European Radiology (2019) 29:1968–1977
https://doi.org/10.1007/s00330-018-5758-7

HEAD AND NECK

Morphological MRI-based features provide pretreatment survival prediction in glioblastoma

Julián Pérez-Beteta[1] · David Molina-García[1] · Alicia Martínez-González[1] · Araceli Henares-Molina[1] · Mariano Amo-Salas[1] · Belén Luque[1] · Elena Arregui[2] · Manuel Calvo[2] · José M. Borrás[3] · Juan Martino[4] · Carlos Velásquez[4] · Bárbara Meléndez-Asensio[5] · Ángel Rodríguez de Lope[6] · Raquel Moreno[7] · Juan A. Barcia[8] · Beatriz Asenjo[9] · Manuel Benavides[10] · Ismael Herruzo[11] · Pedro C. Lara[12] · Raquel Cabrera[12] · David Albillo[13] · Miguel Navarro[14] · Luis A. Pérez-Romasanta[15] · Antonio Revert[16] · Estanislao Arana[17] · Víctor M. Pérez-García[1]

Received: 23 May 2018 / Revised: 19 August 2018 / Accepted: 12 September 2018 / Published online: 15 October 2018

Abstract

Objectives We wished to determine whether tumor morphology descriptors obtained from pretreatment magnetic resonance images and clinical variables could predict survival for glioblastoma patients.

Methods A cohort of 404 glioblastoma patients (311 discoveries and 93 validations) was used in the study. Pretreatment volumetric postcontrast T1-weighted magnetic resonance images were segmented to obtain the relevant morphological measures. Kaplan-Meier, Cox proportional hazards, correlations, and Harrell's concordance indexes (c-indexes) were used for the statistical analysis.

Results A linear prognostic model based on the outstanding variables (age, contrast-enhanced (CE) rim width, and surface regularity) identified a group of patients with significantly better survival ($p < 0.001$, HR = 2.57) with high accuracy (discovery c-index = 0.74; validation c-index = 0.77). A similar model applied to totally resected patients was also able to predict survival ($p < 0.001$, HR = 3.43) with high predictive value (discovery c-index = 0.81; validation c-index = 0.92). Biopsied patients with better survival were well identified ($p < 0.001$, HR = 7.25) by a model including age and CE volume (c-index = 0.87).

Conclusions Simple linear models based on small sets of meaningful MRI-based pretreatment morphological features and age predicted survival of glioblastoma patients to a high degree of accuracy. The partition of the population using the extent of resection improved the prognostic value of those measures.

Key Points

- *A combination of two MRI-based morphological features (CE rim width and surface regularity) and patients' age outperformed previous prognosis scores for glioblastoma.*
- *Prognosis models for homogeneous surgical procedure groups led to even more accurate survival prediction based on Kaplan-Meier analysis and concordance indexes.*

Keywords Glioblastoma · Prognosis · Biomarkers · Survival analysis · Multivariate analysis

Estanislao Arana and Víctor M. Pérez-García are joint senior authors.

Julián Pérez-Beteta and David Molina-García contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-018-5758-7) contains supplementary material, which is available to authorized users.

✉ David Molina-García
david.molina@uclm.es

Extended author information available on the last page of the article

Abbreviations

3D	Three-dimensional
c-index	Concordance index
CE	Contrast-enhanced
CoI	Confidence interval
DICOM	Digital imaging and communication in medicine
GBM	Glioblastoma
HR	Hazard ratio
MA	Morphology- and age-based
MAB	Morphology- and age-based prognosis score for biopsied patients

MASR	Morphology- and age-based prognosis score for subtotally resected patients
MATR	Morphology- and age-based prognosis score for totally resected patients
MRI	Magnetic resonance images
OS	Overall survival
P	*p* value
TCIA	The Cancer Image Archive
WHO	World Health Organization

Introduction

The computation of quantitative—radiomic—features from magnetic resonance images (MRI) in order to develop imaging biomarkers has gained attention in recent years, due to its potential practical applications. The main premise of radiomics is that clinical endpoints could be more associated with quantitative voxel-based features than with the more qualitative radiological and clinical data used today [1, 2].

Glioblastoma (GBM) is the most common and lethal malignant primary brain tumor. Many authors have recently investigated the use of different types of radiomic quantifiers (volumetric, textural, and geometrical MRI-based features) as predictors of survival in GBM [3–15].

The most common treatment for newly diagnosed GBM consists of surgery, if possible, followed by radiotherapy plus adjuvant and concomitant temozolomide [16]. Several studies have related the extent of GBM resection with patient survival [17–21].

Many studies have looked at prognostic parameters obtained from pretreatment volumetric contrast-enhanced (CE) T1-weighted MRIs [3, 5–15] for GBM. Tumor volumes (total, CE, and necrotic) in GBM studies have been identified as prognostic factors by some studies [6, 13, 17] but not by others [4, 9, 14]. Contradictory results have been reported for other features, such as tumor surface [8, 9, 14]. A recent study has found volume sizes before the second line of treatment to be relevant for survival [15].

In this study, we set out to assess the prognostic value of small sets of clinical variables and meaningful morphological MRI-based imaging biomarkers, some based on biologically-inspired mathematical models. Our study was performed on a large set of 404 GBM patients. We further intended to compare this mechanistic approach with more standard "radiomic" methods.

Materials and methods

Patients

The study was approved by the Institutional Review Board of the participating institutions. Informed consent was waived as patients had previously provided authorization for use of their medical records for research.

A total of 1155 GBM patients from ten local medical institutions and diagnosed in the period 2006–2017 with pathologically proven GBM according to the 2007 World Health Organization (WHO) Classification of Tumors of the Central Nervous System were retrospectively reviewed for this study. Inclusion criteria were unifocal primary tumors; availability of the relevant clinical variables: age, extent of resection performed (total, partial, or biopsy), treatment scheme, and survival information; and availability of a pretreatment volumetric CE T1-weighted MRI sequence (slice thickness ≤ 2.00 mm, no gap, pixel spacing ≤ 1.20 mm). Exclusion criteria were substantial imaging artifacts in the images, tumors without contrast-enhancement, and diffuse images. Eight hundred forty-four patients (73.07%) were excluded due to no clinical data available (182), no volumetric T1-weighted sequence (352), multifocal tumors (97), diffuse T1 images (67), faint or no contrast enhancement (83), or presence of artifacts (63).

