Radiological Physics and Technology
https://doi.org/10.1007/s12194-026-01068-4

RESEARCH ARTICLE

A data-driven scientific approach to explore the relationship between MGMT methylation and imaging phenotype in glioblastoma

Sakura Sugyo[1] · Mio Ishii[2] · Mazen Soufi[3] · Yoshikazu Uchiyama[3]

Received: 21 January 2026 / Revised: 1 May 2026 / Accepted: 13 May 2026

Abstract

Glioblastoma is a highly malignant brain tumor, and patient survival remains short despite multimodal treatment. One important prognostic factor is the methylation status of the O⁶-methylguanine DNA methyltransferase (MGMT) gene, as MGMT promoter methylation is associated with a better response to temozolomide chemotherapy. The aim of this study was to evaluate the association between noninvasively obtained imaging features and MGMT methylation status. Magnetic resonance (MR) images and MGMT methylation data for 55 glioblastoma cases (33 methylated, 22 unmethylated) were obtained from the publicly available TCGA-GBM database. After brain morphological standardization, tumor centroid coordinates were calculated and used as inputs for a quadratic discriminant classifier to assess the separability of MGMT methylation status based on anatomical location. In addition, 279 radiomic features, including histogram- and texture-based metrics, were extracted from tumor regions. Using Lasso regression, nine features were selected and subsequently analyzed using linear discriminant analysis. The areas under the receiver operating characteristic curve (AUCs) were 0.71 for anatomical location–based classification and 0.85 for radiomic feature–based classification. Because the correlation between anatomical location and radiomic features was low, integrating both feature sets improved discrimination performance to an AUC of 0.90. These results suggest that imaging examinations capture complementary information regarding tumor morphology and anatomical origin. The anatomical location of glioblastoma may therefore provide useful clues for classifying MGMT gene methylation status.

Keywords Anatomical location · Radiomics · Glioblastoma · MGMT methylation · MR images

1 Introduction

Glioblastoma is the most aggressive malignant brain tumor, and even with multimodal treatment combining surgery, radiotherapy, and chemotherapy, it remains a highly lethal cancer with a median survival of less than 15 months [1]. In recent years, the epigenetic status of O⁶-methylguanine DNA methyltransferase (MGMT) gene promoter methylation has attracted considerable attention as a factor influencing the prognosis of glioblastoma. The MGMT gene encodes a DNA repair enzyme that counteracts the cytotoxic effects of temozolomide chemotherapy; however, glioblastomas with a methylated MGMT promoter have been reported to respond favorably to temozolomide treatment [2]. Consequently, assessment of MGMT promoter methylation status has become an important indicator for determining treatment strategies.

However, evaluation of MGMT promoter methylation requires the invasive acquisition and analysis of tumor tissue, imposing a substantial burden on patients. In addition, current analytical methods are complex and require considerable time and cost for result interpretation. To address these limitations, research has increasingly focused on estimating MGMT promoter methylation status from noninvasive magnetic resonance (MR) images in order to predict the efficacy of temozolomide chemotherapy in advance.

✉ Yoshikazu Uchiyama
y-uchiyama@miyazaki-u.ac.jp

[1] Department of Radiology, Chiba University Hospital, 1-8-1 Inohana, Chuo-ku, Chiba-shi, Chiba 260-8677, Japan

[2] Radiology Division, University of Miyazaki Hospital, 5200 Kihara, Kiyotake-cho, Miyazaki-shi, Miyazaki 889-1692, Japan

[3] Information and Communication Technology Program, Faculty of Engineering, University of Miyazaki, 1-1 Gakuen Kibanadai-nishi, Miyazaki-shi, Miyazaki 889-2192, Japan

Published online: 25 May 2026

Specifically, previous studies have proposed approaches based on deep learning for MGMT methylation prediction [3–5], transfer learning–based methods [6], graph neural networks that treat image patches as graph nodes [7], classification using image features extracted by deep learning in combination with radiomics features [8], machine learning methods using radiomics features as input [9, 10], and systematic investigations of methodologies for MGMT methylation prediction [11]. These studies suggest that noninvasive imaging-based estimation of MGMT methylation status has the potential to support clinical decision-making in the treatment of glioblastoma.

In radiomics research, it is essential to clarify whether imaging findings of glioblastoma contain information related to MGMT promoter methylation. Therefore, rather than relying on complex and less interpretable methods such as deep learning, it is desirable to evaluate this relationship using methods that are as simple and interpretable as possible. Moreover, in the practice of genomic medicine, it is necessary to clearly define the respective roles of imaging and genetic testing and to fully exploit information that can be obtained exclusively from imaging. A key advantage of imaging examinations lies not only in providing morphological information, such as tumor size and shape, but also in capturing anatomical information regarding the location at which the tumor arises.

