Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

An integrated explainable machine learning model for post-treatment peritumoral edema formation in Gamma Knife radiosurgery of meningioma

Matteo Chieregato

matteo.chieregato@poliambulanza.it

Fondazione Poliambulanza Istituto Ospedaliero

Giorgio Spatola

Fondazione IRCCS San Gerardo dei Tintori

Mauro Morassi

Manerbio Hospital - ASST Garda

Karol Migliorati

Fondazione Poliambulanza Istituto Ospedaliero

Milena Cobelli

Fondazione Poliambulanza Istituto Ospedaliero

Lodoviga Giudice

Fondazione Poliambulanza Istituto Ospedaliero

Nicola Redolfi

Fondazione Poliambulanza Istituto Ospedaliero

Chiara Bassetti

ASST Cremona

Corrado D'Arrigo

Fondazione Poliambulanza Istituto Ospedaliero

Cesare Giorgi

Fondazione Poliambulanza Istituto Ospedaliero

Alessandra Huscher

Fondazione Poliambulanza Istituto Ospedaliero

Rosaria Maio

Fondazione Poliambulanza Istituto Ospedaliero

Mario Bignardi

Fondazione Poliambulanza Istituto Ospedaliero

Marco Galelli

Fondazione Poliambulanza Istituto Ospedaliero

Alberto Franzin

Fondazione Poliambulanza Istituto Ospedaliero

Research Article

Keywords:

Posted Date: May 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9640465/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

An integrated explainable machine learning model for post-treatment peritumoral edema formation in Gamma Knife radiosurgery of meningioma.

Matteo Chieregato[1*], Giorgio Spatola[2], Mauro Morassi[3], Karol Migliorati[4], Milena Cobelli[5], Lodoviga Giudice[4], Nicola Redolfi[4], Chiara Bassetti[6], Corrado D'Arrigo[4], Cesare Giorgi[4], Alessandra Huscher[7], Rosaria Maio[1], Mario Bignardi[7], Marco Galelli[1], Alberto Franzin[4]

[1*]Unit of Medical Physics, Fondazione Poliambulanza Istituto Ospedaliero, Via Bissolati 57, Brescia, 25124, Italy.
[2]Department of Neurosurgery, Fondazione IRCCS San Gerardo dei Tintori, Via Pergolesi 33, Monza (MB), 20900, Italy.
[3]Radiology Unit, Manerbio Hospital - ASST Garda, Via Lungomella Valsecchi 2, Manerbio (BS), 25025, Italy.
[4]Gamma Knife Unit, Department of Neurosurgery, Fondazione Poliambulanza Istituto Ospedaliero, Via Bissolati 57, Brescia, 25124, Italy.
[5]Neuroradiology Unit, Department of Radiology, Fondazione Poliambulanza Istituto Ospedaliero, Via Bissolati 57, Brescia, 25124, Italy.
[6]Unit of Medical Physics, ASST Cremona, Largo Priori Emilio 1, Cremona, 26100, Italy.
[7]Department of Radiation Oncology, Fondazione Poliambulanza Istituto Ospedaliero, Via Bissolati 57, Brescia, 25124, Italy.

*Corresponding author(s). E-mail(s): matteo.chieregato@poliambulanza.it;
Contributing authors: giorgio.spatola@irccs-sangerardo.it; mauro.morassi@asst-garda.it; karol.migliorati@poliambulanza.it; milena.cobelli@poliambulanza.it; lodoviga.giudice@poliambulanza.it; nicola.redolfi@poliambulanza.it; chiara.bassetti@asst-cremona.it;

c.darrigo@quipo.it ; giorgi.cesare@gmail.com;
alessandra.huscher@poliambulanza.it; rosaria.maio@poliambulanza.it;
mario.bignardi@poliambulanza.it; marco.galelli@poliambulanza.it;
alberto.franzin@poliambulanza.it;

Abstract

Purpose: Peritumoral edema is a frequent complication of meningiomas and may be exacerbated or newly induced following stereotactic radiosurgery (SRS). We developed an integrated machine learning model incorporating radiomic, clinical, and dosimetric features to predict de novo edema formation after Gamma Knife radiosurgery (GKRS), and applied explainable artificial intelligence (XAI) techniques to enhance model interpretability.
Methods: A cohort of 131 meningiomas in 105 patients treated with GKRS was retrospectively analyzed. All lesions were edema-free at baseline and had a minimum imaging follow-up of 6 months. A total of 107 radiomic features were extracted from planning MRI. An XGBoost classifier was trained within a repeated stratified nested cross-validation framework for radiomic feature selection, integrated model hyperparameter optimization, and decision threshold tuning. Model interpretability was assessed using three complementary XAI approaches: permutation importance, XGBoost built-in gain importance, and genetic algorithm-based counterfactual aggregation.
Results: The model achieved mean sensitivity of $\mathbf{68.9 \pm 12.6\%}$, mean specificity of $\mathbf{71.6 \pm 8.4\%}$, and a mean area under the receiver operating characteristic curve (AUC-ROC) of $\mathbf{0.729 \pm 0.016}$, across repeated cross-validation folds. The Informational Measure of Correlation emerged as the most predictively important radiomic feature. Counterfactual-based feature importance identified a heterogeneous set of radiomic, clinical, and dosimetric variables as primary prediction drivers.
Conclusion: An integrated machine learning model combining MRI radiomics with clinical and dosimetric data predicted post-radiosurgical edema formation with moderate effectiveness and robustness. XAI techniques illuminated model decision-making.

Keywords: keyword1, Keyword2, Keyword3, Keyword4

1 Introduction

Meningiomas are brain tumors originating from arachnoid cap cells of the inner surface of the dura. They can arise in any part of the brain, as well as the spinal cord and ventricles, and account for 37.6% of all intracranial tumors according to the most recent US tumor registry [1], making them the most common brain tumors in adults. Meningiomas typically follow a slow-growing clinical course [2]. Most carry a favorable prognosis: non-malignant meningiomas have a ten-year survival rate of 84%, that falls to 62% [1] for malignant variants.

