Organomics

A Concept Reflecting the Importance of PET/CT Healthy Organ Radiomics in Non-Small Cell Lung Cancer Prognosis Prediction Using Machine Learning

Salimi, Yazdan; Hajianfar, Ghasem; Mansouri, Zahra; Sanaat, Amirhosein; Amini, Mehdi; Shiri, Isaac; Zaidi, Habib

Published in:
Clinical Nuclear Medicine

DOI:
10.1097/RLU.0000000000005400

Publication date:
2024

Document version:
Final published version

OPEN

Organomics

A Concept Reflecting the Importance of PET/CT Healthy Organ Radiomics in Non–Small Cell Lung Cancer Prognosis Prediction Using Machine Learning

Yazdan Salimi, MSc, Ghasem Hajianfar, MSc,* Zahra Mansouri, MSc,* Amirhosein Sanaat, PhD,* Mehdi Amini, MSc,* Isaac Shiri, PhD,*† and Habib Zaidi, PhD*‡§||*

Purpose: Non–small cell lung cancer is the most common subtype of lung cancer. Patient survival prediction using machine learning (ML) and radiomics analysis proved to provide promising outcomes. However, most studies reported in the literature focused on information extracted from malignant lesions. This study aims to explore the relevance and additional value of information extracted from healthy organs in addition to tumoral tissue using ML algorithms.

Patients and Methods: This study included PET/CT images of 154 patients collected from available online databases. The gross tumor volume and 33 volumes of interest defined on healthy organs were segmented using nnU-Net deep learning–based segmentation. Subsequently, 107 radiomic features were extracted from PET and CT images (Organomics). Clinical information was combined with PET and CT radiomics from organs and gross tumor volumes considering 19 different combinations of inputs. Finally, different feature selection (FS; 5 methods) and ML (6 algorithms) algorithms were tested in a 3-fold data split cross-validation scheme. The performance of the models was quantified in terms of the concordance index (C-index) metric.

Results: For an input combination of all radiomics information, most of the selected features belonged to PET Organomics and CT Organomics. The highest C-index (0.68) was achieved using univariate C-index FS method and random survival forest ML model using CT Organomics + PET Organomics as input as well as minimum depth FS method and CoxPH ML model using PET Organomics as input. Considering all 17 combinations with C-index higher than 0.65, Organomics from PET or CT images were used as input in 16 of them.

Conclusions: The selected features and C-indices demonstrated that the additional information extracted from healthy organs of both PET and CT imaging modalities improved the ML performance. Organomics could be a step toward exploiting the whole information available from multimodality medical images, contributing to the emerging field of digital twins in health care.

Key Words: segmentation, radiomics, machine learning, Organomics, survival prediction

(*Clin Nucl Med* 2024;49: 899–908)

Lung cancer is the second most common cancer in all genders, whereas the most common subtype of lung cancer is non–small cell lung cancer (NSCLC), a leading cause of death among other malignancies.[1,2] Knowledge of prognosis prior to treatment and during the treatment can be useful to change or optimize the treatment strategy or prevent other posttreatment. Radiomics analysis aims to convert medical images to high-dimensional data, which could be connected to a desired target, such as biopsy results for clinical diagnosis and patient outcome for prognostic and predictive models.[3–5] Radiomics information coupled with machine learning (ML) algorithms showed potential to predict the prognosis for NSCLC patients after treatment,[6,7] whereas most of the available studies using artificial intelligence (AI)[7–9] focused on radiomic features extracted from the tumoral region and used clinical information, such as age, gender, and blood tests as additional information. Amini et al[9] developed ML models to predict survival using different image fusion strategies and radiomics extracted from the gross tumor volume (GTV) on the same population.[8] Lee et al[10] extracted peritumoral image features and reported gain in classification performance, which depends on tumor size. Hosny et al[11] showed that deep learning classification algorithms emphasized the importance of peritumoral tissue in patient risk estimation. Perez-Morales et al[12] used peritumoral and intratumoral radiomic features to detect a vulnerable subset of lung cancer patients associated with poor survival outcomes who may require aggressive follow-up and/or adjuvant therapy. Mattonen et al[13] reported the importance of metabolic tumor volume penumbra extended by 1 cm in NSCLC recurrence. Guo et al[14] evaluated the predictive value of dosiomics and CT radiomics of esophageal tumor GTV and whole esophagus for predicting complications after radiotherapy. They reported the combination of GTV and whole esophagus as the best predictor using ML models. Lam et al[15] used multiomics data including radiomics and dosiomics extracted from 8 volumes of interest irradiated around the nasopharyngeal GTV to predict the adaptive radiotherapy eligibility in nasopharyngeal cancer patients. They reported the best performance for radiomics plus dosiomics extracted from these 8 regions plus the GTV. They did not compare the GTV only versus added value of the surrounding organs. Girum et al[16]

Received for publication May 8, 2024; revision accepted May 29, 2024.
From the *Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, Geneva, Switzerland; †Department of Cardiology, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland; ‡Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands; §Department of Nuclear Medicine, University of Southern Denmark, Odense, Denmark; and ||University Research and Innovation Center, Óbuda University, Budapest, Hungary.
Conflicts of interest and sources of funding: none declared.
This work was supported by the Euratom research and training programme 2019–2020 Sinfonia project under grant agreement No. 945196.
Compliance with ethical standards: This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the local ethics committee. Consent forms were waived given the retrospective nature of the study.
Correspondence to: Habib Zaidi, PhD, Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, CH-1211 Geneva, Switzerland. E-mail: habib.zaidi@hug.ch.
Supplemental digital content is available for this article. Direct URL citation appears in the printed text and is provided in the HTML and PDF versions of this article on the journal's Web site (www.nuclearmed.com).

ISSN: 0363-9762/24/4910–0899
DOI: 10.1097/RLU.0000000000005400

reported that the position of lesions relative to spleen has additional predictive value in lymphoma patients treated with radiopharmaceutical therapy. During the COVID-19 pandemic, few studies reported the importance of gastrointestinal finding in predicting patient prognosis.[17,18] Szabo et al[19] reported the importance of pericardiac fat in the prognostic prediction of patients with heart failure.