Thus, 311 patients (26.93%) from local institutions made up the discovery cohort. A validation cohort of 93 GBM patients meeting the inclusion/exclusion criteria were obtained from The Cancer Image Archive (TCIA) [22], so a total of 404 GBM patients were finally included in the study. Section S1 contains a list of the TCIA patients included in the study.

The extent of resection was determined from the CE T1-weighted MRI within 72 h after surgery. Gross total resection was defined as the absence of visible CE and subtotal resection as any remnant nodular tumor enhancement after surgery. Overall survival (OS) was computed from the date of the preoperative MRI, up to death or last follow-up examination date (censored events).

Table 1 summarizes the main patient characteristics, showing a standard GBM patient population with median survival of 12.76 months and median age of 63 years.

Image acquisition

Postcontrast T1-weighted sequence was gradient echo using 3D spoiled-gradient recalled echo or 3D fast-field echo after intravenous administration of a single-dose of gadobenate dimeglumine (0.10 mmol/kg) with a (6–8)-min delay.

All MR images were acquired in the axial plane with a 1.5-T ($n = 278$) or 3-T ($n = 33$) imager with the following range of parameters: repetition time/echo time, 6–25 ms/3–10 ms. Regarding the MRI manufacturer, General Electric ($n = 136$), Philips ($n = 108$) and Siemens ($n = 67$) scanners were used. Other image and patient characteristics are summarized in Table 1.

Table 1 Summary of patient characteristics, MR imaging, and morphological parameters for the groups of patients considered

		Discovery cohort	Validation cohort
Patient characteristics	Number of patients (censored)	311 (27)	93 (12)
	Age (years) median (range)	63 (19–86)	62 (14–86)
	Sex (male (M), female (F))	44% F; 56% M	47% F; 53% M
	Survival (months) median (range)	12.76 (0.13–82.97)	11.77 (0.72–59.20)
	Extent of resection (total, subtotal, or biopsy)	149 Total (47.91%)	17 Total (18.28%)
		113 Subtotal (36.33%)	–
		49 biopsy (15.76%)	8 biopsy (8.60%)
		–	68 unknown (73.12%)
	Type of treatment (chemotherapy (CT) and radiotherapy (RT))	241 CT + RT (77.49%)	74 CT + RT (79.57%)
		27 RT alone (8.68%)	4 RT alone (4.30%)
		5 CT alone (1.61%)	4 CT alone (4.30%)
		38 no treatment (12.22%)	11 no treatment (11.83%)
MRI characteristics	Pixel spacing (mm) mean (range)	0.81 (0.46–1.09)	0.90 (0.45–1.06)
	Slice thickness (mm) mean (range)	1.48 (1.00–2.00)	1.41 (0.90–2.00)
	Number of slices mean (range)	174 (80–360)	150 (72–305)
Volumetric parameters	Tumor volume (cm^3) mean (range)	33.14 (0.48–132.54)	41.82 (2.47–116.12)
	CE volume (cm^3) mean (range)	19.64 (0.44–90.06)	24.90 (2.46–90.95)
	Necrotic volume (cm^3) mean (range)	13.50 (0.03–89.31)	16.92 (0.00–69.20)
	CE rim width (cm) mean (range)	0.57 (0.22–1.65)	0.63 (0.24–1.25)
	Maximum diameter (cm) mean (range)	5.11 (1.30–11.09)	5.71 (2.55–9.80)
	Total surface (cm^2) mean (range)	67.27 (3.00–226.32)	83.13 (13.27–196.04)
	Surface regularity mean (range)	0.62 (0.24–0.99)	0.57 (0.30–0.83)

Image analysis

All MRIs were retrospectively analyzed by the same image expert (JPB, 5 years of expertise on tumor segmentation). Since we wished to extract morphological features from images, they were not normalized on their gray-level values, as higher variability could be induced. The DICOM files were imported into the scientific software package Matlab (R2017b, The MathWorks, Inc., Natick, MA, USA). Tumors were semi-automatically delineated using a gray-level threshold chosen to identify the CE tumor volume. Then, segmentations were corrected manually slice by slice. Necrotic tissue was defined as hypointense tumoral regions inside CE tissue. An in-house software was developed by the image expert (JPB) under Matlab software based on its image processing toolboxes, allowing the segmentations to be corrected on a tablet using a digital pencil. An advisory board formed by three medical specialists, each with more than 10 years MR imaging reading of GBM (a radiologist, a neurosurgeon, and a radio-oncologist) revised and validated the segmentations performed. The results were discussed with the advisory board and difficult cases segmented in consensus readings. A reproducibility study for the methodology was performed in [9], showing its reliability.

Geometrical measures

We computed several 3D quantitative morphological measures on the segmented tumors: total volume, CE volume, necrotic volume, CE rim width, maximum 3D diameter, total surface, and surface regularity. Section S2 describes the measures computed in the study.

Statistical analysis

Kaplan-Meier analysis was used to identify parameters associated with prognosis, using both Log-Rank (non-crossing curves) and Breslow (crossing curves) tests to assess the significance of the results. A two-tailed significance level of p value (p) lower than 0.05 was applied. For each parameter, we searched for every threshold value splitting the sample into two different subgroups. Then, we chose the non-isolated significant value obtaining the lowest Log-Rank P [6]. Univariate Cox proportional hazards regression analysis was used to obtain the hazard ratio (HR) and an adjusted 95% confidence interval (CoI) for each threshold.