The traditional treatment of choice for meningioma is gross total resection. SRS may be proposed as a primary treatment modality for an asymptomatic or mildly symptomatic meningioma, and should be considered when a complete surgical excision cannot be achieved or is not amenable [3], [4]. In this context, SRS as an adjuvant tool has been shown to enhance progression-free survival and local control across meningiomas of all grades [5],[6],[7]. These favorable outcomes have prompted proposals to establish radiosurgery as the primary treatment option [8], even for grade II meningiomas, with promising results [9]. Gamma Knife is the first and arguably the most established form of stereotactic radiosurgery. Recent Gamma Knife models (Icon and Esprit) are equipped with cone-beam CT and are therefore suitable for fractionated treatment, using a thermoplastic mask in place of the Leksell frame. At our center, meningiomas abutting the brainstem, optic chiasm, or optic nerves are routinely treated in three or five fractions, to exploit the radiobiological advantages of fractionation.

The development of peritumoral edema is common in meningiomas (more than 60% of cases [10]), although its pathogenesis remains incompletely understood. Evidence points to a close link with factors promoting angiogenesis, vascular permeability, and blood–brain barrier disruption, with the VEGF-A ligand playing a particularly prominent role. Clinical factors such as tumor volume, location, and histological subtype have been associated with peritumoral edema, though no single factor has proved definitively conclusive [10],[11]. One proposed pathway begins with compression of adjacent brain tissue and vessels — a consequence of the above clinical factors — leading to local hypoxia. Hypoxia upregulates VEGF-A through a cascade of mediators, thereby promoting angiogenesis and increasing vascular permeability. Concurrently, other mediators impair blood–brain barrier integrity, extracellular matrix stability, and water–electrolyte balance, collectively contributing to peritumoral edema formation.

Radiosurgery can exacerbate pre-existing edema or induce de novo edema formation [12],[13]. Predicting — and ultimately preventing — post-radiosurgical edema would therefore be of considerable clinical importance in the management of meningioma patients.

Machine learning methods have been applied to predict meningioma histological grade using either radiomics-based tabular classifiers [14],[15],[16], [17] or convolutional neural networks [18]. Other studies have predicted Ki-67 expression, typically binarized to frame the task as a classification problem [19]. More recently, additional endpoints have been investigated, notably brain invasion [20],[21]. Finally, two research groups have developed predictors of post-radiosurgical edema in meningiomas: one using an integrated deep learning and clinical features model [22], and the other applying survival machine learning methods on radiomic, semantic, and clinical features [23]. Notably, neither study distinguishes between newly developed edema — the endpoint of the present work — and progression of pre-existing edema.

A fundamental prerequisite for the adoption of artificial intelligence in general, and machine learning in particular, is trust. This is true especially when the risks involved are substantial, as is the case in clinical contexts [24]. Model explainability is widely regarded as a necessary condition for building that trust [25]. Feature importance analysis is arguably the most widely used and most intuitive method for providing

explainability in tabular machine learning models. However, multiple approaches exist for computing feature importance, both model-agnostic and model-intrinsic, and these often yield discordant results [26].

The aim of this study is to develop a machine learning model to predict de novo edema formation following Gamma Knife radiosurgery, integrating radiomic, clinical, and dosimetric features. In addition, the model is equipped with explainability through feature importance analysis. Two complementary types of feature importance are computed: predictive feature importance, which identifies the features with the greatest discriminative power between classes, and what-if (counterfactual) feature importance, which identifies the features that would need to change in order to invert the prediction of the model. In particular, we built an integrated predictive XGBoost model [27] for classification. Predictive feature importance was assessed using two independent methods: permutation importance [28] and XGBoost built-in gain importance. What-if feature importance was derived by aggregating information from counterfactuals generated with the DiCE library [29].

2 Methods

2.1 Dataset

We included adult patients treated for meningioma with Gamma Knife radiosurgery at Fondazione Poliambulanza between 25 September 2017 and 15 December 2021, with both a planning MRI scan and a follow-up MRI scan at six months or one year post-treatment (n = 204 meningiomas in N = 155 patients). Cavernous sinus meningiomas were excluded as they are not clinically relevant to edema formation (n = 27, N = 27). Patients with WHO grade III meningiomas or early recurrence suggestive of aggressive biological behavior were also excluded (n = 15, N = 3). Treatments in which two or more adjacent targets were treated as a single target were further excluded, as this would distort dosimetric parameters (n = 15, N = 4). Finally, since the primary endpoint of this study is de novo peritumoral edema formation, patients with edema already present on T2-weighted pretreatment MRI were excluded, as assessed independently by two expert readers: the first drawn from a panel comprising one neurosurgeon and two neuroradiologists, and the second a independent neurosurgeon (n = 16, N = 16). The final cohort comprised 131 meningiomas in 105 patients, of whom 18 (13.7%)developed de novo peritumoral edema following radiosurgery. For each patient, clinical and demographic data and GammaPlan-derived dosimetric parameters were collected (Table 1). Dosimetric equivalent quantities for fractionated treatments were calculated as described in 2.4. Planning MRI scans were acquired for radiomic feature extraction, as described in Section 2.3. De novo edema development after radiosurgery was assessed on follow-up MRI by the same panel of expert readers, following a two-step process: an initial evaluation was performed by one member of the panel (either a neurosurgeon or one of the two neuroradiologists), followed by independent review by a second neurosurgeon.

Table 1 Clinical and dosimetric characteristics of the study cohort.

