

---

This is the **published version** of the :

Páez Rodríguez, Adriana Milena; Ramos Terrades, Oriol (tut.). *Modelo de machine learning para la predicción de la progresión del glioblastoma multiforme utilizando datos radiómicos y clínicos*. (Universitat Autònoma de Barcelona), 2025

---

This version is available at <https://ddd.uab.cat/record/322962>

under the terms of the  license.

# MODELO DE MACHINE LEARNING PARA LA PREDICCIÓN DE LA PROGRESIÓN DEL GLIOBLASTOMA MULTIFORME UTILIZANDO DATOS RADIÓMICOS Y CLÍNICOS

Adriana Milena Páez Rodríguez

## RESUMEN

**Objetivo:** Desarrollar y validar un modelo de aprendizaje automático para predecir progresión en glioblastoma multiforme (GBM) mediante características radiómicas de resonancia magnética preoperatoria y variables clínicas. **Materiales y Métodos:** estudio retrospectivo que incluyó 11 pacientes con GBM confirmado histopatológicamente. Se extrajeron 159 características radiómicas de imágenes T1 con contraste siguiendo estándares IBSI. La progresión tumoral se definió mediante criterios RANO. El estudio se adhirió a las guías TRIPOD-AI. La selección de características identificó las variables más predictivas, y se evaluaron nueve clasificadores mediante validación cruzada anidada leave-one-out con bootstrap (1000 iteraciones). Las métricas incluyeron exactitud balanceada, área bajo la curva ROC (AUROC) y error de Brier. **Resultados:** Tres características de la matriz GLRLM fueron consistentemente seleccionadas. El clasificador Extra Trees demostró rendimiento superior (AUROC = 0,833, IC 95%: 0,712-0,927; exactitud balanceada = 0,817; error de Brier = 0,188), superando a otros clasificadores. El modelo mostró calibración adecuada. **Conclusión:** Un modelo compacto de aprendizaje automático basado en tres características texturales GLRLM predice con buen rendimiento la progresión del GBM. Este estudio demuestra que la estandarización metodológica permite modelado predictivo confiable incluso en cohortes limitadas, estableciendo bases para validación externa multicéntrica.

**Palabras clave:** Glioblastoma; Radiómica; Aprendizaje Automático; Resonancia Magnética; Progresión Tumoral; TRIPOD-AI.

## ABSTRACT

**Objective:** to develop and validate a machine learning model for predicting early progression in glioblastoma multiforme (GBM) using preoperative MRI radiomic features and clinical variables. **Materials and Methods:** this retrospective study included 11 patients with histopathologically confirmed GBM. From preoperative contrast-enhanced T1-weighted MRI, 159 radiomic features were extracted following Image Biomarker Standardization Initiative (IBSI) guidelines. Tumor progression was defined using RANO criteria. The study adhered to TRIPOD-AI reporting guidelines. Feature selection identified the most predictive variables, and nine machine learning classifiers were evaluated using nested leave-one-out cross-validation with bootstrap resampling (1000 iterations). Performance metrics included balanced accuracy, area under the receiver operating characteristic curve (AUROC), and Brier score. **Results:** Three Gray-Level Run-Length Matrix (GLRLM) features were consistently selected. The Extra Trees Classifier demonstrated superior performance (AUROC = 0.833, 95% CI: 0.712-0.927; balanced accuracy = 0.817; Brier score = 0.188), outperforming other classifiers. The model showed favorable calibration. **Conclusion:** a compact machine learning model based on three GLRLM textural features robustly predicts GBM progression. This TRIPOD-AI compliant study demonstrates that methodological standardization enables reliable predictive modeling even in limited cohorts, providing a foundation for multicenter validation.

**Keywords:** Glioblastoma; Radiomics; Machine Learning; Magnetic Resonance Imaging; Tumor Progression; TRIPOD-AI.



## 1 INTRODUCCIÓN

El glioblastoma multiforme (GBM), clasificado por la Organización Mundial de la Salud (OMS) como un glioma difuso de grado IV, constituye el tumor cerebral primario más agresivo en adultos [1]. Se caracteriza por un comportamiento altamente infiltrante, rápida proliferación celular y marcada heterogeneidad genética e histológica, factores que complican su manejo y confieren un pronóstico adverso [2, 3].

A pesar del tratamiento estándar multimodal, que combina la resección quirúrgica máxima con radioterapia y temozolomida concomitante y adyuvante según el protocolo

de Stupp [4], el pronóstico continúa siendo desfavorable. La supervivencia global media oscila entre 12 y 15 meses, y más del 90 % de los pacientes presenta recurrencia en los dos años posteriores al diagnóstico [4, 5]. Esta limitada respuesta terapéutica se explica, en gran medida, por la compleja heterogeneidad intratumoral del GBM, que da lugar a clones celulares con distinta sensibilidad al tratamiento y un microambiente tumoral dinámico y resistente [2, 3, 6].

En el seguimiento de estos pacientes, uno de los principales retos clínicos es la diferenciación entre progresión tumoral verdadera y pseudoprogresión, fenómeno inflamatorio transitorio inducido por la quimiorradioterapia que puede simular una recidiva tumoral en las imágenes postoperatorias [7-9]. Esta distinción es importante para evitar

- E-mail de contacto: [adrianap2511@yahoo.com](mailto:adrianap2511@yahoo.com)
- Trabajo tutorizado por: Oriol Ramos Terrades
- Curso: 2025

decisiones terapéuticas inapropiadas y se ha abordado mediante los criterios RANO, que definen la progresión como un aumento mayor o igual al 25 % en el área bidimensional o 40 % en el volumen tumoral medible, la aparición de nuevas lesiones o un deterioro clínico inequívoco atribuible al tumor, considerando el periodo crítico de 12 semanas tras la quimiorradioterapia, durante el cual se requiere confirmación clínica o radiológica de persistencia [10].

En los últimos años, la radiómica ha emergido como una herramienta no invasiva capaz de cuantificar la heterogeneidad tumoral a partir de imágenes médicas mediante la extracción automatizada o semi-automatizada de características de textura, forma e intensidad [11-13]. Estas métricas reflejan propiedades estructurales y fisiológicas del tumor asociadas con proliferación, necrosis y angiogénesis, procesos vinculados con la agresividad tumoral [14, 15]. En este contexto, se ha propuesto que la radiómica actúa como un puente entre la imagen médica y la medicina personalizada [6].