A Spearman's correlation coefficient was used to assess the dependences between every pair of variables. Normality of the variables was assessed by the Kolmogorov-Smirnov test. Significant correlation coefficients over 0.7 or below -0.7 were regarded as strong correlation between variables.

Multivariate Cox proportional hazards analysis with the stepwise Wald method was used to construct prognostic models. This method evaluates a set of variables and discards, step by step, the one with the worst statistical significance until every remaining variable shows statistical significance. SPSS software (v. 22.0.00) was used for the statistical analysis.

The Harrell's concordance index (c-index) was computed to evaluate the model's performance [23]. To compute the c-index, we first constructed the prognosis score using the continuous Cox regression model using the discovery groups. The prognosis score was then computed for each patient. We next computed the c-index for each possible threshold splitting the patient population into two groups and searched for its optimal value using the optimization procedure described above. When a validation group was available, we used both the model and the threshold found for the discovery group to test the accuracy of the predictions. We considered c-index values ranging from 0.75 (i.e., a correct ordering in the survival of 3 out of 4 patients) to 1 (the best possible result for perfect classification) as indicators of good predictors.

Results

Results obtained for the discovery and validation cohorts

Kaplan-Meier analysis showed no differences between the discovery and validation cohorts ($p = 0.578$, HR = 1.07).

Parameters achieving statistical significance in the Kaplan-Meier analysis for the discovery cohort were age, CE rim width, and surface regularity. The favorable subgroups of patients were those with lower age, thinner CE rims, and larger surface regularity. No statistically significant robust thresholds were found for the remaining variables. Table 2 shows the best results obtained for the univariate analysis.

All size-related measures (volumes, diameter, and surface) showed high and significant correlations with each other. Correlations between other parameters were low (Fig. S1A).

For the validation cohort, the thresholds obtained for the discovery group were used to test the robustness of the results. Age, CE rim width, and surface regularity were confirmed to be uncorrelated (Fig. S1B) and the only statistically significant variables (Table 2).

A prognosis score was constructed for the discovery cohort using the multivariate proportional-hazard Cox model using the significant parameters: the morphology- and age-based (MA) prognosis score:

$$\mathrm{MA} = 0.035 \times \text{age} - 0.271 \times \text{CE rim width} - 1.371 \times \text{surface regularity}$$

This score was highly significant in the Kaplan-Meier analysis ($p < 0.001$, HR = 2.57), with a substantial difference in median survival between the two groups (9.83 months, Fig. 1a).

The MA model provided a c-index value of 0.741, substantially improving the results for the measures considered individually: 0.704 for age, 0.590 for the CE rim width, and 0.585 for the surface regularity. The same threshold obtained for the discovery group and the MA prognosis separated the validation cohort into two groups with high significance in the Kaplan-Meier analysis (Fig. 1b) and a c-index value of 0.769.

Prognosis prediction for totally resected patients

In the discovery cohort, 149 patients received a macroscopically complete resection. The Kaplan-Meier analysis developed on this population provided the same three significant uncorrelated (Fig. S1C) parameters as in the full discovery cohort: age, CE rim width, and surface regularity. Table 3 summarizes the results for this patient subgroup.

The MA prognosis score obtained for totally resected patients (MATR) was

$$\mathrm{MATR} = 0.036 \times \text{age} - 0.228 \times \text{CE rim width} - 1.396 \times \text{surface regularity}$$

This score provided well-separated groups according to the Kaplan-Meier analysis (Fig. 1c).

The best c-indexes on the univariate analysis for this patient group were achieved by age (0.700), CE rim width (0.649), and surface regularity (0.630) respectively. The MATR prognosis model obtained an outstanding c-index value of 0.805 on the discovery cohort and 0.917 when applied to the TCIA validation group using the same threshold.

Table 2 Results of the univariate Cox and Kaplan-Meier analyses for the discovery and validation populations. Significant results and best c-index are italicized

Variables	Threshold	Discovery cohort			Validation cohort		
		HR (CI-95%)	*p*	c-index	HR (CI-95%)	*p*	c-index
Age (years)	65.00	*2.09 (1.65, 2.66)*	*< 0.001*	0.704	*2.58 (1.59, 4.17)*	*< 0.001*	0.768
Total volume (cm^3)	28.55	1.12 (0.96, 1.53)	0.107	0.556	1.21 (0.77, 1.90)	0.419	0.508
CE volume (cm^3)	17.12	1.25 (0.99, 1.58)	0.059	0.557	1.24 (0.79, 1.97)	0.353	0.491
Necrotic volume (cm^3)	5.71	1.16 (0.91, 1.47)	0.223	0.559	1.36 (0.84, 2.20)	0.212	0.461
CE rim width (cm)	0.4157	*1.37 (1.06, 1.77)*	*0.015*	0.590	*2.17 (1.12, 4.23)*	*0.019*	0.699
Maximum diameter (cm)	5.15	1.24 (0.98, 1.57)	0.068	0.550	1.03 (0.65, 1.62)	0.911	0.582
Total surface (cm^2)	56.60	1.19 (0.94, 1.50)	0.156	0.554	1.28 (0.80, 2.03)	0.307	0.468
Surface regularity	0.629	*1.48 (1.17, 1.87)*	*0.001*	0.585	*1.79 (1.12, 2.87)*	*0.014*	0.644
MA prognosis score	1.25	*2.57 (1.97, 3.21)*	*< 0.001*	*0.741*	*2.31 (1.47, 3.63)*	*< 0.001*	*0.769*

Prognosis prediction for subtotally resected patients

A total of 113 patients from the discovery cohort received subtotal resections. The Kaplan-Meier analysis provided age and CE rim width as significant parameters (Table 3). The MA prognosis score obtained for subtotally resected patients (MASR) was

$$\mathrm{MASR} = 0.023 \times \mathrm{age} - 0.273 \times \mathrm{CE\ rim\ width}$$

This score was also significant in the Kaplan-Meier analysis (Fig. 1d) and gave a c-index value of 0.702.