	Patients (n=105)		**Meningiomas (N=131)**	
	Number	**%**	**Number**	**%**
Sex				
Female	80	76.2		
Male	25	23.8		
Previous surgery				
Yes	23	21.9		
No	82	78.1		
Number of fractions				
1	87	82.9		
3	5	4.8		
5	13	12.4		
WHO Grade				
Grade I			111	84.7
Grade II			20	15.3
Location				
Supratentorial skull base			42	32.1
Convexity			41	31.3
Falcine and parafalcine			25	19.1
Infratentorial			23	17.6
Multiple targets in same session				
Yes			39	29.8
No			92	70.2
Abutting a venous structure				
Yes			74	56.5
No			57	43.5
Relapse after radiosurgery				
Yes			15	11.5
No			116	88.5

	Mean ± SD	**Median [P10–P90]**
Age (years)	60.7 ± 12.3	61 [46–77]
Volume (cm^3)	3.38 ± 3.69	2.36 [0.32–7.79]
Marginal dose equivalent (Gy)	14.08 ± 0.85	14 [13–15]
Mean dose equivalent (Gy)	20.10 ± 1.75	20.17 [17.70–22.01]
Maximum dose equivalent (Gy)	27.10 ± 2.21	27.35 [23.85–29.47]
Minimum dose equivalent (Gy)	12.21 ± 1.59	12.43 [10.10–13.99]
Coverage	99.49 ± 0.74	100 [99–100]
Selectivity	0.64 ± 0.14	0.67 [0.48–0.79]
Treatment duration (min)	45.05 ± 24.91	40.5 [16.0–77.5]
Number of shots	14.9 ± 10.14	13 [5–24]
V_{10Gy} equivalent (cm^3)	7.87 ± 7.43	5.69 [1.09–16.53]
V_{5Gy} equivalent (cm^3)	21.28 ± 20.64	14.18 [3.08–45.32]

n = number of patients; N = number of meningiomas. Number of fractions and patient-level variables are reported as proportions of n=105. Meningioma-level variables are reported as proportions of N=131. Dose equivalent values are single-fraction equivalents calculated using the linear-quadratic model (α/β = 3.7 Gy for meningioma, α/β = 2 Gy for brain parenchyma). SD = standard deviation; P10–P90 = 10th–90th percentile range.

2.2 MRI scan characteristics

All planning MRI scans were acquired on the same scanner, a 1.5T GE Signa HDxt, using a volumetric T1-weighted gradient recalled echo sequence (spatial resolution 0.977 × 0.977 × 1.0 mm) with gadolinium-based contrast agent (0.2 ml/kg administered as an intravenous bolus 10 minutes prior to acquisition). In a substantial proportion of cases, MRI was acquired with the stereotactic Leksell frame in place; in the remaining cases, either a thermoplastic immobilization mask was used (all fractionated treatments) or the planning MRI was acquired prior to frame mounting. Acquisition parameters were otherwise unchanged across all cases.

2.3 MRI preprocessing and radiomics feature extraction

Bias field correction was first applied to planning MRI images using the N4 algorithm [30]. A brain mask was subsequently computed using ROBEX (Robust Brain Extraction [31]). A z-score normalized image [32] was derived from skull-stripped images and the corresponding transformation was then applied to complete images (Supplementary Figure S1). Radiomic feature extraction was performed using PyRadiomics v3.0.1 [33]. Tumor contours, already delineated for treatment planning purposes, were used as regions of interest. Images were resampled to an isotropic resolution of 1 x 1 x 1 mm. Radiomic features were then extracted using a fixed bin size of 0.25. The following feature classes were extracted: Shape, First Order, GLCM, GLDM, GLRLM, GLSZM, and NGTDM, yielding a total of 107 radiomic features. MeshVolume and VoxelVolume were removed from radiomic features; tumor volume — as calculated by GammaPlan — was instead retained among clinical features in subsequent analysis. The final feature set comprised 105 radiomic features and 20 clinical, demographic and dosimetric features (including tumor volume).

2.4 Dosimetric equivalent quantities

To enable comparison between fractionated and single-fraction treatments, doses from fractionated treatments were converted to single-fraction equivalent values using a linear-quadratic radiobiological model. An α/β ratio of 3.7 Gy was assumed for meningioma [34] (see Discussion), and an α/β ratio of 2 Gy for brain parenchyma [35]. Specifically, target dosimetric parameters (marginal dose, maximum dose, mean dose, and minimum dose) were converted to single-fraction equivalent doses using $\alpha/\beta = 3.7$ Gy, while V_5 and V_{10} were computed by converting the reference dose levels of 5 Gy and 10 Gy to their fractionated equivalents using $\alpha/\beta = 2$ Gy.

2.5 Model building, training, validation, evaluation

The predictive model was an XGBoost classifier [27], trained on integrated radiomic, clinical, and dosimetric features. The validation strategy chosen was repeated stratified nested three-fold cross-validation with patient-level group splitting ensuring that no two meningiomas belonging to the same patient could appear in both the training and test subsets [36]. Stratification ensured that the proportion between positive (de novo edema development) and negative (no edema) instances remained constant across

splits- a critical requirement when validating models on imbalanced datasets. Nesting the cross-validation allowed feature selection, hyperparameter optimization and binary threshold selection via the Youden criterion to be performed in the inner loop, without information leakage from the test set. The number of folds was chosen as a compromise between the number of instances available for the training and the minimum number of positive cases in the inner validation subsample, which should not be overly sensitive to individual variations. The entire splitting procedure was repeated with twenty different random seeds, and mean performance metrics were reported to ensure robustness of validation estimates. The overall model pipeline was as follows (Figure 1):

1. Min-Max normalization of radiomic features.
2. Removal of mutually correlated radiomic features (correlation threshold 0.9, empirically determined).
3. Recursive feature elimination with cross-validation using an XGBoost model in the inner loop, to retain only informative radiomic features.
4. Concatenation of the selected radiomic features with clinical and dosimetric features.
5. Hyperparameter optimization of the XGBoost model on the concatenated feature set using Bayesian Tree-Structured Parzen Estimator (TPE) [37], again within the inner cross-validation loop.
6. Selection of the optimal binary classification threshold using the Youden criterion, estimated from inner loop cross validation.
7. Evaluation of the optimized XGBoost model (with the selected threshold) on the three outer-loops test sets.
8. Repetition of the entire pipeline twenty times with different random seeds.

