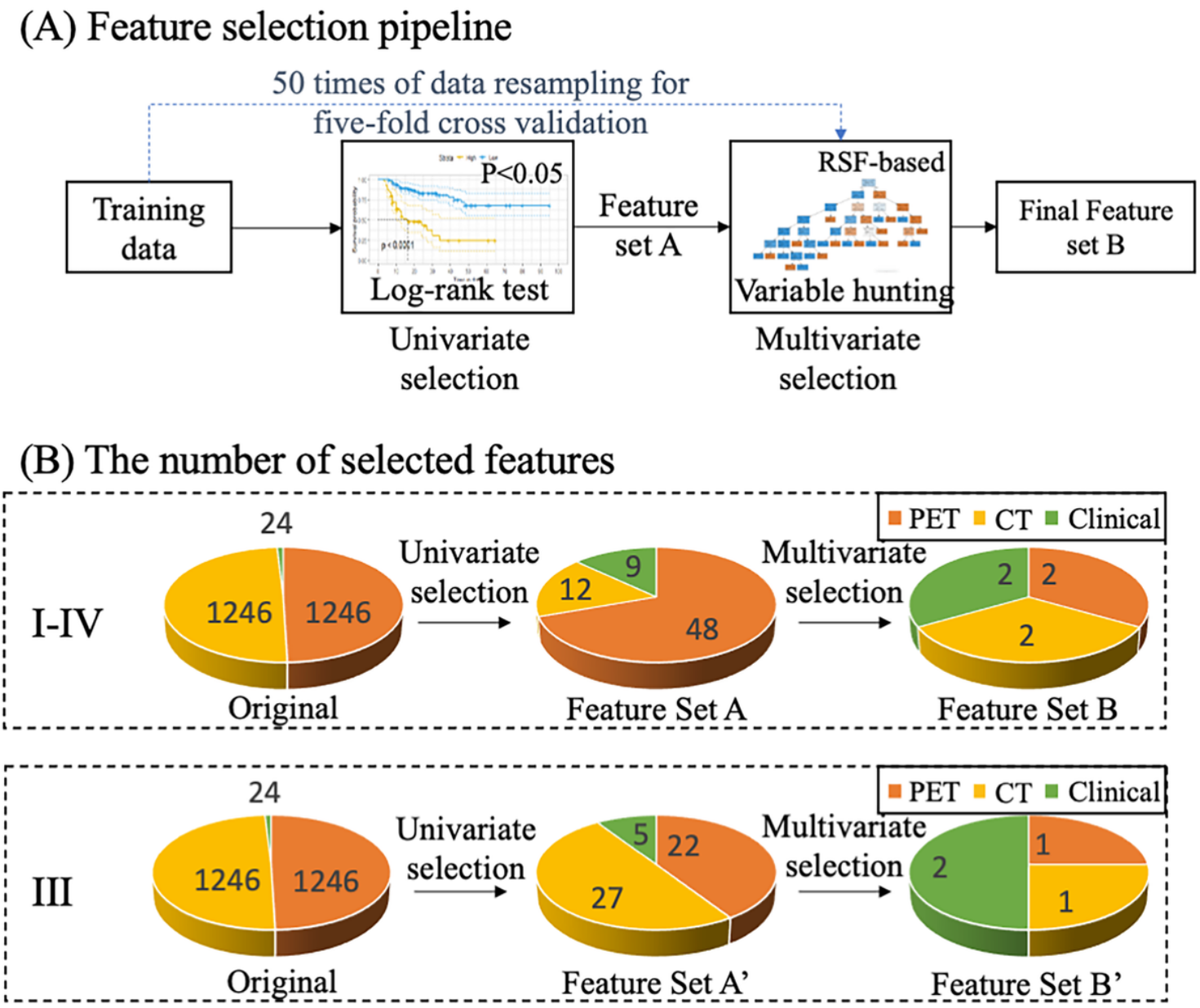

Figure 2

Methodology and result of feature selection. (A) Feature selection pipeline; (B) the number of selected features in the feature selection process.