Prognosis prediction for biopsied patients

Forty-nine patients in the discovery cohort underwent a biopsy. The Kaplan-Meier analysis of this patient sample gave age, all of the volumes, CE rim width, and total surface as significant parameters (Table 3). The correlation pattern found for these variables was similar to that in the other groups (Fig. S1D). The Cox model highlighted age and CE volume as the only significant parameters and the MA prognosis score obtained for biopsied patients (MAB) was

$$\mathrm{MAB} = 0.047 \times \mathrm{age} + 0.021 \times \mathrm{CE\ volume}$$

This score was highly significant in the Kaplan-Meier analysis (Fig. 1e) and achieved a c-index of 0.866, outperforming those obtained from the individual variables (Table 3).

Figure 2 summarizes the different patient subgroups studied in this work and the c-indexes obtained using the prognostic metrics.

Figure 3 summarizes the results of our prognostic models in comparison with other recent approaches.

Discussion

The results described in this paper support the hypothesis that simple quantitative baseline morphological features in addition to age are key biomarkers for OS in GBM, and for the different groups of patients with the same surgical procedure. Morphological features have recently been studied as prognostic factors, and they outperform established biomarkers [15, 24]. This study represents the most comprehensive evidence yet for baseline morphological tumor properties as prognostic factors regardless of clinical scenarios, when the extent of resection is accounted for. The consideration that baseline CE rim width and surface regularity allow prognostic models to be built in GBM, independently of therapy, has important implications for clinical trials and practice.

We used a large dataset of 311 GBM patients diagnosed in the period 2006–2017 with pathologically proven GBM according to the 2007 World Health Organization (WHO) Classification of Tumors of the Central Nervous System. MGMT status or IDH1 mutations were not available for most patients included.

We identified the most relevant geometrical measures obtained from CE T1-weighted MRIs. Our focus was on morphological measures, since many textural features used in radiomic studies have been reported to be dependent on both the spatial and gray-level discretizations, hampering their inclusion in multicenter studies [25].

Age, CE rim width, and surface regularity were identified as the outstanding prognostic parameters of the study. Age has been known for a long time to be one of the parameters with higher prognostic value in GBM [5]. CE rim width and surface regularity are two recently proposed T1-based parameters, confirmed here to have a relevant role as prognostic variables in a substantially larger dataset. These parameters were not substantially correlated, reflecting different underlying biological processes. Mathematical models have substantiated that CE rim width may be related to macroscopic tumor

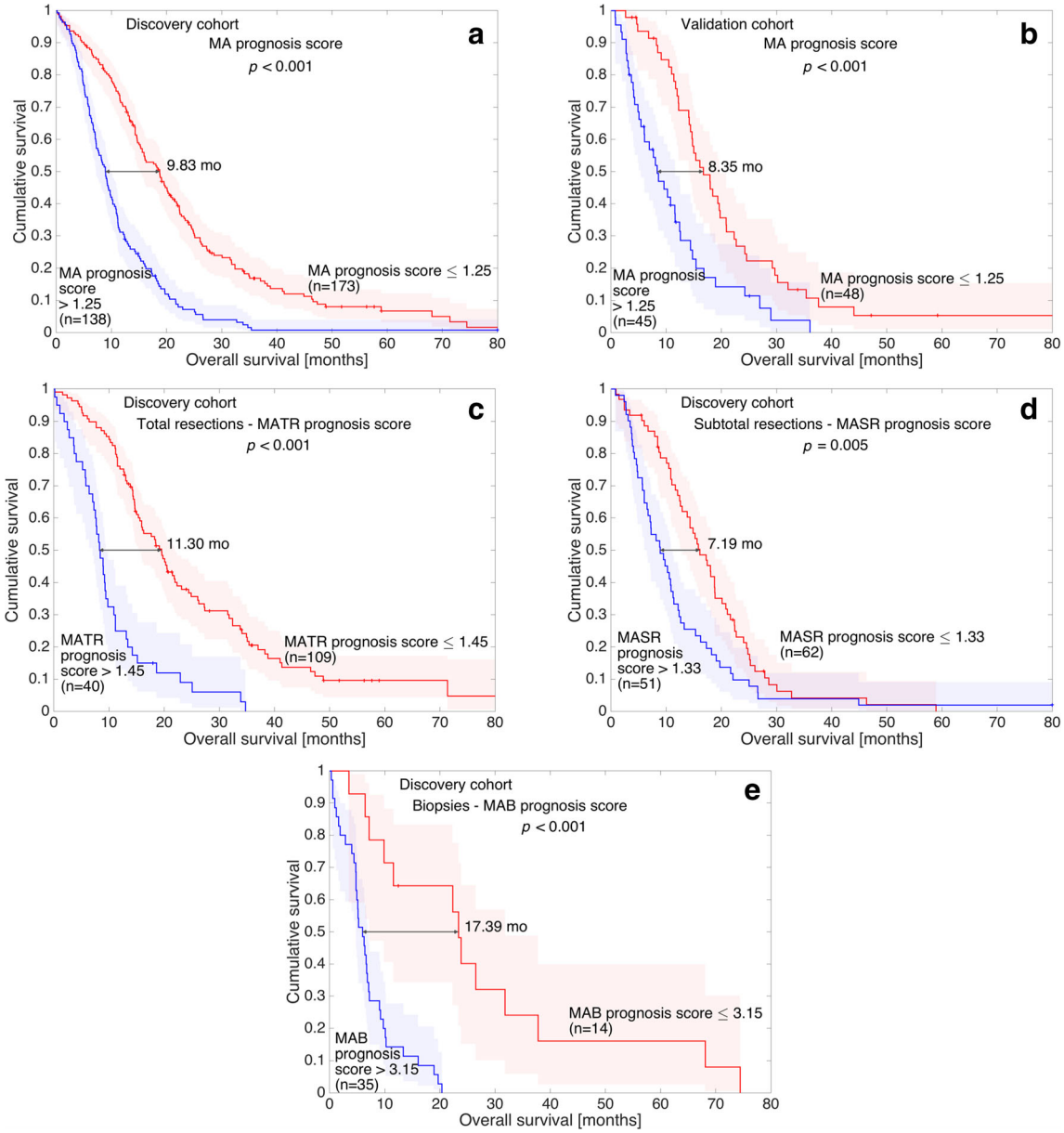

Fig. 1 Kaplan-Meier analysis of the prognosis scores for each population of patients considered: **a** Discovery cohort (MA). **b** Validation cohort (MA). **c** Totally resected patients (MATR). **d** Subtotally resected patients (MASR). **e** Biopsied patients (MAB). The differences in median survival and number of patients in each subgroup are indicated within the plots