Steps 1 and 2 were applied to the combined training and validation subsets; the resulting transformations were subsequently applied to the test subsets, to prevent data leakage. To address data imbalance (13.7% positive cases), hyperparameter optimization included the class balancing parameter `scale_pos_weight`, and the area under the precision-recall curve (AUPRC) was used as the objective metric for both feature selection and hyperparameter tuning. All analysis and coding were implemented in Python 3.8. The Scikit-Learn 1.3.0 library [38] was used for the machine learning pipeline. XGBoost 1.7.6 [27] was used as the classifier. Bayesian hyperparameter optimization was performed with Optuna 3.2.0 [37].

2.6 Feature importance

Explainability was added to the model by computing feature importance with three different methods. The first method is permutation feature importance [28], which analyzes how model performance degrades when the values of a given feature are randomly shuffled across instances. The greater the performance deterioration, the higher the assigned feature importance. This score was computed with best models evaluated on test data, across all folds and repetitions, to avoid optimistic bias. The second method is XGBoost built-in gain importance, which is based on the reduction in node impurity achieved by a given feature across all splits and trees. Both methods measure predictive importance, defined as the influence of each features on model prediction.

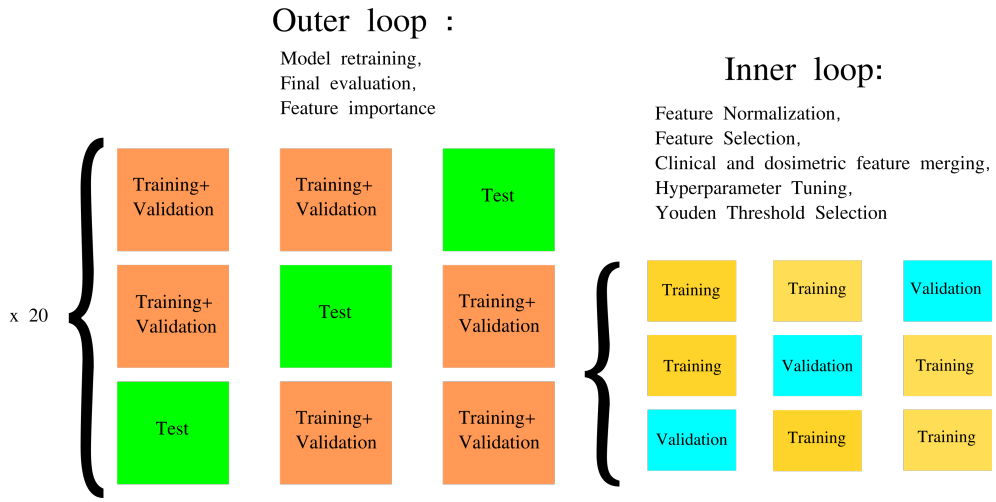


Fig. 1 Schematic representation of the repeated stratified nested three-fold cross-validation scheme. Both outer and inner splits were performed at the patient level and stratified, ensuring a constant proportion of de novo edema cases across all subsets. Random seeds were varied independently for both the outer and inner loops across repetitions.

The third method is based on counterfactual explanations and yields what-if feature importance, quantifying the actionable feature changes required to invert a model prediction. In machine learning, counterfactuals are minimally modified instances that would lead to a different model output; formally, a counterfactual explanation specifies which features of a given data point should change, and by how much, to alter the outcome of the model [39]. Three counterfactuals were generated for each true positive instance in the test set using the genetic algorithm implementation of the DiCE library (Diverse Counterfactual Explanations [29]). What-if feature importance was then derived by aggregating the frequency and magnitude of feature changes across all generated counterfactuals, and a composite importance score was obtained by multiplying these two quantities.

3 Results

The performance of the predictive model is summarized as follows: area under the ROC curve = 0.729 ± 0.016, area under the precision recall curve = 0.287 ± 0.033, Brier score=0.187± 0.015, Matthews correlation coefficient = 0.183 ± 0.049, sensitivity = 68.9% ±12.6 % and specificity = 71.6 % ±8.4 % (mean ± standard deviation over twenty repetitions; binary predictions derived using the optimal Youden threshold; Figure 2). Feature importance are shown in Figure 3. Both predictive feature importance methods consistently identified the Informational Measure of Correlation (Imc1) as the most important feature. Strength was also ranked as important by the XGBoost gain method, but not by permutation importance. What-if feature importance, derived by aggregating counterfactual explanations, revealed a large magnitude of change in Busyness, although this modification appeared in fewer than 10% of

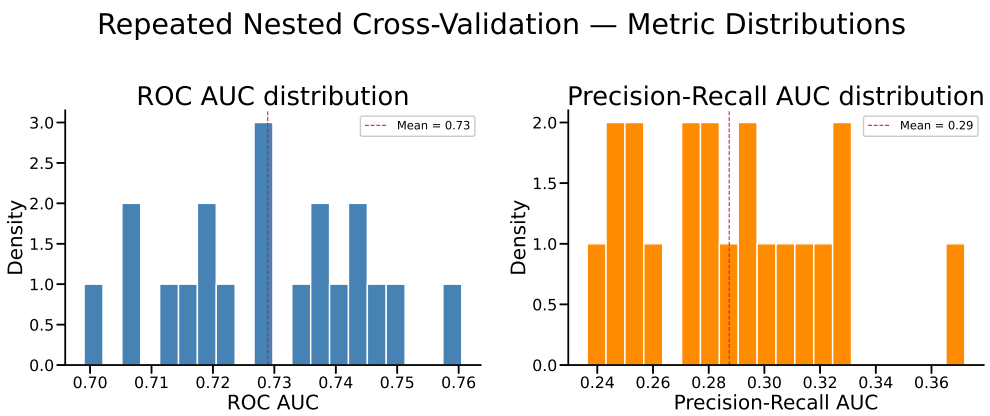


Fig. 2 Histograms of area under the ROC curve and area under the precision-recall curve across the 20 repetitions. Data are shown without rebinning. Area under the ROC curve values are consistent across repetitions. Area under the precision-recall curve values should be interpreted in relation to the positive class prevalence of 13.7%, which represents the expected performance of a random classifier.

