의학석사 학위논문

교모세포종으로 새롭게 진단받은 환자에서 임상적 요인, 유전체학 정보 및 자기공명영상의학적 소견의 라디오믹스 조합을 이용한 노모그램

Revised nomogram combining clinical parameters, genomics, and advanced MR imaging radiomics in patients with newly diagnosed glioblastoma

울산대학교 대학원
의학과
김민재

교모세포종으로 새롭게 진단받은 환자에서 임상적 요인, 유전체학 정보 및 자기공명영상의학적 소견의 라디오믹스 조합을 이용한 노모그램

지도교수 김호성

이 논문을 의학석사학위 논문으로 제출함

2018 년 12 월

울산대학교 대학원
의학과
김민재

김민재의 의학석사학위 논문을 인준함

심사위원 이 승 수 인
심사위원 김 호 성 인
심사위원 이 상 민 인

울산대학교 대학원
2018 년 12 월

영문요약

Revised nomogram combining clinical parameters, genomics, and advanced MR imaging radiomics in patients with newly diagnosed glioblastoma

Author: Minjae Kim, MD. Department of Radiology and the Research Institute of Radiology, University of Ulsan College of Medicine, Asan Medical Center.

Background: To develop and validate a nomogram combining multiparametric MRI radiomic score and clinical predictors including treatment options for individualized prognostication in patients with IDH-wildtype glioblastoma.

Methods: The prognostication model was developed in 158 patients with IDH-wildtype glioblastoma from March 2012 to November 2016. A total of 6472 radiomic features were extracted from contrast-enhanced T1-weighted imaging, fluid-attenuated inversion recovery, diffusion-weighted imaging, and dynamic susceptibility contrast imaging. After radiomic feature selection using LASSO regression, individualized radiomic score was calculated. The nomogram was built incorporating radiomic score, O6-methylguanine-DNA-methyltransferase (MGMT) gene methylation status, and clinical predictors. The model performance was assessed using the C-index and integrated Brier score (IBS) and calibrated. The model was externally validation in 58 patients with different imaging scheme.

Results: The significant clinical predictors were age, KPS, MGMT methylation status, extent of surgery, history of concurrent concurrent chemoradiation therapy, and history of adjuvant temozolomide use. A radiomic score using 6 selected MR features significantly discriminated overall survival in the training (C-index, 0.691; IBS, 0.169) and the validation set (C-index, 0.617; IBS, 0.196). The combined nomogram significantly improved prognostication compared to the radiomic score (P =.002), or baseline clinical predictors and MGMT (P = .035). The nomogram showed good discrimination in both training (C-index, 0.78; IBS, 0.149) and validation (C-index, 0.68; IBS, 0.158) sets with good calibration.

Conclusions: This glioblastoma nomogram including multiparametric MRI radiomics and treatment options enables individualized prognostication and improves prognostication compared to established clinical models.

Keywords: glioblastoma; radiomics; magnetic resonance imaging; nomogram

목차

표 및 그림 차례

서론

Glioblastoma, especially IDH-wildtype glioblastoma, is notorious for its aggressiveness with a median survival period of only 14 – 16 months, and a 2-year survival rate of 26 – 33% [1], despite the standard treatment of surgery with concurrent chemoradiation (CCRT) following adjuvant temozolomide (TMZ) therapy. Immunohistochemistry and genomic sequencing analysis are regarded as gold standard methods for the identification of specific genetic mutations in patients with glioblastoma which aids to predict clinical outcome [2,3]. However, they are obtained from an invasive procedure and a single biopsy may lead to an incorrect result due to intra-tumoral heterogeneity [4]. This intra-tumoral genetic heterogeneity reduces the value of invasive tissue-based genomic analysis, but offers opportunities to medical imaging modalities that can depict the entire tumor in a noninvasive and repeatable way. Magnetic resonance imaging (MRI) is commonly used for the pre-operative work up of patients with glioblastoma for guidance of surgical resection and decisions regarding treatment strategies. A recently introduced radiomics approach extracts high-dimensional features from MRI using an automated data-mining algorithm [5,6] and has shown great promise in surrogating the intra-tumoral heterogeneity of genetic features [3] and in predicting prognosis [7,8].

So far, most of radiomics studies have utilized standard of care imaging, so-called conventional MRI including T1- or T2-weighted imaging, fluid-attenuated inversion recovery (FLAIR), and T1- or FLAIR- contrast enhanced imaging. However, the MRI signal itself greatly varies according to imaging protocol used. As evidence, published radiomic research currently available is derived from single-center studies where internal validation and generalizability has been limited. On the other hand, advanced MRI techniques, including apparent diffusion coefficient (ADC) from diffusion-weighted imaging or cerebral blood volume (CBV) from dynamic susceptibility contrast imaging are relatively underutilized for radiomic

analysis, even though they have demonstrated tumor aggressiveness [9-11] and regional heterogeneity and their associated prognostic relevance [12,13] in glioblastoma. Since quantitative maps obtained from advanced MRI represent normalized value, they would be suitable tools for radiomic analysis at a multicenter level.

A nomogram is a tool that incorporates important variables to provide an easy-to-use application, centered on individualized prediction of risk. High-throughput features from radiomic analysis can be incorporated into a nomogram, using selected individualized variables derived from the pre-operative multiparametric MRI studies. Also, combining treatment options into a nomogram may become a useful tool to assess post-operative prognostication in patients who subsequently undergo standard treatment. Therefore, the aim of this study was to develop and validate a nomogram combining multiparametric MRI radiomic score and clinical predictors including treatment options for individualized prognostication in patients with IDH-wildtype glioblastoma.

연구대상 및 연구방법

1. Patient cohort

Our institutional (Asan Medical Center, Seoul, Korea) review board approved this retrospective study, and the requirement for informed consent was waived. We searched the electronic database of the Department of Radiology at our tertiary hospital and retrospectively reviewed patient records between March 2012 and March 2016. We identified 248 consecutive patients with pathologically confirmed IDH-wildtype glioblastoma, according to the 2016 World Health Organization Classification of Tumors of the Central Nervous System.[14] All patients underwent pretreatment multiparametric MRI including contrast-enhanced T1-weighted imaging (CE-T1), fluid-attenuated inversion recovery (FLAIR), diffusion-weighted imaging (DWI), and dynamic susceptibility contrast (DSC) imaging. Patients were excluded if they had a prior history of surgical treatment (n = 20), had insufficient clinical information (n = 40), or any of multi-parametric imaging data was missing (n = 25) or unreadable (because of an artifact) (n = 5). These steps yielded 158 consecutive patients (mean age, 59.5 years old; male - female ratio, 96:62). This cohort was used as a training set to develop a radiomics model for prognostication in patients with glioblastoma. Identical inclusion criteria were used to identify 58 novel patients with glioblastoma treated at another tertiary center (Seoul National University Hospital, Seoul, Korea) between October 2014 and November 2016 and was used for external validation of the model. The patient inclusion process is shown in Figure 1.

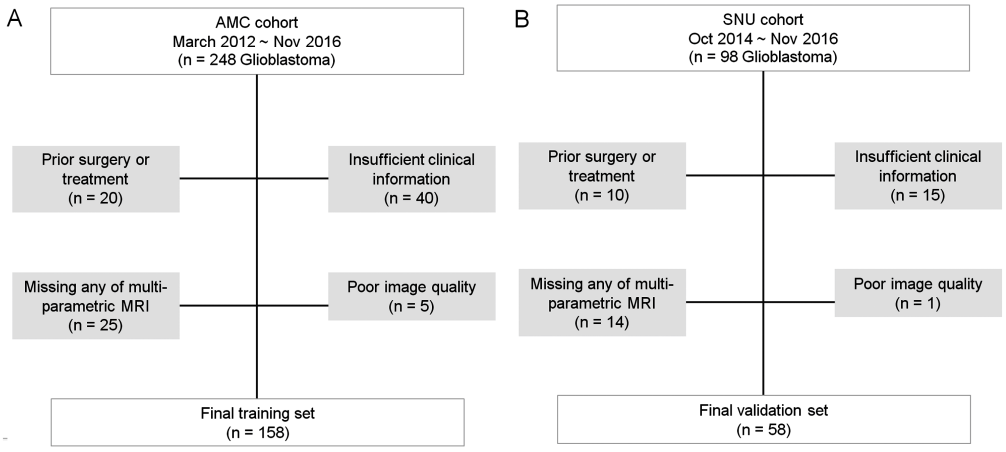


Figure 1. Flow chart of the inclusion criteria. (A) training set and (B) validation set. Abbreviations: AMC, Asan Medical Center; SNU, Seoul National University Hospital; MRI, magnetic resonance imaging.

2. Outcome definition

The primary endpoint of the study was overall survival (OS). Patients who were alive at the time of analysis (n = 36, 22.8 % in the training set and n = 8, 13.9 % in the validation set) were right-censored data and included in the analysis. All patients were followed up every 3 – 6 months after surgical treatment. The minimum follow up time to ascertain survival was 1.8 year.