Diversos estudios han demostrado que la integración de características radiómicas con variables clínicas mediante modelos de aprendizaje automático mejora la predicción de supervivencia, progresión y respuesta terapéutica en gliomas de alto grado [16-18]. Sin embargo, persisten limitaciones relacionadas con la falta de estandarización en la extracción de características, validación interna insuficiente y heterogeneidad de los diseños, lo que ha obstaculizado su implementación clínica [19, 20].

Bajo este escenario, el presente estudio busca desarrollar y validar internamente un modelo predictivo de progresión en GBM, combinando características radiómicas derivadas de resonancias magnéticas T1 con contraste preoperatorias y variables clínicas. El objetivo es contribuir al avance hacia herramientas reproducibles, transparentes y clínicamente aplicables, alineadas con las iniciativas IBSI y TRIPOD-AI, que favorezcan una estratificación pronóstica precisa y una planificación terapéutica personalizada en pacientes con esta patología.

## 2 MATERIALES Y MÉTODOS

El estudio se condujo conforme a la Declaración de Helsinki y la normativa nacional vigente. Fue aprobado por el Comité de Ética en Investigación de la Universidad Tecnológica de Pereira (acta 59 - 260525). Todos los datos clínicos e imagenológicos se anonimizaron antes del análisis, eliminando identificadores directos y llaves de re-identificación. La información se almacenó en Google Cloud Platform con cifrado AES-256 en reposo y TLS en tránsito, acceso restringido y registro de auditoría. El estudio no implicó modificación alguna del manejo clínico de los pacientes.

### 2.1 Diseño

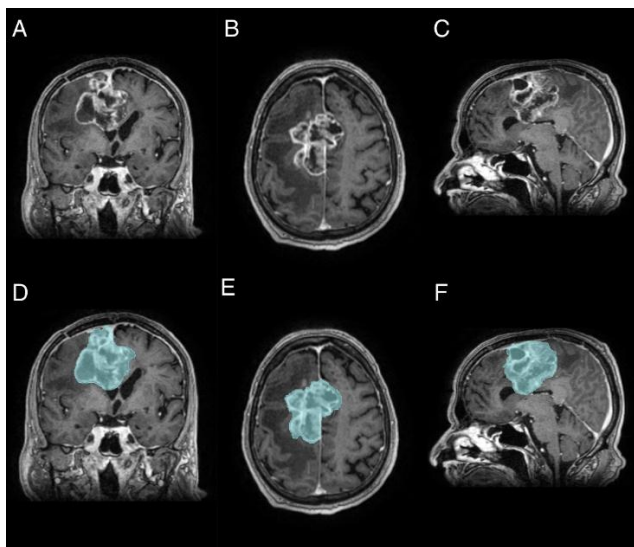
Se trata de un estudio observacional, analítico y retrospectivo predictivo, cuyo objetivo fue desarrollar un modelo de aprendizaje automático para estimar la progresión del GBM en un contexto asistencial real. Se incluyeron pacientes adultos con diagnóstico histopatológico de GBM entre 2019 y 2024 en el Hospital Universitario San Jorge de Pereira. Aunque algunos estudios iniciales se realizaron mediante tomografía, solo se analizaron los casos con resonancia magnética T1 con contraste técnicamente adecuada, garantizando consistencia radiológica.

El desenlace primario fue la progresión tumoral verdadera, definida según los criterios del grupo RANO [10] para gliomas con realce de contraste. Se consideró progresión cuando existió un aumento mayor o igual 25 % del área bidimensional o 40 % del volumen tumoral medible, la aparición de nuevas lesiones o un deterioro clínico atribuible al tumor, bajo dosis estables o crecientes de corticosteroides. Para asegurar trazabilidad y reducir sesgos derivados de subjetividad clínica, los criterios clínicos evaluables se condensaron en tres variables: cefalea, crisis convulsivas y déficit neurológico focal. Los episodios observados durante las primeras 12 semanas tras quimiorradioterapia se verificaron mediante seguimiento clínico-radiológico, y solo se clasificaron como progresión cuando la persistencia o empeoramiento fue documentado en una evaluación posterior. Fuera de ese intervalo, la progresión se estableció con base en criterios clínico-radiológicos sin confirmación adicional.

### 2.2 Radiómica

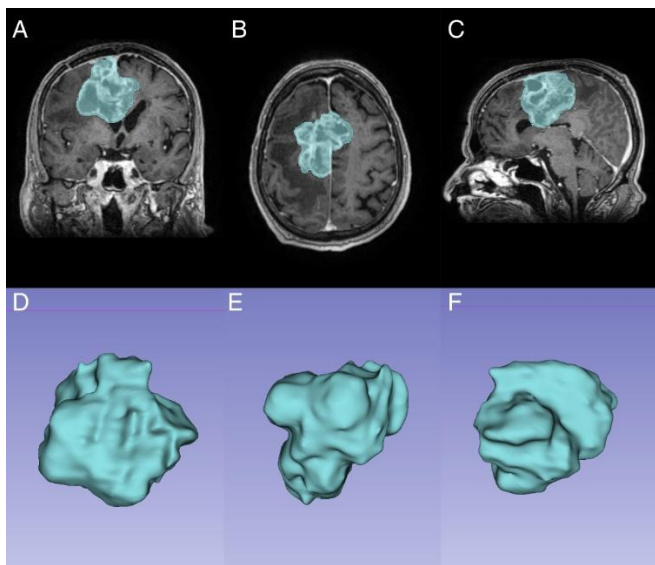
Las imágenes IRM T1 con contraste preoperatoria se descargaron en formato DICOM y se segmentaron en 3D Slicer v5.6 mediante herramientas semiautomáticas con revisión manual y consenso entre lectores entrenados en neuroimagen (Figura 1) (Figura 2). El preprocesamiento se alineó con la IBSI [21] para garantizar reproducibilidad. Se aplicó resampleo isotrópico 1×1×1 mm (interpolación B-spline), normalización z-score (normalizeScale = 100), ancho de bin = 31 y matrices GLCM simétricas. La extracción se realizó con PyRadiomics v3.1.0, empleando un único archivo de configuración (.yaml) para trazabilidad. Se obtuvieron características de primer orden (intensidad), forma (volumen, superficie, esfericidad, elongación, compactness) y textura (GLCM, GLRLM, GLSZM, NGTDM).