counterfactuals. of counterfactuals. Two-dimensional shape parameters and Skewness followed, again exhibiting large magnitudes of change but appearing in a minority of counterfactuals. A summary of feature importance is shown in Figure 3.

4 Discussion

Predicting de novo edema formation after radiosurgery is not a straightforward task, and integrated models may be better suited to reflect the inherent complexity of meningioma patients, who present with a wide spectrum of WHO histological grades, Ki-67 values, tumor locations, and sizes. In this study, we developed a machine learning model combining clinical, radiomic, and dosimetric features, validated through a rigorous repeated nested cross-validation framework, with moderate predictive performance.We acknowledge several limitations. First, the sample size is relatively small, particularly given the class imbalance. Second, the data originate from a single institution, with no external validation cohort available. Third, the analysis was conducted at the meningioma level, with some patients contributing more than one lesion; while this was addressed statistically through patient-level splitting, a hidden bias cannot be excluded due to the potentially distinctive pathological traits of patients with a particular predisposition to developing multiple meningiomas. Finally, the heterogeneity of fractionation regimens was addressed through linear-quadratic-based dose equivalence conversion.

However, uncertainty remains regarding the appropriate α/β coefficients. While there is general consensus on a value of approximately 2 Gy for brain parenchyma, the α/β for meningioma is considerably less well established. The value adopted in this study, $\alpha/\beta = 3.7$ Gy, is substantially lower than the typical value of 10 Gy assumed for most tumors, and consistent with the majority of the literature ($\alpha/\beta = 3.76$ Gy,

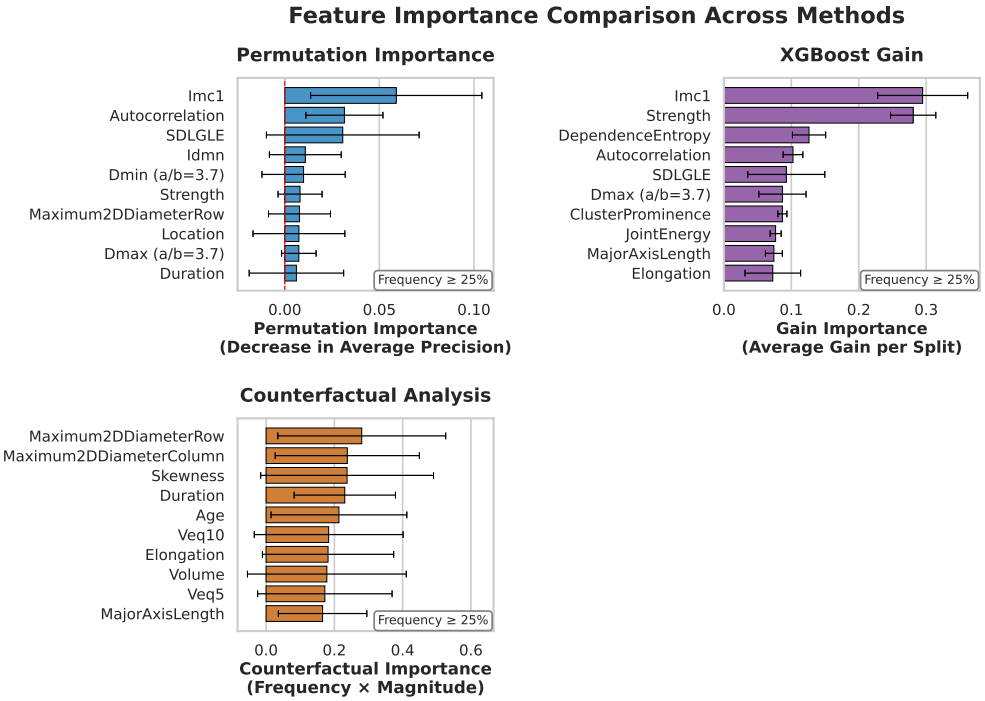


Fig. 3 Top ten features ranked by three feature importance methods: permutation importance, XGBoost gain importance, and counterfactual (what-if) importance. A minimum frequency threshold was applied in each case: for permutation and gain importance, features were required to appear in at least 25 % of repeated folds (15/60); for counterfactual importance, features were required to appear in at least 25 % of all computed counterfactuals (192/765). Permutation importance is calculated from the decrease in average precision resulting from randomly permuting the values of a feature. Gain importance is computed as the mean gain per split averaged across all XGBoost trees. Counterfactual importance is defined as the product of frequency and magnitude of feature changes. Notably, Imc1 is consistently identified as the most important feature by both permutation and gain methods. In the counterfactual analysis, the highest-ranked features comprise two shape descriptors and one first-order feature, despite each appearing in a minority of counterfactuals.

[34], $\alpha/\beta = 3.28$ Gy, [40]). In contrast, Ko et al. [41] found an optimal α/β of approximately 14 Gy for the specific task of predicting radiation-induced peritumoral edema in meningioma. Theoretical concerns also exist regarding the validity of the linear-quadratic model in the context of stereotactic radiosurgery, although experimental evidence supports its robustness at least up to doses of approximately 20 Gy [42], [43], [44].