3. Imaging data acquisition

In the training set, all MRI studies were performed on the same 3-T unit (Achieva; Philips Medical Systems, Best, The Netherlands), using an eight-channel head coil. The brain-tumor imaging protocol includes the following sequences: T2-weighted imaging, FLAIR imaging, T1-weighted imaging, DWI, CE-T1WI, and DSC perfusion MRI. DWI was acquired in three orthogonal directions, and the images were combined into a trace image. DWI was obtained using the following parameters: repetition time (TR)/echo time (TE), 3000/56 ms; diffusion gradient encoding, b = 0, 1000 s/mm2; field of view (FOV), 25 cm; slice thickness/gap, 5 mm/2 mm; matrix, 256 × 256; and acquisition time, 39 s. A contrast-enhanced high-resolution anatomical three-dimensional (3D) volume image was obtained using a gradient-echo T1-weighted sequence with the following parameters: TR/TE, 9.8/4.6 ms; flip angle, 10°; FOV, 256 mm; matrix, 512 × 512; and slice thickness, 1 mm with no gap. DSC perfusion MRI was performed using a gradient-echo, echo-planar sequence during the administration of a standard dose of 0.1 mmol/kg gadoterate meglumine (Dotarem; Guerbet, Paris, France) at a rate of 4 mL/s using a MRI-compatible power injector (Spectris; Medrad, Pittsburgh, PA, USA). The bolus of contrast material was followed by a 20 mL bolus of saline administered at the same injection rate. The parameters for DSC MRI were as follows: TR/TE, 1808/40 msec; flip angle, 35°; FOV, 24 cm; slice thickness/gap, 5 mm/2 mm; and matrix, 128 × 128. The total acquisition time for DSC MRI was 1 min 54 s.

The brain tumor imaging protocol in the validation set was performed on the 3-T system scanner, included the sequences of T2WI, FLAIR, T1WI, DWI, CE-T1WI, and DSC perfusion imaging. The acquisition protocol used is shown in the Supplementary Information.

4. Imaging post-processing and segmentation

The apparent diffusion coefficient map was calculated using the b values of 0 and 1000 s/mm^2, using a two-point estimate of signal decay: ADC = -ln ($S[b]/S[0]$)/b, where b indicates the b value and S(0) and S(b) are the signal intensities of images with b values at 0 and 1000, respectively. The post-processing of DSC imaging was performed using commercial software (NordicICE; NordicNeuroLab, Bergen, Norway). After correction for contrast agent leakage, the whole-brain relative CBV was calculated using the numerical integration of the time concentration curve. Next, we normalized the relative CBV (nCBV) images with the mean intensity of the contralateral normal-appearing cerebral white matter at the centrum semiovale, which was manually selected by a researcher (Y.H.J., with 2 years' experience in neuroimaging processing). The diameter of the selected region of interest (ROI) was 4 mm. The nCBV maps were created by dividing each CBV value by the contralateral ROI on a pixel-by-pixel basis.

Calculated ADC and nCBV maps were then co-registered to the 3D contrast-enhanced T1-weighted image using SPM software *(www.fil.ion.ucl.ac.uk/spm/)*. The co-registration process includes the generation of a brain mask from a 3-dimensional (3D) CE-T1WI and transformation to ADC and nCBV maps for each patient. Images were registered on the brain-extracted 3D CE-T1WI volume using affine transformation with normalized mutual information as a cost function [15], with 12 degrees of freedom and tri-linear interpolation.

For 3D CE-T1WI and FLAIR data, signal intensity normalization was used to remove noise and reduce variance of the T1-based signal intensity of the brain. We applied the hybrid white-stripe method [16] for intensity

normalization using the ANTsR and WhiteStripe packages [17,18] in the R software package (R Foundation for Statistical Computing, Vienna, Austria, URL: *http://www.R-project.org*, 2016). This incorporates processes of the statistical principles of image normalization, preserving ranks among tissue and matching the intensity of tissues without upsetting the natural balance of the tissue intensities [18]. Before feature extraction, we excluded outliers from the image intensities of ADC and nCBV maps by excluding ± 3 standard deviation inside the ROI [19].

Segmentation was performed in the enhancing tumor region by a neuroradiologist (with 4 years of experience in neuro-oncological imaging) who semi-automatically defined on the 3D CE-T1WI using a segmentation threshold and region-growing segmentation algorithm that was implemented using software (MITK, *www.mitk.org,* German Cancer Research Center, Heidelberg, Germany) [16]. All segmented images were validated by an experienced neuroradiologist (with 18 years of experience in neuro-oncologic imaging). Finally, we resampled the images into a uniform voxel size of 1 × 1 × 1 mm across all images.

5. Radiomic feature extraction

The accrual process used for developing the model is summarized in the Supplementary Figure S1. Radiomic features were extracted using Matlab R2014b (The Mathworks, Natick, MA), in accordance with previous studies [8,20]. Briefly, the radiomic features consisted of four feature groups: 7 volume and shape features, 17 first-order features, 162 texture features, and 1432 wavelet features. The volume and the shape features were obtained from the segmented mask and the first-order, texture, and wavelet features were estimated using signal intensity. Volume and shape features were used to describe the 3D geometric properties of the tumor, which included compactness, spherical disproportion, sphericity, surface area, and volume. The first-order features were derived from the intensity histogram using first-order statistics, including intensity range, energy, entropy, kurtosis,

maximum, mean, median, uniformity, and variance. Textural features were obtained from a gray-level co-occurrence matrix (GLCM) and gray-level run-length matrix (GLRLM) [21] using 3D analyses of the tumor ROI in 13 directions of the 3D space. For GLCM analyses, texture features were computed for varying distances of 1, 2, 3 voxels in 13 directions. Wavelet transformation was applied with a single-level directional discrete wavelet transform of a high-pass and low-pass filter [22]. In total, eight wavelet-decomposition images were generated from each MRI imaging sequence input: HHH, HHL, HLH, HLL, LHH, LHL, LLH, and LLL images, where 'H' was designated as a high-pass filter and 'L' is a low-pass filter. The first-order features and texture features were then applied to the wavelet-transformed images, (17 first-order features + 162 texture features) and multiplied by 8 images, to yield 1432 wavelet features. Thus, for each patient, 1618 radiomic features were extracted from the T1CE, FLAIR, ADC and CBV data, respectively, which resulted in a total 6472 extracted features. Finally, all radiomic features were z transformed for group comparison. The processing time to extract 6472 features was approximately 8 min per patient. Details of radiomic feature extraction are provided in the Supplementary Information.

6. Radiomoc feature selection

We used L_1-penalized estimation for Cox regression and the least absolute shrinkage and selection operator (LASSO) method to select the radiomic features that are the most important for prognosis [23,24]. Briefly, the LASSO is a data analysis method that selects features by fitting a Cox regression model via penalized maximum likelihood estimation. The regularization path is computed for the tuning parameter (λ), which is often chosen so as to minimize out-of-sample prediction errors. The non-zero coefficients of the covariates are selected by LASSO, while most of the covariates are reduced to zero. The radiomics features were selected based on the prediction-optimal penalty λ that contains the relevant features with

high probability. The R software and “glmnet” package were used for the LASSO Cox regression model analysis.

7. Imaging predictor using radiomic prognostic score

Using the non-zero coefficients of radiomic features, an individualized radiomic prognostic score was developed. The score is calculated as the sum of each radiomic feature multiplied by a non-zero coefficient from LASSO according to the equation below.

Radiomic prognostic score = coefficient of the 1st feature × value of the 1st feature

+ coefficient of the 2nd feature × value of the 2nd feature

+ coefficient of the 3rd feature × value of the 3rd feature

+ …+ coefficient of the nth feature × value of the nth feature

The radiomic prognostic score was calculated separately using both conventional radiomic and multiparametric radiomic features obtained from conventional MRI, ADC, and CBV. Next, the radiomic prognostic score was used as an imaging predictor for the prognostication model.

8. Clinical predictors

The clinical and molecular characteristics of all patients had been retrospectively assessed by a researcher (M.S.L.) who reviewed the medical records but was blinded to the patient outcome. Patient specific preoperative clinical variables included sex, age at diagnosis, KPS, tumor location, and tumor volume (cm^3). Binary category was applied for KPS score; 1) patients with KPS score same or above 70 and 2) whose KPS score below 70. The treatment-related variables of surgical extent (gross total removal, subtotal removal, and biopsy), history of concurrent chemoradiation, and history of adjuvant TMZ were collected. Also, O6-methylguanine-DNA methyltransferase (MGMT) promoter methylation status was also collected.

9. Statistical analysis

Frequencies and proportions were reported for categorical variables, and the mean and standard deviation were calculated for continuous variables. Differences in categorical and continuous variables were assessed using the chi-square test and independent *t*-test, respectively. The median follow-up time was reported only for censored patients [25] and was determined by considering the time of surgery and time to last available contact.

For clinical predictors, univariate Cox proportional hazard regression analysis was used to test the association between OS and clinical predictors including sex, age at diagnosis, KPS, tumor location, and tumor volume. A *P* value less than 0.05 was considered statistically significant.

The performance of the prognostic model combining imaging and clinical predictors was measured with Harrell's concordance probability index (C-index) and the Integrated Brier Score (IBS). Briefly, the IBS calculates prediction errors over time; the predicted probability of a model can range from 0 (a perfect model) to 0.25 (a non-informative model). The prediction error curves were calculated using bootstrapping method ("pec" package in R used with the "Boot632plus" module) [26] to obtain relatively unbiased estimates of the model's performance. The time for calculating prediction error was the median survival time of the cohort. To compare the C-index across different models, the "compareC" package in R was used.

Finally, a nomogram was built with the significant associative features identified including the radiomic prognostic score and clinical predictors. The nomogram performance was quantified with respect to discrimination and calibration. To develop a well-calibrated and exportable nomogram for newly diagnosed glioblastoma, we built a model using a training cohort and validated it with external validation cohort from another institution. Calibration plots assessed the overall extent of over- or underestimation of OS compared with the nomogram-predicted probability of OS, and calibration

was tested using the D'Agostino-Nam version of the Hosmer-Lemeshow test [27].

