

Predicting short-term survival after total resection in glioblastomas by machine learning-based radiomic analysis of preoperative MRI

Santiago Cepeda (✉ cepeda_santiago@hotmail.com)

University Hospital Río Hortega <https://orcid.org/0000-0003-1667-8548>

Angel Perez-Nuñez

University Hospital 12 de Octubre

Sergio Garcia-Garcia

University Hospital Río Hortega

Daniel Garcia-Perez

University Hospital 12 de Octubre

Ignacio Arrese

University Hospital Río Hortega

Luis Jimenez-Roldan

University Hospital 12 de Octubre

Manuel Garcia-Galindo

University of Valladolid

Pedro Gonzalez

University Hospital 12 de Octubre

Maria Velasco-Casares

University Hospital Río Hortega

Tomas Zamora

University Hospital Río Hortega

Rosario Sarabia

University Hospital Río Hortega

Research Article

Keywords: Glioblastoma, Radiomics, Texture analysis, Survival, Machine learning

Posted Date: June 22nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-640526/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Radiomics, in combination with artificial intelligence, has emerged as a powerful tool for the development of predictive models in neuro-oncology. Our study aims to find an answer to a clinically relevant question: is there a radiomic profile that can identify glioblastoma (GBM) patients with short-term survival after complete tumor resection?

Introduction

Glioblastoma (GBM) continues to be the most threatening primary brain neoplasm, with a median survival of approximately 15 months.¹ Currently, despite the standard treatment that includes maximum safe surgical resection followed by adjuvant chemoradiation therapy^{2,3}, its prognosis remains ominous, and our knowledge of this neoplasm is still limited.

Predicting a patient's survival is of vital importance for determining the ideal choice of treatment and management. Currently, several prognostic factors are commonly used to predict the prognosis of these patients, including age, sex, Karnofsky performance status (KPS), molecular profile, extent of resection, preoperative tumor volume, volume of nonenhancing tumor and degree of necrosis.⁴ However, some of these features depend on radiologists' interpretation, which justifies the increasing need for an unbiased and quantitative radiological evaluation.

Magnetic resonance imaging (MRI) plays a fundamental role in neuro-oncology for diagnosing and assessing response to treatment and is being increasingly used as a noninvasive predictive tool. On the other hand, the term "radiomics" refers to the process of obtaining quantitative features based on the intensity, volume, shape and texture variations of the radiological images and creating algorithms that find the association of these variables with the survival and outcome of the patients.⁵ Through radiomics, converting medical images into high-dimensional data allows us to expose the underlying pathophysiology, especially intratumor heterogeneity.⁶ This extraction process captures tumor characteristics undetectable to the human eye and gives added value to clinical visual perception. Radiomics incorporates several essential disciplines, including radiology for image interpretation, computerized vision for extracting quantitative variables, and machine learning for classification and regression tasks.^{7,8} Such integration has been demonstrated to exceed expert human abilities in multiple tasks, including diagnosis and outcome prediction.

Recognizing patients who would not benefit from standard treatment and identifying those who need a more aggressive approach at the time of diagnosis is essential for managing GBM through personalized medicine.⁹ There are several publications that, through the integration of radiomics and artificial intelligence, seek to establish survival prediction models in GBM based on preoperative MRI.⁹⁻¹⁵ In the vast majority of studies, patients are classified according to their survival into two or three categories,

depending on whether they exceed 10 or 15 months of survival. This approach aims to identify medium- and long-term survivors who could theoretically be beneficiaries of aggressive therapies. Furthermore, in most studies, the extent of resection is not used as a discriminatory factor, and biopsies and partial and subtotal resections are included.

This fact precludes the implementation of such predictive models in newly diagnosed GBM patients. Our study aims to use the radiological characteristics from structural preoperative multiparametric magnetic resonance imaging (mpMRI) to construct a predictive model of short-term survival in patients in whom total or near-total resection of the enhancing tumor has been performed followed by standard treatment.

Methods

Study Population

A retrospective collection of patients who underwent surgery with a diagnosis of GBM was carried out in two institutions between January 2019 and January 2020. In addition, a second cohort of patients was selected from available public databases: the BraTS (Multimodal Brain Tumor Segmentation) Challenge 2020¹⁶⁻¹⁸ and three other sources available through The Cancer Image Archive [Ivy Glioblastoma Atlas Project (Ivy -GAP)¹⁹, the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium Glioblastoma Multiforme (CPTAC-GBM) and The Cancer Genome Atlas Glioblastoma Multiforme (TCGA-GBM)²⁰]. The inclusion criteria were pathologically confirmed glioblastomas, availability of preoperative MRI with structural/conventional sequences [T2-weighted images (T2WI), fluid-attenuated inversion recovery (FLAIR), T1-weighted images (T1WI) and contrast-enhanced T1-weighted images (T1CE)] with adequate resolution, known survival status, and clinical information (age and type of surgical resection). Only those cases in which gross total resection (100% of the enhancing tumor volume) or near-total resection (> 95% of the enhancing tumor volume) were included. Patients were randomly allocated into training and testing data sets following a proportion of 70/30. The primary endpoint was overall survival (OS), which was defined as the number of days from the initial pathological diagnosis to death (censored = 1) or the last date that they were known to be alive (censored = 0). Public datasets do not have patient identifiers. Hence, no institutional review board approval was required. Nevertheless, the study was approved by the institutional review boards and ethics committees of the other two participating centers. Additionally, all institutional patients provided written informed consent. The study was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments.

Image data description and preprocessing

All BraTS scans were acquired with different clinical protocols and various scanners from multiple (n = 19) institutions. Details of the protocol acquisition of the scans from TCIA and institutional cases are shown in Supplementary Table 1.

