Radiogenomics of lower-grade gliomas: a radiomic signature as a biological surrogate for survival prediction

Zenghui Qian[1,*], Yiming Li[1,*], Zhiyan Sun[1], Xing Fan[1,2], Kaibin Xu[6], Kai Wang[4], Shaowu Li[1,3], Zhong Zhang[2], Tao Jiang[1,2,5], Xing Liu[1], Yinyan Wang[2]

[1]Beijing Neurosurgical Institute, Capital Medical University, Beijing, China
[2]Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing, China
[3]Department of Neuroradiology, Beijing Tiantan Hospital, Capital Medical University, Beijing, China
[4]Department of Nuclear Medicine, Beijing Tiantan Hospital, Capital Medical University, Beijing, China
[5]Center of Brain Tumor, Beijing Institute for Brain Disorders, Beijing, China
[6]Chinese Academy of Sciences, Institute of Automation, Beijing, China
[*]Equal contribution

Correspondence to: Tao Jiang, Xing Liu, Yinyan Wang; email: taojiang1964@163.com, 15846591696@126.com, tiantanyinyan@126.com

ABSTRACT

Objective: We aimed to identify a radiomic signature to be used as a noninvasive biomarker of prognosis in patients with lower-grade gliomas (LGGs) and to reveal underlying biological processes through comprehensive radiogenomic investigation. Methods: We extracted 55 radiomic features from T2-weighted images of 233 patients with LGGs (training cohort: n = 85; validation cohort: n = 148). Univariate Cox regression and linear risk score formula were applied to generate a radiomic-based signature. Gene ontology analysis of highly expressed genes in the high-risk score group was conducted to establish a radiogenomic map. A nomogram was constructed for individualized survival prediction.

Results: The six-feature radiomic signature stratified patients in the training cohort into low- or high-risk groups for overall survival (P = 0.0018). This result was successfully verified in the validation cohort (P = 0.0396). Radiogenomic analysis revealed that the prognostic radiomic signature was associated with hypoxia, angiogenesis, apoptosis, and cell proliferation. The nomogram resulted in high prognostic accuracy (C-index: 0.92, C-index: 0.70) and favorable calibration for individualized survival prediction in the training and validation cohorts.

Conclusions: Our results suggest a great potential for the use of radiomic signature as a biological surrogate in providing prognostic information for patients with LGGs.

INTRODUCTION

Diffuse lower-grade gliomas (LGGs) are a class of terminal central nervous system tumors, comprising WHO grades II and III astrocytomas, oligodendrogliomas, and mixed oligoastrocytomas [1, 2]. Although maximal safe resection combined with adjuvant radiotherapy and chemotherapy shows prognostic benefits [3], the overall survival of patients with LGGs remains low, ranging from a few months to more than 10 years [4, 5]. This demonstrates the need for in depth investigations of prognosis of patients with LGGs.

Developments in genomic and bioinformatic techniques allow the creation of molecular classifications and signatures based on expression profiles, and provide promising approaches to identify prognostic or therapeutic biomarkers for patient-tailored management [2, 6, 7]. In a recent study, utilization of IDH mutation and 1p/19q codeletion, LGGs could be classified into 3 distinct subgroups (IDH wild-type, IDH mutation and 1p/19q codeletion, and IDH mutation and 1p/19q non-codeletion) that capture the biologic characteristics with greater fidelity than does histological class [2]. Although such genetic characteristics can be informative and are relatively homogeneous within each tumor, there remains an unmet clinical need for less costly and less time consuming noninvasive surrogates able to determine clinical prognostic and guide individual treatment.

MRI can provide a comprehensive view of the entire tumor, and is routinely used as a noninvasive tool to support clinical decision-making, histological grading, and therapeutic monitoring [8, 9]. Radiomics, an emerging field that extracts a large number of quantitative descriptors reflecting textural and morphological variations, has been introduced to ensure more objective and precise study of oncologic tissue beyond established MRI metrics [10-12]. Consequently, clinicians increasingly rely on radiomics to assist in personalizing treatment in clinical practice, particularly in relation to tumor detection, subtype classification, and prognostic estimation [13-15]. Furthermore, linking radiomics with genomic characteristics, i.e., radio-genomics, has become an increasingly popular approach in various tumors and has expanded to create non-invasive imaging biomarkers for genomic aberrations [16-19]. As for LGGs, Mazurowski et al. [20] have done a preliminary radiogenomics study focusing on the relationship between tumor shape and molecular subtype in a single cohort. However, this study was limited by small number of available radiomic features and the lack of an external validation cohort.

In the present study, we used T2-weighted MR images from The Cancer Imaging Atlas (TCIA) to identify a radiomic-based prognostic signature, which we independently validated in the Chinese Glioma Genome Atlas (CGGA) imaging dataset. A radiogenomic map, which integrates radiomic features and genomic data, was further established to identify biological processes underlying this radiomic signature. Our results suggest that radiomics can aid in predicting survival of patients with LGGs, and reveal the prognostic role of radiomic phenotypes using comprehensive radiogenomic methods.

RESULTS

Prognostic value of the radiomic risk score

We found that radiomic features (Autocorrelation, High Gray Level Run Emphasis (HGLRE), Short Run High Gray Level Emphasis (SRHGLE), SumAverage, SumVariance, and Variance) were significantly associated with overall survival (Supplementary Table 1).

Table 1. Variables associated with overall survival in the Cox regression analysis for lower-grade glioma patients from the TCGA dataset.

	Multivariate Cox Regression		
	HR	95% CI	*P* value
Age			
>45 vs. ≤45	5.788	1.024-32.709	**0.047**
Sex			
Male vs. Female	0.500	0.148-1.691	0.265
WHO Grade			
III vs. II	22.499	1.913-264.626	**0.013**
Seizure			
Yes vs. No	0.304	0.089-1.036	0.057
IDH status			
WT vs. MUT	27.578	2.816-270.110	**0.004**
ATRX status			
WT vs. MUT	0.221	0.042-1.164	0.075
1p/19q status			
Non-codel vs. codel	2.117	0.156-28.724	0.573
Radiomic Risk score			
High vs. Low	4.347	1.055-17.922	**0.042**

MUT = mutant. WT, wild type; Non-codel = non-codeletion; Codel = codeletion; HR = hazard ratio; 95% CI = 95% confidence interval.
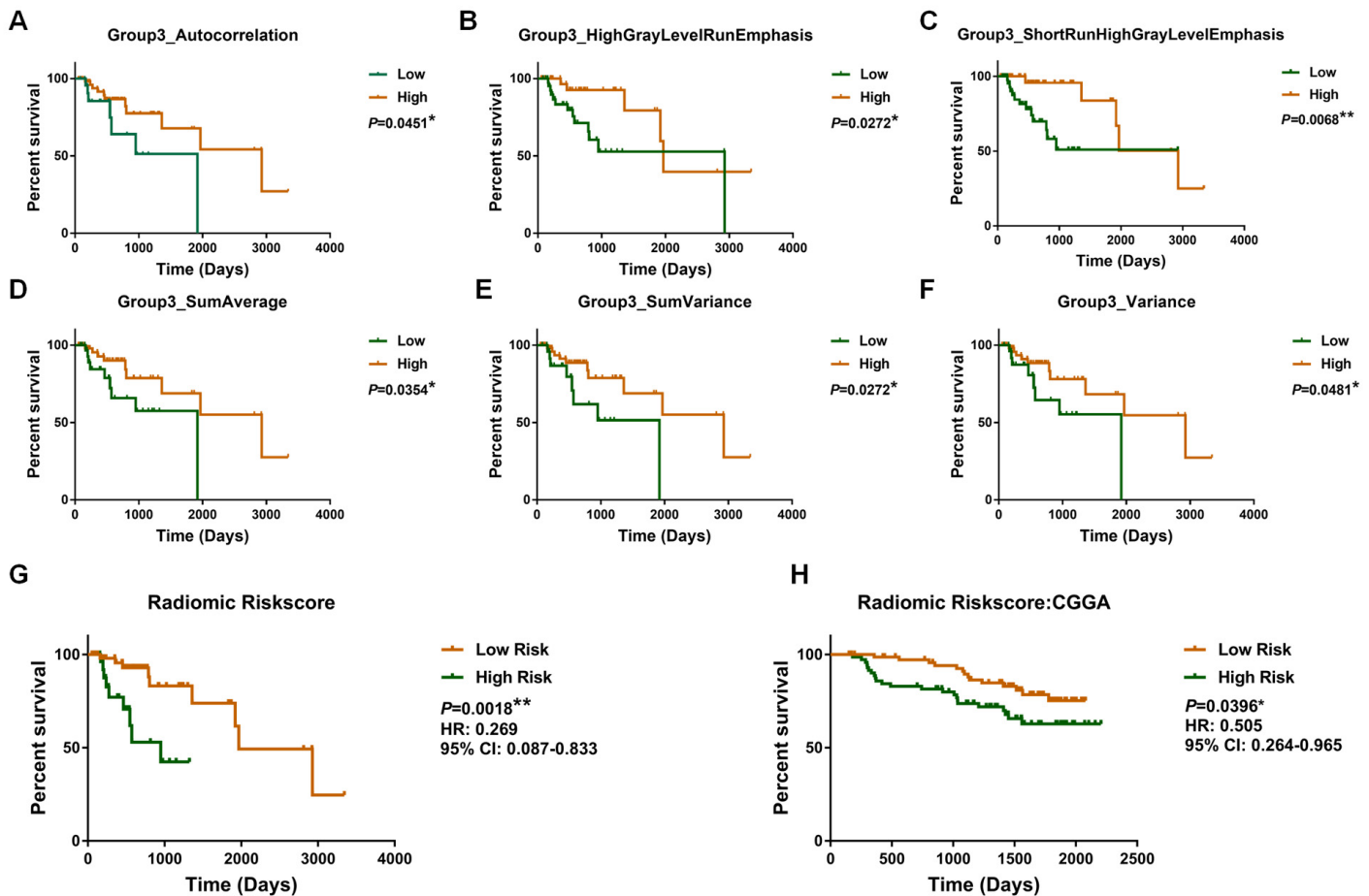
Figure 1. Kaplan–Meier plot for overall survival of patients stratified by the value of each radiomic feature (**A, B, C, D, E, F**) and radiomic risk score (**G**) in the training dataset. The radiomic risk score retained prognostic significance for patients in the validation set (**H**).