The model development and validation methods in our study were adhered to the Transparent Reporting of a multivariable prediction model for the Individual Prognosis or Diagnosis (TRIPOD) statement [28]. Statistical analyses were performed using statistical software (R version 3.3.3, R Core Team, Vienna, Austria).

연구결과

The clinical characteristics of the training and validation cohorts are summarized in Table 1. No differences were found between the training and validation cohorts in terms of sex, age at diagnosis, KPS, tumor location, or tumor volume. The MGMT promoter methylation status was positive in 32.4% and 59.6% of the training and validation sets, and was negative in 67.6% and 40.4% of the training and validation sets among those patients with MGMT status available. The median follow up time was 2.86 years (median) in the training set and 4.47 years in the validation set, which was slightly longer in the validation set. The median survival was 646 days in the training set and 700 days in the validation set.

1. Significant clinical predictors

Among the clinical predictors, older age (P = .039), KPS at treatment initiation (P = .043), less surgical extent (P = .0004), no treatment with standard CCRT (P < .0001), and no treatment with adjuvant TMZ (P <.0001) were significant clinical predictors for shorter survival (Table 2). The performance of baseline clinical predictors before any treatment including age, KPS, volume had a C-index of 0.61 (95% CI, 0.58–0.64) and an IBS of 0.220 in the training set and 0.631 (95% CI, 0.59–0.65) and 0.160, respectively in the validation set..

2. Significant radiomic features and individualized radiomic score

A total of 6472 features extracted from multi-parametric MRI data (each 1618 features from T1CE, FLAIR, ADC, and CBV, respectively) were analyzed to identify significant radiomic features. The 6 significant multiparametric MRI radiomic features were selected using LASSO penalization in the training set (Supplementary Figure S2). From conventional

Table 1. Clinical characteristics of the study patients

Parameter	Training set (n = 158)	External validation set (n = 58)	*P value*
Sex, n			0.52
Male/Female	96/62	38/20	
Age, years			0.27
Median (range)	59.5 (31–83)	57.6 (20–80)	
Primary treatment, n [%]			
Extent of resection			0.12
Gross-total resection	72 (45.6%)	34 (58.6%)	
Subtotal resection	57 (36.1%)	19 (32.8%)	
Biopsy	29 (18.4%)	5 (8.6%)	
Adjuvant treatment			
RT + TMZ	141 (89.2%)	58 (100%)	0.07
Other			
RT only	1 (0.6%)	0	
TMZ only	4 (2.5%)	0	
No RT or TMZ	12 (7.6%)	0	
Location			0.62
Frontal or temporal	73 (46.2%)	29 (50%)	
Others	85 (53.8%)	29 (50%)	
KPS at treatment initiation, n (%)			0.64
≥ 70	138 (87.3%)	52 (89.7%)	
< 70	20 (12.6%)	6 (10.3%)	
MGMT promoter status, n (%)			
Methylated	12 (7.6%)	28 (48.3%)	.13
Unmethylated	25 (15.8%)	19 (32.7%)	
NA	120 (75.9%)	11 (19.0 %)	NA
Median follow-up time, years range)	2.86 (1.06–5.67)	4.47 (3.44–6.18)	.047

Abbreviation: KPS, Karnofsky performance score; CCRT, concurrent chemoradiation therapy; RT, radiation therapy; TMZ, temozolomide; MGMT, O6-methylguanine-DNA-methyltransferase gene methylation status ;NA, information not available

Table 2. Selection of clinical predictors from the training set using a univariate Cox hazard regression model.

Predictors	Hazard ratio	95% Confidence interval	*P*-value
Age	1.02	1.01–1.04	**0.039**
Sex	0.88	0.61–1.30	0.543
KPS at treatment initiation	1.70	1.02–2.86	**0.043**
Location	0.77	0.52–1.12	0.167
Volume	1.00	0.99–1.00	0.767
Surgical extent	0.67	0.53–0.84	**0.0004**
Treated with CCRT	0.23	0.13–0.40	**<0.0001**
Adjuvant TMZ	0.12	0.07–0.21	**<0.0001**

Abbreviation: KPS, Karnofsky performance score; CCRT, concurrent chemoradiation therapy; TMZ, temozolomide. KPS score was binary with 1) score same or above 70 and 2) below 70.

MRI, 1 first order feature (mean absolute deviation), 1 GLCM feature (mean of sum entropy), and 1 GLRLM feature (standard deviation of run emphasis) were selected. From DWI imaging, 1 first order feature (skewness) was selected. From perfusion MRI, 1 GLCM (standard deviation of entropy) and 1 GLRLM (mean of long run emphasis) were selected. Subsequently, the individualized radiomic prognostic score was calculated using the corresponding coefficients of each feature according to the equation described below.

Radiomic prognostic score = -0.07896580× [T1CE_ Sum entropy (mean) LLH GLCM dist=3]
-0.06340327 ×[FLAIR_ Mean absolute deviation LHH first order]
-0.09125977×[FLAIR_ High gray-level run emphasis (std)]
- 0.05745977× [ADC_Skewness HHH first order]
+ 0.03145506 × [CBV_Entropy (std) HHL GLCM dist = 1]
-0.08185888 × [CBV_ Long run high gray-level emphasis (mean) HHH GLRLM]

The individualized, radiomic prognostic score in the training set is shown in Fig. 2. The performance of the radiomic prognostic score was assessed using the C-index and IBS in both the training and validation sets. The performance of the prognostic score was calculated as having a C-index of 0.69 (95% confidence interval [CI], 0.66–0.72) and an IBS of 0.169 in the training set and a C-index of 0.62 (95% CI, 0.57–0.66) and IBS of 0.186 in the validation set, respectively.

As a comparison, the radiomic prognostic score was calculated using conventional MRI only. The performance of the individual radiomic prognostic score using conventional MRI had a C-index of 0.65 (95% CI, 0.60–0.69), which was lower than that of multiparametric MRI. In separate analysis according to each MRI, the conventional MR radiomic features

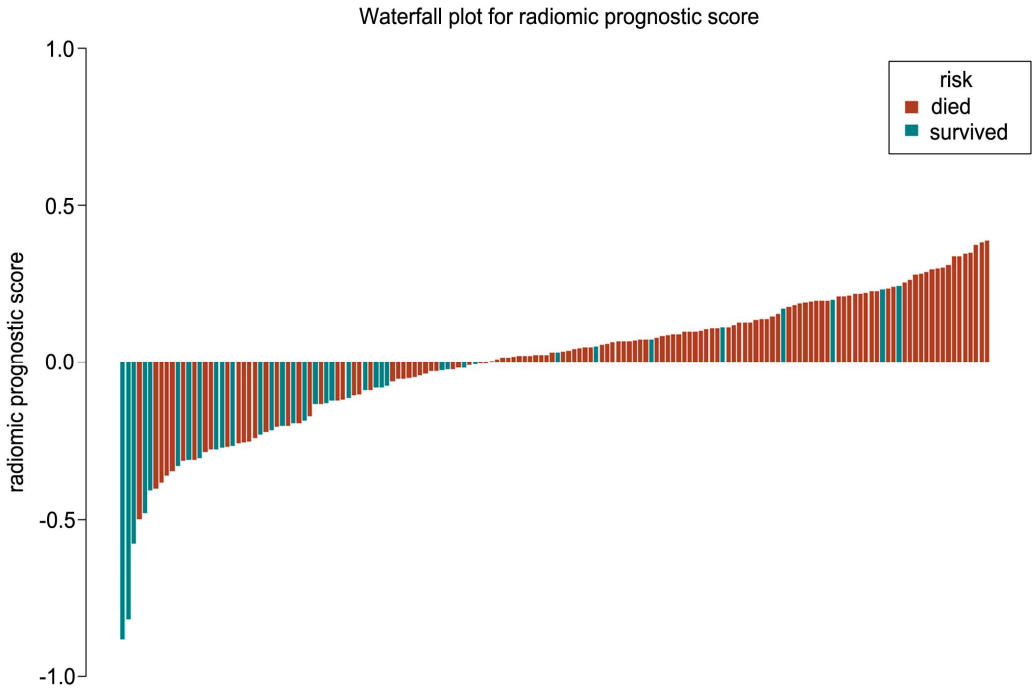


Figure 2. Radiomics prognostication score for each patient in the training set (n = 158). Green bars show scores for patients who survived and were right censored, while red bars show scores of those who died.

showed the highest performance in the training set (C-index 0.65), while the CBV radiomic features showed the highest performance in the validation set (C-index 0.61). The separate performance according to each MRI method is summarized in the Supplementary Information.

3. Building individualized prognostication model

On the basis of the radiomic prognostication score and significant clinical predictors, we constructed a nomogram to predict 1-year and 2-year survival (Fig. 3). The 2-year survival probability was chosen to calculate prediction error and calibration based on the median survival of the study cohort. The performance of our nomogram in the training set had a C-index of 0.78 (95% CI, 0.75–0.81) and an IBS of 0.149 and showed good calibration (Supplementary Figure S3A). The performance of nomogram was significantly better at predicting OS compared to the radiomic prognostic score (C-index, 0.69 [95% CI, 0.66–0.72]; P <.0001) alone, or to baseline clinical predictors alone (C-index, 0.62 [95% CI, 0.60–0.65]; P < .0001), or to even when compared to all clinical predictors included (C-index, 0.72 [95% CI, 0.69–0.75]; P = .007). Table 3 summarizes the performance of our individualized prognostication model in the training and validation set.