Image preprocessing consists of several steps. First, mpMRI scans were converted to Neuroimaging Informatics Technology Initiative (Nifti) format. Then, the scans were placed in a common orientation [“LPS” (left-posterior-superior) in the radiological convention or “RAI” (right-anterior-inferior) in the neurological convention). Later, the scans for every subject were registered to SRI24 anatomical atlas space.²¹ N4 bias correction²² was applied as a temporary step to facilitate optimal registration but was not included at the end of the process since it might obliterate the MRI signal, particularly on the FLAIR modality.²³

The T1W1, T2WI and FLAIR scans were registered to the transformed T1CE scan, resulting in coregistered resampled volumes of 1x1x1 mm isotropic voxels. The brain was then extracted from all coregistered scans using a pretrained deep learning-based model.²⁴ Finally, intensity Z-scoring normalization was carried out. All preprocessing pipelines were generated using The Cancer Imaging Phenomics Toolkit (CaPTk).²⁵

Tumor segmentation and feature extraction

The method used to generate the segmentation labels of the different tumor regions is called GLISTRboost (Boosted GLioma Image SegmenTation and Registration),²⁶ which is defined as a hybrid generative-discriminative tumor segmentation method. This segmentation algorithm comprises a glioma growth model, a discriminative part based on a gradient boosting multiclass classification scheme and a Bayesian strategy.¹⁵ Segmentation labels or volumes of interest (VOIs) were as follows: enhancing tumor (ET), non-enhancing tumor/necrosis (NET), and edema (ED). Segmentations were evaluated by two experts (S.C., S.G.) and corrected manually if necessary.

Using the extraction tool of CaPTk, a total of 15720 characteristics were computed from the tumor subregions (i.e., ET, NET and ED) and the four mpMRI modalities following the Image biomarker standardization initiative (IBSI)²⁷ definitions. These extracted features included intensity features or first-order statistics, histogram-based features, and volumetric, morphologic and textural features, including those based on the gray level cooccurrence matrix (GLCM), gray level run-length matrix (GLRLM), neighborhood gray-tone difference matrix (NGTDM), gray level size-zone matrix (GLSZM) and lattice-based features. A detailed description of these characteristics is shown in Supplementary Table 2.

Data processing and feature selection

After the extraction of the radiomic features, the data were preprocessed by data cleaning (removing the variables with more than 5% missing values) and imputation (using the average value method). Afterward, 13265 features were z-score normalized based on the mean and standard deviation. Then, a feature selection process was necessary to reduce these high-dimensional imaging features to avoid overfitting. Feature selection helps to optimize the generalizability and reproducibility of the models subsequently built. A two-step selection method was used as follows. Spearman’s correlation coefficient was calculated for each pair of radiomic features. Then, the features with Spearman’s correlation

coefficient > 0.95 with each other were discarded, retaining a single feature in each set. Later, features were reduced by including only the variables with a significance < 0.001 with the OS in days. Thus, the number of features was reduced to a set of 260.

Statistical Analysis and Machine Learning

Predicting OS was achieved by two different strategies. The first is a binary classification task between short and long survival. For this purpose and after feature reduction, several classification algorithms were assessed for patient stratification. As a previous step, machine learning (ML)-based filters were used: Gini index (GINI), fast correlation-based filter (FCBF) and information gain (InfoGain). Hence, the top 10 features were selected. Then, five different ML classifiers were trained: naive Bayes, k-nearest neighbors (kNN), random forest (RF), support vector machine (SVM) and a multilayer perceptron algorithm (neural network -NN). The target response for each model was the patient's OS grouped into two classes to distinguish patients who survived < 6 months (short-term survivors) from others. Then, the results were quantitatively validated on the testing data set. The performance of the ML classifiers was measured by the area under the receiver operating characteristic curve (AUC), accuracy, precision, F1 score and recall. All performance metrics were reported as the average value over classes.

The second statistical strategy was conducted using the random survival forest (RSF) approach from the R package "randomForestRSC"^{28,29}, an ensemble-tree method that adapts random forests to right-censored data and survival analysis. RSF does not rely on restrictive assumptions such as proportional hazards and automatically handles nonlinear effects and interactions of high-dimensional data. Features were ranked by positive importance using a variable hunting algorithm as a feature selector. Model hyperparameters were as follows: number of trees = 500, node size = 2, number of splits = 10 and log-rank as the splitting rule. We also evaluated the model's ability to generalize those predictions on the testing group. Training and testing predictions were performed using 5-fold cross-validation.

When the primary outcome is survival (time to event), RSF produces a cumulative hazard function (CHF) from each decision tree that is averaged in an ensemble out-of-bag CHF (OOB-CHF). The predicted ensemble mortality is the mean OOB-CHF estimated by the RSF model for each subject, and it was used to calculate each patient's estimated mortality risk. We used the results from the RSF model to build a mortality risk score and split the sample into high- and low-risk groups. The OOB-CHF cut-off values defining the risk groups were calculated through the "cutp" function of the "survMisc" package.³⁰ The log-rank test was used to compare the survival Kaplan-Meier curves between the patients in the high- and low-risk groups.

Finally, Cox proportional hazard regression models were fitted to the training data set using the dichotomized risk score (high- and low-risk groups) from the RSF model as an explanatory variable, the patient's age and a combination of both. Then, the models were validated in the testing dataset. The performance assessment of the survival models was performed by calculating the prediction errors using

the integrated Brier score (IBS) defined as the average squared distances between the observed survival status and the predicted survival probability by the "pec" package.³¹ Additionally, the discriminatory capacity of the model was evaluated by calculating the concordance index (CI), which refers to the probability that, for a pair of randomly allowed samples, the sample with the highest risk prediction experiences an event before the sample with the smallest risk. Furthermore, the integrated area under the time-dependent ROC curve (iAUC) was calculated for all models using the "risksetROC" package.³² The standard approach of ROC curve analysis considers event (death) status and predictor value for an individual as fixed over time. Because the status and explanatory variables change over time, we used the risksetROC package that estimates the iAUC under incident sensitivity and dynamic specificity definition and produces accuracy measures for censored data under proportional or nonproportional hazard assumption of Cox regression estimator.³³ Following the objective of our study, we also calculated the iAUC at six months for all models.