We believe that overall patients' health condition may play a role in prognosis. Besides, we hypothesize that it may contain some information reflecting overall patients' health in the radiomic features space from structural (CT) and metabolic (PET) images acquired from these regions. Deep learning–based segmentation enables fast and reliable delimitation of healthy organs and hence evaluation of any organ separately.[20,21] To the best of our knowledge, the contribution of healthy organs is always overlooked, and studies exploring the importance of healthy organs to estimate overall patient characteristics in survival prediction in NSCLC patients are lacking.

The aim of this study was to use as much as possible image information available from PET/CT images to predict the prognosis in terms of overall survival prediction in patients with NSCLC malignancies. We used radiomic features extracted from 33 organs and tumoral tissues and evaluated the added value of healthy organs radiomics in a comprehensive study using multiple feature selection (FS) and ML models. The primary question addressed was whether the incorporation of total body organ information could enhance the accuracy of AI-based predictions of overall survival.

PATIENTS AND METHODS

Dataset

This study used the RadioGenomics NSCLC dataset downloaded from the TCIA public database.[22] Cases where PET/CT imaging data are available were separated, and the DICOM images converted to NIFTI format. From 211 cases, there were 166 cases with PET/CT, and after preprocessing and excluding images with any kind of processing error or missing data, a total number of 154 PET/CT images was included for training/testing. A detailed description of the demographics, acquisition, and reconstruction parameters is summarized in Table 1. We calculated the time difference between the PET acquisition date and the date of the last follow-up recorded on the dataset description. It should be mentioned that the PET/CT acquisition date was not available for a few cases in the metadata provided by TCIA. For these cases, the DICOM acquisition date information was used. PET images were converted to SUV prior to feature extraction.

TABLE 1. Demographic Description of the Dataset Included in This Study Summarizing Patient Information, PET, and CT Acquisition/Reconstruction Parameters

Demographics	Age (y)	67.2 ± 11.29
	Height (m)	1.69 ± 0.17
	Weight (kg)	76.26 ± 18.51
	Gender	Male (#97), female (#57)
	Affiliation	Stanford (#87), VA (#67)
	Survival status	Alive (#110), deceased (#44)
PET	Manufacturer	Siemens (#10), GE (#144)
	PET spacing (mm)	4.37 ± 0.84
	PET injected activity (MBq)	453.16 ± 90.46
	Time per bed (minutes)	2.33 ± 0.85
	Scatter correction method	Model-based, convolution subtraction
	PET reconstruction method	OSEM, 3D IR, VPFX, OSEM PSF, VPHDS
CT	kVp	80, 100, 120, 130, 140
	Pitch factor	1.08 ± 0.29
	Average tube current (mA)	267.58 ± 163.93

TABLE 2. List of Segmented Organs for 3 Subgroups of Soft, Lung, and Bony Tissues

Boney structures	1	Clavicles
	2	Hips
	3	Sacrum
	4	Ribs
	5	Vertebrae
	6	Femoral heads
Soft tissue	7	Adrenal glands
	8	Aorta
	9	Brain
	10	Colon
	11	Esophagus
	12	Eyeballs
	13	Whole cardiac
	14	Cardiac right atrium
	15	Cardiac left atrium
	16	Cardiac left ventricle cavity
	17	Cardiac right ventricle
	18	Cardiac left myocardium
	19	Kidneys
	20	Liver
	21	Pancreas
	22	Rectum
	23	Rectus lumborum muscles
	24	Small intestine
	25	Spleen
	26	Stomach
	27	Urinary bladder
Lung tissue	28	Whole lungs
	29	Lung LLL
	30	Lung RLL
	31	Lung RML
	32	Lung LUL
	33	Lung RUL

LLL, left lower lobe; RLL, right lower lobe; RML, right middle lobe; LUL, left upper lobe; RUL, right upper lobe.

Organs Segmentation

We used extended and upgraded versions of previously trained deep learning–based segmentation models in our department[20] to segment 28 volumes of interest in healthy organs on the CT images. Those models were trained using nnU-Net[23] segmentation pipeline using 5-fold data split and ensembling all 5 folds inferred on the RadioGenomics CT compartment of PET/CT dataset. The 3Dfullress training model was continued using 2000 epochs, and initial learning rate of 3e-5 decreased after each epoch.

TABLE 3. Summary of All 19 Combinations of Input Data Used in This Study

	Inputs	PET Organomics	CT Organomics	PET GTV	CT GTV	Clincal Information
Clinics Only	Clinics	✗	✗	✗	✗	✓
GTV Only	PET GTV	✗	✗	✓	✗	✗
	CT GTV	✗	✗	✗	✓	✗
	PET GTV + CT GTV	✗	✗	✓	✓	✗
	PET GTV + Clinics	✗	✗	✓	✗	✓
	CT GTV + Clinics	✗	✗	✗	✓	✓
	PET GTV + CT GTV + Clinics	✗	✗	✓	✓	✓
Organomics Only	PET Organomics	✓	✗	✗	✗	✗
	CT Organomics	✗	✓	✗	✗	✗
	CT Organomics + PET Organomics	✓	✓	✗	✗	✗
	PET Organomics + Clinics	✓	✗	✗	✗	✓
	CT Organomics + Clinics	✗	✓	✗	✗	✓
	PET Organomics + CT Organomics + Clinics	✓	✓	✗	✗	✓
Single Modalities	PET GTV + PET Organomics	✓	✗	✓	✗	✗
	CT GTV + CT Organomics	✗	✓	✗	✓	✗
	PET GTV + PET Organomics + Clinics	✓	✗	✓	✗	✓
	CT GTV + CT Organomics + Clinics	✗	✓	✗	✓	✓
	PET Organomics + CT Organomics + PET GTV + CT GTV	✓	✓	✓	✓	✗
All included	PET Organomics + CT Organomics + PET GTV + CT GTV + Clinics	✓	✓	✓	✓	✓

Red cross sign means not used, whereas the green thick sign means using that input. For better readability, they were classified in 4 subgroups and all included means using all 5 inputs as predictors.