Moreover, we observed that each of the selected radiomic features could be used to stratify patients into high-risk and low-risk groups (Autocorrelation, $P = 0.0451$; HGLRE, $P = 0.0272$; SRHGLE, $P = 0.0068$; SumAverage, $P = 0.0354$; SumVariance, $P = 0.0272$; and Variance, $P = 0.0281$; Fig. 1A-F).

Subsequently, a radiomic risk score were calculated: risk score = Autocorrelation × (-0.007) + HGLRE × (-0.003) + SRHGLE × (-0.005) + SumAverage × (-0.115) + SumVariance × (-0.002) + Variance × (-0.007). The radiomic risk score was associated with overall survival in the training dataset ($P = 0.00018$; HR = 0.269, 95% confidence interval [CI]: 0.087–0.833; Fig. 1G).

Consistently, we confirmed the prognostic value of selected radiomic features in the validation dataset (Autocorrelation, $P = 0.0081$; HGLRE, $P = 0.0120$; SRHGLE, $P = 0.0085$; SumAverage, $P = 0.0168$; SumVariance, $P = 0.0058$; and Variance, $P = 0.0063$;

Supplementary Fig. 1), as well as confirming the prognostic value of the radiomic risk score ($P = 0.0396$; HR = 0.505; 95%CI: 0.264–0.965; Fig. 1H).

We next conducted multivariate Cox regression analyses in TCGA database, which indicated that the radiomic risk score was an independent prognostic factor (P = 0.042). Other independent prognostic factors were age, WHO grade, and IDH status. The prognostic value of all clinical characteristics in the multivariate Cox regression analyses are shown in Table 1.

Functional annotation of different prognoses

To explore the genetic background of prognostic differences, relevant transcriptomic profiles were analyzed. The radiogenomic analysis of high-risk positively associated genes ($n = 239$, Fig. 2) further revealed that biological processes associated with prognosis included hypoxia, angiogenesis, and stem cell

proliferation-related oncogenic functions (Fig. 3). Specifically, genes in the "multicellular organism development" group are the ones that are most significantly associated to the radiomic risks score.

Further investigation revealed that SPRED1 and SPRED2 were the most correlated genes involved in "multicellular organism development" (Supplementary Table 2).
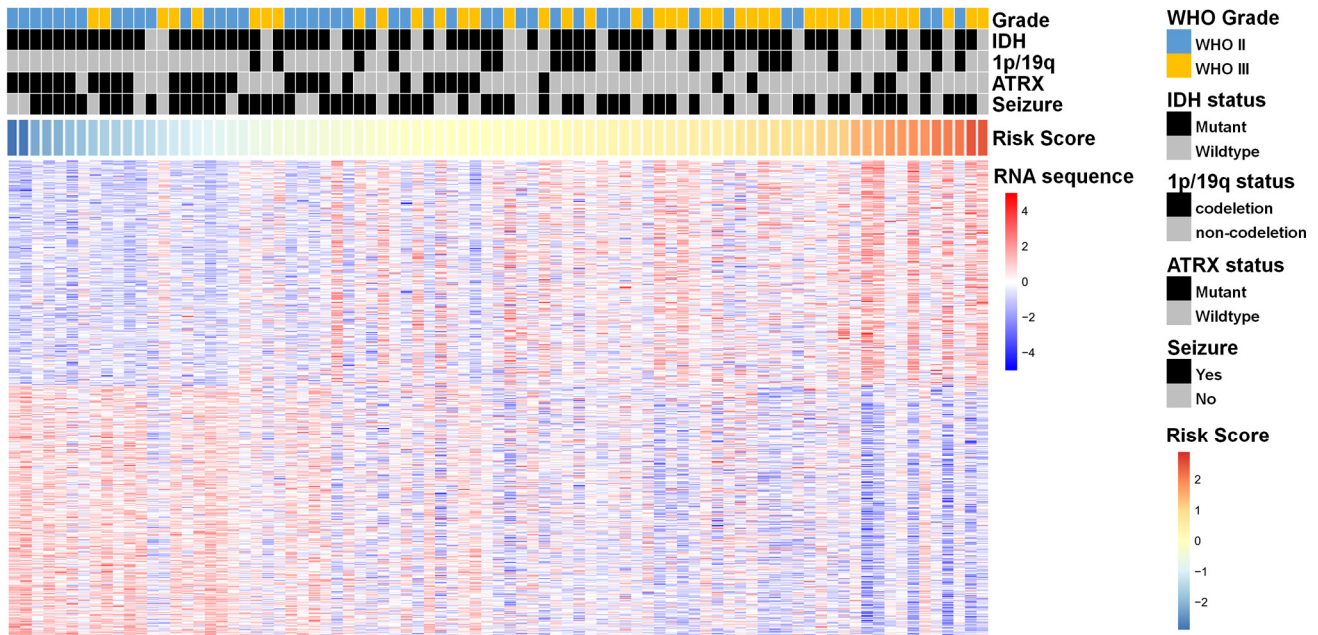
Figure 2. A heat map of the top 200 genes that were positively associated with the radiomic risk score (upper half part) and the top 200 genes that were negatively associated with the radiomic risk score (lower half part) from 85 LGGs samples in the training dataset. "RNA sequence" refers to the overall expression levels of the genes. Associations of clinicopathological characteristics with radiomic features are illustrated.
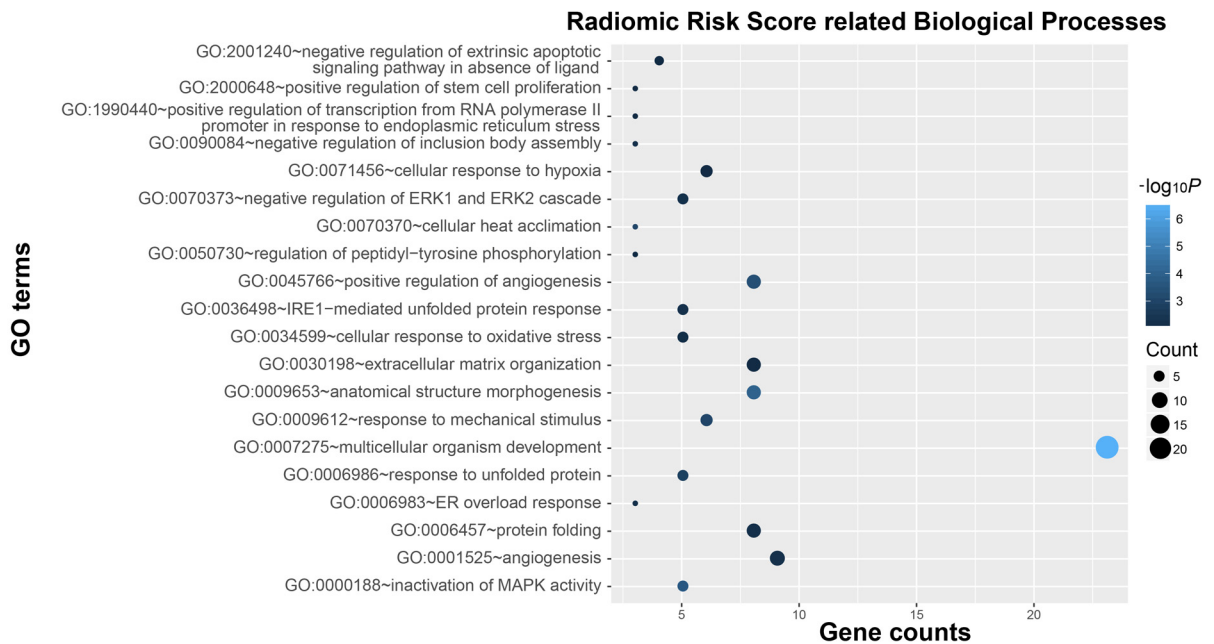