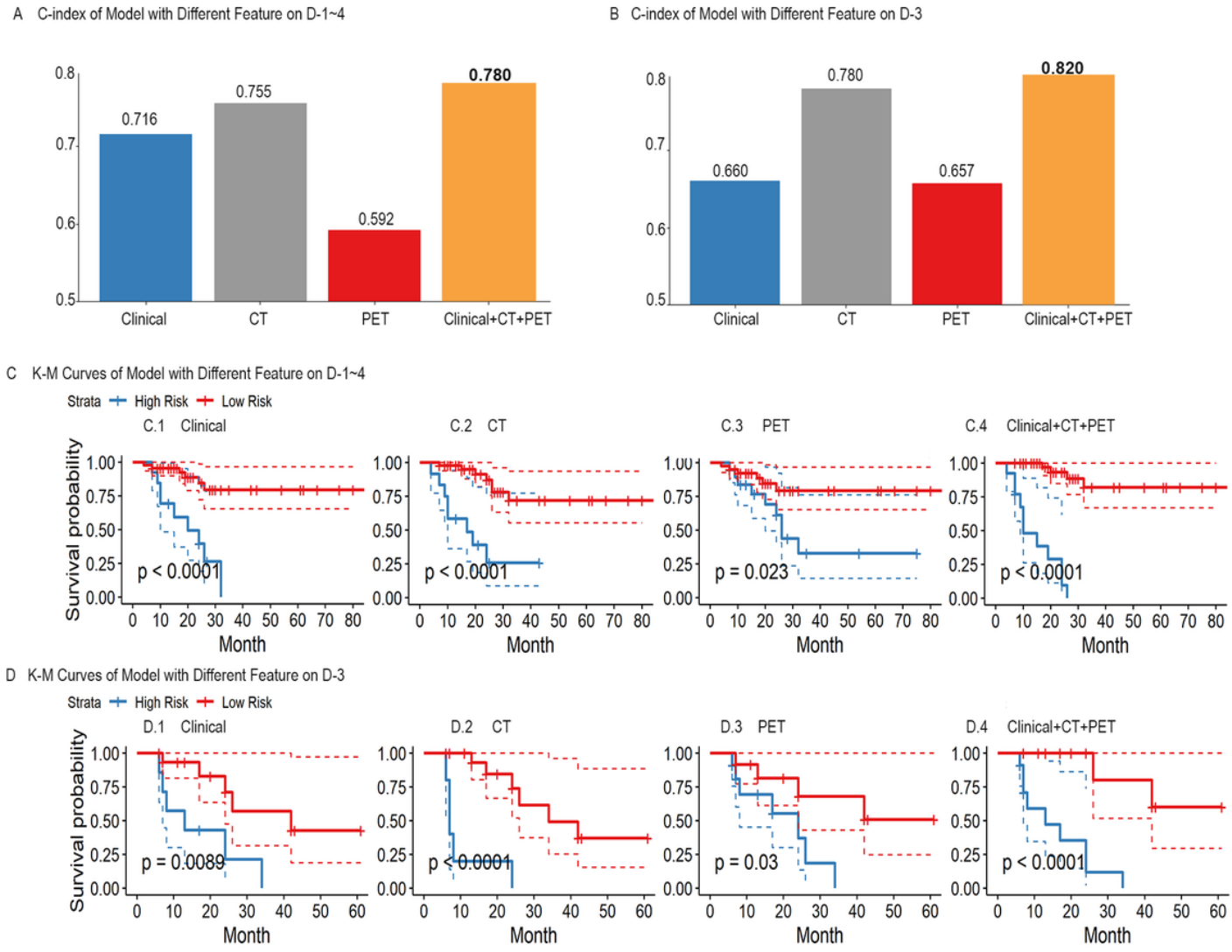

Figure 3

The performance of prognostic models. Figure A and B showed the comparison of the prognostic performance of different feature types on D-1~4 and D-3 respectively. Figure C and D were K-M Curves for different feature types on D-1~4 and D-3 respectively.