growth infiltration speed [26]. It is noteworthy that this feature was a significant predictor of survival in every cohort of patients considered. Also, the shape of the tumor boundary at the microscopic level has been proposed in mathematical studies to be related to the dynamics of the most aggressive cellular phenotypes of which the tumor is made up [27] and to the tumor's invasiveness [28].

In our study, the CE volume was close to being statistically significant in the discovery group. It was not relevant either in the validation group or for totally resected patients. However, tumor surface and total, necrotic, and CE volume were highly prognostic on biopsied patients. Thus, contradictory results published in the literature [13, 14, 17, 29] could be due to different proportions of patients undergoing different types of surgical resections.

Many studies have confirmed the impact of resection in GBM patients [17–21], although it is not an ideal measure of the success of surgery [17]. However, the use of baseline imaging parameters to infer survival prognosis has not been studied thoroughly [15].

For totally resected patients, the same three parameters were found to be significant. These patients are expected to have only residual infiltrative disease after the surgery. Thus, pre-operative measures reflecting the aggressiveness of the tumor (and not its "size"), such as the CE rim width and surface regularity of the tumor before resection, were useful

Table 3 Results of the univariate Cox and Kaplan-Meier analyses, and the concordance indexes for the discovery cohorts of totally resected patients, subtotally resected patients, and biopsied patients. Significant results and best c-index of each group are italicized

Variables	Threshold	HR (CI-95%)	*p*	c-index
Totally resected patients				
Age (years)	60.00	*1.98 (1.38, 2.83)*	*< 0.001*	0.700
Total volume (cm^3)	31.60	1.16 (0.81, 1.66)	0.422	0.544
CE volume (cm^3)	22.50	1.22 (0.84, 1.78)	0.287	0.550
Necrotic volume (cm^3)	13.20	1.22 (0.84, 1.75)	0.292	0.449
CE rim width (cm)	0.367	*1.73 (1.13, 2.64)*	*0.010*	0.649
Maximum diameter (cm)	4.22	1.18 (0.83, 1.69)	0.356	0.486
Total surface (cm^2)	73.8	1.20 (0.82, 1.75)	0.342	0.564
Surface regularity	0.629	*1.61 (1.13, 2.29)*	*0.008*	0.630
MATR prognosis score	1.45	*3.43 (2.30, 5.10)*	*< 0.001*	*0.805*
Subtotally resected patients				
Age (years)	61.00	*1.70 (1.15, 2.50)*	*0.007*	0.692
Total volume (cm^3)	21.60	1.41 (0.94, 2.13)	0.096	0.449
CE volume (cm^3)	13.80	1.32 (0.87, 2.01)	0.189	0.456
Necrotic volume (cm^3)	10.90	1.16 (0.79, 1.69)	0.455	0.477
CE rim width (cm)	0.452	*1.66 (1.07, 2.58)*	*0.022*	0.664
Maximum diameter (cm)	6.31	1.35 (0.90, 2.02)	0.149	0.401
Total surface (cm^2)	89.50	1.32 (0.89, 1.95)	0.166	0.415
Surface regularity	0.57	1.17 (0.80, 1.71)	0.419	0.525
MASR prognosis score	1.33	*1.73 (1.18, 2.53)*	*0.005*	*0.702*
Biopsied patients				
Age (years)	64.00	*2.89 (1.45, 5.77)*	*0.002*	0.675
Total volume (cm^3)	21.20	*1.91 (1.04, 3.49)*	*0.033*	0.710
CE volume (cm^3)	15.40	*2.06 (1.13, 3.76)*	*0.016*	0.687
Necrotic volume (cm^3)	5.60	*1.81 (1.01, 3.25)*	*0.043*	0.714
CE rim width (cm)	0.529	*2.13 (1.12, 4.03)*	*0.018*	0.664
Maximum diameter (cm)	4.40	1.75 (0.96, 3.18)	0.066	0.676
Total surface (cm^2)	33.70	*2.14 (1.08, 4.21)*	*0.025*	0.721
Surface regularity	0.686	1.47 (0.77, 2.80)	0.247	0.538
MAB prognosis score	3.15	*7.25 (2.73, 19.20)*	*< 0.001*	*0.866*

in inferring patient prognosis. Regarding the cohort of biopsied patients, tumor volume was not significantly reduced by the surgical procedure and thus it had a substantial effect on survival.

The extent of resection, as a qualitative variable, combined with the morphological variables, allowed for a substantially better survival prediction than published in the literature. The models constructed for totally resected and biopsied patients achieved c-indexes of 0.805 (MATR) and 0.866 (MAB) respectively, outperforming the result of the "uninformed" MA model (0.741). However, the lack of data on the extent of resection regarding the subtotally resected patient cohort led to a lower c-index (0.702). In comparison with some of the best recent representative studies, Cui et al [8] constructed a model using five imaging features, obtaining a c-index of 0.674 for their validation cohort. Pérez-Beteta et al [9] obtained a prognostic model based on two parameters with a c-index of 0.735 in their validation cohort. However, their study was limited only to high-resolution MRIs. Ingrisch et al [14] obtained a concordance index score of 0.677 in their cross-validation analysis using a model based on 20 parameters, and Lao et al [12] obtained a c-index of 0.710 in their validation dataset using eight features. Kickingereder et al obtained a c-index of 0.696 for their training dataset using 13 parameters [11]. Cui et al [10] obtained 0.653 with two parameters on their validation dataset. Finally, the best prognosis indicators developed using clinical variables achieved a c-index of only 0.582 in their validation cohort [30].