Figure 3. Functional annotation of radiomic risk score groups. Gene ontology analysis revealed a significant association among genes with increased expression in the high-risk radiomic risk score group and twenty main pathways. Column size: gene counts; point color: enrichment *P* value.

Similar findings were obtained during the assessment of genetic alterations underlying the six texture features (Supplementary Fig. 2). As shown in Supplementary Fig. 3, the radiomics-based evaluation may stand for patients with different expression profiles and biological functions among the three molecular classification, therefore serving as a supplementary approach for tailored medicine of LGGs.

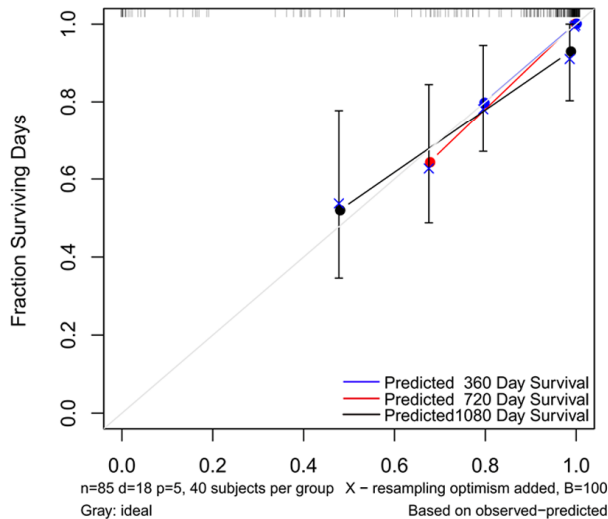## Construction of individualized prediction models

The independent prognostic parameters for overall survival in the training cohort, including WHO grade, age at diagnosis, IDH, seizure, ATRX, and radiomic risk score, were integrated into the nomogram (Supplementary Fig. 4). The C-index of the nomograms for overall survival was 0.934. Meanwhile, the calibration
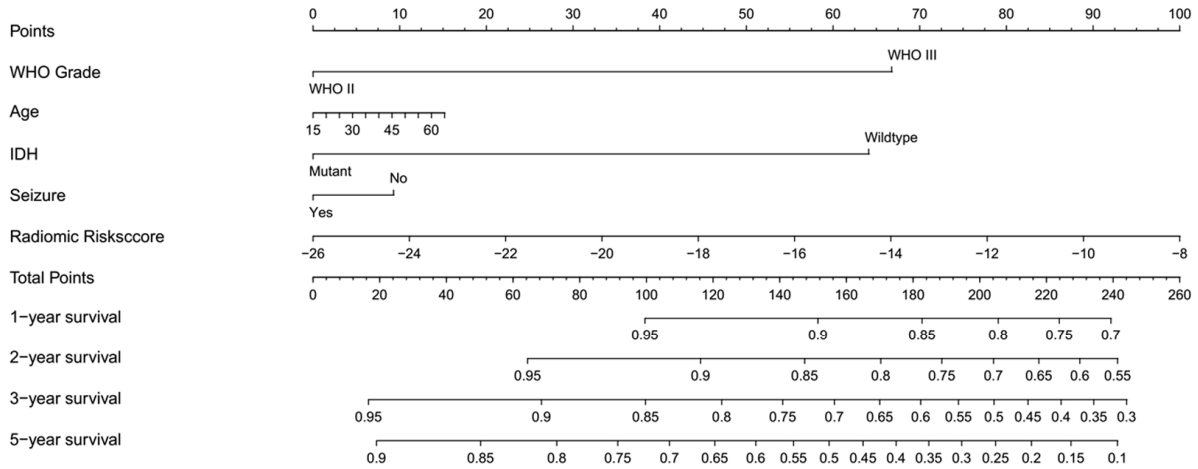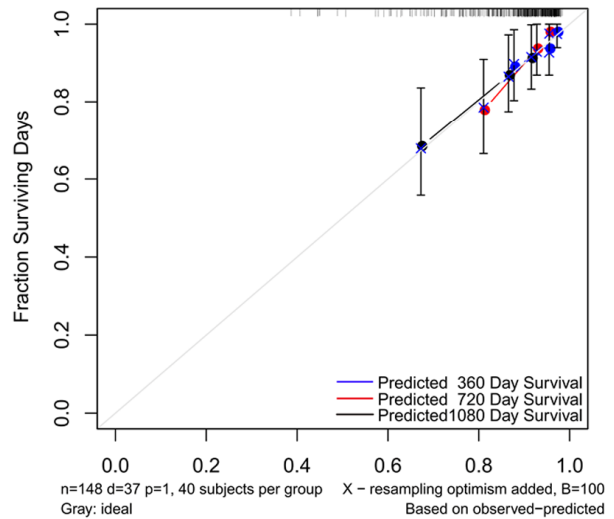
Figure 4. A nomogram for predicting overall survival of patients with LGGs (**A**), along with the assessment of model calibration in the training cohort (**B**) and validation cohort (**C**). After final model selection, radiomic signature, WHO grade, age, IDH status, and seizure were included in the nomogram. The line determines the number of points received for the value of each variable. The sum of these numbers is presented on the total axis, while the line drawn down to the survival axis determines the likelihood of a 1-, 2-, 3-, or 5-year survival rate. The calibration curve of the nomogram is also shown. Three colored lines (blue, red, and black) represent the performance of the nomogram, with a closer fit to the diagonal line representing a better estimation.

plot for the probability of survival showed optimal agreement. Since the ATRX status for patients with LGG was not available in the validation cohort, a prognostic nomogram that integrated all factors except for ATRX was constructed in the training cohort and independently validated in the validation cohort. The C-indices were 0.92 and 0.70 in the training cohort and validation cohort, respectively, indicating satisfactory concordance. Moreover, the calibration plots of the probability of actual survival were concordant with survival outcomes predicted by the nomograms at 1, 2, and 5 years for overall survival (Fig. 4).

DISCUSSION

In the present study, we revealed a radiomic-based signature that noninvasively predicted survival in both the training (TCIA) and validation (CGGA) cohorts. Integrative analysis of radiomic and transcriptomic profiles suggested that a high-risk phenotype indicated by the radiomic analysis could be attributed to several malignant biological processes. Moreover, the combination of radiomic, clinical, and molecular risk factors into a nomogram provided an effective approach for individual survival estimation.

Radiomics is a promising paradigm for extending clinical imaging into comprehensive and quantitative features, and has attracted much interest in personalized medicine [21, 22]. Few studies have identified radiomics could be reliable prognostic biomarkers for stratification of patients across many fields of oncology, however, the currently available studies are typically characterized by smaller sample sizes or lack of using independent public database as validation cohort [12, 23, 24]. Our analyses expand on the work of several recent studies that have uncovered novel associations among radiomic features and several clinical endpoints. The identified radiomic signature in our study consisted of the following features: Autocorrelation, HGLRE, SRHGLE, SumAverage, SumVariance, and Variance, which were derived from the group (iii) parameters (gray level co-occurrence and gray level run-length texture matrices-based parameters). We believe that these features describe textural differences based on gray-tone spatial dependencies, as opposed to relationships or patterns between pixels derived from first-order statistics. Such features provide further insight into tissue microstructure and the local environment of the tumor. Being consistent with our findings, a previous study reported that the most dominant prognostic features are derived from group (iii) [25].