Statistical and survival analyses were performed with R version 4.0.5 (R Foundation for Statistical Computing, Vienna, Austria). The differences in age, OS, the proportion of right-censored cases and short-term (< 6 months) survival cases were assessed using Student's t-test, Mann-Whitney U test and two-proportions Z-test, respectively. For the binary classification model, we used Orange version 3.28.0 (University of Ljubljana, Slovenia).³⁴ The radiomics quality score (RQS) was calculated for this study according to the recommendations by Lambin et al.³⁵ A p value < .05 was considered to indicate a statistically significant difference. The image processing and statistical analysis workflow are shown in Figure 1.

Results

Patient population

Two hundred and three patients were enrolled in this study. The mean age was 61.49 ± 11.76 (range 27.81 – 86.65). The median OS was 407 days [interquartile range (IQR) = 351.5]. A total of 7.9% (16) of patients were right-censored cases, and 17.24% (35) registered an OS of less than six months. The patients were randomly assigned to a training data set of 143 patients and a testing data set of 60 patients. There were no significant differences in age, OS, or proportion of right-censored and short-term survival cases between the training and testing data sets. (Table 1)

Table 1. Patient characteristics					
Dataset	n	Age (SD)	OS (IQR)	Censored (%)	OS < 6 m (%)
BraTS	119	62 ± 12	374 (364)	0	21.8 % (26)
TCIA	34	60.3 ± 10	414 (482)	5.9 % (2)	17.6 % (6)
Institution 1	22	65.3 ± 10	451 (307)	22.7 % (5)	13.6 % (3)
Institution 2	28	57.9 ± 13	466 (217)	32.1 % (9)	0
<i>After random split (70/30)</i>					
Training	143	62.1 ± 11	409 (311)	7.7 % (11)	17.5 % (25)
Testing	60	60.1 ± 13	404 (392)	8.3 % (5)	16.7 % (10)
Statistical comparison between cohorts		$t = 0.10, p = .273$	$U = .33, p = .743$	$c^2 = 0.02, p = .877$	$c^2 = 1.17, p = .279$
BraTS = Brain Tumor Segmentation Challenge 2020, TCIA = The Cancer Image Archive, SD = standard deviation, IQR = interquartile range					

Classification task and survival groups

The variable selection filters made it possible to determine the top 10 radiomic features. (Supplementary Table 3). Based on these characteristics and using the ML classifiers, patients were classified into short-term survivors (< 6 months).

The optimal results were obtained by applying the information gain as a feature selector. Thus, in the training cohort, AUC values were achieved with a range between 0.802 and 0.978, a classification accuracy between 81.8% and 94.4%, and a precision between 82.8% and 94.8%. (Supplementary Table 4)

In the testing cohort, the naive Bayes classifier obtained the best results, with an AUC of 0.769, a classification accuracy of 80%, and a precision of 81%. (Table 2 and Figure 2)

Classifier	Filter	AUC	CA	Precision	F1
Naive Bayes	Information Gain	0.769	0.800	0.812	0.805
	Gini Index	0.784	0.767	0.810	0.781
	FCBF	0.743	0.783	0.803	0.791
k-Nearest Neighbor	Information Gain	0.600	0.817	0.806	0.776
	Gini Index	0.639	0.800	0.771	0.763
	FCBF	0.670	0.767	0.713	0.724
Neural Network	Information Gain	0.691	0.800	0.771	0.763
	Gini Index	0.682	0.767	0.691	0.705
	FCBF	0.722	0.817	0.806	0.776
Random Forest	Information Gain	0.574	0.733	0.683	0.701
	Gini Index	0.666	0.750	0.696	0.713
	FCBF	0.700	0.783	0.738	0.735
Support Vector Machine	Information Gain	0.709	0.750	0.608	0.671
	Gini Index	0.630	0.750	0.608	0.671
	FCBF	0.730	0.800	0.777	0.747

AUC: area under the curve; CA: classification accuracy; FCBC: fast correlation based filter.

Random Survival Forest to predict OS

In the radiomic model based on RSF, the variable-hunting algorithm selected 17 radiomic features to predict OS in the training dataset. Based on these characteristics, the mortality risk score was calculated using the OOB-CHF. The cut-off point used was 0.684. This cutoff point allowed patients to be stratified into low-risk and high-risk groups [HR = 2.19, (95% CI: 1.54 - 3.12), log-rank p = <.001, C-Index = 0.61, IBS = 0.096]. In the testing dataset, patients were also stratified using the same cutoff point [HR 2.16, (95% CI: 1.21 - 3.89), log-rank p = .008, C-Index = 0.61, IBS = 0.123]. The multivariate Cox regression models in which age was incorporated as an explanatory variable are shown in Table 3.

Table 3. Univariate and Multivariate Cox regression analysis								
Univariate Cox Regression Analysis								
Variable	b	HR	95% CI	p	C - Index	IBS	iAUC	6m - iAUC
<i>Training dataset</i>								
Age	0.03	1.03	1.02 - 1.05	< .001	0.61	0.089	0.604	0.599
Radiomic RSF Score (High risk)	0.78	2.19	1.54 - 3.12	< .001	0.61	0.096	0.591	0.712
<i>Testing dataset</i>								
Age	0.03	1.03	1 - 1.06	0.023	0.60	0.128	0.592	0.643
Radiomic RSF Score (High risk)	0.77	2.16	1.21 - 3.89	0.009	0.61	0.123	0.568	0.761
Multivariate Cox Regression Model								
Model	Likelihood Ratio Test			C - Index	IBS	iAUC	6m - AUC	
	c ²	df	p					
<i>Training dataset</i>								
Age + Radiomic RSF Score (High risk)	36.93	2	< .001	0.66	0.084	0.650	0.730	
<i>Testing dataset</i>								
Age + Radiomic RSF Score (High risk)	11.35	2	.003	0.68	0.118	0.6278	0.769	
HR= Hazard Ratio, RSF = Random Survival Forest, IBS = Integrated Brier Score, iAUC = Integrated area under the curve, 6m - AUC = six months - area under the curve								

The iAUC of the radiomic model was 0.591 in the training cohort and 0.568 in the testing cohort. By incorporating age as a variable in the model, the iAUC increased to 0.650 in the training cohort and 0.627 in the testing dataset.