Although previous studies [3–11] have primarily proposed methods for estimating MGMT methylation status based on tumor morphological features, the potential relationship between tumor anatomical location and MGMT methylation remains largely unexplored. Clinically, tumor location can influence surgical planning, functional outcomes, and potentially prognosis, suggesting that tumor site may provide biologically and clinically relevant information. However, these relationships have not yet been systematically examined using a data-driven scientific approach that enables objective and quantitative exploration of imaging–genomic associations. In this study, we specifically investigate the anatomical location of glioblastoma to determine whether it offers complementary information for predicting MGMT promoter methylation. Furthermore, by integrating location information with radiomics features, we assess whether these complementary imaging biomarkers can enhance the noninvasive estimation of MGMT methylation status. Specifically, we demonstrate that anatomical location has a low correlation with radiomics features and that each provides distinct information for classification. Furthermore, we show that integrating these complementary sources of information leads to improved classification performance for MGMT methylation status.

2 Methods

2.1 Dataset

In this study, we used The Cancer Genome Atlas Glioblastoma Multiforme (TCGA-GBM) dataset provided by The Cancer Imaging Archive (TCIA) [12]. Although this database contains information from 262 patients with glioblastoma, not all cases include both imaging and genomic data. Therefore, we conducted experiments using 55 cases for which both preoperative Fluid Attenuated Inversion Recovery (FLAIR) images and MGMT gene methylation status were available. The inclusion and exclusion criteria were as follows. From the TCGA-GBM dataset, cases were included if they had (1) preoperative FLAIR MR images and (2) corresponding MGMT promoter methylation status. Cases lacking either imaging or genomic data were excluded. In addition, cases with insufficient image quality that precluded reliable tumor delineation were excluded from the analysis.

The imaging parameters of the preoperative FLAIR images used in this study were as follows: repetition time (TR) of 8000–11000 ms, echo time (TE) of 73–155 ms, and inversion time (TI) of 2000–2700 ms. The patients' ages ranged from 21 to 84 years (mean: 57.4 years), including 27 males and 28 females. Among these cases, 33 patients had MGMT methylation, while 22 patients did not.

In this study, experiments were performed using only FLAIR images in order to facilitate comparison of results obtained using anatomical location information and radiomics features. FLAIR images were selected because previous studies have reported that FLAIR images are useful for discriminating MGMT methylation status [3], and because preliminary experiments (see Supplementary Materials for details) using T1-weighted, T2-weighted, and FLAIR images demonstrated that FLAIR images achieved the best classification performance.

2.2 Measurement of tumor anatomical location

Figure 1 illustrates an overview of the proposed method. Because brain MR images vary in size and acquisition orientation across cases, anatomical standardization is required to measure tumor locations using a unified coordinate system. In this study, Statistical Parametric Mapping (SPM12) was adopted for this standardization process [13]. Although SPM12 provides several normalization options, only rigid-body transformation was applied in order to avoid deformation of brain morphology. In this rigid transformation, images were aligned to the Montreal Neurological Institute (MNI) template provided in SPM12. As a result, differences in brain size and anatomical location were corrected, and

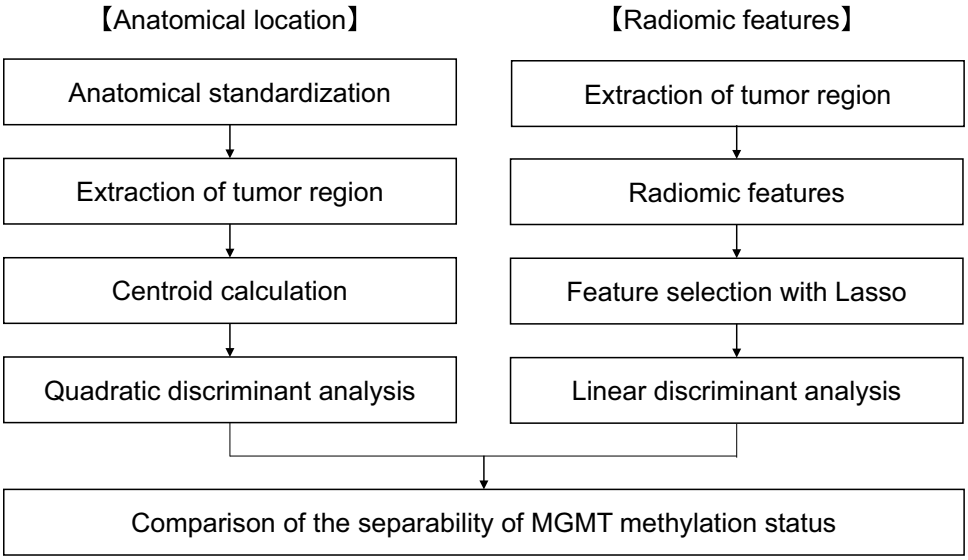

Fig. 1 Exploring genotype–phenotype associations in cancer using a data-driven scientific approach

all MR images were standardized to a size of $79 \times 95 \times 79$ voxels. When the slice spacing of the input images is large and the resolution along the body axis is low, interpolated slices may result in coarse approximations that do not preserve fine intra-tumoral structures. However, this limitation does not pose a problem when only the approximate tumor contour is required.