4. External validation of the prognostication model

In the validation set, the discrimination was good, with a C-index of 0.675 (95% CI, 0.64–0.71) and IBS 0.158, taking into consideration the heterogeneous MRI protocols present in the external validation set. The model was better able to predict OS when compared to the radiomic prognostic score alone, or to baseline clinical predictors alone, or to all clinical predictors combined, though the differences were not statistically significant. The calibration was good with the average difference and maximal difference in the predicted and calibrated probabilities of 2-year survival being 12.4% and 15%, respectively (Supplementary Figure S3B). These results demonstrated that prognostication in an independent data set were excellent and therefore confirmed the exportability of the model.

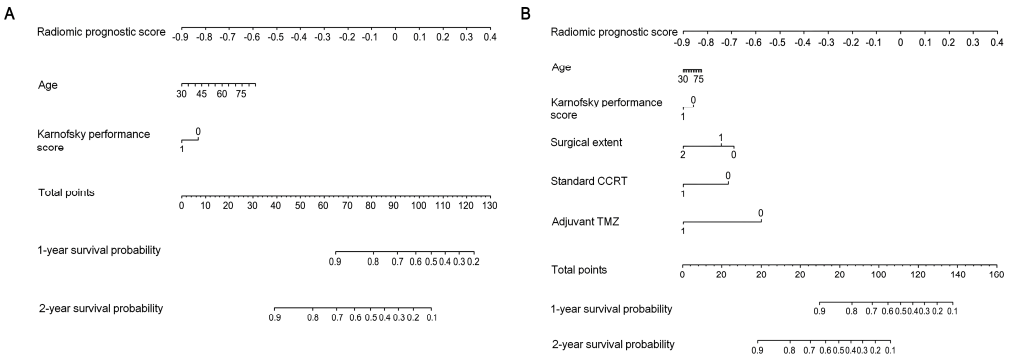


Figure 3. A nomogram predicting the probability of overall survival in patients with glioblastoma. (A) A baseline nomogram including radiomics score, age, and Karnofsky performance score. (B) A post-treatment nomogram after adding treatment including extent of surgery, concurrent chemotherapy, and adjuvant temozolomide treatment.

Table 3. Comparison of prognostic models combining multiparameteric radiomics features for predicting overall survival in the training and the validation set

Model comparison	**Combined (nomogram)**	**Single Model**			
	Radiomic score + all clinical predictors	Radiomic score	Conventional radiomic score	Baseline clinical predictors	All clinical predictors
Training set					
C-index	**0.780**	0.691	0.648	0.614	0.716
	Difference *P-value*	0.089 **<0.0001**	0.132 **<0.0001**	0.166 **<0.0001**	0.064 **0.007**
IBS	0.149	0.169	0.207	0.220	0.160
Validation set					
C-index	**0.676**	0.617	0.559	0.631	0.638
IBS	0.158	0.186	0.197	0.170	0.162

Note: baseline clinical predictors are age and Karnofsky performance score. All clinical predictors are age, Karnofsky performance score, extent of surgery, history of concurrent chemotherapy, and history of adjuvant temozolomide treatment. p-value refers to the significance in the difference of C indices between the combined model and the single model using "CompareC" in R statistical package.

Abbreviation: MR = magnetic resonance IBS = integrated brier score

5. Individual prognostication model considering MGMT methylation status

Missing data for MGMT promoter status was 75.9% (n = 120) in the training set and 19 % (n = 11) in the validation set; complete cases were performed as a subsequent subgroup analysis. The model's performance was tested in the 84 patients with available MGMT methylation status (Table 4). The nomogram showed the highest performance with a C-index of 0.73 (95% CI, 0.69–0.77) and an IBS of 0.142, compared to the performance of the radiomic prognostic score alone (C-index, 0.61 [95% CI, 0.58–0.63]; IBS 0.178, P =.002) or to the baseline clinical predictors alone including the MGMT promoter methylation status (C-index, 0.65 [95% CI, 0.62–0.68]; IBS 0.155, P = .035). Fig. 4 demonstrates the prediction error and calibration of our model. The model showed a higher performance than the model containing all the clinical predictors including TMZ treatment (C-index 0.719 [95% CI, 0.68–0.75]; IBS, 0.147), though the difference was not statistically significant. Thus, the performance of our model was robust in the presence of MGMT promoter methylation status.

Table 4. Radiomics nomogram considering MGMT-gene methylation status (n = 84).

Model comparison	**Combined (nomogram)**	**Single Model**			
	Radiomic score + all clinical predictors	Radiomic score	Conventional radiomic score	Baseline clinical predictors + MGMT	All clinical predictors
C-index	**0.729**	0.613	0.556	0.649	0.719
	Difference	0.116	0.173	0.08	0.01
	p-value	**0.002**	**<0.0001**	**0.035**	0.62
IBS	0.142	0.178	0.184	0.155	0.147

Note: baseline clinical predictors are age and Karnofsky performance score. All clinical predictors are age, Karnofsky performance score, extent of surgery, history of concurrent chemotherapy, and history of adjuvant temozolomide treatment. p-value refers to the significance in the difference of C indices between the combined model and the single model using "CompareC" in R statistical package.

Abbreviation: IBS = integrated brier score; MR = magnetic resonance.

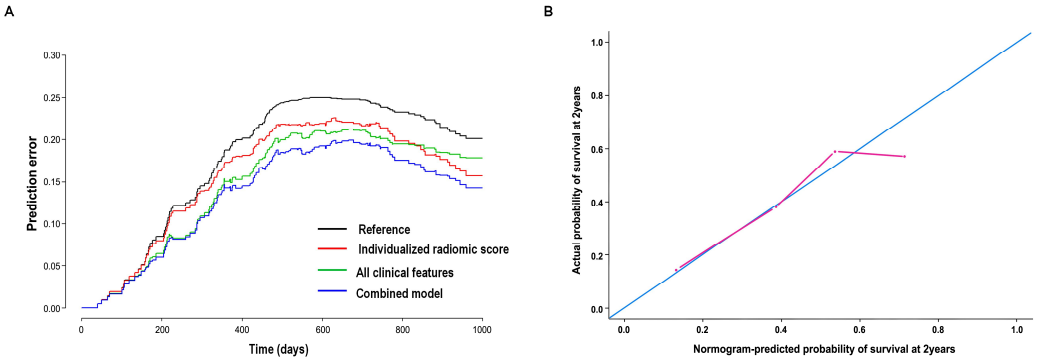


Figure 4. Discrimination and calibration of our nomogram in patients with glioblastoma with available MGMT promoter methylation status. (A) Prediction error curves of the nomogram (blue), clinical predictors (green), and MRI radiomic score (red) are shown. Lower prediction errors indicate higher model accuracy. (B) Calibration curve of the nomogram for prediction of 2-year survival.

고찰

In this study, we developed and validated a nomogram for individualized prognostication in patients with IDH-wildtype glioblastoma. The nomogram incorporates multiparametric MR radiomic score, MGMT methylation status, and important clinical predictors including age, KPS, extent of surgery, concurrent CCRT, and adjuvant TMZ use. The multiparametric MRI radiomic features provided the best performance compared to conventional MRI or single imaging of ADC or CBV alone. Incorporating this radiomic score into the nomogram improved the prognostic value over established clinical models, either when combined with baseline age and KPS as well as with MGMT status and treatment options for glioblastoma. This nomogram provides an easy-to-use prognostication model and facilitates the preoperative and postoperative personalized prediction of outcome in patients with glioblastoma validated in multi-institutional setting.

Our model was designed to use an individualized radiomic score to improve prognostication in patients with IDH-wildtype glioblastoma who will be treated with standard CCRT therapy and adjuvant chemotherapy with TMZ. Despite studies demonstrating MRI data can predict survival in patients with glioblastoma [29,30], the use of MRI in determining prognostication in the clinic is still very limited. One possible explanation is that the existing prediction model based on imaging has not demonstrated its efficacy according to treatment modality in patients who actually undergo surgery, followed by standard CCRT treatment and adjuvant TMZ therapy[31]. In this study, we combined the radiomic score and clinical predictors to identify the best outcome for IDH-wildtype glioblastoma patients. Also, this model was useful for survival prediction when tested in patients having MGMT promotor methylation status, which is a strong predictor for benefit from TMZ therapy [32,33].

The 6 multiparametric MRI radiomic features selected in this study clearly demonstrated discrimination of outcome in the both the training set and validation set. Also, the application of multiparametric MRI radiomic features showed the highest performance compared to standard care of imaging, conventional MRI. Nonetheless, the discrimination ability of the radiomic features varied depending on the imaging modality. A previous CT study [34] showed that radiomic features are reproducible over a wide range of imaging settings, unless smooth and sharp reconstruction algorithms are used. MRI based radiomic features may be more vulnerable to changes in acquisition parameters, wherein margin and signal-to-noise ratio can be easily varied across imaging protocols. Parametric maps such as ADC or normalized cerebral blood volume may be potentially robust across the different acquisition schemes from the application of the same pre-processing method, but this issue needs to be further studied.