After setting the survival time to 6 months, the predictive accuracy of the radiomic model improved to an iAUC of 0.712 in the training dataset and 0.761 in the testing dataset. (Supplementary Figure S1)

The RQS was used to evaluate the methodological quality of our study. We obtained a score of 19/36 (53 %). A detailed report of RQS items is shown in Supplementary Table 7.

Discussion

In the present study, we elaborated a prediction model of short-term survival with high predictive capacity using the radiomic features of structural preoperative multiparametric MRI of GBM patients.

We believe that the main strength of our study is based on a selection of patients who underwent total or near-total resection of the enhancing tumor. We considered this methodologic aspect due to the undeniable link between the extent of resection and survival in these patients.^{36,37} In most previous studies, the extent of resection was not used as a selection criterion, including partial resections and biopsies in their series, without making any adjustment during the analysis phase. The exception is the

studies by Bakas et al.¹⁵ and Fathi et al.³⁸, in which the entire cohort of patients underwent complete resection and standard chemoradiotherapy treatment.

Another crucial point of our work is to set our objective to identify short-term survival patients, in contrast to previously published studies where 10 and 12 - 15 months were used as cutoff points for defining short- and long-term survival, respectively.^{10,11,13,39-44} The only reference we found is in the work of Prasanna et al.⁴⁵, who classified patients in long (> 18 months) versus short-term (< 7 months) survival based on peritumoral region radiomic features. The rationale of our approach lies in the desire to predict the survival of patients diagnosed with GBM by noninvasive methods and to identify those with very short survival. In these patients, the futility of our treatments would lead us to offer patients and their families the option of not taking aggressive measures or, on the contrary, opening new lines of research since those cases would be poor responders to the standard therapies applied currently.

As another strength of our work, we can mention the use of open-source software. The CapTk and Orange programs have a very intuitive yet robust user interface, thanks to which clinicians can access advanced image processing technics and data mining tools. Thanks to these programs, we have performed complex tasks such as automatic tumor segmentation, image processing, radiomic feature extraction, and exploring different ML-based algorithms.

Concerning statistical analysis, we used a dual approach. On the one hand, we have used a binary classification system using different ML-based algorithms. Additionally, we used state-of-the-art survival analysis techniques such as random survival forest and time-dependent ROC curve analysis focused on short-term survival that contribute to corroborating the stability of the models produced here.

We also highlight that the results of our predictive models have been achieved using only structural MRI.¹⁵ These results could even be improved after the inclusion of studies based on diffusion and perfusion sequences.⁴⁶ However, basic MRI is available in most centers, and according to our results, the lack of special sequences is not a limitation in the search for useful radiological patterns in clinical practice.

An important aspect to discuss is the biological correspondence of the variables employed by the prediction models. There is notable variability concerning the radiomic characteristics used by previous studies, which is one of the most significant obstacles in reproducing and validating their results. In our study, most of the selected variables came from the T1CE sequence followed by FLAIR and T2WI, while the different tumor subregions (i.e., ET, NET and ED) were represented in the models in a balanced way. In our series, first-order features and morphological characteristics appeared to be important for OS prediction.

We are aware of the limitations of our work, such as the lack of clinical and molecular data that can be incorporated into predictive models. Even so, age as an explanatory variable was incorporated into our models due to its significant association with the OS of these patients, proving that its mere incorporation

into the analysis improved the performance of the models. Despite having a relatively small sample size, various statistical techniques have been applied to overcome the "curse of dimensionality". Taking into consideration that MRI studies come from numerous sources, the processing method for image standardization that we have chosen aims to be simple and reliable and has been used by several studies.^{38,47,48}

Unquestionably, the combination of texture analysis and artificial intelligence is starting to facilitate knowledge about the biological behaviour of GBMs through the study of their patient-dependent heterogeneity. However, the rapid development of big data tools and the tremendous complexity of advanced medical image analysis dangerously threaten widening the gap between data experts and clinicians. Then, it is a paradox that radiomics, defined by Lambin et al.³⁵ as "the bridge between medical imaging and personalized medicine", is now out of reach of those who treat real patients every day. Therefore, our study arises from a real need and aims to find a solution to a clinically relevant problem: identifying GBM patients with short survival after complete resection. Although our results can be improved, we show that there are currently computer tools and public data sets available to everyone to develop reliable predictive models. Hence, our duty as clinicians is to become immersed in developing these models since our pragmatism can never be replaced even by the most complex algorithm.

Indeed, our results are encouraging, and the precision achieved is similar to the previous literature. However, this article represents an early age of a promising future in which the ultimate link between image, diagnosis and prognosis could finally be decoded to provide instant, useful and precise information to individual patients based on their specific features. Multiinstitutional studies⁴⁹ would allow the generalization of predictive models or even adapt the mechanisms of data preprocessing, extraction, and analysis to the MRI from each center since the standardization of acquisition protocols is not feasible. Finally, we believe that in this catastrophic disease, the quality of life of our patients should be our first consideration, and maximum exploitation of available neuroimaging techniques should be pursued to optimize management strategies avoiding unnecessarily aggressive therapies in those patients who will not benefit from them.