After anatomical standardization, hyperintense regions on FLAIR images were regarded as tumor regions and manually delineated by one of the authors (a radiological technologist with five years of clinical experience). The tumor centroid coordinates (x, y, z) and slice-wise tumor area were then calculated to obtain anatomical location information. Considering the aforementioned issue with interpolated slices, a single slice in which the tumor exhibited the largest diameter was selected; the slice index was used as the z-coordinate. The x- and y-coordinates of the tumor centroid were computed from the manually delineated tumor region on this two-dimensional slice.

In this study, the tumor anatomical location was represented by the centroid of each tumor. The centroid provides a simple and interpretable measure of tumor position in three-dimensional space, allowing straightforward comparison across cases after anatomical standardization. While surface or boundary information could provide additional details regarding tumor morphology, the centroid is sufficient for evaluating the overall spatial distribution of tumors with respect to MGMT promoter methylation. Furthermore, using the centroid reduces computational complexity and minimizes potential noise from irregular tumor shapes, which is particularly important given the limited number of cases in this study.

Figure 2 shows the plotted centroid coordinates of tumors from the 55 cases. The size of each sphere represents the tumor area; however, because area does not directly reflect anatomical location, it was not used in the subsequent analysis. Tumors in cases with MGMT methylation tended to be located in the superior or inferior regions of the brain, whereas tumors in cases without MGMT methylation tended to be located in the central region. Quantitative evaluation of this tendency is presented in the experimental results as the classification performance for MGMT methylation status using only anatomical location features.

2.3 Quantification of the separability of MGMT methylation status using anatomical location

To quantitatively evaluate the separability of MGMT methylation status based on the anatomical location, a quadratic discriminant classifier (Quadratic Discriminant Analysis; QDA) was employed using the tumor centroid coordinates as input features [14]. QDA forms a quadratic decision boundary to classify two groups in a three-dimensional space [14]. As shown in Fig. 2, tumors in cases without MGMT methylation tended to be located in the central region of the brain. When QDA is applied to such data distributions in three-dimensional space, the resulting decision boundary is represented as a quadratic surface. In the present study, graphical visualization of the decision boundary demonstrated that it forms a hyperboloid rather than an ellipsoidal surface, as initially assumed. By setting a cutoff value on the QDA output, it becomes possible to quantitatively evaluate the ability to separate tumors located in different regions of the brain. In this sense, QDA can be regarded as the simplest evaluation function for quantifying whether a tumor is located in regions separated by the quadratic decision boundary defined by QDA.

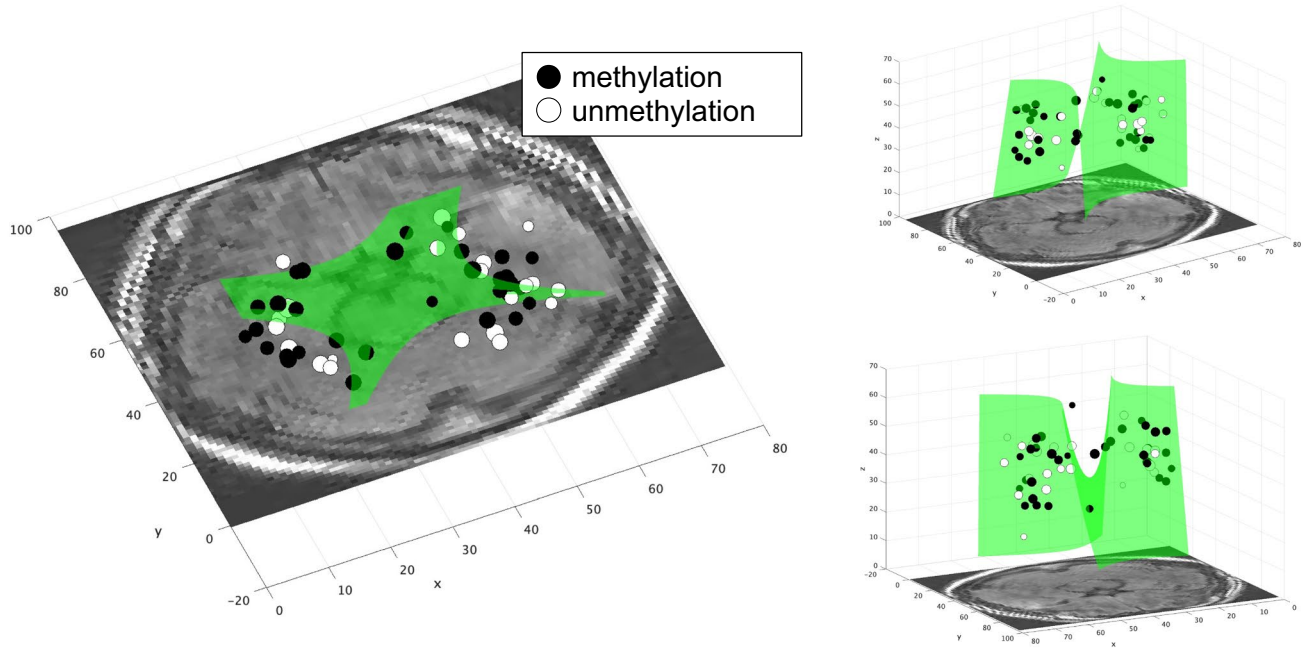

Fig. 2 Anatomical locations of methylated and unmethylated cases plotted on the coordinate axes after brain morphology standardization. The Green curve represents the discrimination boundary and circle sizes represent tumor areas

The performance of separability was evaluated using a receiver operating characteristic (ROC) curve derived from the QDA output values and the corresponding area under the curve (AUC) [15]. ROC analysis was performed using the LABROC algorithm developed at the University of Chicago [15]. This algorithm is capable of estimating the properties of two underlying populations from a limited number of samples and provides 95% confidence intervals for the separability performance.