We strongly advocate for model explainability as a means of illuminating the internal workings of the model and enhancing its value as a decision support tool. Another important aspect of our work is precise task definition, that should be tailored to the specific needs of the clinician. To the best of our knowledge, this study is the first to distinguish between the prediction of de novo edema formation and exacerbation of pre-existing edema. Existing models treat both as a single task, achieving strong

metric performance in at least one case [23]. However, this comes with a paradoxical consequence. Feature importance analysis in the model by Li et al. ranks the presence of pre-treatment edema as the most important predictor of edema formation or exacerbation — an arguably obvious finding that, in our view, offers limited actionable insight for clinical decision-making. We believe that feature importance analysis could enhance the utility of the model as a decision support tool. Predictive importance consistently points to the role of a texture feature, the Informational Measure of Correlation (Imc1), while what-if importance appears to reflect a more complex interplay among features of diverse origin. Imc1 measures the normalized mutual information between rows and columns of the Gray Level Co-occurrence Matrix (GLCM), ranging from -1 (simple, highly structured texture) to 0 (complex, disordered texture). As such, it provides a context-adjusted, information-theoretic index of texture complexity.

It is tempting to speculate on a chain linking texture complexity on MRI, heterogeneity of contrast medium enhancement, microvascular architecture of the meningioma, and VEGF expression. Radiogenomic studies have developed radiomic models to predict VEGF and angiogenic signatures [45],[46], although none have specifically correlated VEGF expression with Imc1. Tumor microvasculature is known to be vulnerable to radiosurgical irradiation. One might further speculate that overdeveloped microvascularity could be disrupted by radiosurgery, shifting the tumor -or part of it- from a normoxic to a hypoxic state, thereby triggering HIF-induced VEGF-A upregulation (as in the pathway discussed in 1) and ultimately driving peritumoral edema formation. Notably, post- radiosurgical VEGF upregulation has already been documented in irradiated meningiomas [47], [48], and such marker elevation has been correlated with peritumoral edema development [47]. While this speculative pathway remains to be validated, it provides a biologically plausible framework for interpreting the predictive role of Im1 (or of any other texture-complexity related) in post-radiosurgical edema formation.

Data availability

The datasets generated and analysed during the current study are not publicly available due to patient privacy restrictions but are available from the corresponding author on reasonable request, subject to institutional and ethical approval.

Ethics Declarations

Ethical approval

This retrospective study was approved by the Ethics Committee of Brescia (approval number: NP 5238). The study was conducted in accordance with the Declaration of Helsinki.

Competing interests

The authors declare no competing interests.

Funding Declaration

This research received no funding.

References

[1] Ostrom, Q., Kruchko, C., Neff, C., Firth, A., Sherman, R.: The central brain tumor registry of the united states histopathological grouping scheme provides clinically relevant brain and other central nervous system categories for cancer registry data. Journal of Registry Management **49**(4), 139 (2022)

[2] Ogasawara, C., Philbrick, B.D., Adamson, D.C.: Meningioma: a review of epidemiology, pathology, diagnosis, treatment, and future directions. Biomedicines **9**(3), 319 (2021)

[3] Marchetti, M., Sahgal, A., De Salles, A.A., Levivier, M., Ma, L., Paddick, I., Pollock, B.E., Regis, J., Sheehan, J., Suh, J.H., *et al.*: Stereotactic radiosurgery for intracranial noncavernous sinus benign meningioma: international stereotactic radiosurgery society systematic review, meta-analysis and practice guideline. Neurosurgery **87**(5), 879–890 (2020)

[4] Lee, C.-C., Trifiletti, D.M., Sahgal, A., DeSalles, A., Fariselli, L., Hayashi, M., Levivier, M., Ma, L., Álvarez, R.M., Paddick, I., *et al.*: Stereotactic radiosurgery for benign (world health organization grade i) cavernous sinus meningiomas—international stereotactic radiosurgery society (isrs) practice guideline: a systematic review. Neurosurgery **83**(6), 1128–1142 (2018)

[5] Wang, C., Kaprealian, T.B., Suh, J.H., Kubicky, C.D., Ciporen, J.N., Chen, Y., Jaboin, J.J.: Overall survival benefit associated with adjuvant radiotherapy in who grade ii meningioma. Neuro-oncology **19**(9), 1263–1270 (2017)

[6] Mason, W.P., Gentili, F., Macdonald, D.R., Hariharan, S., Cruz, C.R., Abrey, L.E.: Stabilization of disease progression by hydroxyurea in patients with recurrent or unresectable meningiomas. Journal of neurosurgery **97**(2), 341–346 (2002)

[7] Bloch, O., Kaur, G., Jian, B.J., Parsa, A.T., Barani, I.J.: Stereotactic radiosurgery for benign meningiomas. Journal of neuro-oncology **107**(1), 13–20 (2012)

[8] Kim, Y.J., Moon, K.-S., Park, S.J., Jung, T.-Y., Kim, I.-Y., Jung, S.: Gamma knife radiosurgery as primary management for intracranial meningioma identified as growing on serial imaging. Medicine **103**(5), 37082 (2024)

[9] Jahanbakhshi, A., Najafi, M., Gomar, M., Ciammella, P., Ruggieri, M.P., Iotti, C., Finocchi Ghersi, S., Serre, A.-A., Bardoscia, L., Sardaro, A., *et al.*: Radiosurgery in grade ii and iii meningiomas: A systematic review and meta-analysis. Journal of Personalized Medicine **14**(8), 802 (2024)

[10] Zhang, G., , Wu, E., Li, Y., Zhang, Y., Lian, M., Geng, D., Zhu, G.: Peritumoral edema in meningiomas: a review of influencing factors, mechanisms, and management. Frontiers in Oncology **16**, 1740332 (2026)