Conclusion

In the present study, we evaluated the capability of the radiomic features of preoperative mpMRI and machine learning-based classification and regression analysis to predict short-term survival in GBM patients. Our model shows an impressive classification accuracy of 80% and an iAUC of 0.76 to predict OS < 6 months in the validation cohort. We believe that these new tools will serve clinicians to understand the biological behaviour of individual GBMs, and we must take advantage of them.

Declarations

Funding

No funding was received for this research.

Conflict of interest: All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements) or nonfinancial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

References

1. Koshy M, Villano JL, Dolecek TA, et al. Improved survival time trends for glioblastoma using the SEER 17 population-based registries. *J Neurooncol.* 2012;107(1):207-212.
2. Stupp R, Hegi ME, Mason WP, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol.* 2009;10(5):459-466.
3. Sanai N, Berger MS. Surgical oncology for gliomas: the state of the art. *Nat Rev Clin Oncol.* 2018;15(2):112-125.
4. Gutman DA, Cooper LAD, Hwang SN, et al. MR Imaging Predictors of Molecular Profile and Survival: Multi-institutional Study of the TCGA Glioblastoma Data Set. *Radiology.* 2013;267(2):560-569.
5. Chaddad A, Kucharczyk MJ, Daniel P, et al. Radiomics in glioblastoma: Current status and challenges facing clinical implementation. *Front Oncol.* 2019;9(MAY):1-9.
6. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology.* 2016;278(2):563-577.
7. Seow P, Wong JHD, Ahmad-Annuar A, Mahajan A, Abdullah NA, Ramli N. Quantitative magnetic resonance imaging and radiogenomic biomarkers for glioma characterisation: a systematic review. *Br J Radiol.* 2018;91(1092):20170930.
8. Vaidya T, Agrawal A, Mahajan S, Thakur MH, Mahajan A. The Continuing Evolution of Molecular Functional Imaging in Clinical Oncology: The Road to Precision Medicine and Radiogenomics (Part II). *Mol Diagn Ther.* 2019;23(1):27-51.
9. Osman AFI. A Multi-parametric MRI-Based Radiomics Signature and a Practical ML Model for Stratifying Glioblastoma Patients Based on Survival Toward Precision Oncology. *Front Comput Neurosci.* 2019;13:58.

10. Kickingereeder 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.
11. Bae S, Choi YS, Ahn SS, et al. Radiomic MRI Phenotyping of Glioblastoma: Improving Survival Prediction. *Radiology*. 2018;289(3):797-806.
12. Shboul ZA, Alam M, Vidyaratne L, Pei L, Elbakary MI, Iftekharruddin KM. Feature-Guided Deep Radiomics for Glioblastoma Patient Survival Prediction. *Front Neurosci*. 2019;13:966.
13. Chaddad A, Daniel P, Sabri S, Desrosiers C, Abdulkarim B. Integration of Radiomic and Multi-omic Analyses Predicts Survival of Newly Diagnosed IDH1 Wild-Type Glioblastoma. *Cancers (Basel)*. 2019;11(8).
14. Baid U, Rane SU, Talbar S, et al. Overall Survival Prediction in Glioblastoma With Radiomic Features Using Machine Learning. *Front Comput Neurosci*. 2020;14(August):1-9.
15. Bakas S, Shukla G, Akbari H, et al. Overall survival prediction in glioblastoma patients using structural magnetic resonance imaging (MRI): advanced radiomic features may compensate for lack of advanced MRI modalities. *J Med imaging (Bellingham, Wash)*. 2020;7(3):031505.
16. Menze BH, Jakab A, Bauer S, et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Trans Med Imaging*. 2015;34(10):1993-2024.
17. Bakas S, Reyes M, Jakab A, et al. Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge. Published online November 5, 2018. <http://arxiv.org/abs/1811.02629>
18. Bakas S, Akbari H, Sotiras A, et al. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci Data*. 2017;4(1):170117.
19. Puchalski RB, Shah N, Miller J, et al. An anatomic transcriptional atlas of human glioblastoma. *Science (80-)*. 2018;360(6389):660-663.
20. Clark K, Vendt B, Smith K, et al. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *J Digit Imaging*. 2013;26(6):1045-1057.
21. Rohlfing T, Zahr NM, Sullivan E V, Pfefferbaum A. The SRI24 multichannel atlas of normal adult human brain structure. *Hum Brain Mapp*. 2010;31(5):798-819.
22. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging*. 2010;29(6):1310-1320.

23. Bakas S, Akbari H, Sotiras A, et al. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci data*. 2017;4:170117.
24. Thakur S, Doshi J, Pati S, et al. Brain extraction on MRI scans in presence of diffuse glioma: Multi-institutional performance evaluation of deep learning methods and robust modality-agnostic training. *Neuroimage*. 2020;220:117081.
25. Davatzikos C, Rathore S, Bakas S, et al. Cancer imaging phenomics toolkit: quantitative imaging analytics for precision diagnostics and predictive modeling of clinical outcome. *J Med imaging (Bellingham, Wash)*. 2018;5(1):011018.
26. Bakas S, Zeng K, Sotiras A, et al. GLISTRboost: Combining Multimodal MRI Segmentation, Registration, and Biophysical Tumor Growth Modeling with Gradient Boosting Machines for Glioma Segmentation. *Brainlesion glioma, Mult sclerosis, stroke Trauma brain Inj BrainLes*. 2016;9556:144-155.
27. Zwanenburg A, Vallières M, Abdalah MA, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*. 2020;295(2):328-338.
28. Ishwaran H, Kogalur UB. Fast unified random forests for survival, regression, and classification (RF-SRC). Available at <https://CRAN.R-project.org/package=randomForestSRC>. Accessed March 31, 2021.
29. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *Ann Appl Stat*. 2008;2(3).
30. Dardis C. survMisc: Miscellaneous Functions for Survival Data. Available at <https://CRAN.R-project.org/package=survMisc>. Accessed July 5, 2018.
31. Mogensen UB, Ishwaran H, Gerds TA. Evaluating Random Forests for Survival Analysis using Prediction Error Curves. *J Stat Softw*. 2012;50(11):1-23.
32. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics*. 2005;61(1):92-105.
33. Kamarudin AN, Cox T, Kolamunnage-Dona R. Time-dependent ROC curve analysis in medical research: current methods and applications. *BMC Med Res Methodol*. 2017;17(1):53.
34. Demšar J, Curk T, Erjavec A, et al. Orange: Data Mining Toolbox in Python. *J Mach Learn Res*. 2013;14(1):2349–2353.
35. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14(12):749-762.