The segmented organs were visually checked searching for potential outliers presenting with significant errors. The list of segmented organs is provided in Table 2.

GTV Segmentation

We used nnU-Net pipeline to train a 3Dfullress deep learning model to segment GTV on CT of PET/CT images. We used 3 online available datasets including LIDC[24] (dataset #1) and NSCLC (dataset #2) and manual segmentations available on RadioGenomics[22] (same patients as PET/CT images, dataset #3) datasets for model training using a 5-fold data split. The RadioGenomics dataset had the same patients whom PET/CT images were used to train the survival ML models. It should be mentioned that the RadioGenomics diagnostic CTs with available manual segmentation (143 pair of CT and GTV segmentations) were used both as part of training set and testing set. We used datasets #1 and #2 to increase the number of training datasets and gain a robust model capable of segmenting CT of PET/CT images with a lower image quality.

These 3 datasets were visually assessed, and cases with presenting with errors were excluded from training. After exclusion, 384 cases from NSCLSC dataset, 143 cases from RadioGenomics dataset, and 787 cases from LIDC dataset (total of 1314 pairs of CT and GTV segmentation) were included. Similar to organ segmentation part, we ensembled the output from all 5 folds inferenced on CT images of PET/CT. The GTV segmentations were visually checked and compared with the available ground truth data provided on the diagnostic CT, which was not coregistered with the PET/CT images in few cases.

Feature Extraction

We used Pyradiomics (version 3.1.0)[25] library to extract 107 radiomic features, including first-order statistics (19 features), shape-based (3D) (16 features), shape-based (2D) (10 features), gray level co-occurrence matrix (24 features), gray level run length matrix (16 features), gray level size zone matrix (16 features), neighboring gray tone difference matrix (5 features), and gray level dependence matrix (14 features). We clipped the images prior to feature extraction depending on organ composition for organs and used a predefined clipping value for malignant lesions. We manually classified organs in 1 of 3 subgroups, namely, lung, soft tissue, and bony structures. Then, for each category, prior to extracting the radiomic features, the images were clipped between empirical

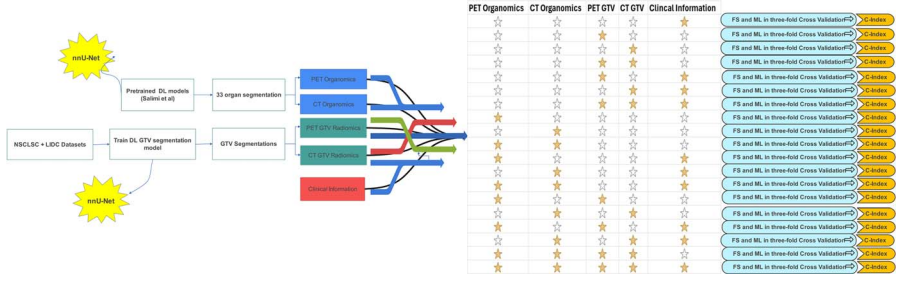

FIGURE 1. Flowchart summarizing the different steps involved in the study protocol. All 19 input combinations were trained using 3-fold cross-validation data split. Filled yellow star means using that input, whereas blank (white) star means that input was not used.

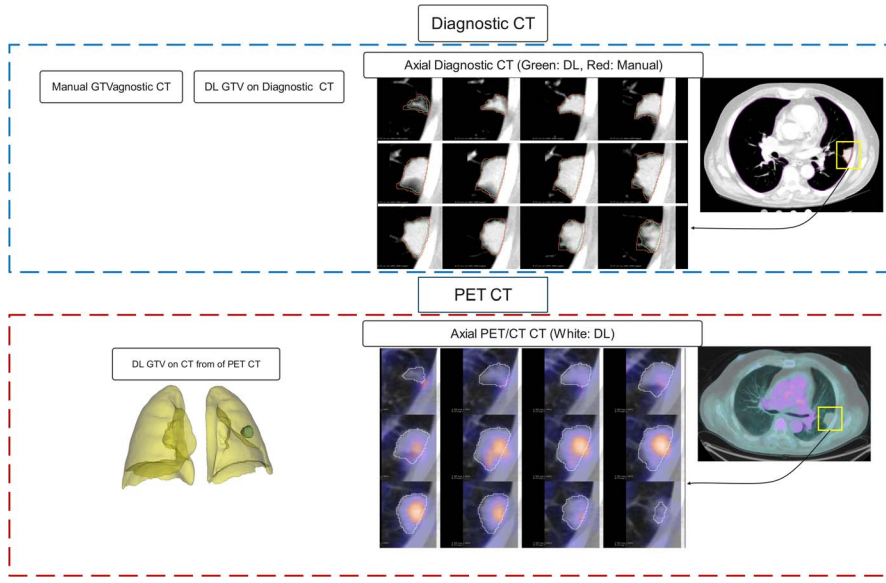

FIGURE 2. GTV segmentations for a case with Dice coefficient of 0.87 on diagnostic CT images. The top row shows a pair of manual (ground truth) and deep learning (DL) segmentation output on a diagnostic CT image where the axial magnified slices compare the manual (red) and DL (green) contours. The bottom row shows the corresponding axial slices segmented using DL on CT of a PET/CT image. The 3D visualization shows the whole lung and GTV segmented.

minimum and maximum values to emphasize the image histogram on the heterogeneities inside these tissues. The clipping values were −900 to 0, −300 to 300, and 0 to 800 HU for lung, soft tissue, and bony structures, respectively. PET images were clipped between 0 and 40 SUV before feature extraction for all segmentation masks.

We extracted features using bin width equal to 10 HUs and 0.4 SUV for CT and PET images, respectively. PET and CT images were resampled to $4 \times 4 \times 4$ mm^3 and $1.5 \times 1.5 \times 1.5$ mm^3, respectively, prior to feature extraction.

Feature Selection and Machine Learning

We considered 19 possible combinations of 5 input data including PET Organomics, CT Organomics, PET GTV, CT GTV, and clinical information. Table 3 summarizes these 19 strategies.