To date, radiomic studies have rarely been validated by an external cohort and the generalizability of studies has been limited. Our model is strengthened by its validation with an external cohort having heterogeneous MR acquisition protocols and vendors. Also, the survival prediction extends to the individual patient using a radiomic prognostication score, which is fitting with the current trend of personalized medicine. Also, the nomogram method could act as a decision-making support tool before treatment using the variables age, KPS, and the radiomic prognostication score, and after treatment by adding treatment options of extent of surgery, CCRT, and adjuvant TMZ therapy.

Study limitations include the small number of patients, especially in the validation set since acquisition of DWI and DSC imaging in a single session is still not widely available. Second, important molecular changes were not considered in this analysis even though radiomic features hold promise in representing genetic heterogeneity. Third, although we tested the influence of MRI radiomics using different scanning parameters on 3.0T,

testing with the 1.5T system must be completed before radiomics can be interchangeably used as a multicenter imaging biomarker. Finally, the radiomics approach consists of a data-driven analysis, and thus the biologic meaning of the radiomic data is often unclear. This can become an obstacle in clinical practice, along with the necessary labor-intensive image processing and data analysis procedures involved. Although the averaged feature extraction time required for each patient was 8 minutes in our study, further effort to reduce time and simplify analysis will be valuable in order to be incorporated in clinic.

결론

In conclusion, a glioblastoma nomogram including multiparametric MRI radiomics and treatment options enables individualized prognostication and improves prognostication compared to established clinical models. This nomogram was validated externally, and further prospective validation of the nomogram will be required to confirm its value.

참고문헌

1. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. *New England Journal of Medicine.* 2005;352(10):987-996.
2. Parker NR, Hudson AL, Khong P, et al. Intratumoral heterogeneity identified at the epigenetic, genetic and transcriptional level in glioblastoma. *Sci Rep.* 2016;6:22477.
3. Hu LS, Ning S, Eschbacher JM, et al. Radiogenomics to characterize regional genetic heterogeneity in glioblastoma. *Neuro Oncol.* 2017;19(1):128-137.
4. Parker NR, Khong P, Parkinson JF, Howell VM, Wheeler HR. Molecular heterogeneity in glioblastoma: potential clinical implications. *Front Oncol.* 2015;5:55.
5. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014;5:4006.
6. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. *Magnetic resonance imaging.* 2012;30(9):1234-1248.
7. Zhou H, Vallieres M, Bai HX, et al. MRI features predict survival and molecular markers in diffuse lower-grade gliomas. *Neuro-oncology.* 2017.
8. Kickingereder P, Burth S, Wick A, et al. Radiomic Profiling of Glioblastoma: Identifying an Imaging Predictor of Patient Survival with Improved Performance over Established Clinical and Radiologic Risk Models. *Radiology.* 2016;280(3):880-889.
9. Choi YS, Ahn SS, Kim DW, et al. Incremental Prognostic Value of ADC Histogram Analysis over MGMT Promoter Methylation Status in Patients with Glioblastoma. *Radiology.* 2016;281(1):175-184.
10. Gupta A, Prager A, Young RJ, Shi W, Omuro AM, Graber JJ. Diffusion-weighted MR imaging and MGMT methylation status in glioblastoma: a

reappraisal of the role of preoperative quantitative ADC measurements. *AJNR Am J Neuroradiol.* 2013;34(1):E10-11.

11. Pope WB, Mirsadraei L, Lai A, et al. Differential gene expression in glioblastoma defined by ADC histogram analysis: relationship to extracellular matrix molecules and survival. *AJNR Am J Neuroradiol.* 2012;33(6):1059-1064.
12. Jain R, Poisson LM, Gutman D, et al. Outcome prediction in patients with glioblastoma by using imaging, clinical, and genomic biomarkers: focus on the nonenhancing component of the tumor. *Radiology.* 2014;272(2):484-493.
13. Lee J, Jain R, Khalil K, et al. Texture Feature Ratios from Relative CBV Maps of Perfusion MRI Are Associated with Patient Survival in Glioblastoma. *AJNR Am J Neuroradiol.* 2016;37(1):37-43.
14. Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol.* 2016;131(6):803-820.
15. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging.* 1997;16(2):187-198.
16. Nolden M, Zelzer S, Seitel A, et al. The Medical Imaging Interaction Toolkit: challenges and advances : 10 years of open-source development. *International journal of computer assisted radiology and surgery.* 2013;8(4):607-620.
17. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage.* 2011;54(3):2033-2044.
18. Shinohara RT, Sweeney EM, Goldsmith J, et al. Statistical normalization techniques for magnetic resonance imaging. *NeuroImage Clinical.* 2014;6:9-19.
19. Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition

protocols and image intensity normalization methods on texture classification. *Magn Reson Imaging.* 2004;22(1):81-91.

20. Kickingereder P, Gotz M, Muschelli J, et al. Large-scale Radiomic Profiling of Recurrent Glioblastoma Identifies an Imaging Predictor for Stratifying Anti-Angiogenic Treatment Response. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2016;22(23):5765-5771.
21. Materka A, Strzelecki M. Texture Analysis Methods–A Review.
22. Wang JZ. Wavelets and imaging informatics: A review of the literature. *Journal of Biomedical Informatics.* 2001;34(2):129-141.
23. Benner A, Zucknick M, Hielscher T, Ittrich C, Mansmann U. High-Dimensional Cox Models: The Choice of Penalty as Part of the Model Building Process. *Biometrical J.* 2010;52(1):50-69.
24. Tibshirani R. The lasso method for variable selection in the cox model. *Stat Med.* 1997;16(4):385-395.
25. Vickers AJ, Sjoberg DD, European U. Guidelines for reporting of statistics in European Urology. *Eur Urol.* 2015;67(2):181-187.
26. Mogensen UB, Ishwaran H, Gerds TA. Evaluating Random Forests for Survival Analysis Using Prediction Error Curves. *J Stat Softw.* 2012;50(11):1-23.
27. D'agostino R, Nam B-H. Evaluation of the performance of survival analysis models: discrimination and calibration measures. *Handbook of statistics.* 2003;23:1-25.
28. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med.* 2015;162(1):55-63.
29. Bosnyak E, Michelhaugh SK, Klinger NV, et al. Prognostic Molecular and Imaging Biomarkers in Primary Glioblastoma. *Clin Nucl Med.* 2017;42(5):341-347.

30. Cui Y, Tha KK, Terasaka S, et al. Prognostic Imaging Biomarkers in Glioblastoma: Development and Independent Validation on the Basis of Multiregion and Quantitative Analysis of MR Images. *Radiology.* 2016;278(2):546-553.
31. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med.* 2005;352(10):987-996.
32. Hegi ME, Diserens AC, Gorlia T, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med.* 2005;352(10):997-1003.
33. Watanabe R, Nakasu Y, Tashiro H, et al. O6-methylguanine DNA methyltransferase expression in tumor cells predicts outcome of radiotherapy plus concomitant and adjuvant temozolomide therapy in patients with primary glioblastoma. *Brain Tumor Pathol.* 2011;28(2):127-135.
34. Zhao B, Tan Y, Tsai WY, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci Rep.* 2016;6:23428.

부록

Supplementary Table 1. Comparison of MR imaging parameters of patients in training set and in external validation set of another tertiary center using the 3-T system.

Pulse Sequence	Training cohort	External validation cohort
Post-contrast T1		
Repetition Time (ms)	9.8 ± 0.2 (9.0–10.1)	9.6 ± 1 (8.6–10.4)
Echo Time (ms)	4.6 ± 0.2 (4.4–4.8)	4.2 ± 0.7 (3.5–4.7)
Pixel size (mm)	0.5	0.45 ± 0.02 (0.43–0.47)
Section thickness (mm)	0.5	1
FLAIR		
Repetition Time (ms)	9250 ± 433 (9000–10000)	8962.9 ± 77.3 (8802–9000)
Echo Time (ms)	128.8 ± 4.8 (125–135)	102.1 ± 10.5 (97–124)
Pixel size (mm)	0.76 ± 0.14 (0.65–0.94)	0.
Section thickness (mm)	4.5 ± 0.3 (4–5)	5
DWI		
Repetition Time (ms)	3750 ± 433 (3000–4000)	8430 (6900–12,000)
Echo Time (ms)	59.4 ± 2.1 (56–61.7)	66.8 ± 14.2 (55–81)
Pixel size (mm)	1.9 ± 0.04 (1.88–1.97)	0.83 ± 0.11 (0.75–0.94)
Section thickness (mm)	4.4 ± 0.5 (4–5)	3.3 ± 1.3 (3–5)
DSC		
Repetition Time (ms)	1781 ± 32.6 (1726–1800)	1500
Echo Time (ms)	40	33 ± 7 (30–40)
Pixel size (mm)	1.98 ± 0.0 (1.96–2)	1.81 ± 0.09 (1.72–1.88)
Section thickness (mm)	5	5.6 ± 0.6 (5–6)

Abbreviation: Data are expressed as mean ± standard deviation. Numbers in parenthesis are range. FLAIR = fluid-attenuated inversion recovery, DWI = diffusion weighted imaging, DSC = dynamic susceptibility contrast imaging

Supplementary Table 2. Selection of clinical predictors from the training set using a univariate Cox hazard regression model.

