# Homological radiomics analysis for prognostic prediction in lung cancer patients

二宮, 健太

Physica Medica 69 (2020) 90–100

Contents lists available at ScienceDirect

Physica Medica

journal homepage: www.elsevier.com/locate/ejmp

Original paper

# Homological radiomics analysis for prognostic prediction in lung cancer patients

Kenta Ninomiya[a], Hidetaka Arimura[b,*]

[a] Division of Medical Quantum Science, Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan

[b] Division of Medical Quantum Science, Department of Health Sciences, Faculty of Medical Sciences, Kyushu University, 3-1-1, Maidashi, Higashi-ku, Fukuoka 812-8582, Japan

ARTICLE INFO

*Keywords:*
Homology
Topologically invariant
Betti number
Radiomics
Lung cancer
Survival prediction
Cox proportional hazard model

ABSTRACT

*Purpose:* This study explored a novel homological analysis method for prognostic prediction in lung cancer patients.

*Materials and methods:* The potential of homology-based radiomic features (HFs) was investigated by comparing HFs to conventional wavelet-based radiomic features (WFs) and combined radiomic features consisting of HFs and WFs (HWFs), using training (n = 135) and validation (n = 70) datasets, and Kaplan–Meier analysis. A total of 13,824 HFs were derived through homology-based texture analysis using Betti numbers, which represent the topologically invariant morphological characteristics of lung cancer. The prognostic potential of HFs was evaluated using statistically significant differences (p-values, log-rank test) to compare the survival curves of high- and low-risk patients. Those patients were stratified into high- and low-risk groups using the medians of the radiomic scores of signatures constructed with an elastic-net-regularized Cox proportional hazard model. Furthermore, deep learning (DL) based on AlexNet was utilized to compare HFs by stratifying patients into the two groups using a network that was pre-trained with over one million natural images from an ImageNet database.

*Results:* For the training dataset, the p-values between the two survival curves were $6.7 \times 10^{-6}$ (HF), $5.9 \times 10^{-3}$ (WF), $7.4 \times 10^{-6}$ (HWF), and $1.1 \times 10^{-3}$ (DL). The p-values for the validation dataset were $3.4 \times 10^{-5}$ (HF), $6.7 \times 10^{-1}$ (WF), $1.7 \times 10^{-7}$ (HWF), and $1.2 \times 10^{-1}$ (DL).

*Conclusion:* This study demonstrates the excellent potential of HFs for prognostic prediction in lung cancer patients.

## 1. Introduction

Lung cancer is the leading cause of cancer mortality worldwide, accounting for more than 1.7 million deaths in 2018 [1]. An association between lung cancer prognosis and tumor traits has been reported, especially in cases where computed tomography (CT) were employed [2–5]. Several radiomics studies have attempted to evaluate lung cancer prognosis by characterizing tumor traits using histograms, texture features, and wavelet-decompositions that quantify intra-tumor heterogeneity [6–8].

Chalkidou et al. [9] demonstrated that the p-values stated in fifteen published papers, which evaluated whether there were associations between patient survivals and texture features, were statistically unreliable when appropriate statistical corrections were applied to original data. Prognosis estimation using automated feature engineering based on deep learning (DL) was slightly more accurate than that employing conventional engineered features, such as wavelet-based radiomic features (WFs), though there was not always a significant performance improvement in two-year survival prediction ($p > 0.05$, permutation test) [10]. These conventional methods (WFs and DL) could not interpret underlying tumor traits related to lung cancer that are present in medical images. The results discussed above necessitate the development of novel mathematical models that are specifically designed to evaluate the prognostic potential of lung cancer.

The prognostic potential of cavitation in lung cancer has been investigated in previous studies at both the macroscale and microscale levels [11–13]. Cavitation is most commonly observed in larger and undifferentiated cancer cases, and probably results from rapid tumor growth that exceeds the available blood supply, leading to central necrosis [14]. Kolodziejski et al. [11] compared two survival curves

* Corresponding author.
*E-mail address:* arimurah@med.kyushu-u.ac.jp (H. Arimura).

https://doi.org/10.1016/j.ejmp.2019.11.026
Received 24 May 2019; Received in revised form 27 November 2019; Accepted 29 November 2019

(Kaplan–Meier curves) representing 100 patients with cavitation and 994 patients with solid tumors (without cavitation) using CT images. Statistically significant differences ($p < 0.05$, log-rank test) were obtained between curves for patients with and without cavitation [11]. Patients with cavitation had lower 5- and 10-year survival probabilities, despite the fact that the patients in both groups were at similar stages [11]. Santos et al. [12] investigated the correlation between patient survival times and ten qualitative features (including cavitation) identified on CT images. The absence or presence of cavitation was recognized as a significant feature that can be used to stratify patients with squamous cell carcinoma (SCC) into good and poor-prognostic groups ($p < 0.001$, log-rank test) [12]. Wang et al. [13] revealed that cavitation in lung pathology images may also be associated with prognosis. The significant prognostic factor was found to be the number of holes (cavitation) using a convolutional neural network based on a univariate Cox proportional hazard model (CPHM) ($p < 0.05$, log-rank test). However, none of these studies quantified the cavitation of lung cancer in CT images.

Homology has been successfully utilized in various image-analysis applications, such as the detection and classification of prognostic cancer lesions in pathological or CT images. This was achieved by quantifying the cavitation derived from pixel value heterogeneity in cancerous regions using Betti numbers [15,16], which are mathematically invariant. The zero- and one-dimensional Betti numbers, which can be calculated from two-dimensional images, indicate the number of connected components (b0) and holes (b1) in an object [16]. The Betti numbers could elucidate the topologically invariant morphological characteristics of cancer, i.e., the intrinsic properties of cancer. However, homology has not been utilized in the evaluation of lung cancer prognosis.

We hypothesized that homology can facilitate both the quantification of heterogeneity, encompassing lung tumor traits such as cavitation, and prognostic prediction. Furthermore, the application of homology and homology-based radiomic features (HFs) could lead to the development of a mathematical model for predicting patient prognoses that compares favorably with conventional WFs. In this study, we introduced a homological radiomics analysis-based approach to prognostic prediction in lung cancer patients. Furthermore, the feasibility of prognostic prediction using engineered features such as HF and WF was compared to that using DL-based automated engineered features.

## 2. Methods

### 2.1. Clinical cases

A total of 205 positron emission tomography (PET)-CT images of lung cancer patients were selected from the publicly available non-small-cell lung carcinoma (NSCLC)-Radiomics database (n = 422; n: number of patients) of The Cancer Imaging Archive (TCIA) (Table S1) [7,17,18]. The selection criteria for patients were as follows: adequate tumor delineation data without defects such as erroneously misplaced or missing contours in the PET-CT images was required for each patient, and their tumor histologic subtype was either adenocarcinoma (ADN), SCC, or large-cell carcinoma (LC). The selected dataset consisted of PET-CT images of NSCLC patients (male: 147; female: 58; age range: 43–92 y; subject to complete follow up: 135; censored: 70) with inoperable, histologically confirmed NSCLCs and tumor stages I (n = 38; 18.5%), II (n = 22; 10.7%), IIIA (n = 52; 25.4%), and IIIB (n = 92; 44.9%), together with unavailable stage data (n = 1; 0.5%). The histologic subtypes of the 205 tumors were ADN (n = 28; 13.6%), SCC (n = 85; 41.5%), and LC (n = 92; 44.9%). The patients had received either radiotherapy or radio-chemotherapy. The case numbers of the selected clinical cases were listed in Table S1 in the Supplemental information. The PET-CT images were acquired using an $^{18}$F-fluorodeoxyglucose (FDG) PET-CT (Biograph SOMATOM Sensation-16 system, Siemens Medical Solutions, Erlangen, Germany) with dimensions of 512 × 512 pixels, an in-plane pixel size of 0.98 mm × 0.98 mm, and slice thickness of 3 mm. The radiation treatment plans were developed by a radiation oncologist using a treatment planning system (XiO, Computerized Medical Systems, St. Louis, Missouri, USA). The gross tumor volumes (GTVs) were defined based on fused FDG PET-CT images. Anisotropic CT images and GTV regions were transformed into isotropic images with an isovoxel size of 0.98 mm, using cubic and shape-based interpolation [19], respectively. In this study, the axial plane of each CT image that contained the maximum GTV area was selected for the calculation of radiomic features [20].

The multiple-segmentation dataset that was used for assessing the reproducibility of features for constructing radiomic signatures (Subsection 2.2.3) consisted of CT images of patients with NSCLC (n = 30) from the Quantitative Imaging Net-work (QIN) multisite collection of lung CT data, with nodule segmentations from the TCIA (Table S2) [21–23]. The patients were selected from the Lung Image Database Consortium-Image Database Resource Initiative database (n = 10; gender unknown; slice thickness: 1.0–2.5 mm), QIN-Lung CT database (n = 10; male: 4; female: 6; slice thickness: 3 mm), and the Reference Image Database to Evaluate Response lung CT database (n = 10; genders unknown; slice thickness: 1.25 mm). The tumor volume in each image was independently segmented by three different institutions: Columbia University Medical Center, Stanford University, and Moffitt Cancer Center/University of South Florida. Each institution performed segmentation using their own custom segmentation algorithms under three different sets of initial conditions. These configurations resulted in nine segmentations of each tumor, for a total of 270 segmentations [21,22].