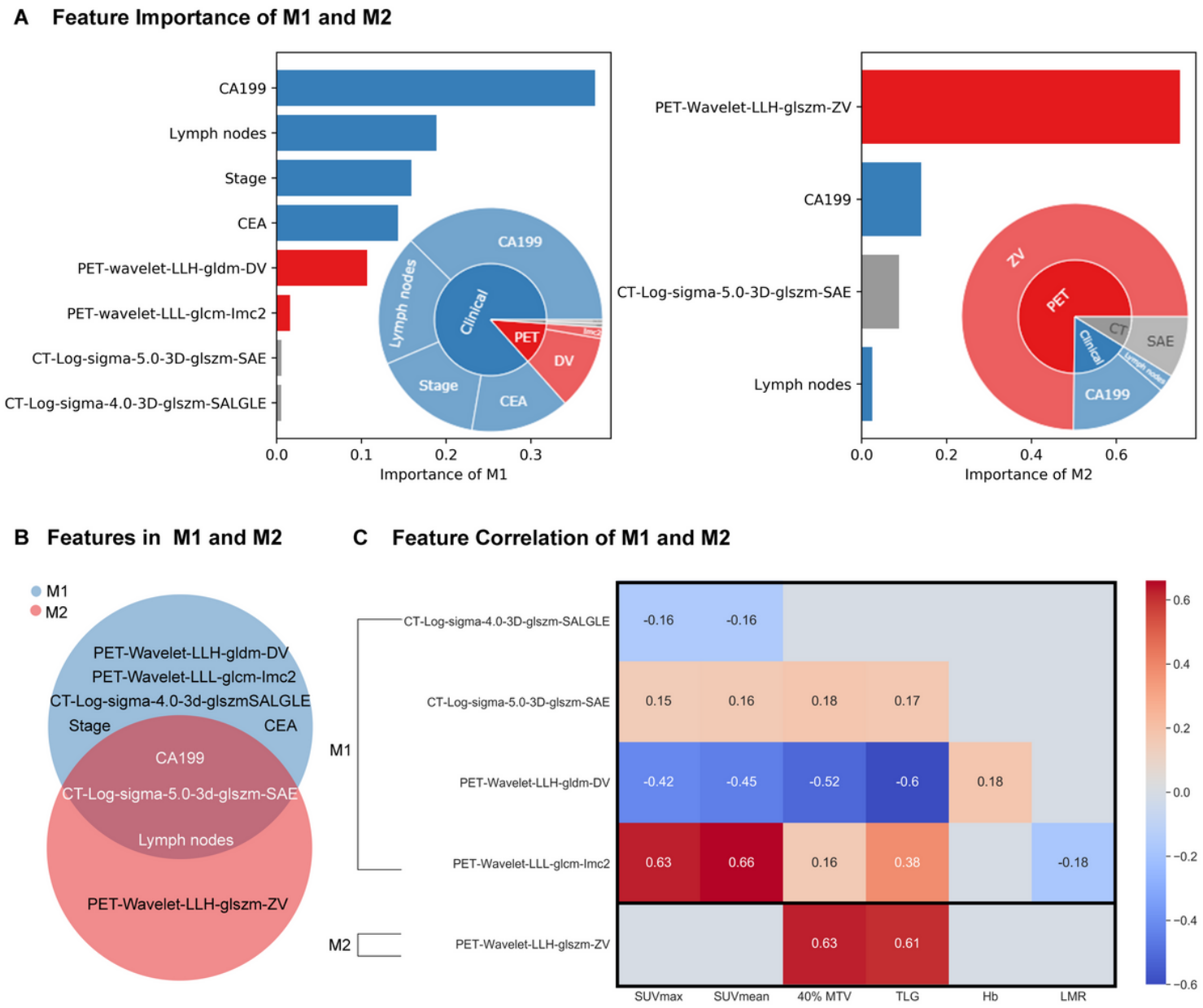

Figure 4

Feature importance and Pearson correlation of M1 and M2. Figure A was the importance of features in models M1 and M2. Clinical, PET, and CT features were represented by using blue, gray, and red bar, respectively. Figure B demonstrated Pearson correlation between selected radiomic features and clinical features in model M1 and M2. Only significant correlation coefficients under Pearson correlation test ($p<0.05$) were shown and others were masked by gray color (Figure C).