Although each of the identified radiomic features in our study is capable of risk stratification, an integrated signature based on these textures achieved a better performance. Pioneering studies supported our findings that a multi-component radiomic signature provides a more statistically robust approach, and have consistently demonstrated the incremental value of the radiomic signature for risk stratification of patients with different cancers [25, 26]. Notably, recent evidence has supported the hypothesis that radiological manifestations are tightly connected with genetic alterations of the tumor [27, 28]. In the present study, the high-risk phenotype was significantly associated with oncogenic biological processes, including stem cell proliferation and angiogenesis, which could partially represent the malignancy of high-risk LGGs. As a classical hallmark of tumors, angiogenesis has emerged as a promising target for individually tailored medicine, although the therapeutic efficacy of anti-angiogenetic agents in gliomas has proved unsatisfactory [29]. This failure was attributed to ambiguities regarding the appropriate population for such treatment. Our results suggest a radiological indicator for vigorous angiogenesis in gliomas, thereby promoting the clinical effectiveness of anti-angiogenic therapy. Therefore, clinical trials exploring the relationship between radiological manifestation and target therapy are urgently needed to elucidate the possible implications of translating radiomic signatures into clinical practice. Additionally, "multicellular organism development" genes were a group of genes that were found to be highly associated with the radiomic risk score. SPRED1 and SPRED2 are the two most relevant genes and are members of the SPRED family of proteins that regulate growth factor-induced activation of the MAP kinase cascade [30]. The mechanisms of the high association of multicellular organism development with the radiomic risk score are still to be investigated.

The objective of precision medicine is to build a model that offers a more tailored approach for individual patients considering their individual prognostic factors. A nomogram is a graphical representation of a statistical model that allows individualized predictions and compares favorably to traditional risk grouping systems. Incorporating the radiomic signature into a nomogram along with clinicopathologic risk factors to estimate disease-free survival for early-stage non–small cell lung cancer has provided better estimation than either the radiomic signature or clinicopathologic nomogram alone [31]. The new WHO classification of the central nervous system has improved the clinical assessment of LGGs [32]. Indeed, our classifier will be more relevant if it incorporates those prognostics genomic subtypes. However, in practice, not all genomic information (e.g., 1p/19q) is available for all patients; hence, classifier should also be designed to accommodate sparse data. In general, we developed a nomogram in the training

cohort by incorporating WHO grade, age at diagnosis, IDH, seizure, and radiomic risk score. Those prognostic factors were independent prognostic factors with using multivariate Cox regression analysis, suggesting their complementary value in predicting LGGs prognosis. The higher degree of predictive precision in our nomogram could be attributed to the integration of different dimensions of information (genetics, clinical, and imaging data), which provides a complementary perspective about a single tumor. In the validation cohort, age and seizure lost their prognostic power in a multivariate model, which may explain for the decreased performance of the prognostic model. Despite this, our prognostic model still displayed good accuracy (CI=0.7) in the external validation cohort, which has strengthened the reliability of the results. Our findings are consistent with the perspective that future research will be most productive if we concentrate on populating a three-domain Venn diagram intersection consisting of imaging, genetics, and clinical data [33].

There are several limitations that should be considered when interpreting the results of this study. First, as a multi-center study, the imaging data used were acquired from multiple MRI systems with varying protocols. Better image quality and consistent protocols will further improve the power of radiomics. Second, the radiomics analysis was performed using T2-weighted MR images in the current study. Although the FLAIR sequence is favorable for lesion delineation it was not used in the radiomics analysis due to limited availability of FLAIR data in the CGGA image database. Third, a slight discrepancy in the manual segmentation of images into the region of interest (ROI) may still exist, even though the image segmentation was carefully controlled by 2 radiologists. The predictive value of radiomic signatures need to be further characterized and validated in prospective study using multi-modal imaging approaches.

CONCLUSIONS

In conclusion, this radiogenomic study established a radiological indicator for prognostic assessment in patients with LGGs. The biological processes associated with the radiomic features was further revealed. The findings of this study may be an important aid to the decision making in personalized clinical management for glioma patients.

MATERIALS AND METHODS

Patient selection

There were a total of 233 patients enrolled in this study (Supplementary Table 3), including 85 from The Cancer

Genome Atlas (TCGA, training cohort) and 148 patients from CGGA (validation cohort). Selection criteria for both cohorts included: (a) histopathologically confirmed grade II or III gliomas, according to WHO classification [32]; (b) minimum age of 18 years; and (c) available preoperative T2-weighted MR images. Specifically, MRI data were available for 199 lower-grade glioma patients in TCGA database. A set of 103 cases without high-quality preoperative T2 images and 11 cases without overall survival data were excluded. Thus, there were 85 lower-grade glioma cases that were included in the current study. Baseline epidemiologic and clinical characteristics of all patients, including age, sex, seizure, WHO grade, and molecular status, are shown in Table 2. This retrospective study protocol was approved by the ethics committee of Beijing Tiantan Hospital. All clinical data and biological information were collected based on published databases.

Image acquisition and tumor segmentation

MR images in the training cohort were collected from TCIA dataset (http://www.cancerimagingarchive.net). MR images in the validation cohort were obtained from the CGGA imaging database (http://www.cgga.org.cn), and the image acquisition was performed as described in our previous publication [34]. Lesions were delineated on T2-weighted images, as T2-weighted MRI is a widely applied sequence for lesion characterization of brain diseases, specifically for brain tumors. In the majority of studies investigating LGGs, T2 MR imaging was well accepted in the identification of abnormal signal intensity representing the involved regions of LGGs.

Images from both cohorts underwent the same preprocessing procedures. To achieve reliable segmentation, manual segmentation was applied in the radiomic analysis in this study [15, 24, 26, 35]. The ROI was manually drawn from the T2-weighted images abnormality on each slice by two experienced neuroradiologists (J.M. and X.C., both with more than 15 years of experience of diagnosis) by using MRIcro (http://www.mccauslandcenter.sc.edu/mricro/). In cases of a discrepancy of more than 5% (Dice index < 95%) in the tumor border outlines between the neuroradiologists, a third senior neuroradiologist (S.L. with more than 20 years of experience) made the final decision.

Extraction of radiomic features

In order to avoid of the bias from data heterogeneity, all MRI data were normalized (Z score transformation with excluding segmented tumor areas) and re-sampled to the same resolution before feature extraction with

using an in-house MATLAB process. Fifty-five quantitative radiomic features were extracted from the ROI using an automated method as previously described [25]. The features can be divided into three groups: (i) first-order statistics; (ii) shape and size based features; and (iii) textural features. In group (i), we estimated 14 texture parameters describing the distribution of voxel intensities within the ROI. In group (ii), we estimated eight 3D features describing 3D size and shape of the tumor region. In group (iii), we estimated 33 textural features describing patterns or spatial distribution of voxel intensities, which can provide information regarding the relative position of various gray levels over the image with gray level co-occurrence and gray level run-length texture matrices. All features were extracted using MATLAB 2014a (Mathworks, Natick, United States), and are listed in Supplementary Table 4.

RNA sequencing and biomarker detection

Whole genome RNA sequencing data and relevant clinical and molecular neuropathological information were downloaded from the TCGA database (http://cancergenome.nih.gov/). The RNA sequencing data was normalized using the fragments per kilobase transcriptome per million reads method [36]. Transcriptome data were collected to identify potential biological processes underlying the radiomic signature. Presence of IDH mutation, 1p/19q codeletion, and ATRX mutation were also collected. For the validation cohort, IDH mutations were detected with pyro-sequencing and Chromosome 1p/19q status were inferred by a Gaussian window smoothing algorithm using the expression values of the genes located on Chr-1p and Chr-19q, which have been described previously [37, 38].

Construction of the radiomic risk score

To obtain prognostic radiomic features in LGGs, we applied univariate Cox regression analyses of 55 features in the training dataset. Subsequently, the selected imaging features (P <0.05) were used to develop a radiomic signature. To investigate the effectiveness of the radiomic signature for clinical outcome prediction, a radiomic risk score was computed for each patient by linearly combining the selected features weighted by their corresponding coefficients as follows:

Risk score = $\text{expr}_{feature1} \times \beta_{feature1} + \text{expr}_{feature2} \times \beta_{feature2} + \ldots + \text{expr}_{feature\ n} \times \beta_{feature\ n}$.

The same β values were applied to the validation cohort.