Figura 1 Segmentación volumétrica del glioblastoma en resonancia magnética preoperatoria



Corte coronal (A, D), axial (B, E) y sagital (C, F) en secuencia T1 con contraste. En la fila inferior se observa la máscara segmentada del tumor (azul) sobrepuesta sobre la imagen radiológica.

Figura 2 Reconstrucción tridimensional del volumen segmentado del glioblastoma.



Imágenes multiplanares (A–C) con la máscara tumoral sobrepuesta y reconstrucción tridimensional correspondiente (D–F).

## 2.2 Análisis estadístico y aprendizaje automático

El conjunto de variables radiómicas se integró con variables clínicas extraídas manual y sistemáticamente de las historias clínicas electrónicas. Posteriormente, se efectuó una fase de depuración y observación exploratoria para identificar valores atípicos y distribuciones asimétricas.

Las variables numéricas se estandarizaron con z-score, mientras que las categóricas se codificaron mediante label encoding. Se eliminaron identificadores, metacampos técnicos y variables posoperatorias sin interpretación clínica directa. Los valores faltantes se imputaron por la mediana de cada variable. La selección de características se basó en su variabilidad y en su contribución a la discriminación de clases según los métodos de ANOVA-KBest, LASSO y Elastic Net y se evaluaron múltiples modelos de predicción y selección de variables.

El modelamiento se desarrolló en Python 3.10 usando Kaggle Notebooks con librerías pandas, numpy, scikit-learn, matplotlib y seaborn. El proceso siguió las recomendaciones de la guía TRIPOD+AI [22]. Debido al tamaño reducido de la muestra, se empleó validación cruzada anidada (leave-one-out) para maximizar el uso de la información disponible y obtener una estimación no sesgada del rendimiento del modelo. La validación anidada permitió evaluar simultáneamente la selección de variables y la capacidad de generalización, manejando el optimismo artificial en la base de datos.

El criterio principal de comparación fue la exactitud balanceada, que pondera equitativamente ambas clases y resulta adecuada cuando la proporción de casos de progresión y no progresión no es simétrica. Adicionalmente se calculó el área bajo la curva ROC (AUROC) como métrica global de discriminación. La calibración se evaluó mediante pendiente e intercepto del modelo, complementada con curvas de calibración gráficas. Para cuantificar la incertidumbre y la estabilidad de las estimaciones se realizó bootstrap con 1000 iteraciones con reemplazo, permitiendo obtener intervalos de confianza del 95 % para todas las métricas. Este número se escogió por equilibrar precisión numérica y eficiencia computacional en cohortes pequeñas.

## 3 RESULTADOS

Se identificaron aproximadamente 30 pacientes con diagnóstico de GBM. Luego de la evaluación de criterios de elegibilidad y exclusión incluyendo calidad de imagen, datos completos en la historia clínica y depuración de inconsistencias, 11 pacientes cumplieron los criterios para hacer parte del análisis. La variable de desenlace fue la progresión tumoral definida acorde a los criterios RANO. La distribución del evento mostró 5 casos de progresión (45,5 %) y 6 (54,5 %) permanecieron libres de progresión durante el seguimiento radiológico y clínico. Las imágenes fueron procesadas de acuerdo con los estándares IBSI, garantizando spacing isotrópico de 1 mm y coherencia espacial entre las máscaras segmentadas.

El conjunto inicial incluyó 159 características radiómicas,

reducidas a 125 tras eliminar campos administrativos, metadatos y variables posoperatorias. La imputación de valores faltantes se realizó mediante la mediana, y las variables fueron estandarizadas por z-score dentro de cada iteración de validación cruzada, evitando fuga de información.

El análisis exploratorio de correlaciones reveló alta colinealidad entre variables morfométricas y texturales derivadas de una misma matriz, con coeficientes superiores a 0,9 en pares como VoxelVolume, MeshVolume y VoxelNum.1, o entre Energy y TotalEnergy. Este hallazgo sustentó la aplicación de un filtro de correlación ( $|r| \leq 0,85$ ) antes de la selección final (Tabla 1). Este patrón era esperable, dado que múltiples descriptores radiómicos cuantifican información redundante de un mismo fenómeno geométrico o de intensidad (Figura 3).

Tabla 1 Correlaciones más altas entre variables radiómicas preoperatorias

| Variable 1          | Variable 2   | Coefficiente r |
|---------------------|--------------|----------------|
| Energy              | TotalEnergy  | 1.000          |
| JointAverage        | SumAverage   | 1.000          |
| Minimum.2           | Minimum.3    | 1.000          |
| Maximum.2           | Maximum.3    | 1.000          |
| VoxelNum.1          | VoxelVolume  | 1.000          |
| Mean.2              | Mean.3       | 1.000          |
| MeshVolume          | VoxelNum.1   | 1.000          |
| LargeAreaEmphasis   | ZoneVariance | 1.000          |
| GrayLevelVariance.2 | Variance     | 0.999          |
| Id                  | Idm          | 0.999          |

En la etapa de selección de características, limitada a tres variables por recomendación TRIPOD-AI para estudios de bajo tamaño muestral, se evaluaron los métodos ANOVA K-Best, LASSO y Elastic Net en combinación con distintos clasificadores. Durante la validación cruzada leave-one-out, tres métricas texturales de la familia GLRLM fueron

identificadas de manera consistente en las once iteraciones: ShortRunHighGrayLevelEmphasis, ShortRunLowGrayLevelEmphasis y LowGrayLevelRunEmphasis. Estas variables presentaron importancia media normalizada de 0.634, 0.627 y 0.597 respectivamente, lo que indica estabilidad interfold y peso comparable en la discriminación del desenlace.