### 2.2. Engineered features

#### 2.2.1. WF calculations

Fig. 1 (right-hand side) presents the overall scheme for the calculation of WFs [24]. A total of 216 WFs were calculated from 54 features, using a histogram and four texture-characterization matrices (i.e., a gray level co-occurrence matrix (GLCM) [25], gray level run-length matrix (GLRLM) [26], gray level size-zone matrix (GLSZM) [27], and neighborhood gray-tone difference matrix (NGTDM) [28]) for four wavelet-decomposed images. The 54 radiomic features are listed in the Supplemental information (Table S3). The matrices were constructed by analyzing eight neighborhoods. Decomposition was performed by applying either a low-pass filter (scaling function, L) or high-pass filter (wavelet function, H) [29] to eight-bit CT images in each direction. Therefore, the four wavelet decomposition filters consisted of a combination of two filters using either a low-pass filter (L) or high-pass filter (H) in each direction (the x or y direction in the two-dimensional images). We acquired eight-bit CT images based on a look-up table ranging from 0 to 255, which correspond to a range of −1000 to 1500 Hounsfield units. Decomposition results in down-sampling the image scale [30,31]. WFs were then calculated from down-sampled GTVs, which were obtained using nearest-neighbor interpolation. The computation of radiomic features was performed using a MATLAB-based radiomics tools package (implemented in MATLAB 2017b, MathWorks) [32,33].

#### 2.2.2. HF calculation

HFs with Betti numbers could indicate the topologically invariant morphological characteristics of cancerous regions, i.e., the intrinsic properties of cancer. Fig. 1 (left-hand side) presents the overall scheme for HF calculation. A total of 13,824 HFs were obtained by applying 54 calculation methods based on histogram and texture features to 256 Betti maps. These maps were obtained from binary images by calculating Betti numbers within local regions of interest (ROIs), which were determined using kernel size, shifting pixels, and Betti number types

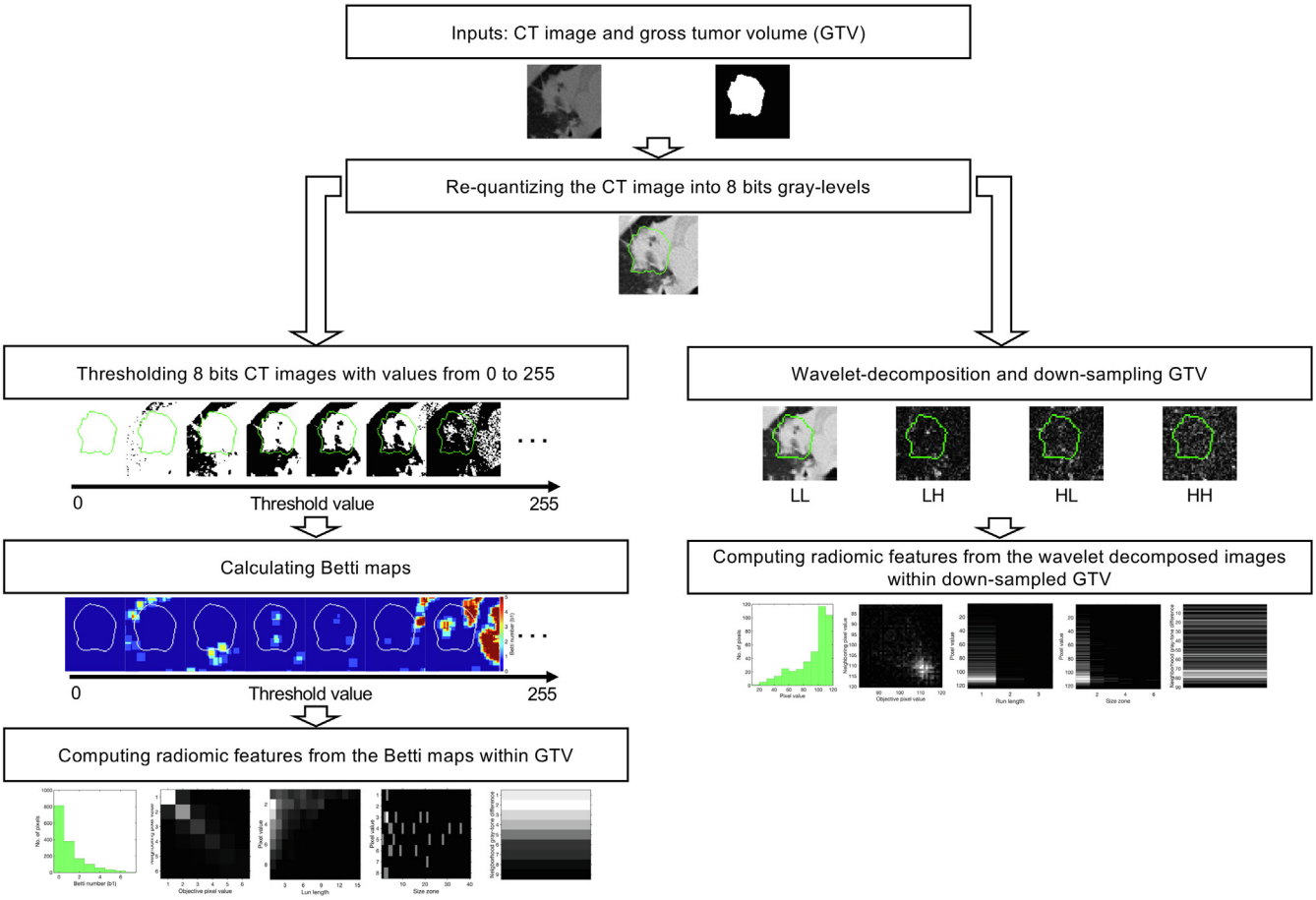


**Fig. 1.** Overall scheme for the calculation of homology-based- (left-hand side) and wavelet-based features (right-hand side).

(Fig. 2). The binary images were derived from eight-bit CT images by thresholding the images with values ranging from 0 to 255. The b0 and b1 values were computed by counting connected components and holes, respectively, that exhibit eight-connectivity within local ROIs. The holes were computed by counting connected components with reversed binary pixels, with exception of connected components that included pixels on the edge of local ROIs with four-connectivity, as shown in Fig. 2. The b1/b0 values were obtained as ratios of b1 to b0, which represents the number of holes per connected component. If b0 was zero, b1/b0 was not calculated, and simply set equal to b1. The kernel sizes and shifting pixels used in the calculation of Betti maps determined the sizes of detectable holes and sampling intervals of Betti

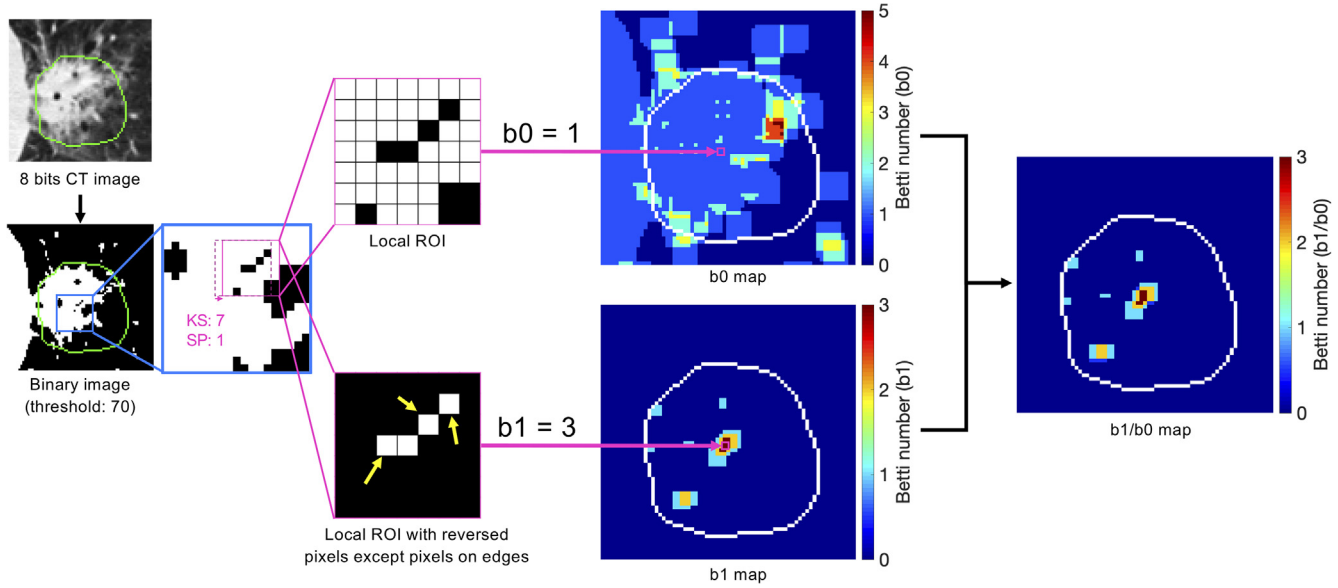


**Fig. 2.** Calculation procedure for the creation of Betti maps (b0, b1, and b1/b0) based on counting the number of connected components within a local region of interest (ROI), determined using the kernel (kernel size (KS): 7 and shifting pixel (SP): 1).