Prediction of survival outcome using radiomic risk scores

Patients with LGG in the training and validation cohorts were divided into high-risk and low-risk groups referring to the median value of the radiomic risk score. The potential association of the radiomic risk score with overall survival was first assessed in the training dataset and then validated in the validation dataset with Kaplan-Meier survival analysis. Similarly, the prognostic value of each feature in the risk score was also evaluated based on Kaplan-Meier survival analysis. Patients were classified into "high" and "low" groups referring to the median value of each radiomic feature. Multivariate Cox regression analysis was performed to identify if the radiomic risk score is an independent prognostic factor.

Radiogenomic analysis

Genes with significant associations from the radiomic-based signature or each radiomic feature were selected with Pearson correlation analysis, conducted with R programming language (http://cran.r-project.org). The candidate genes that positively related to radiomic risk score among the three molecular classification were selected with Pearson correlation analysis. Those with P <0.05 and a Pearson correlation coefficient >0.3 were considered as significant associated candidates. Positively-associated genes were then subjected to DAVID (http://david.ncifcrf.gov/) based gene ontology analysis to identify underlying biological processes. Biological processes with P <0.05 were depicted using the *ggplot2* package of R.

Individualized prediction model construction

To establish a model that can predict an individual patient's overall survival, a nomogram was formulated based on the results of the multivariate analysis with the *rms* package in R [39]. The final model was constructed with a backward stepdown selection process conforming to the Akaike information criterion [40]. Concordance index (C-index) and calibration curves were used to measure the predictive accuracy and discriminative ability of the nomograms. During the external validation of the nomogram, the total score (according to the nomogram) of each patient in the validation cohort were calculated and used as a factor in Cox regression analysis, and the C-index and calibration curve were then obtained [41].

CONFLICTS OF INTEREST

The authors of this manuscript have no conflicts of interest to declare.

REFERENCES

1. Ostrom QT, Gittleman H, Fulop J, Liu M, Blanda R, Kromer C, Wolinsky Y, Kruchko C, Barnholtz-Sloan JS. CBTRUS Statistical Report: Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2008-2012. Neuro-oncol. 2015 (Suppl 4); 17:iv1–62. https://doi.org/10.1093/neuonc/nov189

2. Brat DJ, Verhaak RG, Aldape KD, Yung WK, Salama SR, Cooper LA, Rheinbay E, Miller CR, Vitucci M, Morozova O, Robertson AG, Noushmehr H, Laird PW, et al, and Cancer Genome Atlas Research Network. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. N Engl J Med. 2015; 372:2481–98. https://doi.org/10.1056/NEJMoa1402121

3. Buckner JC, Shaw EG, Pugh SL, Chakravarti A, Gilbert MR, Barger GR, Coons S, Ricci P, Bullard D, Brown PD, Stelzer K, Brachman D, Suh JH, et al. Radiation plus Procarbazine, CCNU, and Vincristine in Low-Grade Glioma. N Engl J Med. 2016; 374:1344–55. https://doi.org/10.1056/NEJMoa1500925

4. van den Bent MJ. Practice changing mature results of RTOG study 9802: another positive PCV trial makes adjuvant chemotherapy part of standard of care in low-grade glioma. Neuro-oncol. 2014; 16:1570–74. https://doi.org/10.1093/neuonc/nou297

5. van den Bent MJ, Brandes AA, Taphoorn MJ, Kros JM, Kouwenhoven MC, Delattre JY, Bernsen HJ, Frenay M, Tijssen CC, Grisold W, Sipos L, Enting RH, French PJ, et al. Adjuvant procarbazine, lomustine, and vincristine chemotherapy in newly diagnosed anaplastic oligodendroglioma: long-term follow-up of EORTC brain tumor group study 26951. J Clin Oncol. 2013; 31:344–50. https://doi.org/10.1200/JCO.2012.43.2229

6. Eckel-Passow JE, Lachance DH, Molinaro AM, Walsh KM, Decker PA, Sicotte H, Pekmezci M, Rice T, Kosel ML, Smirnov IV, Sarkar G, Caron AA, Kollmeyer TM, et al. Glioma Groups Based on 1p/19q, IDH, and TERT Promoter Mutations in Tumors. N Engl J Med. 2015; 372:2499–508. https://doi.org/10.1056/NEJMoa1407279

7. Theeler BJ, Yung WK, Fuller GN, De Groot JF. Moving toward molecular classification of diffuse gliomas in adults. Neurology. 2012; 79:1917–26.
https://doi.org/10.1212/WNL.0b013e318271f7cb

8. Macdonald DR, Cascino TL, Schold SC Jr, Cairncross JG. Response criteria for phase II studies of supratentorial malignant glioma. J Clin Oncol. 1990; 8:1277–80. https://doi.org/10.1200/JCO.1990.8.7.1277

9. Vogelbaum MA, Jost S, Aghi MK, Heimberger AB, Sampson JH, Wen PY, Macdonald DR, Van den Bent MJ, Chang SM. Application of novel response/progresssion measures for surgically delivered therapies for gliomas: Response Assessment in Neuro-Oncology (RANO) Working Group. Neurosurgery. 2012; 70:234–43. https://doi.org/10.1227/NEU.0b013e318223f5a7

10. Bates SE. It's all about the test: the complexity of companion diagnostic co-development in personalized medicine. Clin Cancer Res. 2014; 20:1418. https://doi.org/10.1158/1078-0432.CCR-14-0223

11. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2016; 278:563–77. https://doi.org/10.1148/radiol.2015151169

12. Kim JH, Ko ES, Lim Y, Lee KS, Han BK, Ko EY, Hahn SY, Nam SJ. Breast Cancer Heterogeneity: MR Imaging Texture Analysis and Survival Outcomes. Radiology. 2017; 282:665-75. https://doi.org/10.1148/radiol.2016160261

13. Sottoriva A, Spiteri I, Piccirillo SG, Touloumis A, Collins VP, Marioni JC, Curtis C, Watts C, Tavaré S. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proc Natl Acad Sci USA. 2013; 110:4009–14. https://doi.org/10.1073/pnas.1219747110

14. Yachida S, Jones S, Bozic I, Antal T, Leary R, Fu B, Kamiyama M, Hruban RH, Eshleman JR, Nowak MA, Velculescu VE, Kinzler KW, Vogelstein B, Iacobuzio-Donahue CA. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. Nature. 2010; 467:1114–17. https://doi.org/10.1038/nature09515

15. Kickingereder P, Bonekamp D, Nowosielski M, Kratz A, Sill M, Burth S, Wick A, Eidel O, Schlemmer HP, Radbruch A, Debus J, Herold-Mende C, Unterberg A, et al. Radiogenomics of Glioblastoma: Machine Learning-based Classification of Molecular Characteristics by Using Multiparametric and Multiregional MR Imaging Features. Radiology. 2016; 281:907–18. https://doi.org/10.1148/radiol.2016161382

16. Wang K, Wang Y, Fan X, Wang J, Li G, Ma J, Ma J, Jiang T, Dai J. Radiological features combined with IDH1

status for predicting the survival outcome of glioblastoma patients. Neuro-oncol. 2016; 18:589–97. https://doi.org/10.1093/neuonc/nov239

17. Kim JW, Park CK, Park SH, Kim YH, Han JH, Kim CY, Sohn CH, Chang KH, Jung HW. Relationship between radiological characteristics and combined 1p and 19q deletion in World Health Organization grade III oligodendroglial tumours. J Neurol Neurosurg Psychiatry. 2011; 82:224–27. https://doi.org/10.1136/jnnp.2009.178806

18. Choi YS, Ahn SS, Kim DW, Chang JH, Kang SG, Kim EH, Kim SH, Rim TH, Lee SK. Incremental Prognostic Value of ADC Histogram Analysis over MGMT Promoter Methylation Status in Patients with Glioblastoma. Radiology. 2016; 281:175–84. https://doi.org/10.1148/radiol.2016151913

19. Jamshidi N, Diehn M, Bredel M, Kuo MD. Illuminating radiogenomic characteristics of glioblastoma multiforme through integration of MR imaging, messenger RNA expression, and DNA copy number variation. Radiology. 2014; 270:1–2. https://doi.org/10.1148/radiol.13130078

20. Mazurowski MA, Clark K, Czarnek NM, Shamsesfandabadi P, Peters KB, Saha A. Radio-genomics of lower-grade glioma: algorithmically-assessed tumor shape is associated with tumor genomic subtypes and patient outcomes in a multi-institutional study with The Cancer Genome Atlas data. J Neurooncol. 2017; 133:27–35. https://doi.org/10.1007/s11060-017-2420-1