El modelo Extra Trees Classifier alcanzó el mejor rendimiento global, con exactitud balanceada (BA) = 0.817, área bajo la curva ROC (AUROC) = 0.833 y área bajo la curva precisión-recuperación (AUPRC) = 0.710. Estos valores fueron superiores o comparables a los obtenidos con la regresión logística (BA = 0.817; AUROC = 0.800) y el Random Forest (BA = 0.733; AUROC = 0.800). La Tabla 2 resume los resultados de los nueve clasificadores evaluados, y la Figura 4 presenta la comparación visual de sus métricas principales.

Este modelo también mostró la mejor calibración, con error de Brier = 0.188. La evaluación de curvas ROC, PRC y calibración (Figura 5) evidenció una adecuada correspondencia entre probabilidades predichas y observadas, con pendiente de calibración próxima a 1 e intercepto cercano a 0. Los intervalos de confianza obtenidos por bootstrap mostraron variabilidad amplia pero centrada en estimaciones estables, lo que respalda la consistencia del rendimiento dentro del tamaño muestral disponible.

El mapa de correlación entre variables radiómicas (Figura 3) confirmó interdependencias moderadas entre las tres características seleccionadas ( $r = 0.71-0.76$ ), lo que sugiere complementariedad sin redundancia completa. El análisis univariante de correlación punto-biserial y F-score (Figura 6; Tabla 3) mostró que las mismas variables GLRLM fueron las más asociadas con la progresión tumoral ( $|r| = 0.63-0.60$ ;  $p < 0.05$ ), seguidas por HighGrayLevelRunEmphasis y SumEntropy.

Por último, la comparación multivariada mediante gráfico de radar (Figura 7) evidenció un perfil de desempeño homogéneo del modelo Extra Trees al integrar simultáneamente exactitud balanceada, AUPRC y F1-macro. Este comportamiento fue consistente con los modelos de ensamblado, mientras que los clasificadores lineales penalizados y los de gradiente descendente mostraron menor estabilidad.

Figura 3 Mapa de correlación entre variables radiómicas

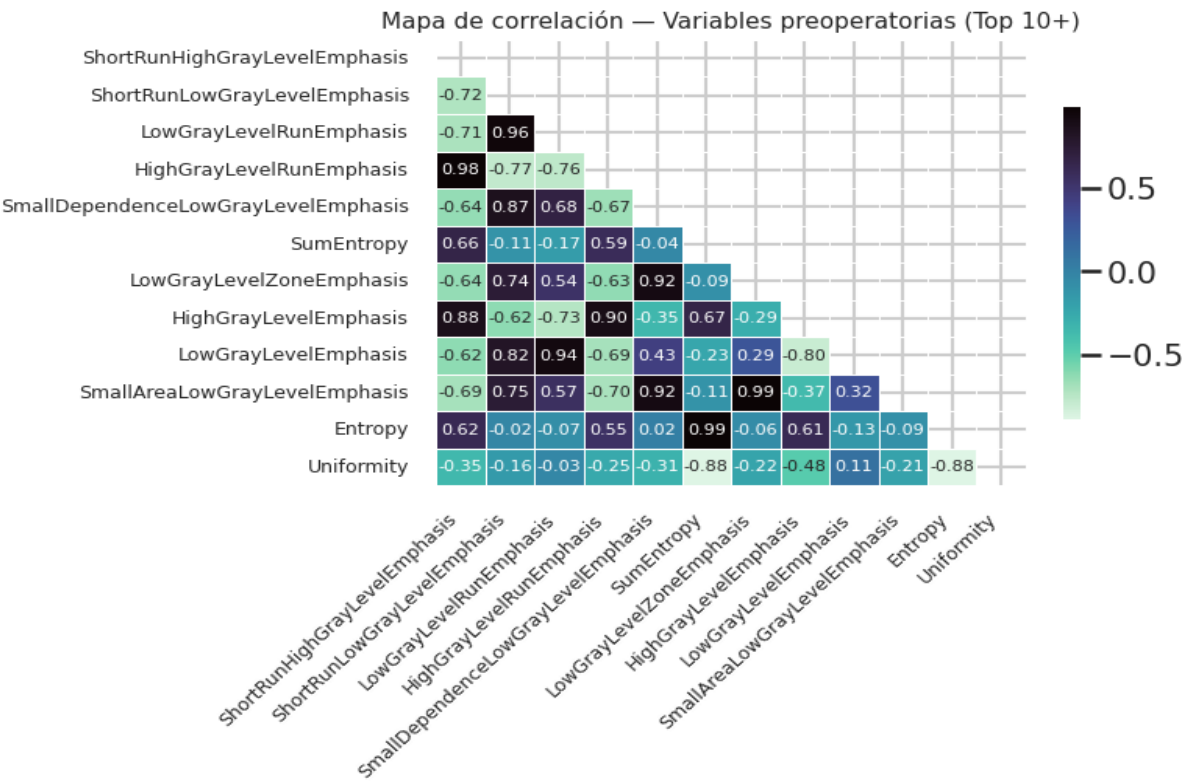


Tabla 2 Rendimiento de los modelos de clasificación

| Modelo                       | Balanced Accuracy | AUROC | AUPRC | F1-Macro | Brier Score | Calibration Error |
|------------------------------|-------------------|-------|-------|----------|-------------|-------------------|
| Extra Trees Classifier       | 0.817             | 0.833 | 0.710 | 0.817    | 0.188       | 0.361             |
| Random Forest                | 0.733             | 0.800 | 0.686 | 0.727    | 0.170       | 0.162             |
| Logistic Regression          | 0.817             | 0.800 | 0.686 | 0.817    | 0.212       | 0.375             |
| Linear Discriminant Analysis | 0.733             | 0.767 | 0.660 | 0.727    | 0.215       | 0.333             |
| Gradient Boosting            | 0.633             | 0.767 | 0.726 | 0.633    | 0.269       | 0.386             |
| Gaussian Naive Bayes         | 0.717             | 0.767 | 0.660 | 0.718    | 0.263       | 0.231             |
| SVM (Linear Kernel)          | 0.733             | 0.667 | 0.563 | 0.727    | 0.262       | 0.408             |
| Ridge Classifier             | 0.733             | 0.667 | 0.563 | 0.727    | 0.213       | 0.283             |
| Stochastic Gradient Descent  | 0.633             | 0.533 | 0.520 | 0.633    | 0.365       | 0.355             |