The MA model for the full cohort of GBM patients using only three variables outperformed the best state-of-the-art methods for the validation groups. The partition of the population using the extent of resection as discrete variable together with the (simple) set of preoperative morphological features allowed for a much better prognosis prediction for the subsets of totally resected and biopsied patients well beyond the best c-indexes previously reported [8–12, 14, 30].

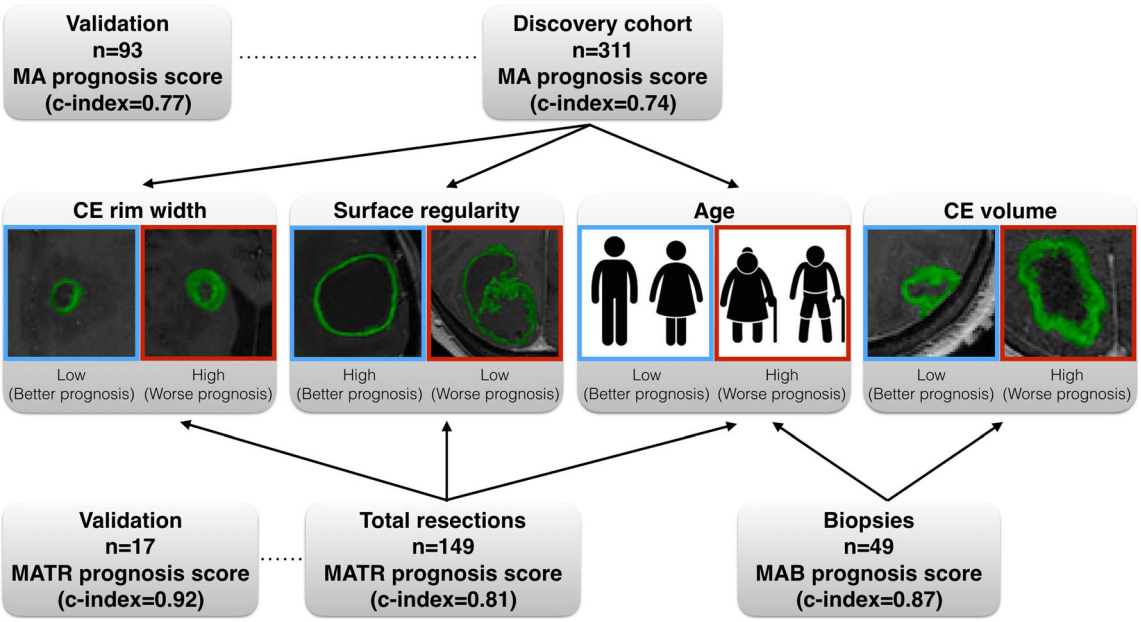

Fig. 2 Scheme showing the different patient cohorts included in the study, and the variables used to construct each model. The results obtained for the c-index for the different models and cohorts are also displayed

The prognostic value of the models for the cohort of patients receiving subtotal resections was only average. Details of the extent of subtotal resections were not available, due to the retrospective nature of this study. The lack of this relevant parameter probably limited the prognostic value of our combined measures for this heterogeneous patient subgroup.

This study has several strengths worth mentioning. Firstly, it encompasses one of the largest series of GBM patients with pretreatment volumetric CE T1-weighted MRIs in the literature [8, 14, 30]. All segmentations were performed semi-automatically and reviewed by experienced radiologists and 3D measures were computed. Secondly, it provides real clinical practice data [13], where the limitations of clinical trials are commonly discovered [31]. Thirdly, it used only morphological features with a simple interpretation that can be obtained straightforwardly from segmented tumors. Fourthly, results were validated using a large and public dataset of patients from TCIA [22], according to the basic statements in radiomics [32]. We think that the combination of these strengths is the reason why the results of this study surpass all previously reported findings.

Regarding the limitations, the first was the lack of clinical, genetic, and molecular information and recurrences for the cohorts. For example, stratification by MGMT or IDH1 genes, which are well-known to have GBM prognosis value, was not available due to the retrospective nature of the study. Previous GBM studies have shown a plethora of imaging parameters [32] or even, no evidence-proven correlations with molecular status [33] without validation groups. Secondly, given the multicenter nature of the study, and although great effort was made to homogenize data, there were differences in the imaging protocols or clinical follow-up. Thirdly, since TCIA datasets reported only 8 biopsied patients satisfying the

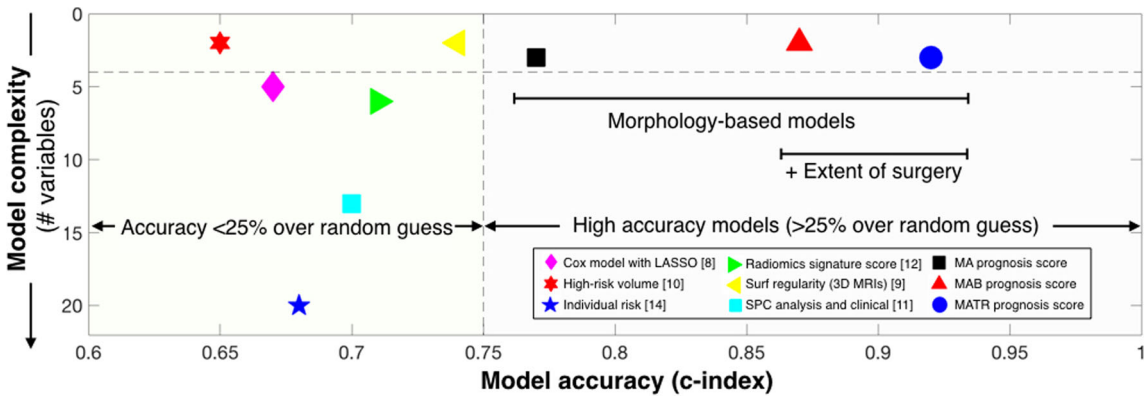

Fig. 3 Predictive value (c-index) and number of variables for representative models from the literature. Also shown for comparison are the results of the morphology-based models developed in this paper. Results are given for the best models in each reference and for the validation groups when available

inclusion criteria, no validation was performed on this patient subset. For totally resected patients, the validation was performed on the set of 17 TCIA patients available so the reliability of this validation might be limited. We were not able to validate the results for the group of subtotally resected patients due to lack of data regarding extent of resection in the TCIA database.