In this study, resubstitution [16] was adopted when applying QDA. In general, machine learning frameworks consist of two distinct components: a feature extraction stage and a classification stage. It is important to emphasize that the purpose of using QDA in this study was not to evaluate the generalization performance of a classifier, but rather to quantify the intrinsic separability of MGMT methylation status based solely on anatomical location features.

Specifically, QDA was used as a mathematical tool to model the spatial distribution of tumor centroids and to construct a quadratic decision boundary that characterizes differences in their spatial distributions. The output of QDA was subsequently used to determine cutoff values for ROC analysis. Under this framework, resubstitution is appropriate because the analysis aims to assess the discriminative capability of the extracted anatomical features themselves, rather than to develop or validate a predictive model. Therefore, the ROC and AUC results obtained in this study should be interpreted as an evaluation of the separability provided by the anatomical location features alone, and not as an assessment of the classification performance or generalization ability of QDA as a classifier.

2.4 Measurement of radiomics features

Next, for comparison with the use of tumor anatomical location, radiomics features were extracted from FLAIR images. For each case, a single slice in which the tumor exhibited the largest diameter was selected from the FLAIR image, and radiomics features were computed in this slice.

To normalize the pixel intensities of the FLAIR images, intensity scaling was applied to the FLAIR images from all 55 cases. Because the FLAIR images contained noise with extremely high pixel intensities, an intensity histogram was first generated, and the top 0.1% of pixel intensities were identified as a threshold. All pixel values above this threshold were set to 1023, after which a linear intensity transformation was applied so that the minimum and maximum pixel values were mapped to 0 and 1023, respectively.

Radiomics feature extraction was performed using the freeware MaZda [17–19]. The tumor region was manually delineated as a hyperintense area on the FLAIR image by one of the authors (a radiological technologist with five years of clinical experience), and subsequently reviewed and corrected, if necessary, by another author. An example of the delineation results is shown in Fig. 3. From the extracted tumor region, a total of 279 radiomics features were calculated, including nine histogram-based features, 245 texture features, five features related to autoregressive

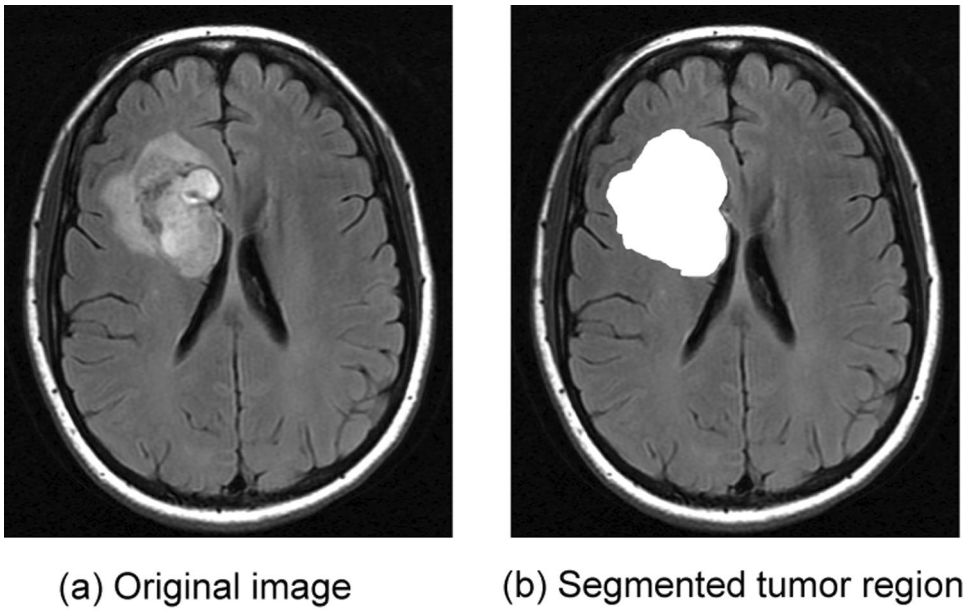

Fig. 3 Example of manually segmented tumor region. **a** original image, **b** segmented tumor region

models, and 20 features related to image resolution. All parameters for radiomics feature extraction were set to the default values provided by MaZda. For example, in the calculation of texture features, gray-level co-occurrence matrices were computed with 16 Gy levels, pixel distances of 1–5, and directions of 0°, 45°, 90°, and 135°. In this study, hyperintense regions on FLAIR images were defined as tumor regions to facilitate manual delineation; however, it should be noted that this approach may include edematous regions that do not exhibit typical glioblastoma characteristics. Furthermore, to minimize the influence of variations in tumor delineation, features that depend on tumor contours were excluded from the analysis.