Tabla 3 Correlación y discriminación de las 10 variables más asociadas con el desenlace

| Variable                            | r punto-biserial | p-valor | F-score | p-valor |
|-------------------------------------|------------------|---------|---------|---------|
| ShortRunHighGrayLevelEmphasis       | 0.634            | 0.036   | 6.06    | 0.036   |
| ShortRunLowGrayLevelEmphasis        | 0.627            | 0.039   | 5.84    | 0.039   |
| LowGrayLevelRunEmphasis             | 0.597            | 0.052   | 4.98    | 0.052   |
| HighGrayLevelRunEmphasis            | 0.571            | 0.067   | 4.34    | 0.067   |
| SmallDependenceLowGrayLevelEmphasis | 0.539            | 0.087   | 3.69    | 0.087   |
| SumEntropy                          | 0.535            | 0.090   | 3.61    | 0.090   |
| LowGrayLevelZoneEmphasis            | 0.522            | 0.100   | 3.37    | 0.100   |
| HighGrayLevelEmphasis               | 0.516            | 0.104   | 3.27    | 0.104   |
| LowGrayLevelEmphasis                | 0.509            | 0.110   | 3.15    | 0.110   |
| SmallAreaLowGrayLevelEmphasis       | 0.491            | 0.125   | 2.86    | 0.125   |

Figura 4 Comparación de métricas de rendimiento entre clasificadores (Balanced Accuracy, AUROC, AUPRC).

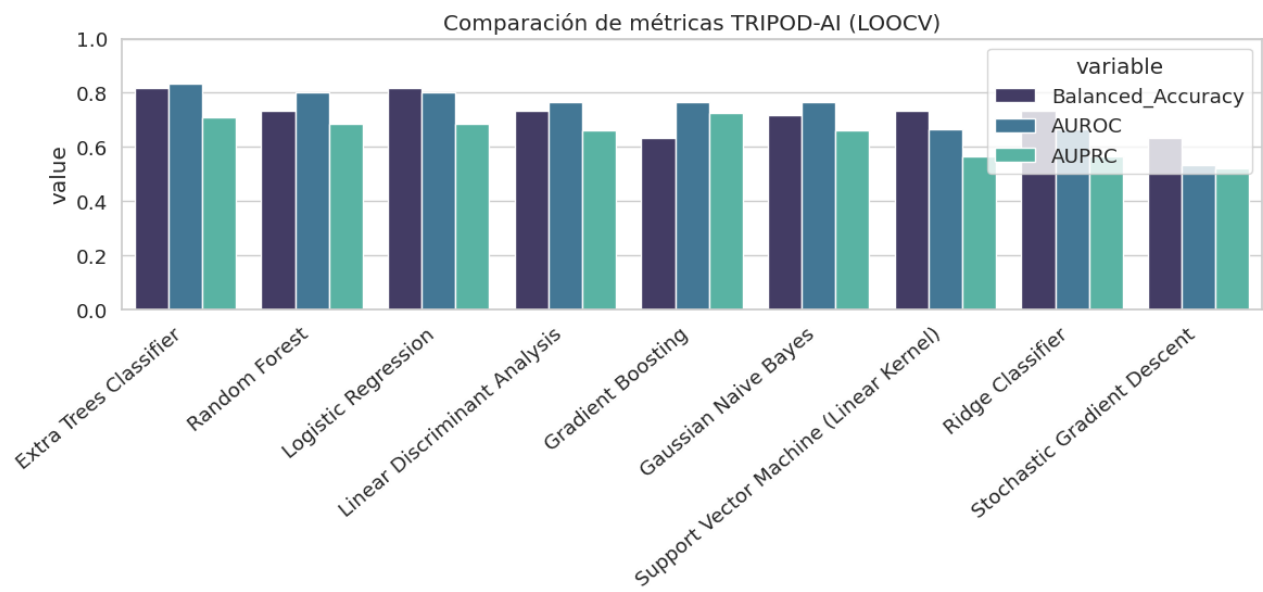


Figura 5 Curvas ROC, PRC y calibración del modelo Extra Trees Classifier.

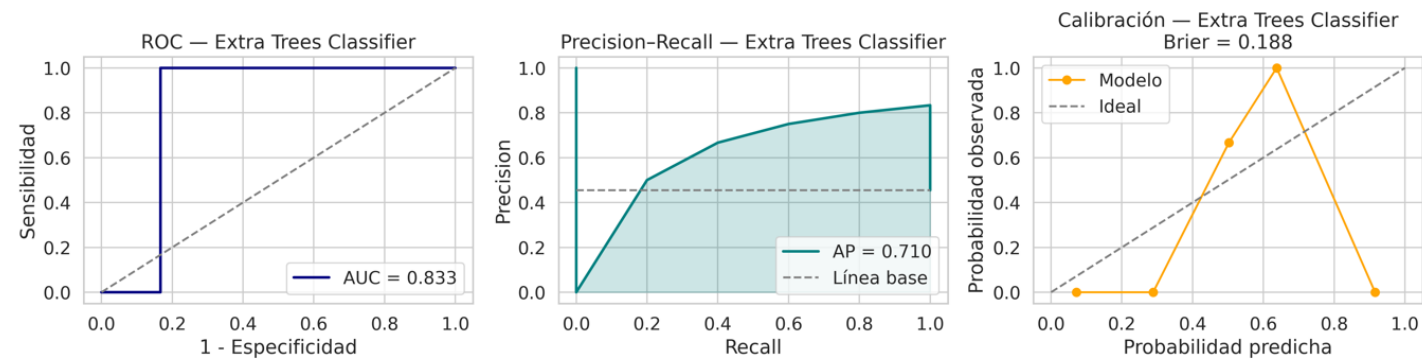


Figura 6 variables radiómicas correlacionadas con el desenlace ( $r$  biserial y F-score).