Figure 1 shows the flowchart of steps followed in this study protocol. We used combinations of 5 FS, 6 ML models, and 19 types of input in 3-fold data split to train overall 570×3 (1710) models and compared the performance in terms of concordance index (C-index). Different FS algorithms were used in this study, including minimal depth (MD), mutual information (MI), univariate C-index (UCI), Boruta, variable hunting (VH), and variable hunting variable importance (VH.VIMP). We implemented 6 ML models, including Cox boost (CB), Cox proportional hazards regression (CoxPH), generalized linear model network (GLMN), GLM boosting (GLMB), random survival forest (RSF), and survival tree (ST). Details about the implemented methods are provided in supplementary material.

First, we applied 3-fold nested cross-validation for each input. In each fold (external fold), we used z-score method to normalize feature values based on train dataset and transformed the values (mean and standard deviation) to test dataset. To remove redundant feature, we used Spearman correlation test with a threshold of 90%. This method removes one of the features that have a Spearman correlation coefficient over 90%. Then, FS algorithms were applied on the train dataset. The best selected features for each FS method were fed to ML algorithms. Internal 3-fold cross-validation with grid search was used for hyperparameter optimization. The detail of these parameters is provided in Supplementary Table 1, http://links.lww.com/CNM/A493. The trained model with best hyperparameter was evaluated on test dataset with 1000 bootstraps. Model evaluation was performed with C-index. Mean and standard deviation of 3000 C-indices were reported for each model. The mlr package version 2.18 in R 4.1.2 was used for model development.

Statistical Analysis

The top performance models with respect to the C-index were selected for Kaplan-Meier (KM) curve analysis. The risk score in the test dataset for each fold KM was extracted and combined for all patients. The risk scores were transformed to high-risk and low-risk groups using the median value as the threshold. The log-rank test was used to show significant differences between 2 groups ($P < 0.05$).

RESULTS

Segmentation Accuracy

Figure 2 shows an example of GTVs segmented on both diagnostic quality CT and CT of a PET/CT image for a case with Dice coefficient equal to 0.87, which is lower than the average value. An average Dice coefficient of 0.92 ± 0.08 was calculated on the 143 diagnostic cases showing excellent segmentation performance on GTV segmentation. Figure 3 presents an example of organs segmented on CT of a PET/CT image showing excellent performance of organ segmentation model as reported in a previous study.[20]

Selected Features

Table 4 shows the number of selected features for every 14 possible combinations of inputs where at least 2 types of inputs were used. In other words, CT GTV, PET GTV, CT Organomics, PET Organomics, and clinical parameters were not included in this

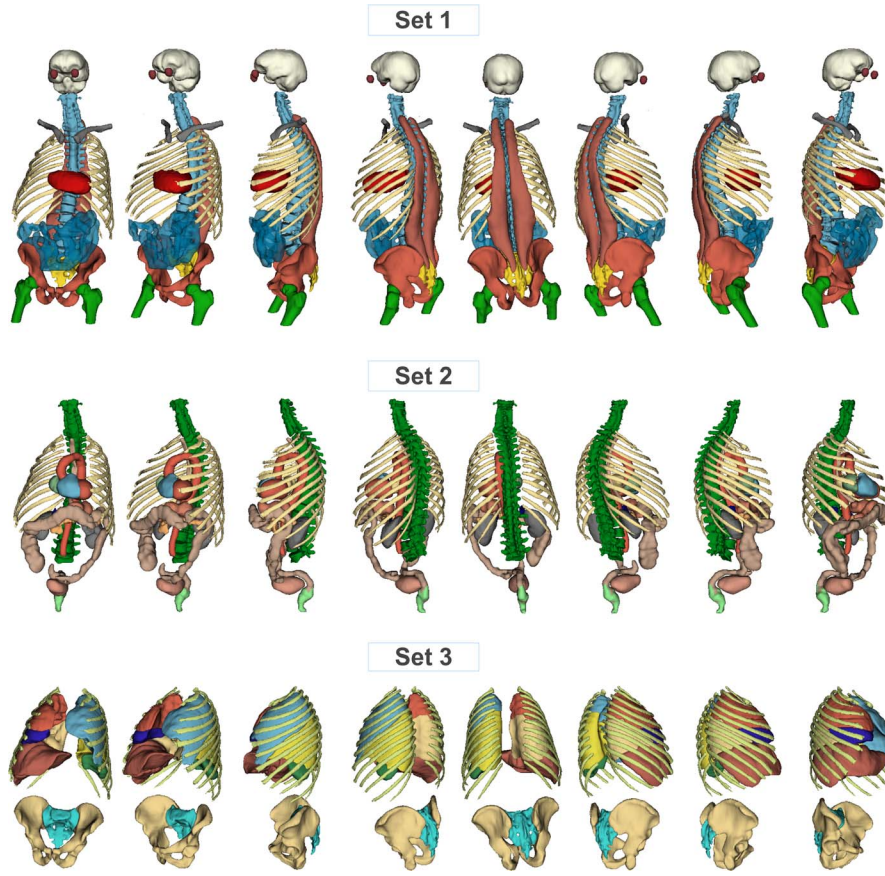

FIGURE 3. 3D visualization of organ segmentations. Set1: Brain, eyeballs, vertebrae, clavicles, ribs, whole heart, rectus lumborum muscle, small intestine, sacrum, hips, and femoral heads. Set2: Vertebrae, esophagus, aorta, heart substructures (LV, RV, LV cavity, RA, LA), stomach, pancreas, colon, rectum, and bladder. Set3: Lung 5 lobes, ribs, sacrum, hips. Some organs are repeated in all 3 sets for better visualization and as anatomical reference.

Table since all the selected features were from the single input data. The most frequently selected features were for PET Organomics. The detailed names of features and organ names selected by all 5 FS models may be found in Supplementary Table 2, http://links.lww.com/CNM/A493.

For inputs of PET Organomics + CT Organomics + PET GTV + CT GTV (all inputs except clinical information), all 5 FS methods selected mostly PET Organomics (86/150), and then for CT Organomics (57/150) features, the most frequent selected organs by all FS methods were aorta, whole lung textures, heart left ventricle myocardium textures, and heart right ventricle textures.