2.5 Quantification of the separability of MGMT methylation status using radiomics features

Although 55 cases were available for the experiments, the number of radiomics features (279) is high-dimensional, necessitating dimensionality reduction. In this study, feature selection was performed using the least absolute shrinkage and selection operator (Lasso) [20]. Lasso is defined as follows:

$$\widehat{\beta}^{lasso} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \quad (1)$$

where y_i represents the MGMT methylation status of the i -th case, x_j denotes the value of the j-th radiomics feature, β_0 is an intercept term, $\lambda \geq 0$ is a regularization parameter controlling the degree of shrinkage, and p is the total number of radiomics features. A fourfold cross-validation procedure was used to determine the optimal value of λ, which resulted in the selection of 11 radiomics features.

Next, to quantify the separability of MGMT methylation status in the 11-dimensional selected feature space, linear discriminant analysis (LDA) was applied to determine a separating hyperplane. Resubstitution was also employed when applying LDA. It should be emphasized that LDA was not used as a predictive classifier, but rather as a tool to generate a boundary for evaluating the separability between the two groups in the selected feature space. By setting a cutoff value on the LDA output, the separability within the feature space can be quantitatively assessed.

The performance of separability was further evaluated using ROC curves and the corresponding AUC. ROC analysis was conducted by applying the LDA output values as inputs, allowing for a quantitative assessment of how well the selected radiomics features distinguish between cases with and without MGMT methylation. This approach provides a standardized metric for comparing the discriminative power of anatomical location features and radiomics features in the present study.

3 Results

Figure 4 shows the results of ROC analysis for MGMT methylation status using anatomical location information alone. When only anatomical location was used, AUC was 0.71 (95% CI: 0.56–0.84). Table 1 lists the radiomics features selected by Lasso. The selected features included one histogram feature, six texture features, an autoregressive model feature, and three resolution-related features, for a total of 11 features. The results of MGMT methylation

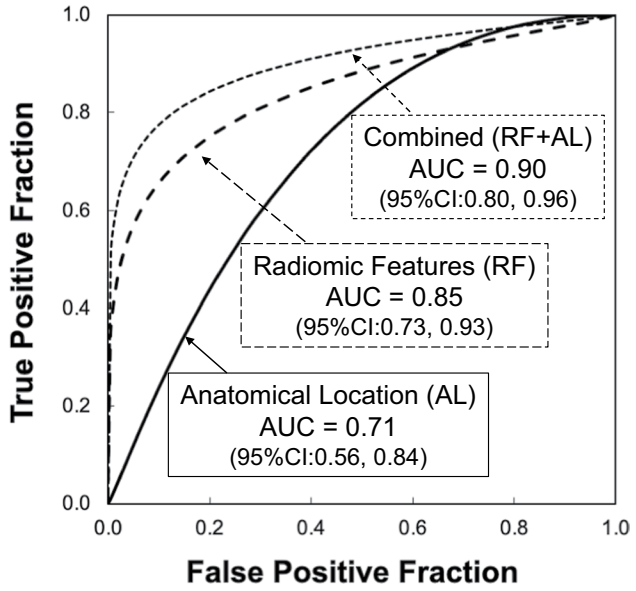

Fig. 4 ROC curves for evaluating the separability between methylated and unmethylated cases

Table 1 Eleven radiomic features selected by Lasso

#	Feature	Category	Description
#1	Skewness	Histogram	Histogram's skewness
#2	S(1,0) Correlat	Texture Cooccurrence matrix	Correlation S(1,0) is the between-pixel distance
#3	S(5,5)Correlat	Texture Cooccurrence matrix	Correlation S(5,5) is the between-pixel distance
#4	S(0,5)AngScMom	Texture Cooccurrence matrix	Angular second moment S(0,5) is the between-pixel distance
#5	S(5,5)DifVarnc	Texture Cooccurrence matrix	Difference variance S(5,5) is the between-pixel distance
#6	45dgr_RLNonUni	Texture Run length	Run length nonuniformity 45dgr is angle
#7	135dr_LngREmph	Texture Run length	Long run emphasis 135dr is angle
#8	Teta1	Autoregressive model	Parameter θ_1
#9	WavEnHH_s-3	Resolution	Wavelet energy HH is frequency band at scale 3
#10	WavEnHL_s-5	Resolution	Wavelet energy HL is frequency band at scale 5
#11	WavEnHH_s-5	Resolution	Wavelet energy HH is frequency band at scale 5