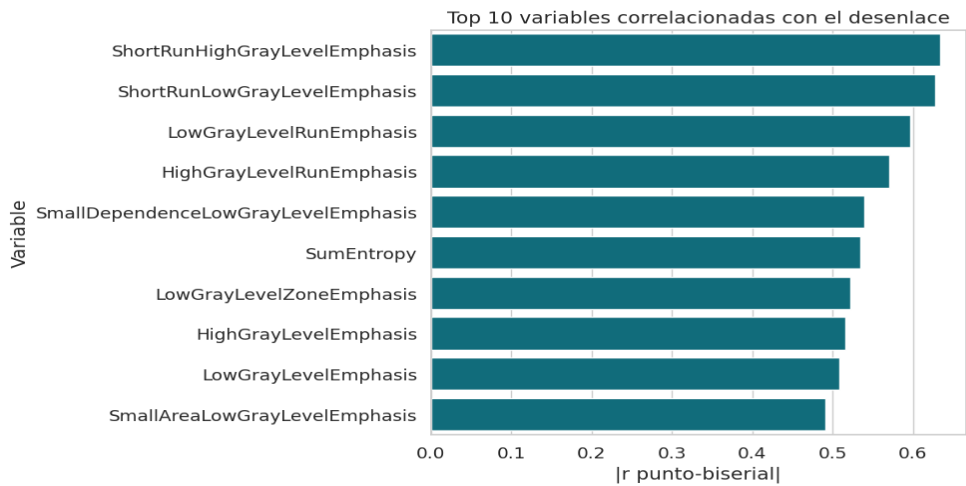
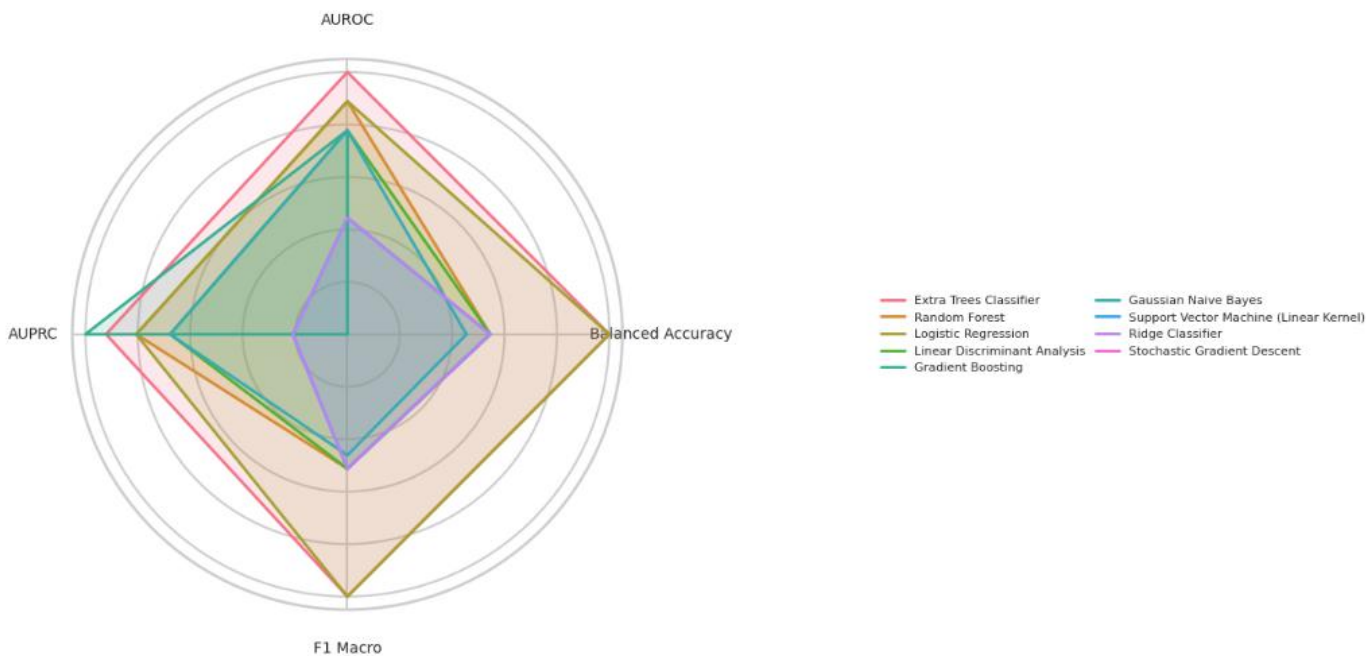


Figura 7 Radar comparativo de rendimiento multivariado entre modelos





## 4 DISCUSIÓN

El presente estudio demuestra que un subconjunto reducido de características radiómicas preoperatorias derivadas de resonancia magnética contrastada procesadas bajo los estándares IBSI y reportadas siguiendo la guía TRI-POD-AI, permite predecir de forma robusta la progresión temprana del glioblastoma multiforme GBM. Este hallazgo refuerza la utilidad de la radiómica como herramienta cuantitativa para la estratificación pronóstica individual en tumores cerebrales caracterizados por su alta heterogeneidad [1, 2, 13].

La reproducibilidad de la radiómica exige una estandarización rigurosa del procesamiento de imágenes y del reporte científico. En este estudio se aplicó *resampling* isotrópico de 1 mm<sup>3</sup>, extracción de características conforme a las últimas guías [21] y validación cruzada anidada *leave-one-out* con *bootstrap*, garantizando así transparencia y estabilidad estadística [23, 24]. El cumplimiento de los principios TRIPOD+AI permite la transparencia en el reporte de la metodología empleada para manejar el tamaño muestral limitado, permitiendo una calibración adecuada de las predicciones. Este enfoque metodológico se alinea con la recomendación de evaluar no solo la discriminación sino también la calibración de los modelos predictivos, un elemento frecuentemente ausente en estudios radiómicos oncológicos [25, 26].