Predictors	Hazard ratio	95% Confidence interval	*P*-value
Age	1.02	1.01–1.04	**0.039**
Sex	0.88	0.61–1.30	0.543
KPS at treatment initiation	1.70	1.02–2.86	**0.043**
Location	0.77	0.52–1.12	0.167
Volume	1.00	0.99–1.00	0.767
Surgical extent	0.67	0.53–0.84	**0.0004**
Treated with CCRT	0.23	0.13–0.40	**<0.0001**
Adjuvant TMZ	0.12	0.07–0.21	**<0.0001**

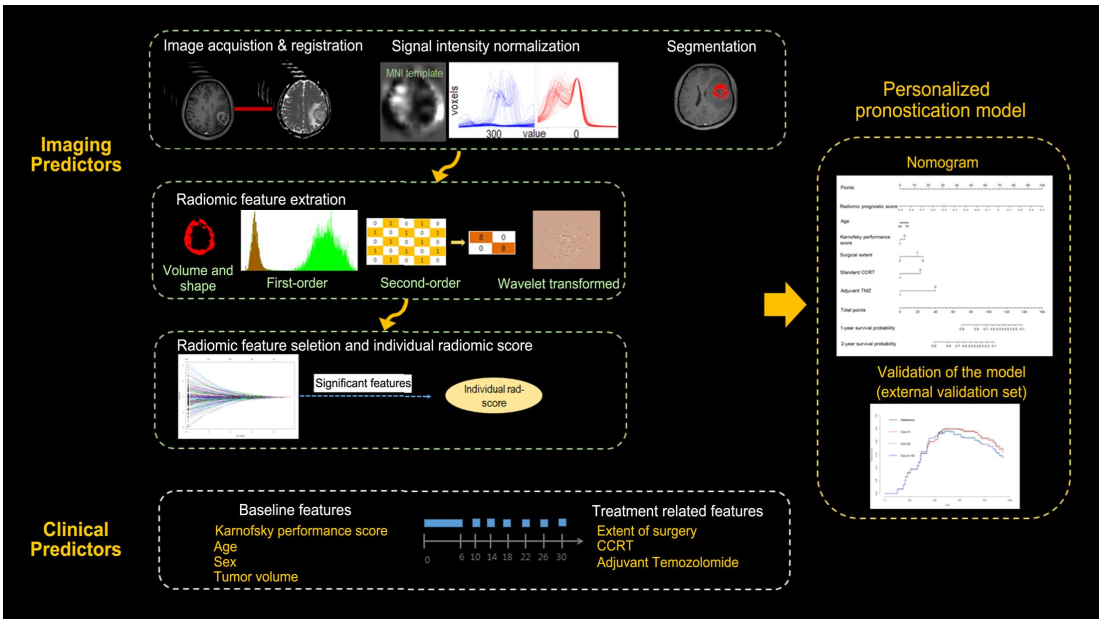
Abbreviation: KPS, Karnofsky performance score; CCRT, concurrent chemoradiation therapy; TMZ, temozolomide. KPS score was binary with 1) score same or above 70 and 2) below 70.

Supplementary Table 3. Performance of radiomic features assessed separately using MR imaging data to predict overall survival in the training and validation sets

	Multiparametric MR radiomic features	Conventional MR radiomic features	ADC radiomic features	CBV radiomic features
Number of selected features	6	3	4	8
C-index				
Training	0.691	0.654	0.641	0.643
Validation	0.617	0.544	0.568	0.609
IBS				
Training	0.169	0.189	0.194	0.190
Validation	0.186	0.204	0.197	0.192

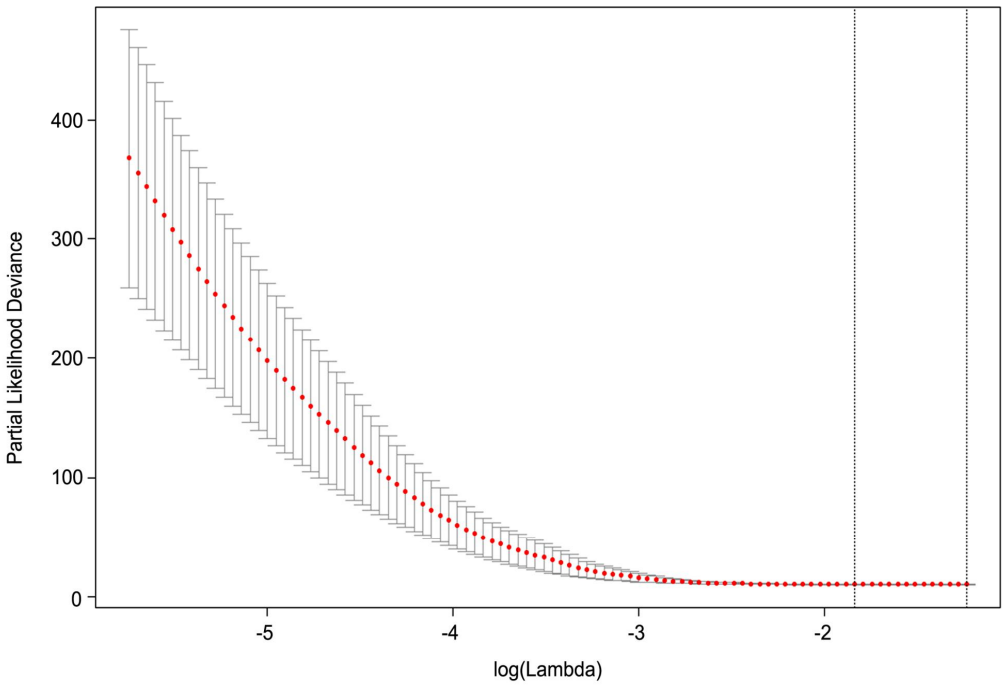
Note: C indices and IBS were calculated separately for each imaging modality with selected features using LASSO.

Abbreviations: MR = magnetic resonance; ADC = apparent diffusion coefficient; CBV = cerebral blood volume; IBS = integrated brier score

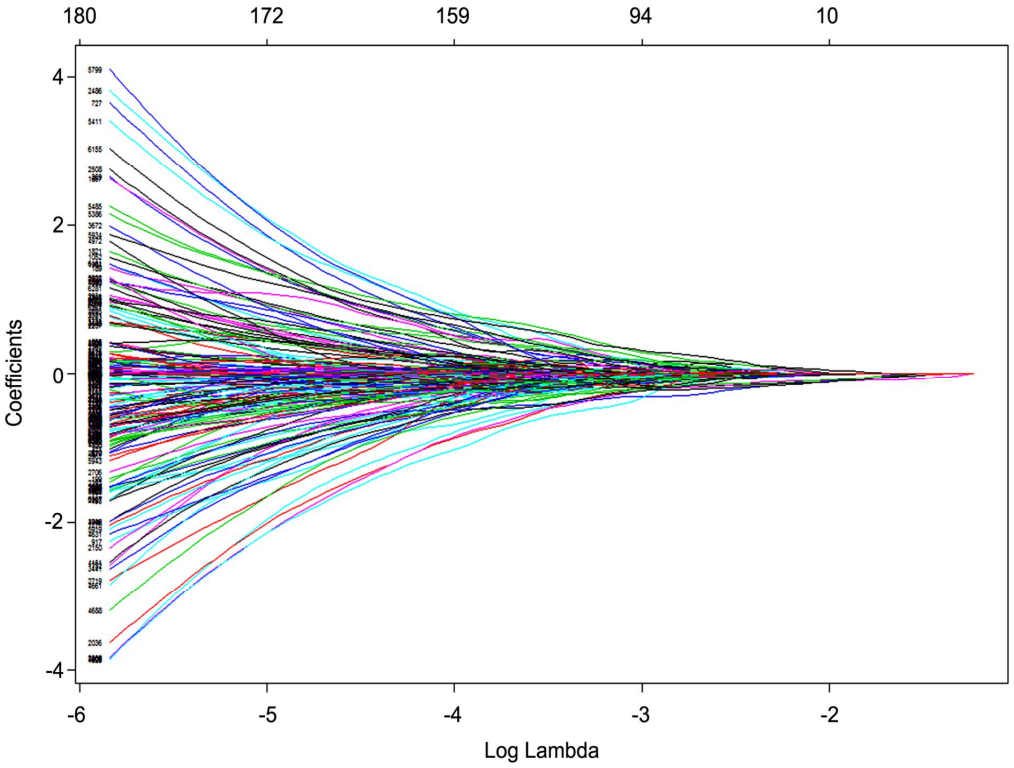


Supplementary Figure 1. Analysis pipeline of this study. Imaging analysis includes acquisition, co-registration, signal intensity normalization for conventional magnetic resonance imaging data, and segmentation. The least absolute shrinkage and selection operator method (LASSO) for Cox regression was applied to select significant radiomic features. The individualized radiomic score is calculated as the sum of each radiomic variable multiplied by a non-zero coefficient from LASSO. Subsequently, a nomogram was built with the significant associative features identified including the individual radiomic prognostic score and clinical predictors.