Model Comparison

Table 5 summarizes the average and the best model C-indices for every 19 combinations of inputs averaged over 3 folds. Supplementary Table 3, http://links.lww.com/CNM/A493, depicts the C-index for every 3 folds and all 570 combinations of FS and models. The highest C-index (0.76) was achieved for a single fold using MD FS method, RSF machine, and PET Organomics input. The resulting C-indices heatmap comparing all the 570 models are depicted in Figure 4.

Table 6 summarizes the inputs for every 30 combinations of FS and models with the highest C-index averaged over all folds. PET Organomics was used as input in 18/30 of those combinations, whereas CT Organomics was used in 14/30 combinations. It should be noted that only 6/30 and 11/30 combinations used CT GTV and PET GTV radiomics.

Figure 5 shows the KM curves for 9 selected models. GTV MD/RSF FS and model using PET Organomics + CT Organomics + PET GTV + CT as input showed the lowest *P* value (0.00074), confirming its ability to separate high-risk patients from the low-risk group.

DISCUSSION

Survival prognosis information may be useful in optimizing treatment plans, risk stratification, and resource allocation. Artificial intelligence has been proven to be promising in predicting the prognosis of patients with various malignancies.[26–28] However, the potential information in regions other than the GTV is often overlooked and was not considered in NSCLC cancer prognosis. Lee et al[10] used peritumoral regions radiomics and demonstrated its importance in 2-year survival prediction. Hosny et al[11] showed the importance of radiomics and dosiomics extracted from areas surrounding the GTV in NSCLC patients in prognosis through explainable deep learning and importance maps. Mattonen et al[13] reported on the importance of metabolic tumor volume penumbra extended by 1 cm in NSCLC recurrence. To the best of our knowledge, this is the first study exploring the added value of information contained in regions other than the treatment planning GTV and its surrounding tissues.

TABLE 4. Frequency of Selected Features by Every Input Data by All 5 Feature Selection Methods

Inputs	FS	GTV_CT	GTVs_PET	organomics_CT	organomics_PET	clinics
CT GTV + CT Organomics + Clinics	UCI	7	not included	23	not included	0
	MI	0	not included	1	not included	29
	VH	0	not included	30	not included	0
	VH.VIMP	0	not included	11	not included	0
	MD	0	not included	29	not included	0
PET GTV + PET Organomics + Clinics	UCI	not included	1	not included	29	0
	MI	not included	0	not included	8	22
	VH	not included	0	not included	30	0
	VH.VIMP	not included	0	not included	17	0
	MD	not included	0	not included	24	0
CT GTV + Clinics	UCI	28	not included	not included	not included	2
	MI	0	not included	not included	not included	30
	VH	15	not included	not included	not included	0
	VH.VIMP	9	not included	not included	not included	1
	MD	18	not included	not included	not included	2
PET GTV + Clinics	UCI	not included	29	not included	not included	1
	MI	not included	0	not included	not included	30
	VH	not included	11	not included	not included	2
	VH.VIMP	not included	10	not included	not included	0
	MD	not included	26	not included	not included	4
PET GTV + CT GTV	UCI	14	16	not included	not included	not included
	MI	13	17	not included	not included	not included
	VH	9	6	not included	not included	not included
	VH.VIMP	6	5	not included	not included	not included
	MD	16	14	not included	not included	not included
PET GTV + CT GTV + Clinics	UCI	11	19	not included	not included	0
	MI	3	5	not included	not included	22
	VH	11	10	not included	not included	1
	VH.VIMP	9	5	not included	not included	0
	MD	10	11	not included	not included	2
CT Organomics + Clinics	UCI	not included	not included	28	not included	2
	MI	not included	not included	1	not included	29
	VH	not included	not included	29	not included	1
	VH.VIMP	not included	not included	13	not included	0
	MD	not included	not included	30	not included	0
PET Organomics + Clinics	UCI	not included	not included	not included	30	0
	MI	not included	not included	not included	8	22
	VH	not included	not included	not included	30	0
	VH.VIMP	not included	not included	not included	16	0
	MD	not included	not included	not included	24	0
CT Organomics + PET Organomics	UCI	not included	not included	10	20	not included
	MI	not included	not included	13	17	not included
	VH	not included	not included	13	17	not included
	VH.VIMP	not included	not included	10	7	not included
	MD	not included	not included	15	11	not included
PET Organomics + CT Organomics + Clinics	UCI	not included	not included	6	24	0
	MI	not included	not included	0	1	29
	VH	not included	not included	9	20	1
	VH.VIMP	not included	not included	2	14	0
	MD	not included	not included	17	13	0
CT GTV + CT Organomics	UCI	4	not included	26	not included	not included
	MI	4	not included	26	not included	not included
	VH	0	not included	30	not included	not included
	VH.VIMP	0	not included	13	not included	not included
	MD	0	not included	28	not included	not included
PET GTV + PET Organomics	UCI	not included	4	not included	26	not included
	MI	not included	5	not included	25	not included
	VH	not included	1	not included	29	not included
	VH.VIMP	not included	0	not included	16	not included
	MD	not included	0	not included	30	not included
PET Organomics + CT Organomics + PET GTV + CT GTV	UCI	1	2	9	18	not included
	MI	1	1	12	16	not included
	VH	0	0	10	20	not included
	VH.VIMP	0	0	7	8	not included
	MD	0	0	15	8	not included
PET Organomics + CT Organomics + PET GTV + CT GTV + Clinics	UCI	0	1	4	25	0
	MI	0	0	0	1	29
	VH	0	1	11	18	0
	VH.VIMP	0	0	5	10	0
	MD	0	0	11	13	0

The blue color bar shows the frequency. In cases where input information was not used, the value was replaced by "not included."