El modelo Extra Trees Classifier alcanzó un AUROC de 0.833 y una exactitud balanceada de 0.817, superando en este estudio a clasificadores lineales y mostrando una mejor calibración. Estos resultados se sitúan en un rango favorable con respecto a lo reportado en la literatura, donde los modelos de predicción de progresión o supervivencia en GBM suelen presentar AUC entre 0.70 y 0.85 [16, 27, 28]. Los métodos basados en árboles de decisión, como *Extra Trees*, ofrecen ventajas en cohortes pequeñas debido a su capacidad para manejar interacciones no lineales y reducir la varianza sin incrementar excesivamente el sesgo [29]. Su rendimiento comparable al de arquitecturas más complejas ha sido documentado en revisiones recientes, que resaltan su interpretabilidad y estabilidad frente a la limitada disponibilidad de datos en neurooncología [13, 30]. El buen ajuste del modelo y la dispersión acotada de los intervalos de confianza obtenidos por *bootstrap* respaldan su rendimiento.

Las tres características radiómicas seleccionadas *ShortRunHighGrayLevelEmphasis*, *ShortRunLowGrayLevelEmphasis* y *LowGrayLevelRunEmphasis* pertenecen a la familia GLRLM y cuantifican la textura fina del tumor, reflejando su

heterogeneidad microestructural. Este patrón es coherente con estudios que vinculan la irregularidad espacial del realce con procesos biológicos clave como la proliferación celular y la neoangiogénesis [15, 31]. Investigaciones previas han demostrado que la complejidad textural predice la supervivencia y la respuesta quirúrgica con independencia del volumen tumoral [27], mientras que otros autores han evidenciado que la heterogeneidad de la señal en MRI se asocia a desenlaces adversos y a un menor control local [32]. Estos hallazgos sustentan que la textura aporta información pronóstica más discriminante que los descriptores volumétricos tradicionales, y explican por qué las variables GLRLM fueron seleccionadas de forma consistente en todos los *folds* de validación. Estudios recientes que emplean *deep learning* para radiogenómica confirman que se pueden predecir firmas moleculares específicas a partir de la imagen, como las mutaciones EGFRvIII o el estado de IDH [33, 34], lo que sugiere una convergencia funcional entre modelos.

La progresión tumoral fue definida operativamente mediante los criterios RANO [10], los cuales establecen umbrales volumétricos y bidimensionales para estandarizar la evaluación de respuesta en gliomas de alto grado. Aunque RANO aborda también la diferenciación entre progresión verdadera y pseudoprogresión, en este estudio la atención se centró exclusivamente en la predicción de la progresión verdadera, evitando así confusiones con cambios postratamiento y garantizando un desenlace clínico claro y reproducible. El uso de RANO permite una comparabilidad directa con la literatura neurooncológica más actualizada y sitúa la metodología dentro del marco de consenso para la evaluación tumoral [35].

Si bien los modelos de deep learning han ganado popularidad en radiología, su superioridad sobre los métodos de machine learning clásico no es concluyente en contextos de datos limitados. Revisiones sistemáticas y metaanálisis recientes muestran que la mayoría de las redes neuronales profundas para la predicción de progresión o supervivencia en GBM alcanzan desempeños comparables (AUC 0.80–0.85) a los de modelos de ensamblado bien calibrados [30, 36]. El modelo *Extra Trees* ofrece ventajas prácticas al ser más interpretable, de menor costo computacional y reproducible con cohortes pequeñas, lo que lo convierte en una alternativa factible para escenarios clínicos con disponibilidad restringida de datos etiquetados [30]. La integración futura de estas características radiómicas con datos genómicos o transcriptómicos podría potenciar su capacidad pronóstica y contribuir a la consolidación de modelos de radiogenómica [37, 38].

## 5 LIMITACIONES Y PERSPECTIVAS

El presente estudio debe interpretarse considerando sus limitaciones. La principal limitación del estudio es el tamaño muestral reducido. Sin embargo, este enfoque tiene precedentes en la literatura. Múltiples estudios recientes en revistas Q1/Q2 han explorado la radiómica en GBM utilizando cohortes de tamaño comparable para el desarrollo inicial de modelos [39 - 43].

Bajo este panorama, la viabilidad depende de la rigurosidad metodológica. Por ello, el presente estudio priorizó la transparencia y la reproducibilidad desde su diseño con la adhesión estricta a los estándares y el reporte conforme a las guías junto con el uso de validación cruzada anidada y bootstrap. Esto mitiga el sesgo de sobreoptimismo inherente al tamaño muestral y aporta validez a los hallazgos.

No obstante, el siguiente paso crítico es la validación externa en cohortes multicéntricas e independientes, una necesidad generalizada en el campo, como lo documentan revisiones sistemáticas [19]. Futuras validaciones deberán abordar la heterogeneidad técnica entre escáneres, idealmente empleando datasets públicos armonizados como BraTS [44]. Las líneas de investigación futuras también deben orientarse a la integración multimodal con datos moleculares [37, 38] y a la implementación de estrategias de IA explicable (XAI) [45, 46] para facilitar la adopción e implementación clínica.

## 6 CONCLUSION

Este estudio desarrolló y validó internamente un modelo predictivo de progresión en glioblastoma mediante la integración de características radiómicas texturales preoperatorias y aprendizaje automático. El modelo Extra Trees, sustentado en un subconjunto de tres características GLRLM, demostró un rendimiento predictivo competitivo (AUROC 0.833) con una calibración adecuada.

Los hallazgos evidencian que la heterogeneidad microestructural tumoral, cuantificable de forma no invasiva mediante radiómica, alberga un valor pronóstico sólido para la progresión de la enfermedad. En segundo lugar, se constata que la aplicación rigurosa de marcos de estandarización (IBSI) y reporte (TRIPOD-AI), junto con estrategias de validación, permite obtener estimaciones fiables incluso en el contexto de cohortes de tamaño limitado.

Estos resultados sientan las bases para futuras investigaciones orientadas a la validación externa multicéntrica. Perspectivas adicionales incluyen la exploración de la integración de estas características con datos histopatológicos y moleculares, lo que podría potenciar significativamente

la capacidad predictiva del modelo dentro de un marco de medicina de precisión en neurooncología.