S2A


Partial Likelihood Deviance
400
300
200
100
0
-5
-4
-3
-2
log(Lambda)

S2B

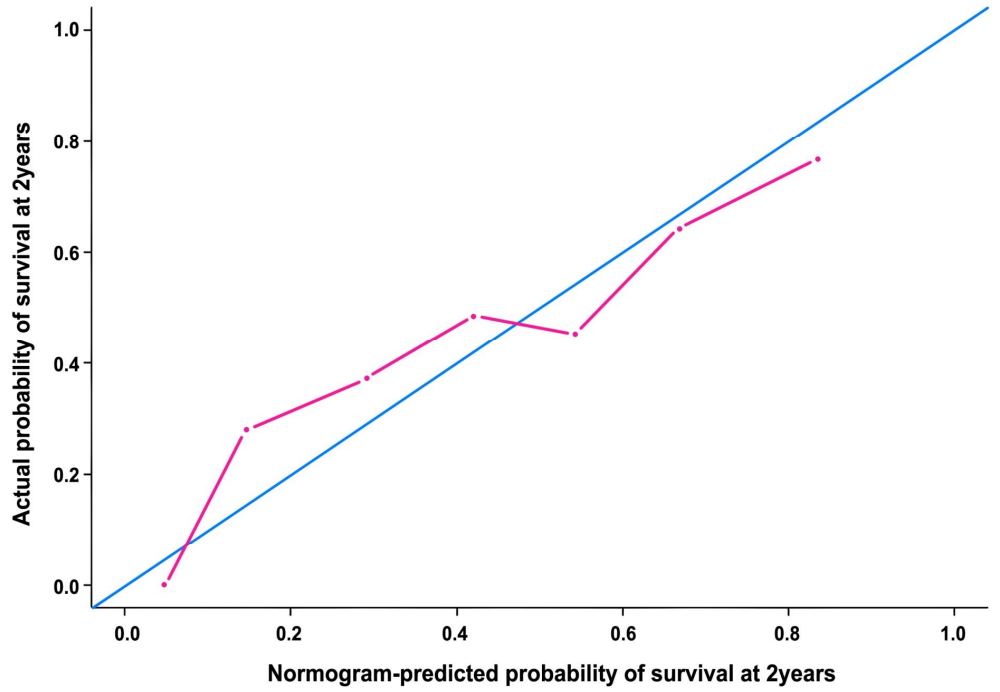

180
172
159
94
10
Coefficients
4
2
0
-2
-4
-6
-5
-4
-3
-2
Log Lambda

Supplementary Figure 2. Multiparametric magnetic resonance radiomics feature selection using the least absolute shrinkage and selection operator (LASSO) Cox regression model.

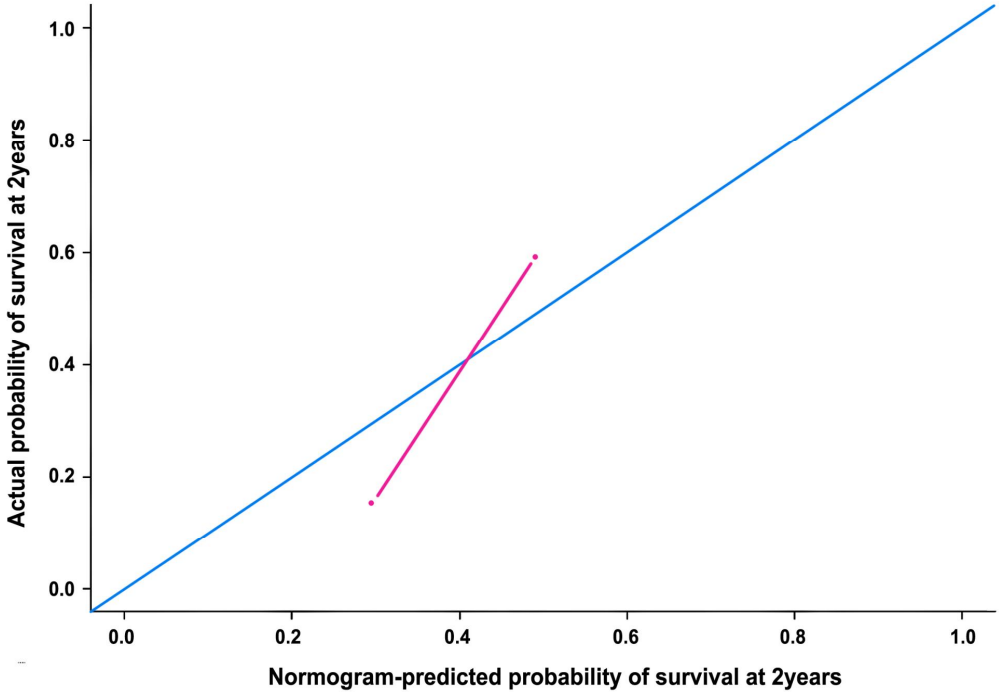
(A) Identification of the optimal regularization parameter (λ) in the LASSO model used 10-fold cross-validation. As a result, a λ value between the minimum criterion of 0.159 and one standard error of 0.292 was selected. The dotted vertical line was plotted at the value selected using 10-fold cross-validation, for which the optimal λ resulted in 6 significant non-zero coefficients.

(B) LASSO coefficient profiles using the 6472 derived multiparametric radiomics features.

S3A


Actual probability of survival at 2years
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Normogram-predicted probability of survival at 2years

S3B


Actual probability of survival at 2years
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Normogram-predicted probability of survival at 2years

Supplementary Figure 3. Calibration curves comparing the radiomics nomogram in the validation sets. The calibration curve for 2-year survival prediction shows good calibration.

Supplementary information: extraction of radiomic features

The imaging features that were calculated were divided into three groups: 17 first-order statistics, 7 volume and shape-based features, and 162 texture features. Wavelet transformation was applied to the ADC and post-contrast T1-weighted images, and eight wavelet-decomposition images were generated from each MRI image that was input. Then first-order features and texture features were calculated from the eight wavelet decomposition images, which further resulted in 1432 wavelet features ([17 + 162] × 8). In total, 1618 features (17 first-order statistics, 7 volume and shape-based features, 162 texture features, and 1432 wavelet features) were obtained from ADC images and post-contrast T1 weighted images.

1. Volume and shape features

These features compute the three-dimensional size and shape of volume of interest $V(x,y,z)$. Let Vol and A denote the volume and surface area, respectively.

Surface area	A	Compactness 2	$36\pi\frac{A^2}{Vol^3}$
Volume	$N * pixel_x * pixel_y * pixel_z$	Spherical disproportion	$\frac{A}{(6\sqrt{\pi} * Vol)^{\frac{2}{3}}}$
Surface to volume ratio	$\frac{A}{Vol}$	Sphericity	$\frac{(6\pi^2 * Vol)^{\frac{2}{3}}}{A}$
Compactness 1	$\frac{Vol}{\sqrt{\pi} * A^{\frac{2}{3}}}$		

2. First-order statistics

First-order statistics were calculated from the histogram of voxel intensities, which represents the distribution of gray values within an image. Let P denote the first-order histogram of a volume of interest V(x,y,z) with isotropic voxel size and Pi describe the number of voxels with gray level i. The number of gray-level bins set for P is represented as Ng. The ith probability vector of the first-order histogram is then defined as follows:

$$p(i) = \frac{P(i)}{\sum_{i=1}^{N_g} P(i)}$$

Let $\boldsymbol{V}_{ROI}$ and $\boldsymbol{V}$ denote the intensity values of all voxels within $V(x,y,z)$ with N voxels and the whole image, respectively. The mean and center gray values within $V(x,y,z)$ are $\overline{V}$ and $\boldsymbol{C}$, respectively.

Number of voxels	$\lvert\boldsymbol{V}_{ROI}\rvert$	Sum of intensities	$\sum_{i}^{N} \boldsymbol{V}_{ROI}(i)$
Range	$\max(\boldsymbol{V}_{ROI}) - \min(\boldsymbol{V}_{ROI})$	Energy	$N^2 \sum_{i}^{N_g} p(i)^2$
Covered imageintensity range	$\frac{\max \boldsymbol{V}_{ROI} - \min \boldsymbol{V}_{ROI}}{(\max \boldsymbol{V} - \min \boldsymbol{V} + eps)}$	Entropy	$-\sum_{i}^{N_g} [p(i) * \log_2(p(i) + eps)]$
Maximum	$\max(\boldsymbol{V}_{ROI})$	Kurtosis	$\frac{\sum_{i}^{N_g} \left[p(i) * \left(\boldsymbol{C}(i) - \overline{V}\right)^4\right]}{\left(\sum_{i}^{N_g} [p(i) * \left(\boldsymbol{C}(i) - \overline{V}\right)] + eps\right)^2}$

intensity value			
Mean intensity value	$\frac{1}{N}\sum_{i}^{N}\boldsymbol{V}_{ROI}(i)$	Skewness	$\frac{\sum_{i}^{N_g}\left[p(i)*\left(\boldsymbol{C}(i)-\overline{V}\right)^3\right]}{\left(\sum_{i}^{N_g}\left[p(i)*\left(\boldsymbol{C}(i)-\overline{V}\right)^3\right]+eps\right)^{\frac{3}{2}}}$
Median intensity value	$\mathrm{med}(\boldsymbol{V}_{ROI})$	Root means square	$\sum_{i}^{N_g}[p(i)*\boldsymbol{C}(i)^2]$
Minimum intensity value	$\min(\boldsymbol{V}_{ROI})$	Variance	$\frac{1}{N-1}\sum_{i}^{N}\left(\boldsymbol{V}_{ROI}(i)-\overline{V}\right)^2$
Mean absolute deviation	$\sum_{i}^{N_g}[p(i)*(\boldsymbol{C}(i)-\overline{V})]$	Standard deviation	$\sqrt{\frac{1}{N-1}\sum_{i}^{N}\left(\boldsymbol{V}_{ROI}(i)-\overline{V}\right)^2}$
Uniformity	$\sum_{i}^{N_g}p(i)^2$		

3. Texture features

Although first-order features provide information on the gray-level distribution of the volume of interest, they do not describe information related to the relative positions of the various gray levels of the volume of interest. The methods most often used for texture analysis are the gray level co-occurrence matrix (GLCM) and the gray level run length matrix (GLRLM). In the use of GLCM, various textural features are extracted, and GLRLM characterizes coarse textures as having many pixels in a constant gray level

run and fine textures as having few pixels in such a run. Both GLCM and GLRLM are matrix-based features as well as being constructed from 3D analysis of a volume of interest with 26-voxel connectivity, which are considered neighbors in all 13 directions in 3D.