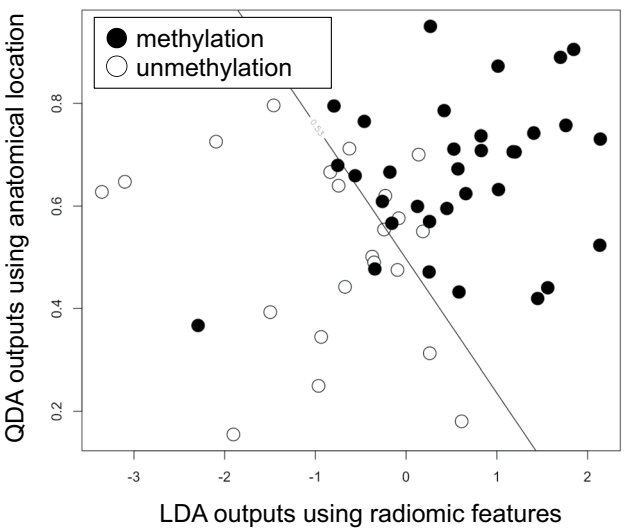

Fig. 5 Relationship between output values of the LDA using radiomic features and the QDA using anatomical locations

status classification using these 11 radiomics features are also shown in Fig. 4. When only radiomics features were used, the AUC was 0.85 (95% CI: 0.73–0.93). To further evaluate the robustness of the observed separability and to address potential optimistic bias associated with resubstitution, additional analysis was performed using leave-one-out cross-validation (LOOCV). For anatomical location features, QDA with LOOCV yielded an AUC of 0.71 (95% CI: 0.56–0.83), which was comparable to the result obtained using resubstitution. In contrast, for radiomics features, LDA with LOOCV resulted in a lower AUC of 0.66 (95% CI: 0.51–0.79), compared to the resubstitution-based AUC of 0.85. These results suggest that the discriminative performance of radiomics features may be overestimated when evaluated using resubstitution, whereas anatomical location features demonstrated relatively stable performance under cross-validation. Overall, the LOOCV analysis provides a more conservative estimate of separability and highlights the importance of accounting for variability in small-sample studies.

Figure 5 presents a scatter plot in which the horizontal axis represents the LDA output based on radiomics features alone, and the vertical axis represents the QDA output based on anatomical location alone. Tumors in cases with MGMT methylation tended to cluster in the upper-right region, whereas tumors in cases without methylation tended to cluster in the lower-left region. To quantitatively evaluate this trend, LDA was applied using the vertical and horizontal outputs as input features. The line in Fig. 5 represents the discriminant boundary obtained by LDA. Cases located above and to the right of this boundary were classified as methylated, while cases below and to the left were classified as unmethylated. The ROC curve for the combined approach

Table 2 Confusion matrix for the integration of radiomics features and anatomical location

		Computer output	
		methylation	unmethylation
Truth	methylation	90.9% (30/33)	9.1% (3/33)
	unmethylation	22.7% (5/22)	77.3% (17/22)

has been added to Fig. 4 alongside the ROC curves for location alone and radiomics alone. This allows a direct visual comparison of sensitivity and specificity across the three methods, confirming the complementary nature of these imaging biomarkers. The AUC for this classification was 0.90 (95% CI: 0.80–0.96), confirming that the integration of anatomical location information and radiomics features improved separability. Additionally, the confusion matrix corresponding to the classification using this discriminant boundary with a threshold is shown in Table 2. The sensitivity was 90.9% (30/33), and the specificity was 77.3% (17/22).

4 Discussion

Radiomics is often equated with texture analysis, but this is a narrow definition. The term Radiomics is a portmanteau of "Radio," referring to imaging, and "Omics," referring to genomics, proteomics, and other molecular data. In its broader sense, Radiomics provides a novel framework for medical imaging that analyzes the relationship between tumor genotype and imaging phenotype, thereby supporting the implementation of cancer genomic medicine. A challenge in Radiomics studies is that classifying tumors based on genotype often results in small sample sizes. Consequently, there are almost no previous studies that have applied deep learning to Radiomics using tens of thousands of cases [3–5], and results from deep learning trained on small sample sizes are generally unreliable. Therefore, for Radiomics studies analyzing the relationship between tumor phenotype and genotype, machine learning approaches capable of learning from limited data are more suitable. In general, machine learning consists of feature extraction and classification stages, but for analyzing genotype–phenotype relationships in cancer, the extraction of image features is particularly crucial. If image features capable of clearly discriminating tumor genotypes can be obtained, it may even be possible for imaging to partially substitute for genetic testing. It should be noted that the present study focused on designing image features to determine MGMT methylation status and did not evaluate the performance of the classifier itself.