The inclusion of additional parameters such as robust textural imaging features, parameters derived from other sequences such as T2/FLAIR, diffusion-weighted imaging, T1DCE perfusion, and genomic data may provide a way to improve our highly prognostic indexes.

In conclusion, in this study, small sets of meaningful MRI-based morphological features obtained from pretreatment T1-weighted images, and age, predicted survival of GBM patients very accurately in combined models. The partition of the population using the extent of resection improved the prognostic value of those measures.

Acknowledgements We would like to thank C. López (Radiology Department, Hospital General de Ciudad Real), M. Claramonte (Neurosurgery Department, Hospital General de Ciudad Real), L. Iglesias (Neurosurgery Department, Hospital Clínico San Carlos), J. Avecillas (Radiology Department, Hospital Clínico San Carlos), J. M. Villanueva (Medical Oncology Department, Hospital Universitario de Salamanca), and J. C. Paniagua (Medical Oncology Department, Hospital Universitario de Salamanca) for their help in the data collection. We would also like to thank J. A. Ortiz Alhambra (Mathematical Oncology Laboratory) and A. Fernández-Romero (Mathematical Oncology Laboratory) for their help in the tumor segmentation tasks.

Funding This research has been supported by the Ministerio de Economía y Competitividad/FEDER, Spain (grant number MTM2015-71200-R), and James S. Mc. Donnell Foundation Twenty-First Century Science Initiative in Mathematical and Complex Systems Approaches for Brain Cancer (Collaborative award 220020450).

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Victor Manuel Pérez-García, full professor and head of Department of Mathematics at Universidad de Castilla-La Mancha (Spain).

Conflict of interest The authors declare that they have no conflict of interest.

Statistics and biometry Complex statistical methods were necessary for this paper. However, Victor M. Pérez-García, Alicia Martínez-González, and David Molina (mathematicians) have significant statistical expertise.

Informed consent Written informed consent was obtained from all subjects (patients) in this study.

Ethical approval Institutional Review Board approval was obtained.

Methodology
- Retrospective
- Observational
- Multicenter study

References

1. Verma V, Simone CB 2nd, Krishnan S, Lin SH, Yang J, Hahn SM (2017) The Rise of Radiomics and Implications for Oncologic Management. J Natl Cancer Inst 109(7). https://doi.org/10.1093/jnci/djx055
2. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: Images Are More than Pictures, They Are Data. Radiology 278(2):563–577
3. Narang S, Lehrer M, Yang D, Lee J, Rao A (2016) Radiomics in glioblastoma: current status, challenges and opportunities. Transl Cancer Res 5(4):383–397
4. Ellingson BM, Bendszus M, Sorensen AG, Pope WB (2014) Emerging techniques and technologies in brain tumor imaging. Neuro Oncol 16(7):12–23
5. Ellingson BM (2015) Radiogenomics and imaging phenotypes in glioblastoma: novel observations and correlation with molecular characteristics. Curr Neurol Neurosci Rep 15(1):506
6. Pérez-Beteta J, Martínez-González A, Molina D et al (2017) Glioblastoma: Does the pretreatment geometry matter? A postcontrast T1 MRI-based study. Eur Radiol 27:1096–1104
7. Abrol S, Kotrotsou A, Salem A, Zinn PO, Colen RR (2017) Radiomic phenotyping in brain cancer to unravel hidden information in medical images. Top Magn Reson Imaging 26(1):43–53
8. Cui Y, Tha KK, Teresaka S et al (2016) Prognostic imaging biomarkers in glioblastoma: Development and independent validation on the basis of multiregion and quantitative analysis of MR images. Radiology 278(2):546–553
9. Pérez-Beteta J, Molina-García D, Ortiz-Alhambra JA et al (2018) Tumor surface regularity at MR imaging predicts survival and response to surgery in patients with glioblastoma. Radiology 288(1):218–225
10. Cui Y, Ren S, Tha KK, Wu J, Shirato H, Li R (2017) Volume of high-risk intratumoral subregions at multi-parametric MR imaging predicts overall survival and complements molecular analysis of glioblastoma. Eur Radiol 27(9):3583–3592
11. Kickingereder P, Burth S, Wick A et al (2016) Radiomic Profiling of Glioblastoma: Identifying an Imaging Predictor of Patient Survival with Improved Performance over Established Clinical and Radiologic Risk Models. Radiology 280(3):880–889
12. Lao J, Chen Y, Li ZC et al (2017) A Deep Learning-Based Radiomics Model for Prediction of Survival in Glioblastoma Multiforme. Sci Rep 7(1):10353
13. Wangaryattawanich P, Hatami M, Wang J et al (2015) Multicenter imaging outcomes study of The Cancer Genome Atlas glioblastoma patient cohort: imaging predictors of overall and progression-free survival. Neuro Oncol 17(11):1525–1537
14. Ingrisch M, Schneider MJ, Nörenberg D et al (2017) Radiomic Analysis Reveals Prognostic Information in T1-Weighted Baseline Magnetic Resonance Imaging in Patients With Glioblastoma. Invest Radiol 52(6):360–366
15. Ellingson BM, Harris RJ, Woodworth DC et al (2017) Baseline pretreatment contrast enhancing tumor volume including central necrosis is a prognostic factor in recurrent glioblastoma: evidence from single and multicenter trials. Neuro Oncol 19(1):89–98
16. Stupp R, Mason WP, van den Bent MJ et al (2005) Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med 352(10):987–996
17. Grabowsky MM, Recinos PF, Nowacki AS et al (2014) Residual tumor volume versus extent of resection: predictors of survival after surgery for glioblastoma. J Neurosurg 121:1115–1123
18. Kuhnt D, Becker A, Ganslandt O, Bauer M, Buchfelder M, Nimsky C (2011) Correlation of the extent of tumor volume resection and patient survival in surgery of glioblastoma multiforme with high-field intraoperative MRI guidance. Neuro Oncol 13(12):1339–1348