We explored the survival prediction capability of different sets of radiomic features extracted from different regions of the GTV and other organs from PET and CT imaging modalities. We also exploited the available clinical information and extensively tested 5 × 6 × 19 models in a 3-fold data split to avoid the effect of random test/train split and invalid results. The aim of this study

TABLE 5. Model Performance Comparison Based on Inputs in Terms of 3-Fold Average C-index

	#	Inputs	C-index (Average Over All Folds)			
			Mean	Std	Minimum	Maximum
Clinical Only	1	Clinical	0.58	0.02	0.52	0.61
GTV Only	2	PET GTV	0.59	0.02	0.55	0.63
	3	CT GTV	0.59	0.02	0.55	0.62
	4	PET GTV + CT GTV	0.59	0.02	0.55	0.63
	5	PET GTV + Clinical	0.60	0.03	0.53	0.65
	6	CT GTV + Clinical	0.59	0.02	0.54	0.63
	7	PET GTV + CT GTV + Clinical	0.60	0.02	0.53	0.63
Organomics Only	8	PET Organomics	0.61	0.03	0.57	**0.68**
	9	CT Organomics	0.61	0.02	0.55	**0.67**
	10	CT Organomics + PET Organomics	0.60	0.03	0.52	**0.68**
	11	PET Organomics + Clinical	0.60	0.02	0.57	0.65
	12	CT Organomics + Clinical	0.60	0.02	0.57	0.63
	13	PET Organomics + CT Organomics + Clinical	0.60	0.02	0.55	0.63
Single Modalities	14	PET GTV + PET Organomics	0.61	0.02	0.57	0.66
	15	CT GTV + CT Organomics	0.60	0.02	0.56	0.65
	16	PET GTV + PET Organomics + Clinical	0.60	0.03	0.54	0.66
	17	CT GTV + CT Organomics + Clinical	0.59	0.02	0.52	0.62
	18	PET Organomics + CT Organomics + PET GTV + CT GTV	0.59	0.02	0.55	**0.67**
All included	19	PET Organomics + CT Organomics + PET GTV + CT GTV + Clinical	0.60	0.02	0.55	0.65

The highest values of 0.68 and 0.67 achieved are highlighted with bold font.

was to investigate the prognostic value of information extracted from different regions. Hence, we used multiple combinations of FS and ML methods to determine the approach achieving the best performance. Our results demonstrated that there is much more information in Organomics that can be used to predict the prognosis with AI. As summarized in Figure 4, all models achieving a C-index more than 0.65 used Organomics, except one. The frequency of the selected features in Table 4 indicates the importance of Organomics in risk stratification, especially for the last 2 input combinations "PET Organomics + CT Organomics + PET GTV + CT GTV" and "PET Organomics + CT Organomics + PET GTV + CT GTV + Clinical" where the whole radiomics information from organs and GTVs was fed into FS algorithm and most of the selected features belong to PET Organomics and CT Organomics inputs for all FS methods, except for MI that selected clinical information instead and not the GTV information. Besides as presented in Table 6, most FS/Model combinations achieved the best results using PET Organomics and CT Organomics information. The most important organs affecting patients' prognosis were the aorta, lungs, and heart substructures.

Our best models using PET Organomics and CT Organomics C-index averaged over 3 folds were 0.68, whereas the highest C-

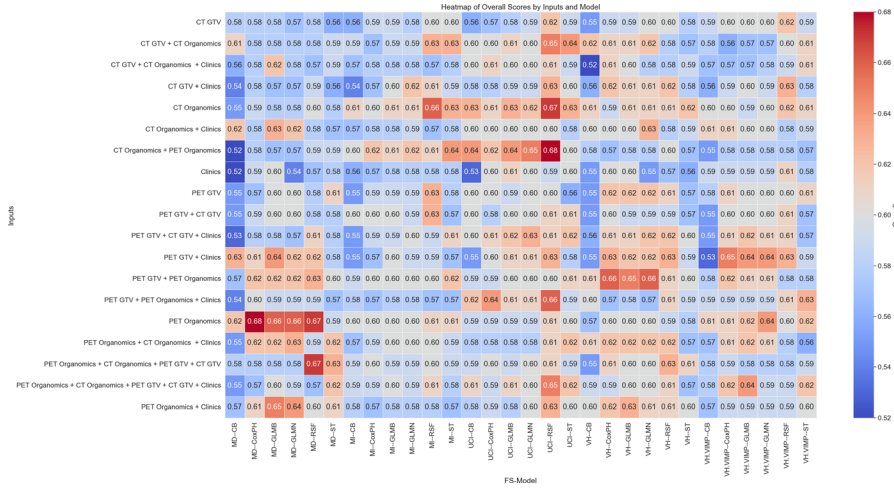

FIGURE 4. C-indices heatmap comparing all the 570 models. The colormap on the right shows the significance of colors. The vertical axis depicts the inputs, whereas the horizontal access depicts the FS/Model combination.

TABLE 6. Highest C-index and the Corresponding Inputs Shown for Every 30 Combinations of FS/Model

FS-Model	Best Inputs	Overall C-index
MD–CB	PET GTV + Clinical	0.63
MD–CoxPH	PET Organomics	0.68
MD–GLMB	PET Organomics	0.66
MD–GLMN	PET Organomics	0.66
MD–RSF	PET Organomics	0.67
MD–ST	PET Organomics + CT Organomics + PET GTV + CT GTV	0.63
MI–CB	CT Organomics	0.61
MI–CoxPH	CT Organomics + PET Organomics	0.62
MI–GLMB	CT Organomics	0.61
MI–GLMN	CT GTV + Clinical	0.62
MI–RSF	CT Organomics	0.66
MI–ST	CT Organomics + PET Organomics	0.64
UCI–CB	CT Organomics + PET Organomics	0.64
UCI–CoxPH	PET GTV + PET Organomics + Clinical	0.64
UCI–GLMB	CT Organomics + PET Organomics	0.64
UCI–GLMN	CT Organomics + PET Organomics	0.65
UCI–RSF	CT Organomics + PET Organomics	0.68
UCI–ST	CT GTV + CT Organomics	0.64
VH–CB	CT GTV + CT Organomics	0.62
VH–CoxPH	PET GTV + PET Organomics	0.66
VH–GLMB	PET GTV + PET Organomics	0.65
VH–GLMN	PET GTV + PET Organomics	0.66
VH–RSF	PET GTV + Clinical	0.63
VH–ST	CT Organomics	0.62
VH.VIMP–CB	PET Organomics	0.61
VH.VIMP–CoxPH	PET GTV + Clinical	0.65
VH.VIMP–GLMB	PET Organomics + CT Organomics + PET GTV + CT GTV + Clinical	0.64
VH.VIMP–GLMN	PET GTV + Clinical	0.64
VH.VIMP–RSF	CT GTV + Clinical	0.63
VH.VIMP–ST	PET GTV + PET Organomics + Clinical	0.63