36. Sanai N, Polley M-Y, McDermott MW, Parsa AT, Berger MS. An extent of resection threshold for newly diagnosed glioblastomas. *J Neurosurg*. 2011;115(1):3-8.
37. Awad A-W, Karsy M, Sanai N, et al. Impact of removed tumor volume and location on patient outcome in glioblastoma. *J Neurooncol*. 2017;135(1):161-171.
38. Fathi Kazerooni A, Akbari H, Shukla G, et al. Cancer Imaging Phenomics via CaPTk: Multi-Institutional Prediction of Progression-Free Survival and Pattern of Recurrence in Glioblastoma. *JCO Clin Cancer Informatics*. 2020;(4):234-244.
39. Kickingereder P, Neuberger U, Bonekamp D, et al. Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma. *Neuro Oncol*. 2018;20(6):848-857.
40. Lu Y, Patel M, Natarajan K, et al. Machine learning-based radiomic, clinical and semantic feature analysis for predicting overall survival and MGMT promoter methylation status in patients with glioblastoma. *Magn Reson Imaging*. 2020;74:161-170.
41. Kickingereder P, Götz M, Muschelli J, et al. Large-scale radiomic profiling of recurrent glioblastoma identifies an imaging predictor for stratifying anti-angiogenic treatment response. *Clin Cancer Res*. 2016;22(23):5765-5771.
42. Ingrisch M, Schneider MJ, Nörenberg D, et al. Radiomic Analysis Reveals Prognostic Information in T1-Weighted Baseline Magnetic Resonance Imaging in Patients With Glioblastoma. *Invest Radiol*. 2017;52(6):360-366.
43. Zhang X, Lu H, Tian Q, et al. A radiomics nomogram based on multiparametric MRI might stratify glioblastoma patients according to survival. *Eur Radiol*. 2019;29(10):5528-5538.
44. Liu Y, Xu X, Yin L, Zhang X, Li L, Lu H. Relationship between Glioblastoma Heterogeneity and Survival Time: An MR Imaging Texture Analysis. *AJNR Am J Neuroradiol*. 2017;38(9):1695-1701.
45. Prasanna P, Patel J, Partovi S, Madabhushi A, Tiwari P. Radiomic features from the peritumoral brain parenchyma on treatment-naïve multi-parametric MR imaging predict long versus short-term survival in glioblastoma multiforme: Preliminary findings. *Eur Radiol*. 2017;27(10):4188-4197.
46. Kim JY, Yoon MJ, Park JE, Choi EJ, Lee J, Kim HS. Radiomics in peritumoral non-enhancing regions: fractional anisotropy and cerebral blood volume improve prediction of local progression and overall survival in patients with glioblastoma. *Neuroradiology*. 2019;61(11):1261-1272.
47. Davatzikos C, Rathore S, Bakas S, et al. Cancer imaging phenomics toolkit: quantitative imaging analytics for precision diagnostics and predictive modeling of clinical outcome. *J Med Imaging*. 2018;5(01):1.

48. Rathore S, Bakas S, Pati S, et al. Brain Cancer imaging phenomics toolkit (brain-CaPTk): An interactive platform for quantitative analysis of glioblastoma. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2018;10670 LNCS(September):133-145.
49. Davatzikos C, Barnholtz-Sloan JS, Bakas S, et al. AI-based prognostic imaging biomarkers for precision neuro-oncology: the ReSPOND consortium. *Neuro Oncol*. 2020;22(6):886-888.