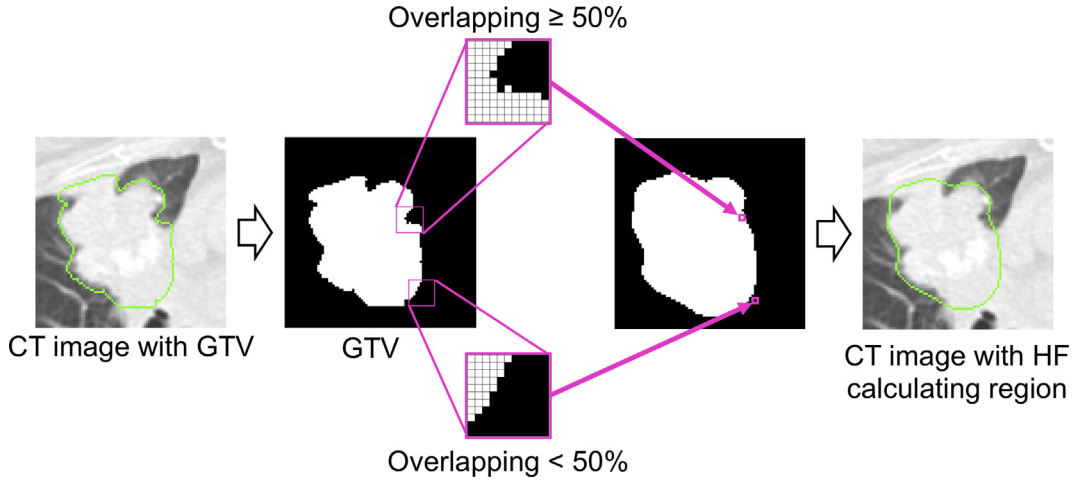


**Fig. 3.** Determination of homology-based feature (HF) calculation regions based on the gross tumor volume (GTV) in computed tomography (CT) images.

numbers, respectively. The HF calculation was performed within the GTVs. Overlapping regions between GTVs and kernels of more than 50% were considered to be within the GTVs (Fig. 3). The computation of Betti numbers was performed using custom-built algorithms for counting connected objects that were developed in MATLAB 2017b.

*2.2.3. Radiomic signature construction*

Radiomic signatures were derived from frequent features that were selected in more than 90% of the iterations (e.g. over 90 times in 100 iterations) and had the greatest sums of absolute coefficients that were obtained using the Cox-net algorithm [34–36]. The Cox-net algorithm selected different features with minimum cross-validation errors at each time, due to the random variables for splitting data for cross-validation in the algorithm. For this reason, frequent features were considered to be useful for prognostic prediction in the majority of cases. In the Cox-net algorithm, the CPHM was optimized by maximizing the partial likelihood of the model using L1 and L2 norm penalties, which were based on a radiomic signature (a feature vector with length $m$) $\boldsymbol{x}_i \in \boldsymbol{R}^m$, where $m$ is the number of radiomic features in real coordinate space$\boldsymbol{R}$, and $i$ is the patient number. The CPHM $h_i(t|\boldsymbol{x}_i)$ for a patient $i$ at a time $t$ can be expressed as

$$h_i(t|\boldsymbol{x}_i) = h_0(t)\exp(\boldsymbol{\beta}^T\boldsymbol{x}_i), \tag{1}$$

where $h_0(t)$ is the baseline hazard function, and $\boldsymbol{\beta}$ is a coefficient vector with length $m$. To maximize the partial likelihood that is subject to a convex combination of L1 and L2 norm penalties (elastic net penalty), a Lagrangian formulation is constructed as follows [35]:

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmax}}[kl(\boldsymbol{\beta}) - \lambda P_\alpha(\boldsymbol{\beta})], \tag{2}$$

where

$$l(\boldsymbol{\beta}) = \sum_{i=1}^{n} d_i\left[\boldsymbol{\beta}^T\boldsymbol{x}_i - \ln\left\{\sum_{t_j \geq t_i}\exp(\boldsymbol{\beta}^T\boldsymbol{x}_j)\right\}\right], \tag{3}$$

$$d_i = \begin{cases} 1, & \text{event occured at } t_i \\ 0, & \text{censored at } t_i \end{cases}, \tag{4}$$

$$P_\alpha(\boldsymbol{\beta}) = \alpha\|\boldsymbol{\beta}\|_1 + \frac{1-\alpha}{2}\|\boldsymbol{\beta}\|_2^2. \tag{5}$$

In Eqs. (2)–(5), $k$ is a scaling factor, $n$ is the number of patients in a training dataset, $t_i$ is the survival time of a patient $i$, $t_j$ is the survival time of a patient $j$ who lived longer than patient $i$, $\|\bullet\|_1$ indicates the $L_1$ norm (least absolute shrinkage and selection operator (LASSO) penalty term), $\|\bullet\|_2$ indicates the $L_2$ norm (ridge regression penalty term), $\lambda$ is the Lagrange multiplier, and $\alpha \in [0, 1]$ is a blending parameter for adjusting the impact of the LASSO and ridge-regression penalties on overall regularization. In this study, a value of 0.5 (elastic net) was used for the blending parameter $\alpha$. A value of $\alpha = 0$ reduces the penalization of ridge regression, which means that a greater number of radiomic features are included in the model. A value of $\alpha = 1$ reduces the penalization of the LASSO, thereby reducing the number of features present by "shrinking" the coefficients of highly correlated features to zero. The parameter $\lambda$ was optimized based on a uniform grid of 100 values (as recommended in Ref. [34]) using a one-step coordinate descent algorithm [35]. The cross-validation error was calculated based on five-fold cross validation, where the training and validation datasets were selected using random variables. The cross-validation error $CE(\lambda)$ of s-fold cross-validation can be expressed as [36]

$$CE(\lambda) = -\frac{1}{n}\sum_{k=1}^{s}\{l(\hat{\boldsymbol{\beta}}_\lambda^k) - l^k(\hat{\boldsymbol{\beta}}_\lambda^k)\} \tag{6}$$

In this case, k is the fold number; $\hat{\boldsymbol{\beta}}_\lambda^k$ is the coefficient vector that was optimized based on the datasets using the Cox-net algorithm with $\lambda$, where the $k^{th}$ subset of patients is removed and used for validation; and $l^k$ is the partial likelihood calculated for the patients, excluding the $k^{th}$ subset.

It was assumed that features with higher absolute $\boldsymbol{\beta}$ values are more strongly associated with prognosis. The number of features that composed a signature was therefore determined, based on features selected more frequently than in 90% of the iterations (over 90 times in 100 iterations) of the Cox-net algorithm. If more than 13 features were selected over 90 ($< n_t/10$; $n_t$: number of patients in a training dataset, see Section 2.4 [37–39]), then the 13 features with the greatest sums of the absolute coefficients for 100 iterations of the Cox-net algorithm were selected for the signature. In addition, the optimal calculation parameters for HFs (kernel size, shifting pixel, and Betti number type) were determined based on average minimum cross-validation errors.

Linearly dependent features were removed from the signatures using a variance inflation factor (VIF) because these can lead to overfitting or unstable model performance [40]. If a signature generated using the procedure detailed above had features with high VIFs ($\geq 10$ [41]) the feature with the lowest absolute value of coefficients over 100 Cox-net algorithm iterations was removed. This procedure was repeated until the VIFs of all the features were lower than 10.

Irreproducible features were also removed from the signatures according to intraclass correlation coefficients (ICCs), which were independently computed using the multiple segmentation dataset to enhance the robustness of the signatures to inter-observer variability in GTVs. Features that yielded ICCs $< 0.8$ [42] were considered to be irreproducible, and eliminated following the VIF-based feature removal procedure discussed above.

#### 2.2.4. Homology and wavelet combined features

Homology and wavelet combined features (HWFs) were generated from HF and WF signatures and fed into the Cox-net algorithm to construct signatures. The feature selection method was the same as that described in Subsection 2.2.3 for the construction of radiomic signatures.

### 2.3. DL-based automated feature engineering

A DL model based on AlexNet pre-trained using natural images from the ImageNet database [43,44] was retrained using training datasets, including CT images of patients whose survival times (training data) were known. We used eight-bit CT images and survival times as training data for transfer learning. The DL model was trained using the Adam algorithm with 100 epochs [45]. Data augmentation including rotation, flipping, translation, scaling, and shearing, was applied during the training procedure [46].

### 2.4. Comparison of prognostic potential

The feasibility of HF-based prognosis was investigated by comparing the proposed approach to conventional WFs and HWFs for both training and validation datasets in a Kaplan–Meier analysis. The prognostic potentials of the radiomic features were evaluated based on statistically significant differences (p-values, log-rank test) between two survival curves (Kaplan–Meier curves) that were stratified using the medians of the rad-scores of radiomic signatures. A rad-score was computed for each patient based on a linear combination of constituent features in the signature, weighted by their respective coefficients, which were determined by constructing CPHMs based on the signature. Fig. 4 presents the workflow for the validation of prognostic potentials. Patients (n = 205) were divided into training (n = 135; subject to complete follow up: 90; censored: 45) and validation datasets (n = 70; subject to complete follow up: 45; censored: 25) based on random variables. For the validation of engineered features, a CPHM was constructed based on the radiomic signature of the training dataset to obtain the $\beta$ coefficients. The rad-scores for the training and validation datasets were then calculated using the $\beta$ coefficients and radiomic signatures. A DL model was constructed based on patients in the training dataset that were monitored until they died to predict the survival times of patients in both the training and validation datasets (see Subsection 2.3). Finally, patients in both the training and validation datasets were stratified into high- and low-risk groups using the medians of the rad-scores or the survival times in the training dataset. If the survival curves for the validation dataset showed significant differences ($p < 0.05$, log-rank test), then the associated prognostic prediction models were considered to be appropriate for stratifying new patients into the two groups. Construction of the CPHMs was performed using the "coxph" function in the "survival" package of R Ver. 3.4.3 [47].

According to a review paper by Chalkidou et al. [9], the correction of threshold values for statistical significance ($p < 0.05$) was not necessary in this study because validation tests were performed.