index in a single fold was 0.76. Our best results using PET GTV, CT GTV, and PET GTV + CT GTV in terms of C-index were 0.63, 0.59, and 0.63, respectively, which is in agreement with results reported by Amini et al[9] using the same inputs (0.63, 0.64, and 0.65, respectively), except CT GTV where the C-index achieved is lower in our study. It should be mentioned that we used 3-fold cross-validation without harmonization, whereas they used 2-fold split strategy and ComBat harmonization. This comparison proved that although we did not have access to the manual GTV segmentations, our deep learning segmentation model provided a comparable GTV segmentation.

One limitation of our study was the lack of ground truth segmentation on PET/CT images. We tried to overcome this issue by using a large training dataset including the diagnostic CT for the same group of patients to train the state-of-the-art nnU-Net model through ensemble learning. We used CT images of PET/CT for the same group of patients as part of the training dataset. It should be clarified that the aim of this study was not to develop a generalizable deep learning segmentation model. This study aimed to test the hypothesis of the presence of important radiomics information in regions other than the GTV and its surrounding tissues. We used the deep learning models to transfer the segmentations from diagnostic CTs available in part of the dataset to PET/CT images. The overall Dice of 0.92 ± 0.08, actually comparable with results reported by Zhang et al[29] and Wang et al,[30] demonstrated the successful transform of the segmentations. However, as we illustrate a case with Dice coefficient equal to 0.87, which is lower than average in Figure 2, there is a good match between the segmentations. We used 2 other datasets for training to overcome the image quality difference between diagnostic CT images and nonenhanced low-dose CT images of PET/CT. It should be mentioned that we cannot claim that organs other than the lungs were healthy organs; it may be additional pathologies in other areas, which may be captured in the radiomics textures. Organomics information might provide a more accurate prediction of patients' prognosis by classifying patients as high or low risk. The added accuracy may be helpful in the decision-making process regarding the selection of treatment plans or in monitoring response to treatment.

CONCLUSIONS

There is important and useful information in terms of radiomic features outside the primary malignancy regions, including organs such as the aorta, heart, and lung, which can improve the performance of AI algorithms. Our study suggests using as much as possible information from medical images toward generating a digital twin of patients with Organomics, GTV information, and clinical data.

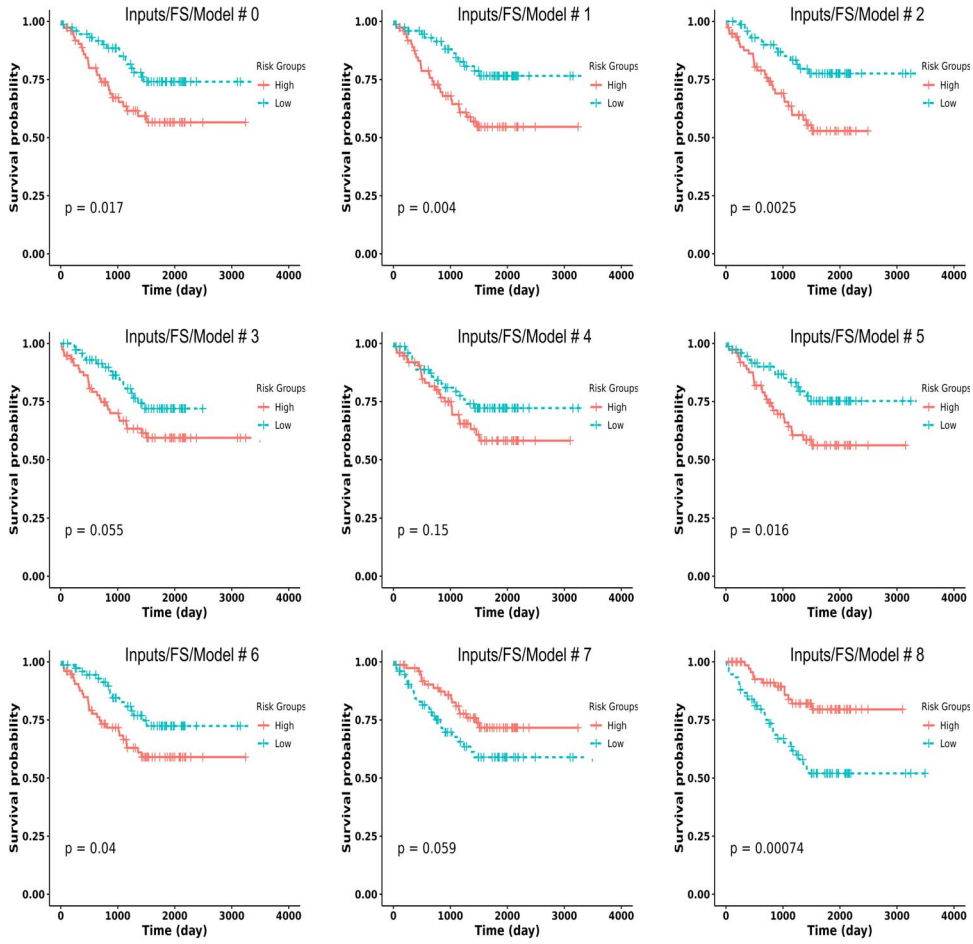

FIGURE 5. KM curves of 9 selected combinations of Inputs/FS/Model. #0: PET GTV + PET Organomics + Clinical/UCI/RSF, #1: PET Organomics + CT Organomics + PET GTV + CT GTV + Clinical/UCI/RSF, #2: CT Organomics/UCI/RSF, #3: CT Organomics/MI/RSF, #4: PET Organomics/MD/Coxph, #5: CT Organomics + PET Organomics/UCI/glmnet, #6: CT GTV + CT Organomics/UCI/RSF, #7: CT Organomics + PET Organomics/VH/Coxph, #8: PET Organomics + CT Organomics + PET GTV + CT GTV/MD/RSF. *P* values shown in the bottom of each curve. *P* values <0.05 are considered statistically significant.