Figures

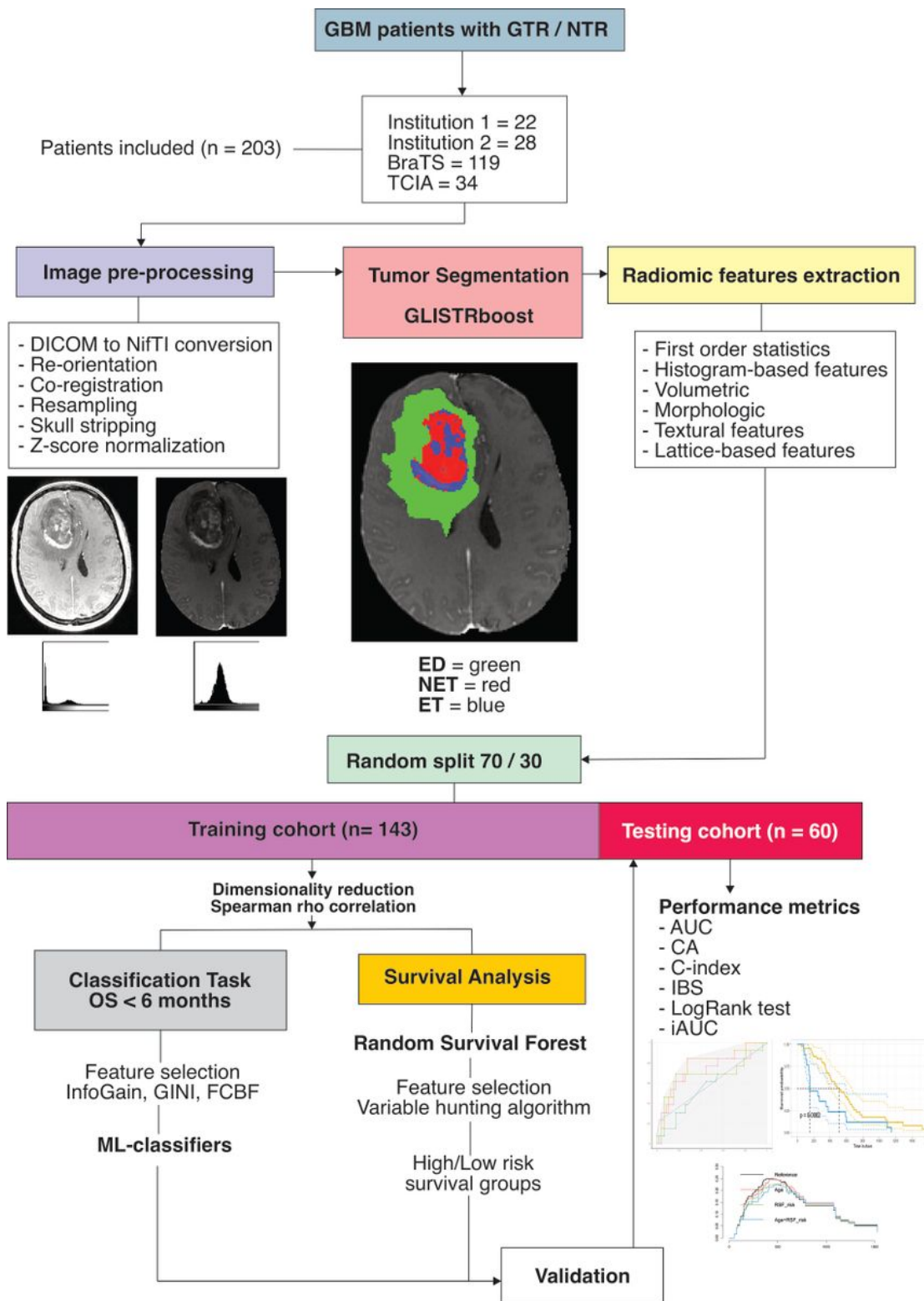


Figure 1

Study workflow. GBM = glioblastoma, GTR = gross-total resection, NTR = near-total resection, BraTS = Brain tumor segmentation challenge 2020, TCIA = The cancer image archive, GLISTRboost = Boosted GLIoma Image SegmenTation and Registration, ED = edema, NET = non-enhancing tumor, ET = enhancing tumor, OS = overall survival, InfoGain = Information gain, GINI = Gini Index, FCBF = Fast

correlation based filter, ML = machine learning, AUC = area under the curve, CA = classification accuracy, IBS = integrated Brier score, iAUC = integrated AUC.