Images provide information that cannot be obtained from genetic tests, such as tumor location and morphological characteristics. Previous studies [3–11] have proposed methods to estimate MGMT methylation status using tumor morphology, but none have focused solely on tumor location. Therefore, this study aimed to explore the potential of Radiomics in its broad sense as a new imaging-based framework supporting cancer genomic medicine and to investigate whether tumor location, as a component of the tumor phenotype, can provide useful information. Next, we discuss the relationship between tumor anatomical location and MGMT methylation status. In the present study, the discriminative performance of MGMT methylation status using only tumor anatomical location was AUC = 0.71 (95% CI: 0.56–0.84), suggesting that the site of glioblastoma occurrence may provide clues for predicting MGMT methylation. Previous studies have suggested that the anatomical location of glioblastoma may be associated with its molecular and biological characteristics. In particular, glioblastoma subtypes have been reported to occupy distinct regions of the brain and to differ in their proximity to the subventricular zone (SVZ) [21]. The SVZ, which is known as a neural stem cell niche, has been proposed as a potential origin of glioblastoma [22], and tumors involving this region have been shown to exhibit distinct molecular profiles and clinical behavior [23]. Furthermore, recent studies have investigated the relationship between SVZ involvement and MGMT promoter methylation status [24]. Although the direct relationship between MGMT promoter methylation and tumor

location has not been fully established, it is plausible that regional differences in the brain microenvironment, cellular origin, and epigenetic regulation contribute to variability in MGMT methylation status. From this perspective, the tendency observed in the present study—that tumors with different MGMT methylation status show distinct spatial distributions—may reflect underlying biological heterogeneity associated with tumor origin and growth patterns. Therefore, anatomical location, even when represented by a simple feature such as the tumor centroid, may provide indirect but meaningful information about tumor genotype. This finding is significant from a data-driven scientific perspective, as the study proposes a hypothesis regarding the relationship between tumor location and MGMT methylation, which should be further investigated with a larger cohort. As such, when machine learning is used to extract latent relationships between tumor phenotype and genotype, the design and evaluation of image features is more critical than the classifier itself. In Radiomics studies, it is not sufficient to simply construct high-accuracy classifiers; identifying which image features contribute to classification performance is essential for establishing a new imaging framework that supports cancer genomic medicine.

Next, the relationship between tumor location and radiomics features is discussed. Figure 5 shows the relationship between the QDA output using anatomical location (from the left-hand processing in Fig. 1) and the LDA output using radiomics features (from the right-hand processing in Fig. 1). These outputs represent the separability based solely on anatomical location and solely on radiomics features, respectively. If anatomical location and radiomics features provided identical information for classifying MGMT methylation status, the scatter plot in Fig. 5 would show a strong correlation, with points aligned along a straight line. However, the scatter plot in Fig. 5 is widely distributed, indicating a low correlation between anatomical location and radiomics features (correlation coefficient = 0.26), and demonstrating that they provide complementary information for MGMT methylation classification. Furthermore, integrating anatomical location and radiomics features improved classification performance to AUC = 0.90 (95% CI: 0.80–0.94). In addition, the LOOCV results demonstrated that the discriminative performance, particularly for radiomics features, decreased compared to the resubstitution results. This finding suggests the presence of optimistic bias and highlights the importance of evaluating variability, especially in small-sample studies. Furthermore, to ensure a fair comparison between anatomical location and radiomics features, QDA was also applied to the selected radiomics features and evaluated using LOOCV. The QDA-based analysis yielded an AUC of 0.72 (95% CI: 0.57–0.84), which was higher than that obtained using LDA (AUC = 0.658) and comparable to the performance achieved using anatomical location features (AUC = 0.71). In addition, the combined analysis integrating anatomical location and radiomics features was evaluated using the same QDA framework under LOOCV. The combined model achieved an AUC of 0.76 (95% CI: 0.61–0.87), which was higher than that obtained using anatomical location alone (AUC = 0.71) and radiomics features alone (AUC = 0.72). These results suggest that the integration provides complementary information, although the degree of improvement under LOOCV is modest. In addition, the findings indicate that nonlinear decision boundaries may better capture the structure of the radiomics feature space. At the same time, the comparable performance between anatomical location and radiomics features indicates that both feature types possess similar levels of discriminative capability when evaluated under matched model conditions. Furthermore, the comparison between anatomical location and radiomics features depends on the choice of classifier. In the initial analysis, QDA was used for anatomical location, whereas LDA was used for radiomics features, which may have introduced an imbalance in model flexibility. In this study, additional analysis using QDA for radiomics features demonstrated improved performance (AUC = 0.72), indicating the presence of nonlinear relationships in the radiomics feature space. Notably, when evaluated under comparable model conditions, anatomical location and radiomics features exhibited similar levels of discriminative performance.

A limitation of this study is that only 55 cases had both imaging and genetic data available. While it is possible that MGMT methylation status could be inferred from imaging, the 95% confidence interval of the AUC is relatively wide, highlighting the need for future studies with larger sample sizes to further validate and refine these findings.

5 Conclusion

Information obtained from medical imaging provides insights into tumor morphology and location that cannot be obtained through genetic testing. This study suggests that the site of glioblastoma occurrence may serve as a clue for predicting MGMT methylation status. Furthermore, it was demonstrated that tumor location provides information distinct from that of radiomics features. These findings offer new insights into which imaging characteristics can be utilized to assess the presence or absence of MGMT methylation in glioblastoma.