REFERENCES

1. Kratzer TB, Bandi P, Freedman ND, et al. Lung cancer statistics, 2023. *Cancer*. 2024;130:1330–1348.
2. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin*. 2024;74:12–49.
3. Tomaszewski MR, Gillies RJ. The biological meaning of radiomic features. *Radiology*. 2021;298:505–516.
4. Zwanenburg A, Vallieres M, Abdalah MA, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*. 2020;295:328–338.
5. Hatt M, Tixier F, Pierce L, et al. Characterization of PET/CT images using texture analysis: the past, the present... any future? *Eur J Nucl Med Mol Imaging*. 2017;44:151–165.
6. Wang T, Deng J, She Y, et al. Radiomics signature predicts the recurrence-free survival in stage I non–small cell lung cancer. *Ann Thorac Surg*. 2020;109:1741–1749.
7. Libling WA, Korn R, Weiss GJ. Review of the use of radiomics to assess the risk of recurrence in early-stage non–small cell lung cancer. *Transl Lung Cancer Res*. 2023;12:1575–1589.
8. Amini M, Nazari M, Shiri I, et al. Multi-level multi-modality (PET and CT) fusion radiomics: prognostic modeling for non–small cell lung carcinoma. *Phys Med Biol*. 2021;66:205017.
9. Amini M, Hajianfar G, Hadadi Avval A, et al. Overall survival prognostic modelling of non–small cell lung cancer patients using positron emission tomography/computed tomography harmonised radiomics features: the quest for the optimal machine learning algorithm. *Clin Oncol*. 2022;34:114–127.
10. Lee S, Jung J, Hong H, et al. Prediction of two-year recurrence-free survival in operable NSCLC patients using radiomic features from intra- and size-variant peri-tumoral regions on chest CT images. *Diagnostics (Basel)*. 2022;12:1313.
11. Hosny A, Parmar C, Coroller TP, et al. Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. *PLoS Med*. 2018;15:e1002711.
12. Perez-Morales J, Tunali I, Stringfield O, et al. Peritumoral and intratumoral radiomic features predict survival outcomes among patients diagnosed in lung cancer screening. *Sci Rep*. 2020;10:10528.
13. Mattonen SA, Davidzon GA, Bakr S, et al. [^{18}F] FDG positron emission tomography (PET) tumor and penumbra imaging features predict recurrence in non–small cell lung cancer. *Tomography*. 2019;5:145–153.
14. Guo W, Li B, Xu W, et al. Multi-omics and multi-VOIs to predict esophageal fistula in esophageal cancer patients treated with radiotherapy. *J Cancer Res Clin Oncol*. 2024;150:39.
15. Lam S-K, Zhang Y, Zhang J, et al. Multi-organ omics-based prediction for adaptive radiation therapy eligibility in nasopharyngeal carcinoma patients undergoing concurrent chemoradiotherapy. *Front Oncol*. 2022;11.
16. Girum KB, Cottereau A-S, Vercellino L, et al. Tumor location relative to the spleen is a prognostic factor in lymphoma patients: a demonstration from the REMARC Trial. *J Nucl Med*. 2024;65:313–319.

17. Bishehsari F, Adnan D, Deshmukh A, et al. Gastrointestinal symptoms predict the outcomes from COVID-19 infection. *J Clin Gastroenterol*. 2022; 56:e145–e148.
18. Jin S, Lu X, Xu C. COVID-19 induces gastrointestinal symptoms and affects patients' prognosis. *J Int Med Res*. 2022;50:3000605221129543.
19. Szabo L, Salih A, Pujadas ER, et al. Radiomics of pericardial fat: a new frontier in heart failure discrimination and prediction. *Eur Radiol*. 2024;34: 4113–4126.
20. Salimi Y, Shiri I, Mansouri Z, et al. Deep learning–assisted multiple organ segmentation from whole-body CT images. *medRxiv*. 2023; 2023.2010. 2020.23297331.
21. Wasserthal J, Breit HC, Meyer MT, et al. TotalSegmentator: robust segmentation of 104 anatomic structures in CT images. *Radiol Artif Intell*. 2023; 5:e230024.
22. Bakr S, Gevaert O, Echegaray S, et al. A radiogenomic dataset of non–small cell lung cancer. *Sci Data*. 2018;5:180202.
23. Isensee F, Jaeger PF, Kohl SAA, et al. nnU-Net: a self-configuring method for deep learning–based biomedical image segmentation. *Nat Methods*. 2021;18:203–211.
24. Armato Iii SG, McLennan G, Bidaut L, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. *Med Phys*. 2011;38: 915–931.
25. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res*. 2017;77: e104–e107.
26. Mansouri Z, Salimi Y, Amini M, et al. Development and validation of survival prognostic models for head and neck cancer patients using machine learning and dosiomics and CT radiomics features: a multicentric study. *Radiat Oncol*. 2024;19:12.
27. Shiri I, Salimi Y, Pakbin M, et al. COVID-19 prognostic modeling using CT radiomic features and machine learning algorithms: analysis of a multi-institutional dataset of 14,339 patients. *Comput Biol Med*. 2022;145:105467.
28. Zaidi H, El Naqa I. Quantitative molecular positron emission tomography imaging using advanced deep learning techniques. *Annu Rev Biomed Eng*. 2021;23:249–276.
29. Zhang F, Wang Q, Fan E, et al. Enhancing non–small cell lung cancer tumor segmentation with a novel two-step deep learning approach. *J Radiat Res Appl Sci*. 2024;17:100775.
30. Wang P, Ge J, Zheng D, et al. Anatomy-guided deep learning model for accurate and robust gross tumor volume segmentation in lung cancer radiation therapy. *Int J Radiat Oncol Biol Physics*. 2023;117:e71.