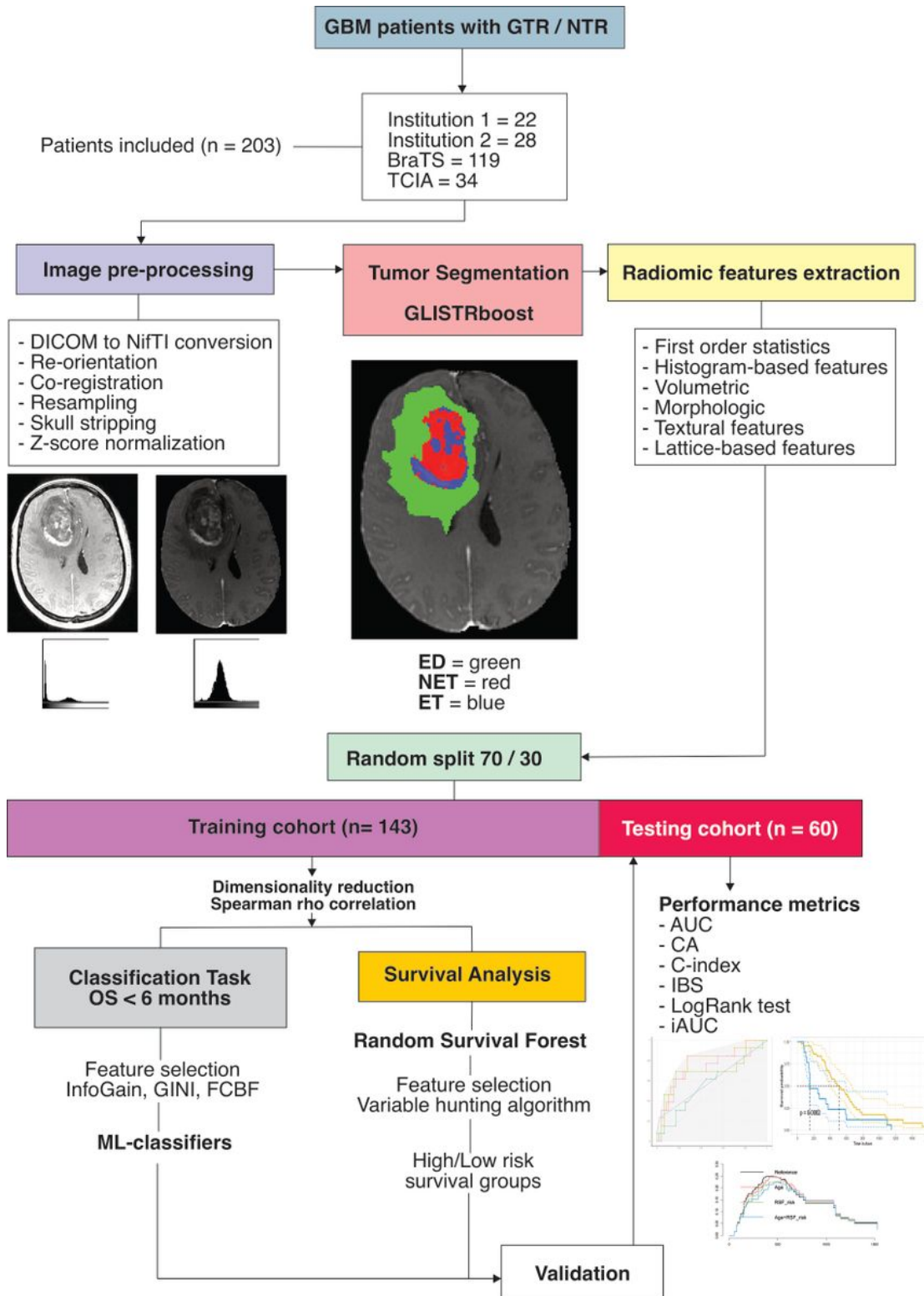


Figure 2

Study workflow. GBM = glioblastoma, GTR = gross-total resection, NTR = near-total resection, BraTS = Brain tumor segmentation challenge 2020, TCIA = The cancer image archive, GLISTRboost = Boosted GLioma Image SegmenTation and Registration, ED = edema, NET = non-enhancing tumor, ET =

enhancing tumor, OS = overall survival, InfoGain = Information gain, GINI = Gini Index, FCBF = Fast correlation based filter, ML = machine learning, AUC = area under the curve, CA = classification accuracy, IBS = integrated Brier score, iAUC = integrated AUC.

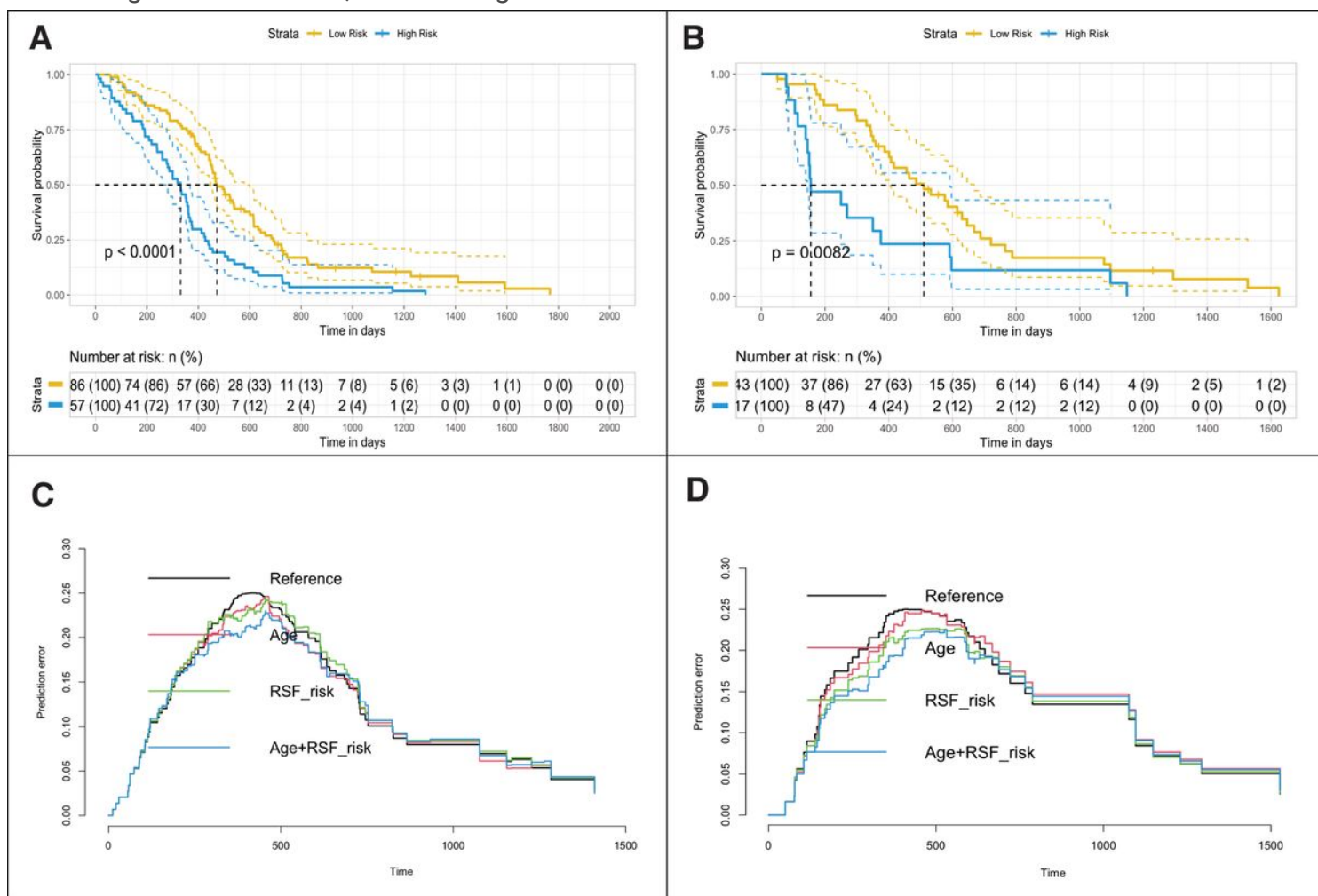


Figure 3

Kaplan-Meier plots showing differences in overall survival for patients in training (A) and testing (B) data sets, stratified in low or high-risk groups by the Random Survival Forest (RSF)-based ensemble mortality. Prediction error curves show Cox regression model performance in training (C) and testing (D) cohorts.

Supplementary Files

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

- [DataSupplement.docx](#)