## 3. Results

The optimal parameters for HF calculation were a kernel size of seven, one shifting pixel, and a Betti number type of b1/b0. Fig. S1 (Supplemental information) presents box plots of the minimum cross-validation errors that were obtained from 100 Cox-net algorithm iterations for HFs calculated using 60 parameters. In Fig. S1, each feature was ordered according to their average minimum cross-validation errors, which are indicated by magenta crosses, from left (smallest) to right (largest). The average minimum cross-validation

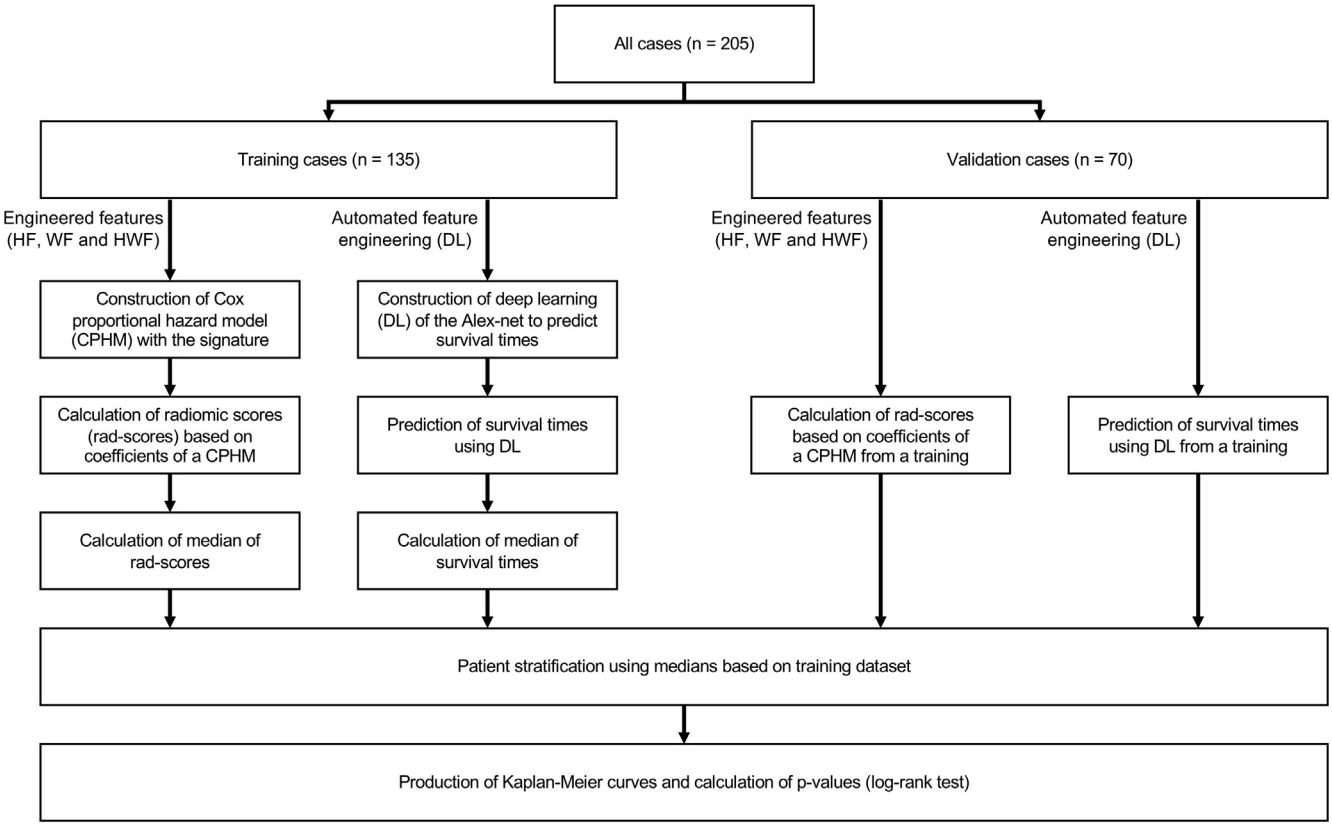


**Fig. 4.** Workflow for training and validation of homology-based (HF), wavelet-based (WF), homology and wavelet combined feature-based (HWF), and deep learning (DL) prognostic potentials.

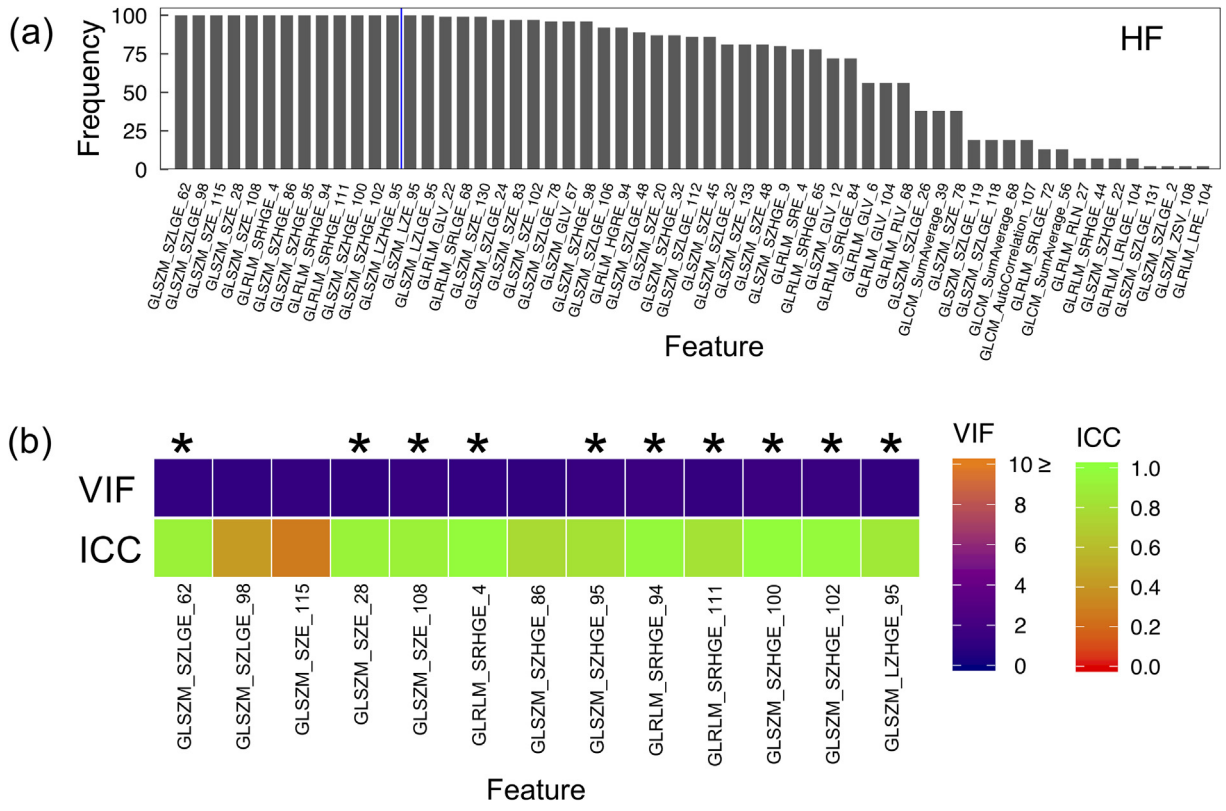


**Fig. 5.** (a) Frequencies of radiomic features with non-zero coefficients optimized over 100 Cox-net algorithm iterations for homology-based radiomic features (HFs) with (b) variance inflation factors (VIFs) and intraclass correlation coefficients (ICCs). Features on the left-hand side of the blue line satisfied the necessary conditions, i.e., selection more than 90 times, and were one of the 13 features with the greatest sums of absolute values of the coefficients $\boldsymbol{\beta}$ ($13 < n_t/10$; $n_t$: number of patients in a training dataset). The ten features indicated by asterisks (*) were included in an HF signature. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

error obtained using the optimal parameters was $5.3 \pm 0.055$. There was a statistically significant difference ($p < 0.001$, *t*-test) between the average minimum cross-validation errors of the first (a kernel size of seven, one shifting pixel, and a Betti number type of b1/b0) and the second parameter settings (kernel size of eleven, four shifting pixels, and a Betti number type of b0).

Ten, six, and nine features were selected for the radiomic signatures of the HFs, WFs, and HWFs, respectively. Figs. 5–7 present (a) the frequencies of radiomic features with non-zero coefficients optimized over 100 Cox-net algorithm iterations with (b) VIFs and ICCs, together with radiomic signatures (denoted by asterisks) for the HFs, WFs, and HWFs. Feature names were expressed in their short forms (see the Supplemental information). In Figs. 5 and 7, digits in the range of 0–255 following the names of the features correspond to the threshold values for obtaining Betti maps. In Figs. 6 and 7, H and L represent low- and high-pass filters for wavelet decomposition, respectively. The first letter indicates high- or low-pass filtration along the x direction, while the second letter represents filtration along the y direction.

The HFs yielded a greater number of frequent features compared to the WFs and HWFs, which is considered to be useful for most patients from the perspective of prognostic prediction (Figs. 5(a), 6(a), and 7(a)). The numbers of frequent features for the HFs, WFs, and HWFs were 26, 10, and 9, respectively. In Fig. 5(a), 6(a), and 7(a), features with the same frequencies were ordered based on the sums of the absolute coefficients from large (left) to small (right). Features on the left-hand side of the blue line satisfied conditions, i.e., those were selected at least 90 times, and were one of the 13 features with the greatest sum of absolute coefficient values ($13 < n_t/10$; $n_t$: number of patients in a training dataset).