19. Chaichana KL, Jusue-Torres I, Lemos AM et al (2014) The butterfly effect on glioblastoma: is volumetric extent of resection more effective than biopsy for these tumors? J Neurooncol 120(3):625–634
20. Li YM, Suki D, Hess K, Sawaya R (2016) The influence of maximum safe resection of glioblastoma on survival in 1229 patients: Can we do better than gross-total resection? J Neurosurg 124(4):977–988
21. Lacroix M, Abi-Said D, Fourney DR et al (2001) A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis, extent of resection, and survival. J Neurosurg 95:109–198
22. Prior FW, Clark K, Commean P et al (2013) TCIA: an information resource to enable open science. Conf Proc IEEE Eng Med Biol Soc 1282–1285. https://doi.org/10.1109/EMBC.2013.6609742.
23. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA (1982) Evaluating the yield of medical tests. JAMA 247(1):2543–2546
24. Cordova JS, Gurbani SS, Holder CA et al (2016) Semi-automated volumetric and morphological assessment of glioblastoma resection with fluorescence-guided surgery. Mol Imaging Biol 18:454–462
25. Molina D, Pérez-Beteta J, Martínez-González A et al (2017) Lack of robustness of textural measures obtained from 3D brain tumor MRIs impose a need for standardization. PLoS One 12(6):e0178843
26. Pérez-García VM, Calvo GF, Belmonte-Beitia J, Diego D, Pérez-Romasanta L (2011) Bright solitary waves in malignant gliomas. Phys Rev E Stat Nonlin Soft Matter Phys 84:021921
27. Anderson AR, Weaver A, Cummings PT, Quaranta V (2006) Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. Cell 127:905–915
28. Martin M (2008) Comparing invasive species to metastatic cancers inspires new insights for modelers. J Natl Cancer Inst 100:88
29. Henker C, Kriesen T, Glass Ä, Schneider B, Piek J (2017) Volumetric quantification of glioblastoma: experiences with different measurement techniques and impact on survival. J Neurooncol 135:391–402
30. Gittleman H, Lim D, Kattan MW et al (2017) An independently validated nomogram for individualized estimation of survival among patients with newly diagnosed glioblastoma: NRG Oncology RTOG 0525 and 0825. Neuro Oncol 19(5):669–677
31. Chung C (2015) Imaging biomarkers in preclinical studies on brain tumors. In: Preedy V, Patel V (eds) Biomarkers in cancer. Biomarkers in disease: methods, discoveries and applications. Springer, Dordrecht
32. Yip SS, Aerts HJ (2016) Applications and limitations of radiomics. Phys Med Biol 61(13):R150–R166
33. Grossmann P, Gutman DA, Dunn WD Jr, Holder CA, Aerts HJ (2016) Imaging-genomics reveals driving pathways of MRI derived volumetric tumor phenotype features in Glioblastoma. BMC Cancer 16:611

Affiliations

Julián Pérez-Beteta[1] · David Molina-García[1] · Alicia Martínez-González[1] · Araceli Henares-Molina[1] · Mariano Amo-Salas[1] · Belén Luque[1] · Elena Arregui[2] · Manuel Calvo[2] · José M. Borrás[3] · Juan Martino[4] · Carlos Velásquez[4] · Bárbara Meléndez-Asensio[5] · Ángel Rodríguez de Lope[6] · Raquel Moreno[7] · Juan A. Barcia[8] · Beatriz Asenjo[9] · Manuel Benavides[10] · Ismael Herruzo[11] · Pedro C. Lara[12] · Raquel Cabrera[12] · David Albillo[13] · Miguel Navarro[14] · Luis A. Pérez-Romasanta[15] · Antonio Revert[16] · Estanislao Arana[17] · Víctor M. Pérez-García[1]

1 Mathematical Oncology Laboratory (MôLAB), Department of Mathematics, Universidad de Castilla-La Mancha, 13071 Ciudad Real, Spain

2 Department of Radiology, Hospital General de Ciudad Real, Ciudad Real, Spain

3 Department of Neurosurgery, Hospital General de Ciudad Real, Ciudad Real, Spain

4 Department of Neurosurgery, Hospital Universitario Marqués de Valdecilla and Fundación Instituto de Investigación Marqués de Valdecilla, Santander, Spain

5 Department of Molecular Biology, Hospital Virgen de la Salud, Toledo, Spain

6 Department of Neurosurgery, Hospital Virgen de la Salud, Toledo, Spain

7 Department of Radiology, Hospital Virgen de la Salud, Toledo, Spain

8 Department of Neurosurgery, Hospital Clínico San Carlos, Madrid, Spain

9 Department of Radiology, Hospital Carlos Haya, Málaga, Spain

10 Department of Medical Oncology, Hospital Carlos Haya, Málaga, Spain

11 Department of Radiation Oncology, Hospital Carlos Haya, Málaga, Spain

12 Department of Radiation Oncology, Hospital Universitario Doctor Negrín, Gran Canaria, Spain

13 Department of Radiology, Hospital Universitario de Salamanca, Salamanca, Spain

14 Department of Medical Oncology, Hospital Universitario de Salamanca, Salamanca, Spain

15 Department of Radiation Oncology, Hospital Universitario de Salamanca, Salamanca, Spain

16 Department of Radiology, Hospital de Manises, Valencia, Spain

17 Department of Radiology, Fundación Instituto Valenciano de Oncología, Valencia, Spain