[11] Berhouma, M., Jacquesson, T., Jouanneau, E., Cotton, F.: Pathogenesis of peritumoral edema in intracranial meningiomas. Neurosurgical review **42**, 59–71 (2019)

[12] Sheehan, J.P., Lee, C.-C., Xu, Z., Przybylowski, C.J., Melmer, P.D., Schlesinger, D.: Edema following gamma knife radiosurgery for parasagittal and parafalcine meningiomas. Journal of neurosurgery **123**(5), 1287–1293 (2015)

[13] Milano, M.T., Sharma, M., Soltys, S.G., Sahgal, A., Usuki, K.Y., Saenz, J.-M., Grimm, J., El Naqa, I.: Radiation-induced edema after single-fraction or multi-fraction stereotactic radiosurgery for meningioma: a critical review. International Journal of Radiation Oncology* Biology* Physics **101**(2), 344–357 (2018)

[14] Coroller, T.P., Bi, W.L., Huynh, E., Abedalthagafi, M., Aizer, A.A., Greenwald, N.F., Parmar, C., Narayan, V., Wu, W.W., Moura, S., *et al.*: Radiographic prediction of meningioma grade by semantic and radiomic features. PloS one **12**(11), 0187908 (2017)

[15] Park, Y.W., Oh, J., You, S.C., Han, K., Ahn, S.S., Choi, Y.S., Chang, J.H., Kim, S.H., Lee, S.-K.: Radiomics and machine learning may accurately predict the grade and histological subtype in meningiomas using conventional and diffusion tensor imaging. European radiology **29**, 4068–4076 (2019)

[16] Hamerla, G., Meyer, H.-J., Schob, S., Ginat, D.T., Altman, A., Lim, T., Gihr, G.A., Horvath-Rizea, D., Hoffmann, K.-T., Surov, A.: Comparison of machine learning classifiers for differentiation of grade 1 from higher gradings in meningioma: A multicenter radiomics study. Magnetic resonance imaging **63**, 244–249 (2019)

[17] Laukamp, K.R., Shakirin, G., Baeßler, B., Thiele, F., Zopfs, D., Hokamp, N.G., Timmer, M., Kabbasch, C., Perkuhn, M., Borggrefe, J.: Accuracy of radiomics-based feature analysis on multiparametric magnetic resonance images for noninvasive meningioma grading. World Neurosurgery **132**, 366–390 (2019)

[18] Zhu, Y., Man, C., Gong, L., Dong, D., Yu, X., Wang, S., Fang, M., Wang, S., Fang, X., Chen, X., *et al.*: A deep learning radiomics model for preoperative grading in meningioma. European journal of radiology **116**, 128–134 (2019)

[19] Khanna, O., Kazerooni, A.F., Farrell, C.J., Baldassari, M.P., Alexander, T.D., Karsy, M., Greenberger, B.A., Garcia, J.A., Sako, C., Evans, J.J., *et al.*: Machine learning using multiparametric magnetic resonance imaging radiomic feature analysis to predict ki-67 in world health organization grade i meningiomas. Neurosurgery **89**(5), 928 (2021)

[20] Zhang, J., Yao, K., Liu, P., Liu, Z., Han, T., Zhao, Z., Cao, Y., Zhang, G., Zhang, J., Tian, J., et al.: A radiomics model for preoperative prediction of brain invasion in meningioma non-invasively based on mri: A multicentre study. EBioMedicine **58** (2020)

[21] Xiao, D., Zhao, Z., Liu, J., Wang, X., Fu, P., Le Grange, J.M., Wang, J., Guo, X., Zhao, H., Shi, J., *et al.*: Diagnosis of invasive meningioma based on brain-tumor interface radiomics features on brain mr images: a multicenter study. Frontiers in Oncology **11**, 708040 (2021)

[22] Kim, K.H., Jung, S., Lee, H.-J., Kwon, H.-J., Choi, S.-W., Koh, H.-S., Youm, J.-Y., Kim, S.-H.: A deep neural network-based model predicting peritumoral edema after radiosurgery for meningioma. World Neurosurgery **164**, 280–289 (2022)

[23] Li, X., Lu, Y., Liu, L., Wang, D., Zhao, Y., Mei, N., Geng, D., Ma, X., Zheng, W., Duan, S., et al.: Predicting peritumoral edema development after gamma knife radiosurgery of meningiomas using machine learning methods: a multicenter study. European Radiology, 1–13 (2023)

[24] Quinn, T.P., Senadeera, M., Jacobs, S., Coghlan, S., Le, V.: Trust and medical ai: the challenges we face and the expertise needed to overcome them. Journal of the American Medical Informatics Association **28**(4), 890–894 (2021)

[25] Freyer, N., Groß, D., Lipprandt, M.: The ethical requirement of explainability for ai-dss in healthcare: a systematic review of reasons. BMC Medical Ethics **25**(1), 104 (2024)

[26] Ewald, F.K., Bothmann, L., Wright, M.N., Bischl, B., Casalicchio, G., König, G.: A guide to feature importance methods for scientific inference. In: Longo, L., Lapuschkin, S., Seifert, C. (eds.) World Conference on Explainable Artificial Intelligence, pp. 440–464. Springer, ??? (2024)

[27] Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, pp. 785–794 (2016)

[28] Altmann, A., Toloşi, L., Sander, O., Lengauer, T.: Permutation importance: a corrected feature importance measure. Bioinformatics **26**(10), 1340–1347 (2010)

[29] Mothilal, R.K., Sharma, A., Tan, C.: Explaining machine learning classifiers through diverse counterfactual explanations. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, pp. 607–617 (2020)

[30] Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C.: N4itk: improved n3 bias correction. IEEE transactions on medical imaging **29**(6), 1310–1320 (2010)

[31] Iglesias, J.E., Liu, C.-Y., Thompson, P.M., Tu, Z.: Robust brain extraction across datasets and comparison with publicly available methods. IEEE transactions on medical imaging **30**(9), 1617–1634 (2011)