All the frequent features of the HFs were independent ($VIF < 1.5$), whereas four of those (Hist_Mean_HL, GLSZM_LZE_LL, GLCM_AutoCorrelation_HL, and GLSZM_LZHGE_LL) of the WFs were highly correlated with other features ($VIF \geq 10$), see Fig. 5(b) and 6(b). The frequent features of the HWFs were independent ($VIF < 1.4$), see Fig. 7(b).

HFs were able to produce larger numbers of reproducible features than WFs. The number of reproducible features ($ICC > 0.8$) for the HFs was 12,211 (88% of 13,824 HFs) and that for the WFs was 106 (49% of 216 WFs). In addition, features comprising the HF signatures tended to be more reproducible than those comprising the WF signatures (Fig. 5(b) and 6(b)). The average ICCs of constituent features in the signatures of HFs and WFs were $0.92 \pm 0.062$ and $0.88 \pm 0.045$ ($p > 0.05$, *t*-test), respectively, see Fig. 5(b) and 6(b).

The HF-based approach exhibited greater prognostic potential than the conventional models that were based on WFs and DL. HWFs showed a similar potential compared to HFs. Fig. 8 presents the Kaplan–Meier curves for the training and validation datasets, where patients were stratified using the medians of the rad-scores or survival times for the training dataset. For the training dataset, the p-values (hazard ratio (HR), confidence interval (CI)) between the two survival curves were $6.7 \times 10^{-6}$ (HR: 0.41, 95% CI: 0.26–0.65) (HF), $5.9 \times 10^{-3}$ (HR: 0.57, 95% CI: 0.37–0.88) (WF), $7.4 \times 10^{-6}$ (HR: 0.41, 95% CI: 0.26–0.65) (HWF), and $1.1 \times 10^{-3}$ (HR: 0.51, 95% CI: 0.33–0.78) (DL) (Table1). For the validation dataset, the p-values (HRs, CIs) between the two survival curves were $3.4 \times 10^{-5}$ (HR: 0.32, 95% CI: 0.16–0.62) (HF), $6.7 \times 10^{-1}$ (HR: 0.88, 95% CI: 0.48–1.6) (WF), $1.7 \times 10^{-7}$ (HR: 0.24, 95% CI: 0.12–0.48) (HWF), and $1.2 \times 10^{-1}$ (HR: 0.63, 95% CI: 0.35–1.1) (DL) (Table1).

## 4. Discussion

Our study produced results similar to a previous work [12] in terms

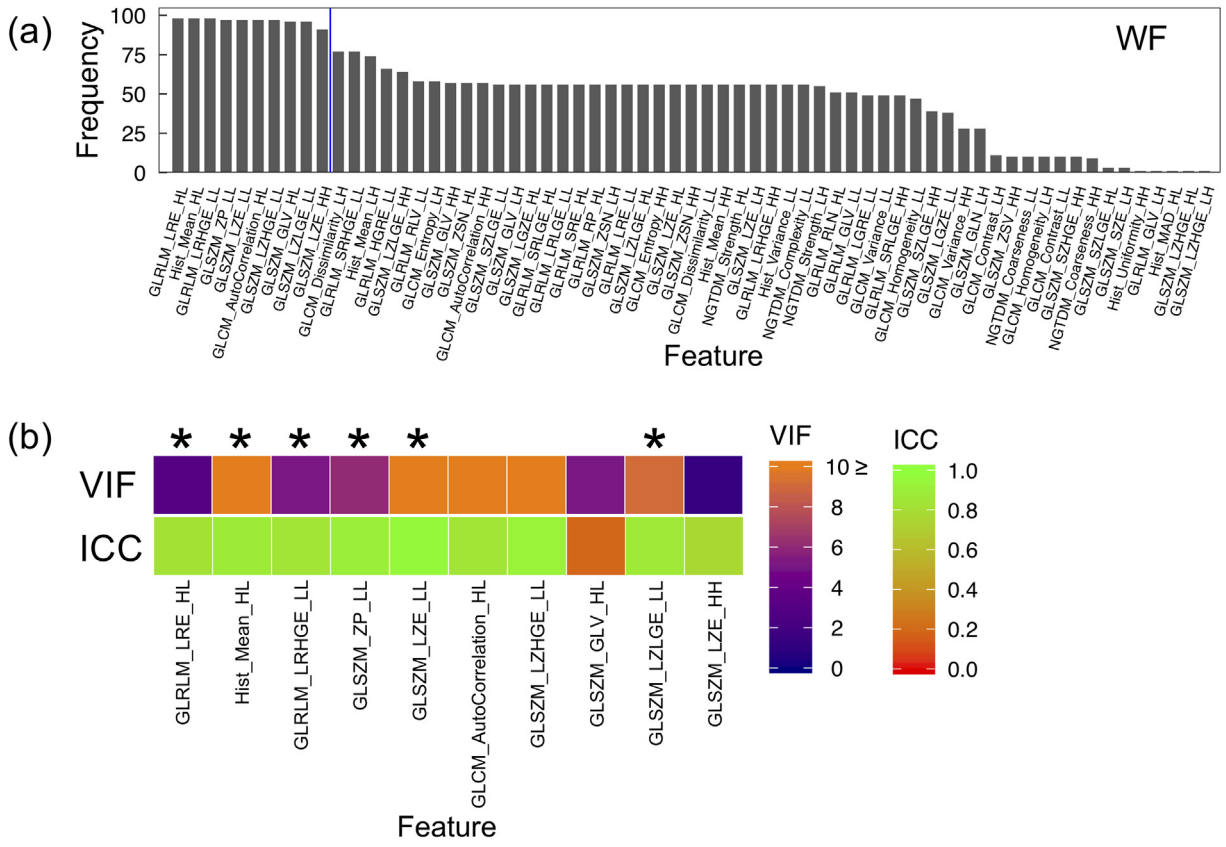


**Fig. 6.** (a) Frequencies of radiomic features with non-zero coefficients optimized over 100 Cox-net algorithm iterations for wavelet-based radiomic features (WFs) with (b) variance inflation factors (VIFs) and intraclass correlation coefficients (ICCs). Features on the left-hand side of the blue line satisfied the necessary conditions, i.e., selection more than 90 times, and were one of the 13 features with the greatest sums of absolute values of the coefficients $\boldsymbol{\beta}$ ($13 < n_t/10$; $n_t$: number of patients in a training dataset). The six features indicated by asterisks (*) were included in an WF signature. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of qualitative features for characterizing cavities in lung tumors. Fig. 9 presents the Kaplan–Meier curves for SCC patients who were stratified into high- and low-risk groups, based on the medians of the rad-scores calculated using the optimal parameters. The high-rad-score group had associated survival curve with the shorter survival expectation. The p-value (HR, CI) between the two survival curves was $7.3 \times 10^{-4}$ (HR: 0.43, 95% CI: 0.24–0.75) (log-rank test). Homological radiomics analysis revealed the same trend as shown in the previous study [12] with higher rad-scores indicating poorer prognosis.

HFs produced greater numbers of frequent features compared to WFs. In this study, feature selection was based on 100 iterations of the Cox-net algorithm and the training and validation datasets for model construction were different in each iteration. Fifteen features were selected 100 times for the HFs (Fig. 5(a)), whereas no features were selected 100 times for the WFs (Fig. 6(a)). The features that were frequently selected for the HFs tended to be the same, regardless of the training dataset.

HF and WFs may quantify different traits of tumors. The VIFs for features in the HWF signature were very low (VIF $< 1.4$), despite the inclusion of both HFs and WFs. This result indicates that the HFs and WFs were not correlated to each other.

HFs represent a geometrically unbiased method for evaluating inhomogeneity in local regions, including cavitation. Both WFs and HFs can be considered to quantify tumor inhomogeneity because cavitation may be derived from variations in voxel values as a result of thresholding. However, differences between WFs and HFs can be observed if there is any bias in the analysis of local regions. Wavelet-based analysis focuses on specific inhomogeneity patterns that are defined by mother wavelets and the directions of application for high- or low-pass filters (x or y axis). Homology-based analysis can extract the topologically invariant morphological characteristics of cancer, i.e., its intrinsic properties, which could be included in patterns of tumor inhomogeneity. These results indicate that HFs more accurately evaluate the characteristics of tumor malignancy and provide better patient stratification compared with conventional image features.

The results of this study indicate that differences in the number of holes within a tumor were associated with the prognoses of lung cancer patients. We assumed that low-density holes and tumor heterogeneities could be characterized by features derived from Betti maps. Low-density holes with low CT-threshold values may be associated with tiny cavities and/or tumor heterogeneity at the CT resolution levels. However, it should be noted that Betti map features may reflect the image noise in CT images as well.

In past studies, tumor volumes were found to be a significant prognostic factor in the treatment of stage I-III NSCLC with radiotherapy [48] or radio-chemotherapy [49]. We investigated correlations between tumor volume and other radiomic features using Spearman's correlation coefficients (CC). The CCs for all the features in the HF-, WF-, and HWF-based signatures were less than 0.2. The correlation of GLSZM_SZE_108 in the HFs-based signature was statistically significant ($p = 0.014$, *t*-test) but that of the other radiomic features was not.