21. Aerts HJ. The Potential of Radiomic-Based Phenotyping in Precision Medicine: A Review. JAMA Oncol. 2016; 2:1636–42. https://doi.org/10.1001/jamaoncol.2016.2631

22. Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuzé S, Schernberg A, Paragios N, Deutsch E, Ferté C. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. Ann Oncol. 2017; 28:1191–206. https://doi.org/10.1093/annonc/mdx034

23. Zhang B, Tian J, Dong D, Gu D, Dong Y, Zhang L, Lian Z, Liu J, Luo X, Pei S, Mo X, Huang W, Ouyang F, et al. Radiomics Features of Multiparametric MRI as Novel Prognostic Factors in Advanced Nasopharyngeal Carcinoma. Clin Cancer Res. 2017; 23:4259–69. https://doi.org/10.1158/1078-0432.CCR-16-2910

24. Zhou H, Vallières M, Bai HX, Su C, Tang H, Oldridge D, Zhang Z, Xiao B, Liao W, Tao Y, Zhou J, Zhang P, Yang L. MRI features predict survival and molecular markers in diffuse lower-grade gliomas. Neuro-oncol. 2017; 19:862–70. https://doi.org/10.1093/neuonc/now256

25. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D, Hoebers F, Rietbergen MM, Leemans CR, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014; 5:4006. https://doi.org/10.1038/ncomms5006

26. Kickingereder P, Burth S, Wick A, Götz M, Eidel O, Schlemmer HP, Maier-Hein KH, Wick W, Bendszus M, Radbruch A, Bonekamp D. Radiomic Profiling of Glioblastoma: Identifying an Imaging Predictor of Patient Survival with Improved Performance over Established Clinical and Radiologic Risk Models. Radiology. 2016; 280:880–89. https://doi.org/10.1148/radiol.2016160845

27. Sala E, Mema E, Himoto Y, Veeraraghavan H, Brenton JD, Snyder A, Weigelt B, Vargas HA. Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. Clin Radiol. 2017; 72:3-10. https://doi.org/10.1016/j.crad.2016.09.013

28. Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, Ye Z, Makrigiorgos M, Fennessy F, Mak RH, Gillies R, Quackenbush J, Aerts HJ. Somatic Mutations Drive Distinct Imaging Phenotypes in Lung Cancer. Cancer Res. 2017; 77:3922–30. https://doi.org/10.1158/0008-5472.CAN-17-0122

29. Kickingereder P, Götz M, Muschelli J, Wick A, Neuberger U, Shinohara RT, Sill M, Nowosielski M, Schlemmer HP, Radbruch A, Wick W, Bendszus M, Maier-Hein KH, Bonekamp D. Large-scale Radiomic Profiling of Recurrent Glioblastoma Identifies an Imaging Predictor for Stratifying Anti-Angiogenic Treatment Response. Clin Cancer Res. 2016; 22:5765–71. https://doi.org/10.1158/1078-0432.CCR-16-0702

30. Siljamäki E, Abankwa D. SPRED1 Interferes with K-ras but Not H-ras Membrane Anchorage and Signaling. Mol Cell Biol. 2016; 36:2612–25. https://doi.org/10.1128/MCB.00191-16

31. Huang Y, Liu Z, He L, Chen X, Pan D, Ma Z, Liang C, Tian J, Liang C. Radiomics Signature: A Potential Biomarker for the Prediction of Disease-Free Survival in Early-Stage (I or II) Non-Small Cell Lung Cancer. Radiology. 2016; 281:947–57. https://doi.org/10.1148/radiol.2016152234

32. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol. 2016; 131:803–20. https://doi.org/10.1007/s00401-016-1545-1

33. Jaffe CC. Imaging and genomics: is there a synergy? Radiology. 2012; 264:329–31. https://doi.org/10.1148/radiol.12120871

34. Wang Y, Qian T, You G, Peng X, Chen C, You Y, Yao K, Wu C, Ma J, Sha Z, Wang S, Jiang T. Localizing seizure-susceptible brain regions associated with low-grade gliomas using voxel-based lesion-symptom mapping. Neuro-oncol. 2015; 17:282–88. https://doi.org/10.1093/neuonc/nou130

35. Zhang B, Chang K, Ramkissoon S, Tanguturi S, Bi WL, Reardon DA, Ligon KL, Alexander BM, Wen PY, Huang RY. Multimodal MRI features predict isocitrate dehydrogenase genotype in high-grade gliomas. Neuro Oncol. 2017; 19:109-17. https://doi.org/10.1093/neuonc/now121

36. Chen HM, Yu K, Tang XY, Bao ZS, Jiang T, Fan XL, Chen XW, Su XD. Enhanced expression and phosphorylation of the MET oncoprotein by glioma-specific PTPRZ1-MET fusions. FEBS Lett. 2015; 589:1437–43. https://doi.org/10.1016/j.febslet.2015.04.032

37. Yang P, Cai J, Yan W, Zhang W, Wang Y, Chen B, Li G, Li S, Wu C, Yao K, Li W, Peng X, You Y, et al, and CGGA project. Classification based on mutations of TERT promoter and IDH characterizes subtypes in grade II/III gliomas. Neuro-oncol. 2016; 18:1099–108. https://doi.org/10.1093/neuonc/now021

38. Hu X, Martinez-Ledesma E, Zheng S, Kim H, Barthel F, Jiang T, Hess KR, Verhaak RG. Multigene signature for predicting prognosis of patients with 1p19q co-deletion diffuse glioma. Neuro-oncol. 2017; 19:786–95. https://doi.org/10.1093/neuonc/now285

39. Feng LH, Dong H, Lau WY, Yu H, Zhu YY, Zhao Y, Lin YX, Chen J, Wu MC, Cong WM. Novel microvascular invasion-based prognostic nomograms to predict survival outcomes in patients after R0 resection for hepatocellular carcinoma. J Cancer Res Clin Oncol. 2017; 143:293–303. https://doi.org/10.1007/s00432-016-2286-1

40. Li J, Zhou J, Yang PH, Xia Y, Shi YH, Wu D, Lv G, Zheng W, Wang K, Wan XY, Lau WY, Wu MC, Fan J, Shen F. Nomograms for survival prediction in patients undergoing liver resection for hepatitis B virus related early stage hepatocellular carcinoma. Eur J Cancer. 2016; 62:86–95. https://doi.org/10.1016/j.ejca.2016.04.011