[32] Reinhold, J.C., Dewey, B.E., Carass, A., Prince, J.L.: Evaluating the impact of intensity normalization on mr image synthesis. In: Elsa D. Angelini, B.A.L. (ed.) Proc SPIE Int Soc Opt Eng., vol. 10949, pp. 890–898. SPIE, ??? (2019)

[33] Van Griethuysen, J.J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R.G., Fillion-Robin, J.-C., Pieper, S., Aerts, H.J.: Computational radiomics system to decode the radiographic phenotype. Cancer research **77**(21), 104–107 (2017)

[34] Vernimmen, F.J., Slabbert, J.P.: Assessment of the $\alpha/ß$ ratios for arteriovenous malformations, meningiomas, acoustic neuromas, and the optic chiasma. International journal of radiation biology **86**(6), 486–498 (2010)

[35] Santacroce, A., Kamp, M.A., Budach, W., Hänggi, D.: Radiobiology of radiosurgery for the central nervous system. BioMed research international **2013**(1), 362761 (2013)

[36] Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V.X., Doshi-Velez, F., Jung, K., Heller, K., Kale, D., Saeed, M., *et al.*: Do no harm: a roadmap for responsible machine learning for health care. Nature medicine **25**(9), 1337–1340 (2019)

[37] Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M.: Optuna: A next-generation hyperparameter optimization framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2623–2631 (2019)

[38] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. Journal of Machine Learning Research **12**, 2825–2830 (2011)

[39] Guidotti, R.: Counterfactual explanations and how to find them: literature review and benchmarking. Data Mining and Knowledge Discovery, 1–55 (2022)

[40] Shrieve, D.C., Hazard, L., Boucher, K., Jensen, R.L.: Dose fractionation in stereotactic radiotherapy for parasellar meningiomas: radiobiological considerations of efficacy and optic nerve tolerance. Journal of neurosurgery **101**(Supplement3), 390–395 (2004)

[41] Ko, S.-W., Won, Y.D., Ha, B.J., Cheong, J.H., Ryu, J.I., Hong, S.W., Min, K.-W., Han, M.-H.: Optimal α/β ratio for biologically effective dose-based prediction of radiation-induced peritumoral brain edema in meningioma. Cancers **18**(3), 448 (2026)

[42] Wang, J.Z., Huang, Z., Lo, S.S., Yuh, W.T., Mayr, N.A.: A generalized linear-quadratic model for radiosurgery, stereotactic body radiation therapy, and high–dose rate brachytherapy. Science translational medicine **2**(39), 39–483948 (2010)

[43] Nahum, A.E.: The radiobiology of hypofractionation. Clinical oncology **27**(5), 260–269 (2015)

[44] Lam, F.C., Byun, J., Guru, S., AbuReesh, D., Hori, Y.S., Rahimy, E., Pollom, E.L., Soltys, S., Park, D.J., Chang, S.D.: Understanding the radiobiology of central nervous system diseases in the golden age of radiosurgery—does it matter? Brain Sciences **15**(6), 649 (2025)

[45] Sun, Z., Li, Y., Wang, Y., Fan, X., Xu, K., Wang, K., Li, S., Zhang, Z., Jiang, T., Liu, X.: Radiogenomic analysis of vascular endothelial growth factor in patients with diffuse gliomas. Cancer Imaging **19**(1), 68 (2019)

[46] Zheng, F., Chen, B., Zhang, L., Chen, H., Zang, Y., Chen, X., Li, Y.: Radiogenomic analysis of vascular endothelial growth factor in patients with glioblastoma. Journal of computer assisted tomography **47**(6), 967–972 (2023)

[47] Kan, P., Liu, J.K., Wendland, M.M., Shrieve, D., Jensen, R.L.: Peritumoral edema after stereotactic radiosurgery for intracranial meningiomas and molecular factors that predict its development. Journal of neuro-oncology **83**(1), 33–38 (2007)

[48] Park, S.-H., Hwang, J.-H., Hwang, S.-K.: Change in plasma vascular endothelial growth factor after gamma knife radiosurgery for meningioma: a preliminary study. Journal of Korean Neurosurgical Society **57**(2), 77 (2015)

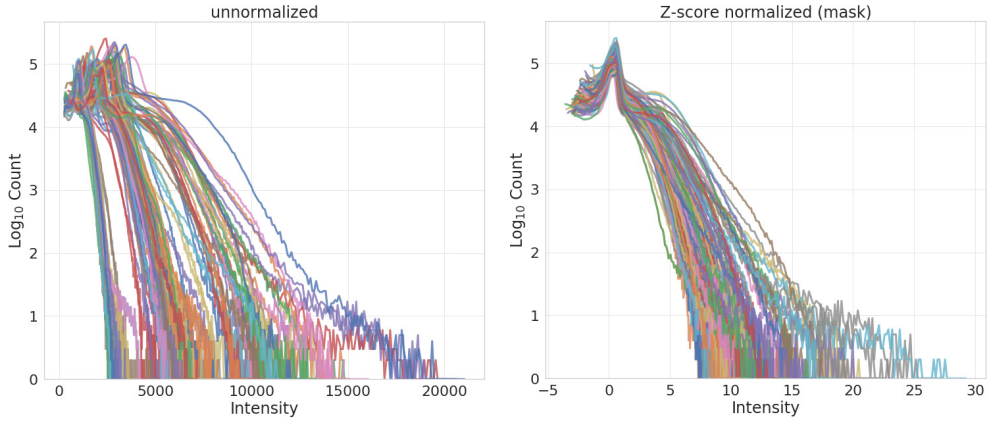


Fig. S1 Histograms of intensity values of MRI planning images before (left) and after (right) z-score normalisation [32]. The histogram of normalised images show greater consistency across cases, although some residual dispersion persists at low voxel counts.