5.1 Supplementary methods and results: preliminary experiment

FLAIR, T1-weighted, and T2-weighted MR images from 23 patients in the TCGA-GBM dataset were analyzed. Fourteen patients had MGMT promoter methylation and nine did not. For each case, the slice with the largest tumor area was selected, and the tumor region was manually segmented on each image. A total of 375 radiomics features were extracted. Using LASSO regression, three radiomics features relevant for discriminating MGMT methylation status were selected each MRI sequence. Classification performance was evaluated using linear discriminant analysis and receiver operating characteristic (ROC) analysis. The area under the curve (AUC) was 0.931 for FLAIR images, 0.901 for T1-weighted images, and 0.892 for T2-weighted images. Based on these results, FLAIR images were selected for subsequent analyses. These preliminary results were presented at the 76th Annual Meeting of the Japanese Society of Radiological Technology (April 2020).

Funding A part of this study was funded by JSPS KAKENHI Grant Number JP 24K15794.

Data availability Data will be made available on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest to disclose.

Ethical approval We used data from a public database in this study. The IRB of our institute allowed us to use data from those cases.

References

1. Report of Brain Tumor Registry of Japan(2001–2004) . 13th Edition Neurologia medico-chirurgica, Suppl. 2014; 54.
2. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Eng J Med. 2005;352:987–96.
3. Saeed N, Ridzuan M, Alasmawi H, et al. MGMT promoter methylation status prediction using MRI scans? An extensive experimental evaluation of deep learning models. Med Image Anal. 2023;90:102989.
4. Kim BH, Lee H, Choi KS, et al. Validation of MRI-based models to predict MGMT promoter methylation in gliomas: BraTS 2021 radiogenomics challenge. Cancers (Basel). 2022;14(19):4827.
5. Capuozzo S, Gravina M, Gatta G, et al. A multimodal knowledge-based deep learning approach for MGMT promoter methylation identification. J Imaging. 2022;8:321.
6. Sakly H, Said M, Seekins J, et al. Brain tumor radiogenomic classification of O^6-Methylguanine-DNA Methyltransferase promoter methylation in malignant gliomas-based transfer learning. Cancer Control. 2023;30:10732748231169148.
7. Hu M, Yang K, Wang J, et al. MGMT promoter methylation prediction based on multiparametric MRI via vision graph neural network. J Med Imaging (Bellingham). 2024;11:014503.
8. Qureshi SA, Hussain L, Ibrar U, et al. Radiogenomic classification for MGMT promoter methylation status using multi-omics fused feature space for least invasive diagnosis through mpMRI scans. Sci Rep. 2023;13:3291.
9. Chen S, Xu Y, Ye M, et al. Predicting MGMT promoter methylation in diffuse gliomas using deep learning with radiomics. J Clin Med. 2022;11:3445.
10. Do DT, Yang MR, Lam LHT, et al. Improving MGMT methylation status prediction of glioblastoma through optimizing radiomics features using genetic algorithm-based machine learning approach. Sci Rep. 2022;12:13412.
11. Samartha MVS, Dubey NK, Jena B, et al. AI-driven estimation of O6 methylguanine-DNA-methyltransferase (MGMT) promoter methylation in glioblastoma patients: a systematic review with bias analysis. J Cancer Res Clin Oncol. 2024;150:57.
12. TCGA-GBM, https://www.cancerimagingarchive.net/collection/tcga-gbm/ (Accessed 8 January 2025).
13. SPM, https://www.fil.ion.ucl.ac.uk/spm/. Accessed 17 Jan 2026.
14. Duda RO, Hart PE, Stork DG. Pattern classification. New York: John Wiley & Sons; 2001.
15. Metz CE. Sonic practical issues of experimental design and data analysis in radiological ROC studies. Invest Radiol. 1989;24:234–45.
16. Theodoridis S, Koutroumbas K. Pattern recognition. London: Academic Press; 1999.
17. MaZda, https://qmazda.p.lodz.pl/. Accessed 17 Jan 2026.
18. Szczypinnski PM, Strzelecki M, Materka A, et al. MaZda–a software package for image texture analysis. Comput Methods Programs Biomed. 2009;94:66–76.
19. Strzelecki M, Szczypinski P, Materka A, et al. A software tool for automatic classification and segmentation of 2D/3D medical images. Nucl Instrum Methods Phys Res Sect A. 2013;702:137–40.
20. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning, data minimg, inference and prediction. second edition. New York: Springer; 2009.
21. Steed TC, Treiber JM, Patel K, et al. Differential localization of glioblastoma subtype: implications on glioblastoma pathogenesis. Oncotarget. 2016;7(18):24899–907.
22. Sanai N, Tramontin AD, Quinones-Hinojosa A, et al. Unique astrocyte ribbon in adult human brain contains neural stem cells but lacks chain migration. Nature. 2004;427(6976):740–4.
23. Lim DA, Alvarez-Buylla A. The adult ventricular–subventricular zone (V-SVZ) and olfactory bulb (OB) neurogenesis. Cold Spring Harb Perspect Biol. 2016;8:a018820.
24. Liu R, Liu Q, Zhao K, et al. MR imaging, MGMT promoter methylation features and prognostic analysis of subventricular zone contacting IDH wild-type glioblastoma. Curr Med Imaging. 2023;19(12):1378–86.