3-1. GLCM features

Let G denote the GLCM of a quantized volume $V(x,y,z)$ with isotropic voxel size and let $G_{\alpha,\delta}(i,j)$ represent the number of times that voxels of gray level i were neighbors with voxels of gray level j in $V(x,y,z)$ in one of 13 directions of α and at a distance $\delta = 1,2,3$. GLCM is the size of $N_g \times N_g$ where N_g describes a pre-defined number of quantized gray level sets in $V(x,y,z)$. For each direction α and distance δ, the normalized GLCM is obtained as follows:

$$g_{\alpha,\delta}(i,j) = g(i,j) = \frac{G(i,j)}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} G(i,j)}.$$

Frequently used feature quantities for each direction and distance are also defined as follows:

- μ is the mean of $g(i,j)$
- σ is the standard deviation of $g(i,j)$
- $g_x(i)$ is the marginal row probability of $g(i,j)$: $g_x(i) = \sum_j^{N_g} g(i,j)$,
- μ_x is the mean of $g_x(i)$
- σ_x is the standard deviation of $g_x(i)$

Mean and standard deviation of the followings for 13 directions and 3 distances			
autocorrelation	$\sum_{i}^{N_g}\sum_{j}^{N_g} i*j*g(i,j)$	Haralick correlation	$\frac{1}{\sigma_x}\sum_{i}^{N_g}\sum_{j}^{N_g}(i*j*g(i,j)) - \mu_x$
cluster prominence	$\sum_{i}^{N_g}\sum_{j}^{N_g}(i+j-2\mu)^4 * g(i,j)$	inverse difference	$\sum_{i}^{N_g}\sum_{j}^{N_g}\frac{g(i,j)}{1+\lvert i-j\rvert}$
cluster shade	$\sum_{i}^{N_g}\sum_{j}^{N_g}(i+j-2\mu)^3 * g(i,j)$	inverse difference normalized	$\frac{1}{N_g}\sum_{i}^{N_g}\sum_{j}^{N_g}\frac{g(i,j)}{1+\lvert i-j\rvert}$
cluster tendency	$\sum_{i}^{N_g}\sum_{j}^{N_g}(i+j-2\mu)^2 * g(i,j)$	inverse difference moment	$\sum_{i}^{N_g}\sum_{j}^{N_g}\frac{g(i,j)}{1+(i-j)^2}$
contrast	$\sum_{i}^{N_g}\sum_{j}^{N_g}(i-j)^2 * g(i,j)$	inverse difference moment normalized	$\frac{1}{{N_g}^2}\sum_{i}^{N_g}\sum_{j}^{N_g}\frac{g(i,j)}{1+(i-j)^2}$
correlation	$\frac{1}{\sigma}\sum_{i}^{N_g}\sum_{j}^{N_g}(i-\mu)(j-\mu)*g(i,j)$	inverse variance	$\sum_{i}^{N_g}\sum_{j}^{N_g}\frac{g(i,j)}{(i-j)^2}$
difference average	$\sum_{k}^{N_g} k * g_{x-y}(k)$	maximum probability	$max(g(i,j))$
difference entropy	$-\sum_{k}^{N_g} g_{x-y}(k) * log_2(g_{x-y}(k)+eps)$	sum average	$\sum_{k}^{2N_g} i * g_{x+y}(k)$
difference variance	$\sum_{k}^{N_g}(k-\overline{g_{x-y}})^2 * g_{x-y}(k)$	sum entropy	$-\sum_{k}^{2N_g} g_{x+y}(k) * log_2(g_{x+y}(k)+eps)$
dissimilarity	$\sum_{i}^{N_g}\sum_{j}^{N_g}\lvert i-j\rvert * g(i,j)$	sum variance	$\sum_{k}^{2N_g}(k-\overline{g_{x+y}})^2 * g_{x+y}(k)$
energy	$\sum_{i}^{N_g}\sum_{j}^{N_g} g(i,j)^2$	variance	$\sum_{i}^{N_g}\sum_{j}^{N_g}(i-\mu)^2 g(i,j)$

entropy	$-\sum_{i}^{N_g}\sum_{j}^{N_g} g(i,j) * log_2(g(i,j) + eps)$		

3-2. GLRLM features

Let Q denote the GLRLM of a quantized volume $V(x,y,z)$ with isotropic voxel size and $Q(i,j)$ represent the number of runs of gray level i with j consecutive voxels in 1 of the 13 directions of α. GLRLM is the size of $N_g \times N_l$ where N_g describes the pre-defined number and N_l represents the length of the longest run of quantized gray level sets in $V(x,y,z)$. N_p is the number of voxels in $V(x,y,z)$.

Mean and standard deviation of the followings for 13 directions and 3 distances			
number of runs	N_{run}	low gray level run emphasis	$\frac{1}{N_{run}}\sum_{i}^{N_g}\sum_{j}^{N_l}\frac{1}{i^2} * Q(i,j)$
gray level nonuniformity	$\frac{1}{N_{run}}\sum_{i}^{N_g}\left(\sum_{j}^{N_l} Q(i,j)\right)^2$	run length nonuniformity	$\frac{1}{N_{run}}\sum_{j}^{N_l}\left(\sum_{i}^{N_g} Q(i,j)\right)^2$
high gray level run emphasis	$\frac{1}{N_{run}}\sum_{i}^{N_g}\sum_{j}^{N_l} i^2 * Q(i,j)$	run percentage	$\frac{N_{run}}{N_p}$
long run emphasis	$\frac{1}{N_{run}}\sum_{i}^{N_g}\sum_{j}^{N_l} j^2 * Q(i,j)$	short run emphasis	$\frac{1}{N_{run}}\sum_{i}^{N_g}\sum_{j}^{N_l}\frac{1}{j^2} * Q(i,j)$
long run high gray level emphasis	$\frac{1}{N_{run}}\sum_{i}^{N_g}\sum_{j}^{N_l} i^2 * j^2 * Q(i,j)$	short run high gray level emphasis	$\frac{1}{N_{run}}\sum_{i}^{N_g}\sum_{j}^{N_l}\frac{i^2}{j^2} * Q(i,j)$
long run low gray level emphasis	$\frac{1}{N_{run}}\sum_{i}^{N_g}\sum_{j}^{N_l}\frac{j^2}{i^2} * Q(i,j)$	short run low gray level emphasis	$\frac{1}{N_{run}}\sum_{i}^{N_g}\sum_{j}^{N_l}\frac{1}{i^2 * j^2} * Q(i,j)$

국문요약

연구제목: 교모세포종으로 새롭게 진단받은 환자에서 임상적 요인, 유전체학 정보 및 자기공명영상의학적 소견의 라디오믹스 조합을 이용한 노모그램

연구배경: IDH-wildtype 교포세포종으로 새롭게 진단받은 환자에서 환자의 개인 예후 예측을 위하여 임상적 요인, 유전체학 정보 및 자기공명영상의학적 소견의 라디오믹스 조합을 이용한 노모그램을 개발하고 검증하고자 한다.

연구방법: 2012년 3월부터 2016년 11월까지 IDH-wildtype 교모세포종으로 새롭게 진단받은 158명의 환자에서 예후 예측 모델을 개발하였다. 조영증강 T1 강조 영상, 액체감쇠역전회복 영상, 확산강조 영상 및 관류 영상에서 총 6472개의 라디오믹스 인자들이 추출되었다. Lasso 회귀분석을 통해 인자를 뽑아낸 후 환자 개개인의 라이오믹스 점수를 계산하였다. 환자의 라디오믹스 점수와 O6-methylguanine-DNA-methyltransferase (MGMT) 메틸화 여부 및 임상적 요인을 조합한 노모그램을 만들었다. 모델의 성능은 C-index와 integrated Brier score (IBS)를 통해 측정하였고 이후 보정하였다. 모델은 다른 영상 환경에서 촬영된 58명의 환자를 이용하여 외적 타당도 검증을 거쳤다.

연구결과: 임상적으로 중요한 인자는 환자의 나이, KPS, MGMT 메틸화 여부, 수술의 범위, 동시항암방사선치료 여부 및 추가 temozolomide 치료 여부로 확인되었다. 6개의 자기공명영상 인자를 이용한 라디오믹스 점수는 훈련 집합 (C-index, 0.691; IBS, 0.169)과 검증 집합 (C-index, 0.617; IBS, 0.196) 모두에서 전체 생존율을 유의하게 예측하는 것으로 나타났다. 통합노모그램은 라디오믹스 점수 혹은 임상적인 요인의 이용한 모델보다 환자 예후 관련 유의하게 더 높은 예측력을 보였다. 통합노모그램은 훈련 집합 (C-index, 0.78; IBS, 0.149) 및 검증 집합 (C-index, 0.68; IBS, 0.158) 모두에서 좋은 성능 및 보정도를 보였다.

결론: 교모세포종으로 새롭게 진단받은 환자에서 임상적 요인, 유전체학 정보 및 자기공명영상의학적 소견의 라디오믹스 조합을 이용한 노모그램은 환자의 예후와 관련하여 기존 임상적인 요인을 이용한 모델보다 높은 예측력을 보였다.