41. Wang Y, Li J, Xia Y, Gong R, Wang K, Yan Z, Wan X, Liu G, Wu D, Shi L, Lau W, Wu M, Shen F. Prognostic nomogram for intrahepatic cholangiocarcinoma after partial hepatectomy. J Clin Oncol. 2013; 31:1188–95. https://doi.org/10.1200/JCO.2012.41.5984
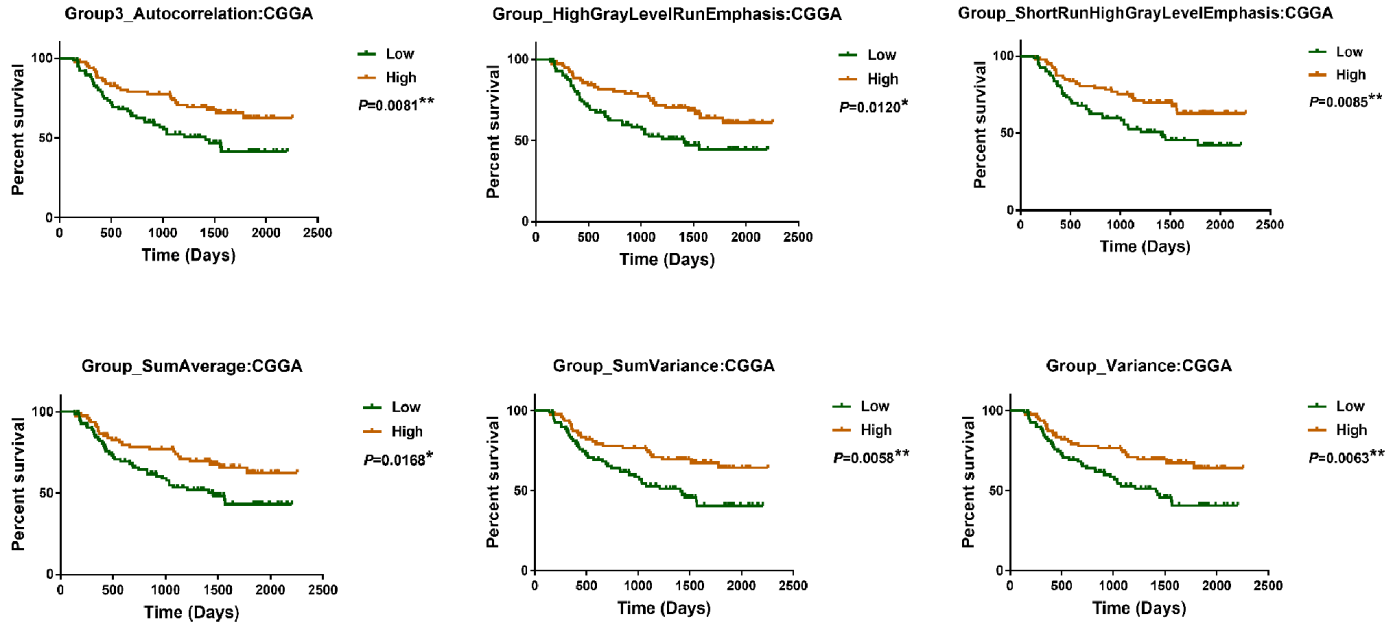
Supplementary Figure 1. Kaplan–Meier plot for overall survival of patients stratified by low- and high-value of each radiomic feature in the validation set.
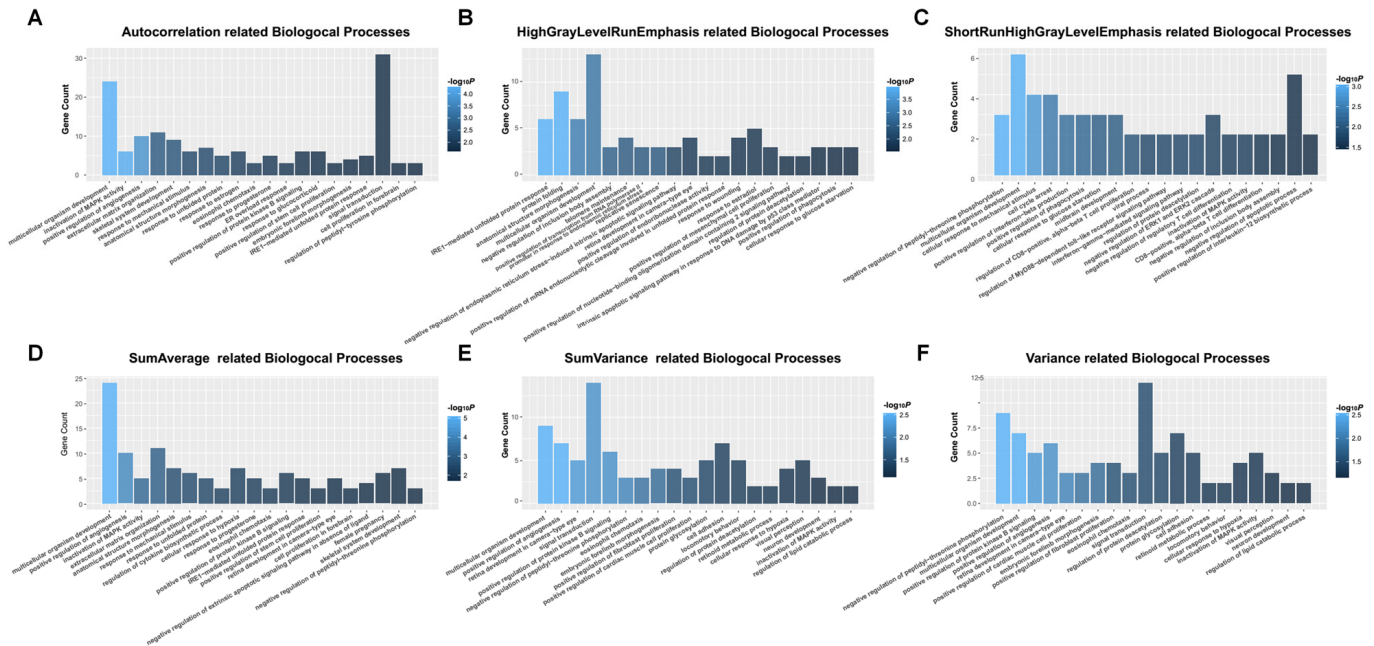
Supplementary Figure 2. Functional annotation of significant radiomic features. Gene ontology analysis revealed a significant association among genes with increased expression in each high-risk radiomic feature and twenty main pathways. Column height: gene counts; point color: enrichment *P* value.
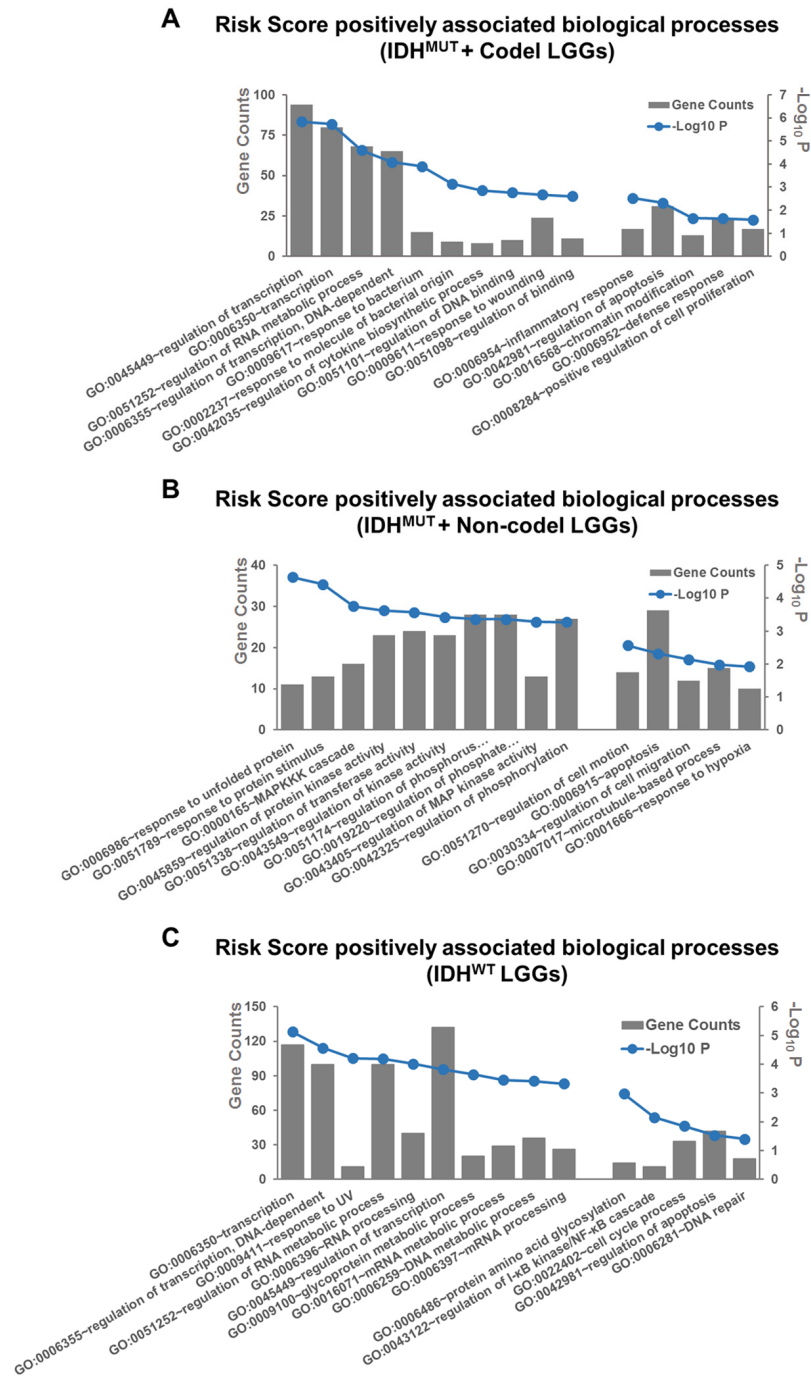
Supplementary Figure 3. The risk score positively associated biological processes among the three molecular classification. IDH mutation and 1p/19q codeletion (**A**), IDH mutation and 1p/19q non-codeletion (**B**), and IDH wild-type (**C**). Column height: gene counts; point color: enrichment *P* value.
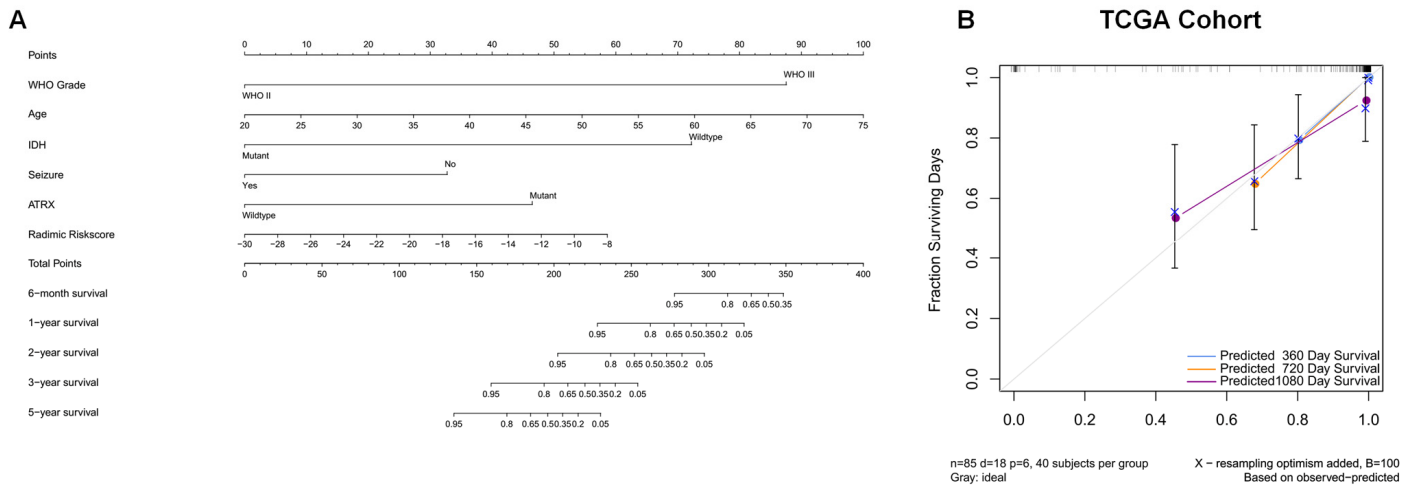
Supplementary Figure 4. A nomogram for predicting overall survival of patients with LGGs (**A**), along with assessment of the model calibration in the training cohort (**B**). After final model selection, radiomic signature, WHO grade, age, IDH status, ATRX status, and seizure were included in the nomogram. The line represents the number of points received for the value of each variable. The sum of these numbers is presented on the total axis, while line drawn down to the survival axis determines the likelihood of a 1-, 2-, 3-, or 5-year survival rate. The calibration curve of the nomogram is also shown. Three colored lines (blue, yellow, and purple) present the performance of the nomogram, with a closer fit to the diagonal line representing a better estimation.