Prognostic potential of tumor volumes was assessed by stratifying the patients into high- and low-risk groups based on a median tumor volume. The p-values (HR, CI) between the survival curves for the training dataset and validation dataset were $1.3 \times 10^{-1}$ (HR: 0.73, 95% CI: 0.48–1.1) and $3.4 \times 10^{-1}$ (HR: 1.4, 95% CI: 0.74–2.5) for only the tumor volume, respectively (Table 2). On the other hand, the HFs showed smaller p-values than the tumor volume, that are $6.7 \times 10^{-6}$

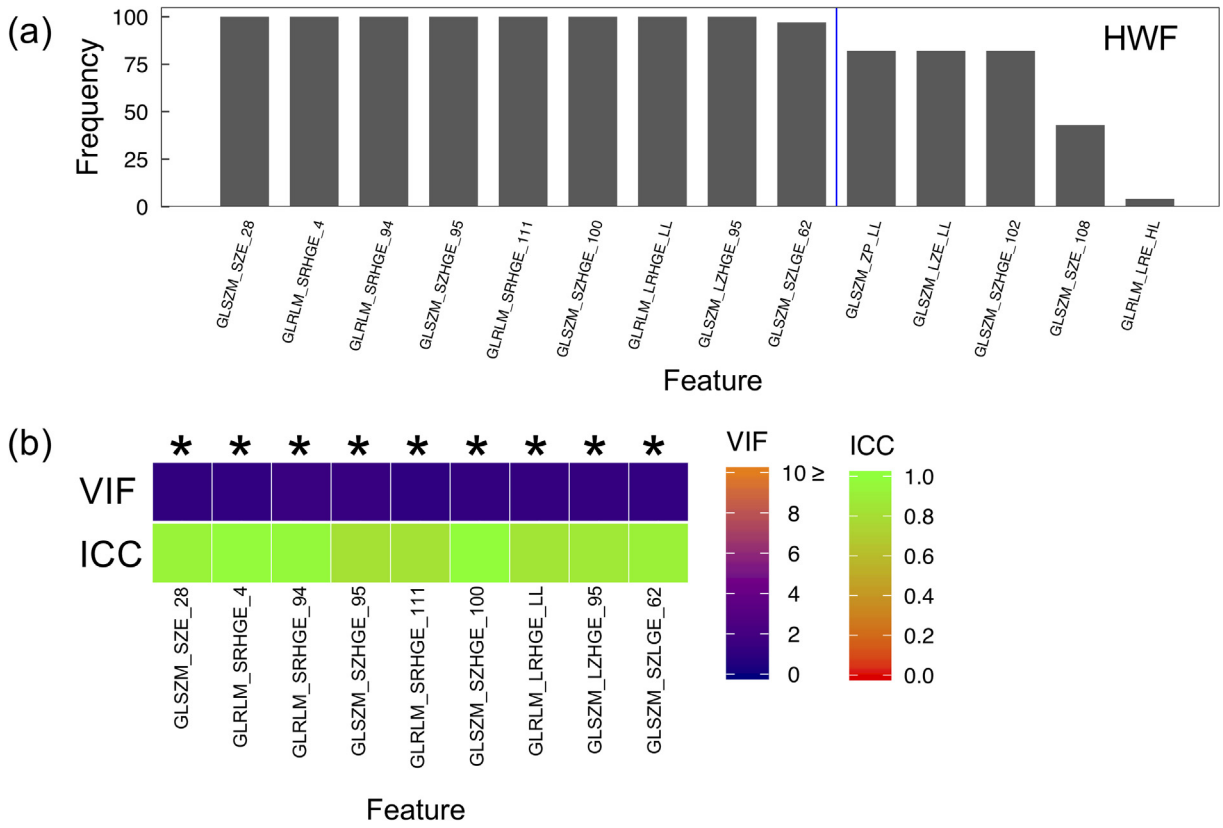


**Fig. 7.** (a) Frequencies of radiomic features with non-zero coefficients optimized over 100 Cox-net algorithm iterations for homology and wavelet combined features (HWFs) with (b) variance inflation factors (VIFs) and intraclass correlation coefficients (ICCs). Features on the left-hand side of the blue line satisfied the necessary conditions, i.e., selection more than 90 times, and were one of the 13 features with the greatest sums of absolute values of the coefficients $\boldsymbol{\beta}$ ($13 < n_t/10$; $n_t$: number of patients in a training dataset). The nine features indicated by asterisks (*) were included in an HWF signature. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(HR: 0.41, 95% CI: 0.26–0.65) (training) and $3.4 \times 10^{-5}$ (HR: 0.32, 95% CI: 0.16–0.62) (validation) (Table 1). This indicates that HFs are a more feasible and stable prognostic prediction tool than tumor volumes.

Furthermore, to investigate the impact of the tumor volume on the prognostic power of a radiomic signature for the rad-scores, the tumor volume was used with HFs, WFs, and HWFs in a radiomic signature to classify patients into high- and low-risk groups using the rad-scores. For the training dataset, the p-values (HR, CI) between the two survival curves of two groups based on the rad-scores with the tumor volume were $2.3 \times 10^{-6}$ (HR: 0.39, 95% CI: 0.24–0.62) (HF), $5.6 \times 10^{-4}$ (HR: 0.50, 95% CI: 0.32–0.77) (WF) and $1.3 \times 10^{-8}$ (HR: 0.33, 95% CI: 0.20–0.53) (HWF) (Table2). For the training dataset, the p-values (HR, CI) between the two survival curves were $1.4 \times 10^{-2}$ (HR: 0.49, 95% CI: 0.26–0.91) (HF), $4.4 \times 10^{-1}$ (HR: 0.80, 95% CI: 0.44–1.4) (WF) and $4.1 \times 10^{-8}$ (HR: 0.23, 95% CI: 0.11–0.47) (HWF) (Table 2). The tumor volume enhanced the prognostic power of the radiomic signatures with WFs and HWFs.

In our training dataset, there were significant differences in terms of the mean volumes of lung cancer between high- and low-risk groups that were stratified based on the medians of the rad-scores of HFs, WFs, HWFs, and DL. For the training dataset, the p-values (Mann–Whitney *U* test) between the mean volumes for these groups were $2.5 \times 10^{-1}$ (HF), $8.8 \times 10^{-1}$ (WF), $8.4 \times 10^{-1}$ (HWF), and $4.8 \times 10^{-1}$ (DL). For the validation dataset, the p-values (Mann–Whitney *U* test) between the mean volumes were $7.6 \times 10^{-2}$ (HF), $5.9 \times 10^{-2}$ (WF), $1.9 \times 10^{-2}$ (HWF) and $5.4 \times 10^{-1}$ (DL).

Variations in the number of quantization levels affected the accuracy of the CPHMs. The cross-validation error of the feature selection was employed to assess the impact of the number of quantization levels on the HFs. The mean cross-validation errors (see Subsection 2.2.3) with seven- and nine-bit images were $5.5 \pm 0.023$ and $5.5 \pm 0.041$, respectively. Both errors differed significantly ($p < 0.001$, *t*-test) from the error obtained when using eight-bit CT images ($5.3 \pm 0.055$).

There were four limitations to this study. First, given that the local ROIs used to calculate Betti numbers were positioned according to shifting pixels and kernel sizes, there may have been overlaps or gaps between neighboring local ROIs. Based on the optimal parameters identified in this study (a kernel size of seven, one shifting pixel, and a Betti number type of b1/b0), overlaps between two local ROIs and several holes could have been counted multiple times. Furthermore, in cases where there were gaps between two local ROIs, some holes may have been missed, which could have affected the subsequent feature calculations.

The second limitation was that patient prognosis evaluation results may change when three-dimensional image analysis is applied. In this study, only two-dimensional homological radiomic analysis was performed because Shen et al. demonstrated the superior prognostic potential of radiomic features when two-dimensional calculations were utilized in place of three-dimensional calculations [20]. However, tumor traits that represent patient malignancy may reside throughout the volume of a tumor [50]. Further experiments involving three-dimensional homological image analysis should be performed in future studies.

Third, signatures constructed using data for patients who are subject to a single treatment method (e.g. only radiotherapy) may differ from those constructed from the database used in this study. In the database, we included patients who received radiotherapy treatment and patients who had been treated using radio-chemotherapy to increase the total number of patients. Although the results of validation (Fig. 8)