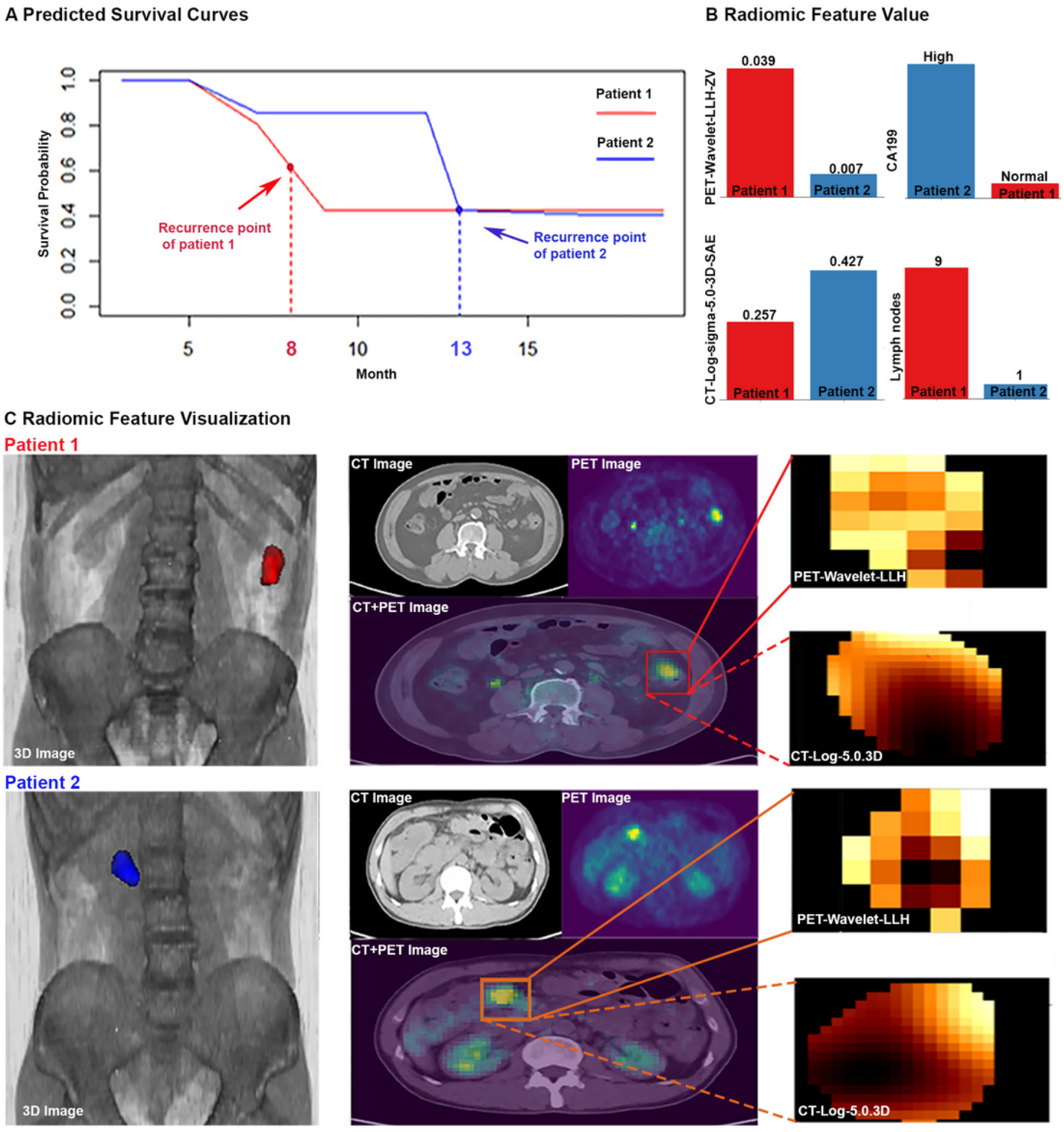

Figure 5

Case study for individualized result interpretation. Figure A showed the predicted survival curves of individual patients yielded by the model M2. Figure B showed the values of radiomic features of patients in the case study. Figure C was the visualization of radiomic features.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementary.docx