SUPPLEMENTARY TABLES

Supplementary Table 1. Univariable Cox of radiomic features in TCGA lower grade glioma patients.

Feature name	Interpretation	β	95%CI Lower	95%CI Upper	HR	*P*
Group3_Autocorrelation	A texture feature derived from Gray-Level Co-Occurrence Matrix based features. Autocorrelation evaluates the linear spatial relationship between texture primitives and measures the coarseness of an image.	-0.007	0.988	0.999	0.993	**0.024**
Group3_HighGrayLevelRunEmphasis	One of the texture features derived from Gray-Level Run-Length matrix based features. High Gray-Level Run Emphasis measures the distribution of high gray scale values. This feature is high for the image with high gray scale values.	-0.003	0.996	0.999	0.997	**0.008**
Group3_ShortRunHighGrayLevelEmphasis	A texture feature derived from Gray-Level Run-Length matrix based features. Short Run High Gray Level Emphasis describes the complementary metric to the previous one for high gray levels.	-0.005	0.992	0.998	0.995	**0.002**
Group3_SumAverage	A texture feature derived from Gray-Level Co-Occurrence Matrix based features. Sum Average measures overall image brightness.	-0.115	0.808	0.983	0.891	**0.022**
Group3_SumVariance	One of the texture features derived from Gray-Level Co-Occurrence Matrix based features. It describes the overall voxel intensity value variability of the tumor.	-0.002	0.996	0.999	0.998	**0.024**
Group3_Variance	A texture feature derived from Gray-Level Co-Occurrence Matrix based features. Variance is a measure of how much the gray levels differ from the mean.	-0.007	0.987	0.999	0.993	**0.023**

Supplementary Table 2. Genes involved in the "multicellular organism development" group.

Order	GO term	Gene	Pearson correlation coefficient*
1	GO:0007275~multicellular organism development	SPRED1	0.514802938
2	GO:0007275~multicellular organism development	SPRED2	0.479288376
3	GO:0007275~multicellular organism development	SPRY4	0.468034887
4	GO:0007275~multicellular organism development	EDA2R	0.463660169
5	GO:0007275~multicellular organism development	HILS1	0.443597755
6	GO:0007275~multicellular organism development	CREB3L1	0.402457048
7	GO:0007275~multicellular organism development	INVS	0.381701917
8	GO:0007275~multicellular organism development	GCNT2	0.361145274
9	GO:0007275~multicellular organism development	GGN	0.360364448
10	GO:0007275~multicellular organism development	LBH	0.351541368
11	GO:0007275~multicellular organism development	SPRY2	0.349641312
12	GO:0007275~multicellular organism development	EPHA2	0.337424369
13	GO:0007275~multicellular organism development	NFE2	0.336535386
14	GO:0007275~multicellular organism development	HMGA2	0.33571953
15	GO:0007275~multicellular organism development	INSL4	0.333116639
16	GO:0007275~multicellular organism development	OTX1	0.322126755
17	GO:0007275~multicellular organism development	CATSPER3	0.315592995
18	GO:0007275~multicellular organism development	TXNDC2	0.311809689
19	GO:0007275~multicellular organism development	TP53	0.309901514
20	GO:0007275~multicellular organism development	TNFRSF10B	0.309109194
21	GO:0007275~multicellular organism development	EYA4	0.306919931
22	GO:0007275~multicellular organism development	SALL4	0.306853623
23	GO:0007275~multicellular organism development	LIF	0.301635627

*The Pearson correlation coefficient was calculated using the radiomic risk score and the expression of the genes.

Supplementary Table 3. Clinical characteristics of lower grade gliomas in TCGA and CGGA datasets.

	TCGA cohort (n=85)	CGGA cohort (n=148)
Age (range, median)	20-74(43)	18-63(38)
Sex		
Female	49	54
Male	36	94
WHO Grade		
WHO II	45	105
WHO III	40	43
Seizure		
Yes	56	89
No	29	59
IDH status		
Mutant	65	109
Wildtype	20	39
ATRX		
Mutant	34	NA
Wildtype	51	NA
1p/19q		
Codeletion	21	22
Non-codeletion	64	47
NA	0	79

NA = Not Available; TCGA = the Cancer Genome Atlas; CGGA = Chinese Glioma Genome Atlas

Supplementary Table 4. Radiomics features extracted from T2-weight magnetic resonance image (n=55).

Groups

I. First order statistics (n=14)

Energy	Kurtosis	Mean absolute deviation
Entropy	Maximum	Median
Standard deviation	Mean	Minimum
Uniformity	Root mean square (RMS)	Range
Variance	Skewness	

II. Shape and size based features (n=8)

Compactness 1	Spherical disproportion	Surface area
Compactness 2	Sphericity	Surface to volume ratio
Maximum 3D diameter	Volume	

III. Textural features (n=33)

Gray level co-occurrence matrix (GLCM) (n=22)

Autocorrelation	Energy	Inverse Difference Moment
Cluster Prominence	Entropy (H)	Normalized (IDMN)
Cluster Shade	Homogeneity 1	Inverse Difference Normalized
Cluster Tendency	Homogeneity 2	(IDN)
Contrast	Informational measure of	Inverse variance
Correlation	correlation 1 (IMC1)	Maximum Probability
Difference entropy	Informational measure of	Sum average
Dissimilarity	correlation 2 (IMC2)	Sum entropy
Sum variance	Variance	

Gray level run-length matrix (GLRLM) (n=11)

Short Run Emphasis (SRE)	Short Run Low Gray Level Emphasis (SRLGLE)
Long Run Emphasis (LRE)	Short Run High Gray Level Emphasis (SRHGLE)
Gray Level Non-Uniformity (GLN)	Long Run Low Gray Level Emphasis (LRLGLE)
Run Length Non-Uniformity (RLN)	Long Run High Gray Level Emphasis (LRHGLE)
Run Percentage (RP)	Low Gray Level Run Emphasis (LGLRE)
	High Gray Level Run Emphasis (HGLRE)