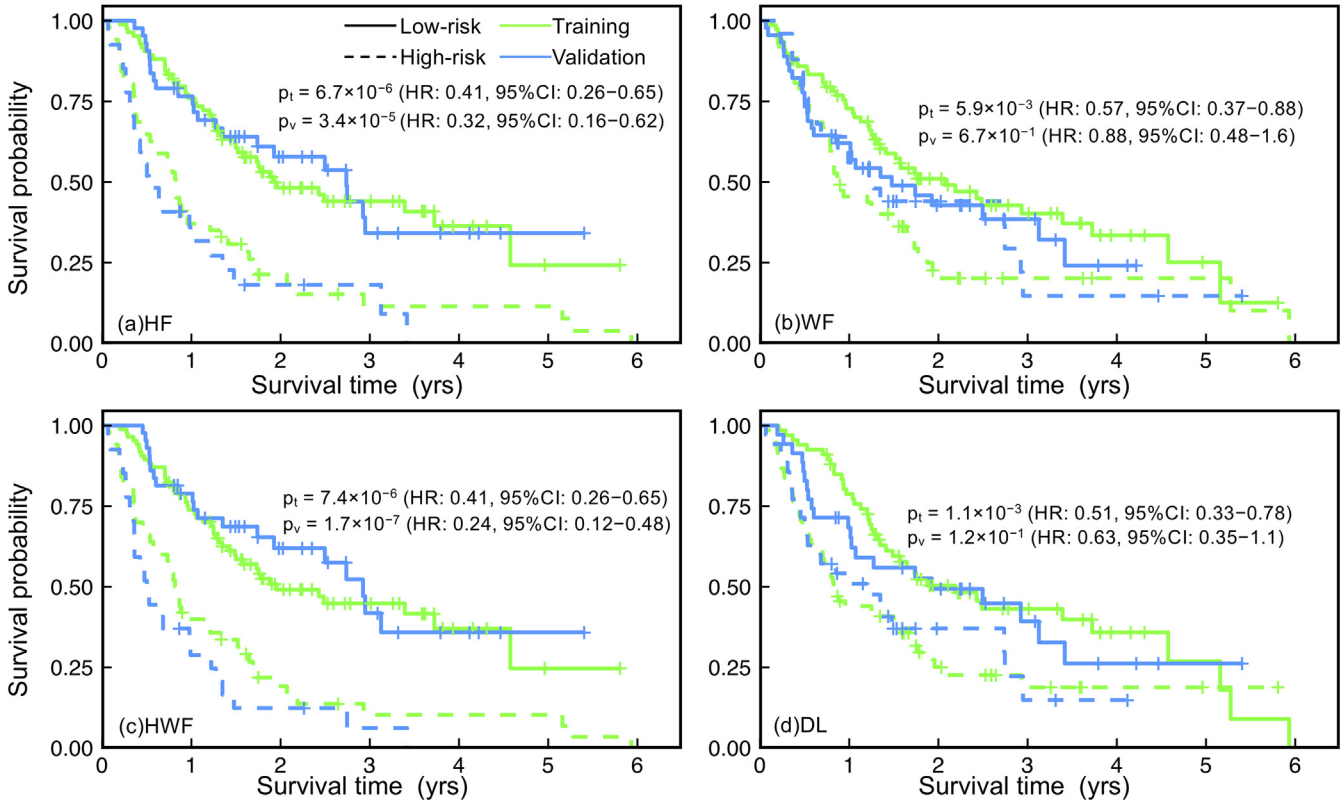


**Fig. 8.** Kaplan–Meier curves for two patient groups that were stratified using the median of the radiomic-scores based on (a) homology (HFs), (b) wavelets (WFs), (c) homology and wavelet combined features (HWFs), and (d) predicted survival time based on deep learning (DL) using AlexNet with p-values calculated using the log-rank test. $p_t$ and $p_v$ indicate the p-values of the survival curves for the training and validation datasets, respectively. HR and CI refer to the hazard ratio and confidence interval, respectively.

**Table 1**
p-values for the two survival curves of high- and low-risk patients using homology-based radiomic features (HFs), wavelet-based radiomic features (WFs), combined radiomic features consisting of HFs and WFs (HWFs) and deep learing (DL).

| | P-value (Hazard ratio, 95% confidence interval) | |
|---|---|---|
| | Training (n = 135) | Validation (n = 70) |
| HF | $6.7 \times 10^{-6}$ (0.41, 0.26–0.65) | $3.4 \times 10^{-5}$ (0.32, 0.16–0.62) |
| WF | $5.9 \times 10^{-3}$ (0.57, 0.37–0.88) | $6.7 \times 10^{-1}$ (0.88, 0.48–1.6) |
| HWF | $7.4 \times 10^{-6}$ (0.41, 0.26–0.65) | $1.7 \times 10^{-7}$ (0.24, 0.12–0.48) |
| DL | $1.1 \times 10^{-3}$ (0.51, 0.33–0.78) | $1.2 \times 10^{-1}$ (0.63, 0.35–1.1) |

HF: homology-based radiomic features, WF: wavelet-based radiomic features, HWF: combined radiomic features consisting of HFs and WFs, DL: deep learning, n: number of cases.

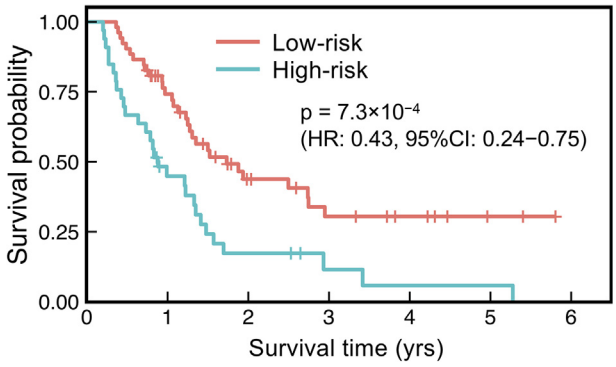


**Fig. 9.** Kaplan–Meier curves for squamous cell carcinoma patients who were stratified into two groups using the median of the radiomic scores for homology-based radiomic features, with p-values for the two curves (log-rank test). HR and CI refer to the hazard ratio and confidence interval, respectively.

demonstrated that the trained prediction models were effective for use with mixed patient populations containing both radiotherapy and radio-chemotherapy patients, the prediction performance may vary if only one specific treatment method was considered in the model construction process.

Finally, this study did not assess the reproducibility of HFs for different scanner/scanning protocols or patient populations. The signatures of the HFs in this study were composed of reproducible features, which are robust in terms of inter-observer variability in GTVs, because irreproducible features (ICCs $< 0.8$) were removed during the construction of signatures (Subsection 2.2.3). In contrast, conventional features computed from CT images were likely to be irreproducible, due to variations in scanner/scanning protocols [51,52] or populations. Mackin et al. demonstrated that the inter-scanner variability of conventional features calculated from CT images should be considered based on the characteristics inherent to scanners [51]. Li et al. reported that using CT images with thicker slices significantly reduces the accuracy of support vector machines when using conventional features in the identification of gene-mutated NSCLC patients [52]. Furthermore, the impact of variations in patient populations on accuracy is unclear because this study focused on a single database. Therefore, the reproducibility of the HFs used in this study must be investigated using different scanner/scanning protocols and patient populations.

## 5. Conclusions

In this study, we explored the feasibility of a novel homological radiomics analysis method for prognostic prediction in lung cancer

**Table 2**
p-values for the two survival curves of high- and low-risk patients using tumor volumes alone and with the radiomic signatures of homology-based radiomic features (HFs), wavelet-based radiomic features (WFs) and combined radiomic features consisting of HFs and WFs (HWFs).

| | P-value (Hazard ratio, 95% confidence interval) | |
|---|---|---|
| | Training (n = 135) | Validation (n = 70) |
| Tumor volume | $1.3 \times 10^{-1}$ (0.73, 0.48–1.1) | $3.4 \times 10^{-1}$ (1.4, 0.74–2.5) |
| HF with tumor volume | $2.3 \times 10^{-6}$ (0.39, 0.24–0.62) | $1.4 \times 10^{-2}$ (0.49, 0.26–0.91) |
| WF with tumor volume | $5.6 \times 10^{-4}$ (0.50, 0.32–0.77) | $4.4 \times 10^{-1}$ (0.80, 0.44–1.4) |
| HWF with tumor volume | $1.3 \times 10^{-8}$ (0.33, 0.20–0.53) | $4.1 \times 10^{-8}$ (0.23, 0.11–0.47) |

HF: homology-based radiomic features, WF: wavelet-based radiomic features, HWF: combined radiomic features consisting of HFs and WFs, n: number of cases.

patients. Radiomic signatures consisting of HFs that were calculated using optimal parameters (a kernel size of seven, one shifting pixel, and a Betti number type of b1/b0) showed a more promising prognostic potential than both conventional methods that employ WFs, which are currently used in most radiomics studies, and DL models. These findings indicate that the proposed homological analysis method is a feasible option for evaluating the prognoses of lung-cancer patients, with the potential to provide direct clinical benefits.

## Acknowledgments

The authors are grateful to Dr. Kazuaki Nakane, Osaka University and all the members of the Arimura Laboratory (http://web.shs.kyushu-u.ac.jp/~arimura), whose comments and suggestions made enormous contributions to this study. We would also like to thank Editage (www.editage.jp) for their English language editing service.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmp.2019.11.026.

## References

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;2018. https://doi.org/10.3322/caac.21492.
[2] Ma J, Yang T-L, Wang Y, Zhang X-W, Gu X-S, Wang Z-C. Relationship between computed tomography morphology and prognosis of patients with stage I non-small cell lung cancer. Onco Targets Ther 2017;10:2249–56. https://doi.org/10.2147/ott.s114960.
[3] Fu JY, Wan YL, Huang TY, Wu CF, Liu YH, Hsieh MJ, et al. Correlation between image characteristics and pathologic findings in non small cell lung cancer patients after anatomic resection. PLoS One 2018;13:1–14. https://doi.org/10.1371/journal.pone.0206386.
[4] Wang X, Lv L, Zheng Q, Huang X, Li B. Differential diagnostic value of 64-slice spiral computed tomography in solitary pulmonary nodule. Exp Ther Med 2018;15:4703–8. https://doi.org/10.3892/etm.2018.6041.
[5] Avanzo M, Stancanello J, El Naqa I. Beyond imaging: the promise of radiomics. Phys Medica 2017;38:122–39. https://doi.org/10.1016/j.ejmp.2017.05.071.
[6] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Van Stiphout RGPM, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441–6. https://doi.org/10.1016/j.ejca.2011.11.036.
[7] Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Cavalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014. https://doi.org/10.1038/ncomms5006.
[8] Arimura H, Soufi M, Kamezawa H, Ninomiya K, Yamada M. Radiomics with artificial intelligence for precision medicine in radiation therapy. J Radiat Res 2019;60(1):150–7. https://doi.org/10.1093/jrr/rry077.
[9] Chalkidou A, O'Doherty MJ, Marsden PK. False discovery rates in PET and CT studies with texture features: a systematic review. PLoS One 2015;10. https://doi.org/10.1371/journal.pone.0124165.
[10] Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, et al. Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. PLoS Med 2018;15:e1002711https://doi.org/10.1371/journal.pmed.1002711.
[11] Kolodziejski LS, Dyczek S, Duda K, Góralczyk J, Wysocki WM, Lobaziewicz W. Cavitated tumor as a clinical subentity in squamous cell lung cancer patients. Neoplasma 2003;50:66–73. https://doi.org/10.1080/10615800802272251.
[12] Koenigkam Santos M, Muley T, Warth A, De Paula WD, Lederlin M, Schnabel PA, et al. Morphological computed tomography features of surgically resectable pulmonary squamous cell carcinomas: impact on prognosis and comparison with adenocarcinomas. Eur J Radiol 2014;83:1275–81. https://doi.org/10.1016/j.ejrad.2014.04.019.
[13] Wang S, Chen A, Yang L, Cai L, Xie Y, Fujimoto J, et al. Comprehensive analysis of lung cancer pathology images to discover tumor shape and boundary features that predict survival outcome. Sci Rep 2018;8. https://doi.org/10.1038/s41598-018-27707-4.
[14] Mooi WJ. Common lung cancers. In: Hasleton PS, editor. Spencer's pathology of the lung. 5th ed.New York, NY: McGraw-Hill; 1996. p. 1009–64.
[15] Adcock A, Rubin D, Carlsson G. Classification of hepatic lesions using the matching metric. Comput Vis Image Underst 2014;121:36–42. https://doi.org/10.1016/j.cviu.2013.10.014.
[16] Nakane K, Takiyama A, Mori S, Matsuura N. Homology-based method for detecting regions of interest in colonic digital images. Diagn Pathol 2015;10:1–5. https://doi.org/10.1186/s13000-015-0244-x.
[17] Aerts HJWL, Rios-Velazquez E, Leijenaar RTH. Data from NSCLC-radiomics. The Cancer Imaging Archive (TCIA). https://doi.org/10.7937/K9/TCIA.2015.PF0M9REI.
[18] Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The cancer imaging archive (TCIA): maintaining and operating a public information repository. J Digit Imaging 2013;26:1045–57. https://doi.org/10.1007/s10278-013-9622-7.
[19] Herman GT, Zheng J, Bucholtz CA. Shape-based interpolation. IEEE Comput Graph Appl 1992;12:69–79. https://doi.org/10.1109/38.135915.
[20] Shen C, Liu Z, Guan M, Jiangdian S, Yucheng L. 2D and 3D CT radiomics features prognostic performance comparison in non-small cell lung cancer. Transl Oncol 2017;10:886–94. https://doi.org/10.1016/j.tranon.2017.08.007.
[21] Kalpathy-Cramer J, Zhao B, Goldgof D, Gu Y, Wang X, Yang H, et al. A comparison of lung nodule segmentation algorithms: methods and results from a multi-institutional study. J Digit Imaging 2016;29:476–87. https://doi.org/10.1007/s10278-016-9859-z.
[22] Kalpathy-Cramer J, Mamomov A, Zhao B, Lu L, Cherezov D, Napel S, et al. Radiomics of lung nodules: a multi-institutional study of robustness and agreement of quantitative imaging features. Tomography 2016;2:430–7. https://doi.org/10.18383/j.tom.2016.00235.
[23] Kalpathy-Cramer J, Napel S, Goldgof D, Zhao B. Multi-site collection of lung CT data with nodule segmentations. Cancer Imaging Arch (TCIA) 2015. https://doi.org/10.7937/k9/tcia.2015.1buvfjr7.
[24] Soufi M, Arimura H, Nakamoto T, Hiroseaki T, Ohga S, Umezu Y, et al. Exploration of temporal stability and prognostic power of radiomic features based on electronic portal imaging device images. Phys Med 2018;46:32–44. https://doi.org/10.1016/j.ejmp.2017.11.037.
[25] Haralick RM, Dinstein I, Shanmugam K. Textural features for image classification. IEEE Trans Syst Man Cybern 1973. https://doi.org/10.1109/TSMC.1973.4309314.
[26] Galloway MM. Texture analysis using gray level run lengths. Comput Graph Image Process 1975;4:172–9. https://doi.org/10.1016/S0146-664X(75)80008-6.
[27] Thibault G, Fertil B, Navarro C, Pereira S. Texture indexes and gray level size zone matrix application to cell nuclei classification. Pattern Recognit Inf Process 2009.
[28] Amadasun M, King R. Textural features corresponding to textural properties. IEEE Trans Syst Man Cybern 1989;19:1264–74. https://doi.org/10.1109/21.44046.
[29] Soufi M, Arimura H, Nagami N. Identification of optimal mother wavelets in survival prediction of lung cancer patients using wavelet decomposition-based radiomic features. Med Phys 2018;45:5116–28. https://doi.org/10.1002/mp.13202.
[30] Strang G, Nguyen T. Wavelets and filter banks. Wellesley: Wellesley-Cambridge Press; 1997.
[31] Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 1989. https://doi.org/10.1109/34.192463.
[32] Vallières M, Freeman CR, Skamene SR, El Naqa I. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys Med Biol 2015. https://doi.org/10.1088/0031-9155/60/14/5471.
[33] Vallières M. Radiomics: MATLAB programming tools for radiomics analysis.

GitHub; 2015. https://github.com/mvallieres/radiomics; 2018 [accessed 18.11.30].
[34] Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010;33:1–22. https://doi.org/10.1016/j.expneurol.2008.01.011.
[35] Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for Cox's proportional hazards model via coordinate descent. J Stat Softw 2011;39:1–13. https://doi.org/10.18637/jss.v039.i05.
[36] Sun H, Lin W, Feng R, Li H. Network-regularized high-dimensional Cox regression for analysis of genomic data. Stat Sin 2014;24:1433–59. https://doi.org/10.5705/ss.2012.317.
[37] Arimura H, Soufi M, Ninomiya K, Kamezawa H, Yamada M. Potentials of radiomics for cancer diagnosis and treatment in comparison with computer-aided diagnosis. Radiol Phys Technol 2018;11:365–74. https://doi.org/10.1007/s12194-018-0486-x.
[38] Peduzzi P, Concato J, Kemper E, Holford TR, Feinstem AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373–9. https://doi.org/10.1016/S0895-4356(96)00236-3.
[39] Jain A, Rpw D, Mao J. Statistical pattern recognition: a review. IEEE Trans Pattern Anal Mach Intell 2000;22:4–37.
[40] Ninomiya K, Arimura H, Sasahara M, Kai Y, Hirose T, Ohga S. Feasibility of anatomical feature points for the estimation of prostate locations in the Bayesian delineation frameworks for prostate cancer radiotherapy. Radiol Phys Technol 2018;11:434–44. https://doi.org/10.1007/s12194-018-0481-2.
[41] Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate data analysis. J Pharm 2010. https://doi.org/10.1016/j.ijpharm.2011.02.019.
[42] Pavic M, Bogowicz M, Würms X, Glatz S, Finazzi T, Riesterer O, et al. Influence of inter-observer delineation variability on radiomics stability in different tumor sites. Acta Oncol (Madr) 2018;57:1070–4. https://doi.org/10.1080/0284186X.2018.1445283.
[43] Alexnet (Pretrained AlexNet convolutional neural network). Mathworks. MATLAB. https://jp.mathworks.com/help/deeplearning/ref/alexnet.html?lang=en; 2018 [accessed 18.11.30].
[44] ImageNet. http://www.image-net.org; 2018 [accessed 18.11.30].
[45] Kingma D, Ba J Adam. a method for stochastic optimization. Proc Int Conf Learn Represent. 2014:1–13.
[46] ImageDataAugmenter (Configure image data augmentation). Mathworks. MATLAB. https://jp.mathworks.com/help/deeplearning/ref/imagedataaugmenter.html?lang=en; 2018 [accessed 18.11.30].
[47] Survival v2.43-3. https://www.rdocumentation.org/packages/survival; 2018 [accessed 18.11.30].
[48] Dehing-Oberije C, De Ruysscher D, van der Weide H, Hochstenbag M, Bootsma G, Geraedts W, et al. Tumor volume combined with number of positive lymph node stations is a more important prognostic factor than TNM stage for survival of non-small-cell lung cancer patients treated with (chemo)radiotherapy. Int J Radiat Oncol Biol Phys 2008;70:1039–44. https://doi.org/10.1016/j.ijrobp.2007.07.2323.
[49] Koo RR, Moon HH, Lim JJ, Kim YY, Kim Y, Kim HH, et al. The effect of tumor volume and its change on survival in stage III non-small cell lung cancer treated with definitive concurrent chemoradiotherapy. Radiat Oncol 2014;9:1–8. https://doi.org/10.1186/s13014-014-0283-6.
[50] Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 2012;366:883–92. https://doi.org/10.1016/j.eururo.2013.04.025.
[51] Mackin D, Fave X, Zhang L, Fried D, Yang J, Brian T, et al. Measuring CT scanner variability of radiomics features. Invest Radiol 2015;50:757–65. https://doi.org/10.1097/RLI.0000000000000180.Measuring.
[52] Li Y, Lu L, Xiao M, Dercle L, Huang Y, Zhang Z, et al. CT slice thickness and convolution kernel affect performance of a radiomic model for predicting EGFR status in non-small cell lung cancer: a preliminary study. Sci Rep 2018;8:1–10. https://doi.org/10.1038/s41598-018-36421-0.