## REFERENCIAS

1. Louis DN, Perry A, Wesseling P, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol.* 2021;142(2):163-230. DOI: 10.1007/s00401-021-02303-y
2. Weller M, Wick W, Aldape K, et al. Glioblastoma. *Nat Rev Dis Primers.* 2015; 1:15017. DOI: 10.1038/nrdp.2015.17
3. Alexander BM, Cloughesy TF. Adult glioblastoma. *J Clin Oncol.* 2017;35(21):2402-9. DOI: 10.1200/JCO.2017.73.0119
4. Stupp R, Mason WP, Van Den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med.* 2005;352(10):987-96. DOI: 10.1056/NEJMoa043330
5. Ostrom QT, Price M, Neff C, et al. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2015-2019. *Neuro Oncol.* 2022;24(Suppl 5): v1-v95. DOI: 10.1093/neuonc/noac202
6. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol.* 2017;14(12):749-62. DOI: 10.1038/nrclinonc.2017.141
7. Taal W, Brandsma D, de Bruin HG, Bromberg JE, Sminia P, van den Bent MJ. Incidence of early pseudoprogression in malignant glioma treated with temozolomide. *Cancer.* 2008;113(2):405-10. DOI: 10.1002/cncr.23562
8. Huang RY, Young RJ. Differentiating true progression from pseudoprogression in glioblastoma. *Neuro-Oncology Advances.* 2021;3(1): vdab083. DOI: 10.1093/no-ajnl/vdab083
9. Thomas AA, et al. Machine learning for differentiation of pseudoprogression and true progression in glioblastoma. *Cancers.* 2020;12(1):123. DOI: 10.3390/cancers12010123
10. Wen PY, Macdonald DR, Reardon DA, et al. Updated response assessment criteria for high-grade gliomas: RANO group. *J Clin Oncol.* 2010;28(11):1963-72. DOI: 10.1200/JCO.2009.26.3541
11. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. *Magn Reson Imaging.* 2012;30(9):1234-48. DOI: 10.1016/j.mri.2012.06.010
12. Chaddad A, Kucharczyk MJ, Daniel P, et al. Radiomics in glioblastoma: current status and challenges. *Front Oncol.* 2019; 9:374. DOI: 10.3389/fonc.2019.00374
13. Binczyk F, Prazuch W, Bozek P, Polanska J. Radiomics and artificial intelligence in glioma: update on current applications and future directions. *Cancers.* 2021;13(4):719. DOI: 10.3390/cancers13040719
14. Perez-Beteta J, Molina-Garcia D, Ortiz-Alhambra JA, et al. Tumor surface regularity at MRI predicts survival and surgical response in GBM. *Radiology.* 2018;288(1):218-25. DOI: 10.1148/radiol.2018171521
15. Choi Y, Jang J, Kim B, Ahn KJ. Impact of multiparametric MR radiomics for predicting GBM prognosis: meta-analysis. *Eur J Radiol.* 2023; 168:111130. DOI: 10.1016/j.ejrad.2023.111130
16. Kickingeder P, Burth S, Wick A, et al. Radiomic Profiling of Glioblastoma: Identifying an Imaging Predictor of

- Patient Survival with Improved Performance over Established Clinical and Radiologic Risk Models. *Radiology*. 2016;280(3):880–889. DOI: 10.1148/radiol.2016160845
17. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine Learning methods for Quantitative Radiomic Biomarkers. *Sci Rep*. 2015; 5:13087. DOI: 10.1038/srep13087
18. Zhang L, Chen B, Zhang L, et al. Machine learning-based integration of radiomics and clinical data for prognosis of glioblastoma. *Neuro Oncol*. 2021;23(7):1160–1169. DOI: 10.1093/neuonc/noab013
19. Park JE, Kim HS, Kim D, et al. A systematic review reporting quality of radiomics research in neuro-oncology: toward clinical utility and quality improvement using high-dimensional imaging features. *BMC Cancer*. 2020;20(1):29. DOI: 10.1186/s12885-020-6525-0
20. Kotrotsou A, Zinn PO, Colen RR. Radiomics in brain tumors: an emerging technique for characterization of tumor environment. *Magn Reson Imaging Clin N Am*. 2016;24(4):719–729. DOI: 10.1016/j.mric.2016.06.006
21. Zwanenburg A, Vallières M, Abdalah MA, et al. The Image Biomarker Standardization Initiative (IBSI). *Radiology*. 2020;295(2):328–38. DOI: 10.1148/radiol.2020191145
22. Collins GS, Dhiman P, Riley RD, et al. TRIPOD+AI: updated guidance for AI prediction models. *BMJ*. 2024;385: e078378. DOI: 10.1136/bmj-2023-078378
23. Sanduleanu S, Woodruff HC, de Jong EEC, et al. Tracking tumor biology with radiomics: A systematic review utilizing a radiomics quality score. *Radiother Oncol*. 2018;127(3):349–360. DOI: 10.1016/j.radonc.2018.03.033
24. Traverso A, Wee L, Dekker A, Gillies R. Repeatability and Reproducibility of Radiomic Features: A systematic Review. *Int J Radiat Oncol Biol Phys*. 2018;102(4):1143–1158. DOI: 10.1016/j.ijrobp.2018.05.053
25. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35(29):1925–1931. DOI: 10.1093/eurheartj/ehu207
26. Van Calster B, McLernon DJ, van Smeden M, et al. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17(1):230. DOI: 10.1186/s12916-019-1466-7
27. Bae S, Choi YS, Ahn SS, et al. Radiomic MRI phenotyping of glioblastoma: improving survival prediction. *Radiology*. 2018;289(3):797–806. DOI: 10.1148/radiol.2018180200
28. Benedict SH, Lyo P, Kim H, et al. Artificial Intelligence in Neuro-Oncology: Current Status and Future Directions. *Radiology*. 2021;300(2):266–280. DOI: 10.1148/radiol.2021203755
29. Biau G, Scornet E. A random forest guided tour. *Test*. 2016; 25:197–227. DOI: 10.1007/s11749-016-0481-7
30. Vamvakas A, Williams SC, Theodorou K, Kappas C, Tsougos I. Imaging Biomarker Analysis of Advanced Machine Learning and Deep Learning in Glioma: A Systematic Review. *Cancers*. 2022;14(20):5118. DOI: 10.3390/cancers14205118
31. Lee J, Wang N, Turk S, et al. Discriminating survival outcomes in patients with glioblastoma using a radiomic approach. *Sci Rep*. 2017;7(1):17774. DOI: 10.1038/s41598-017-17731-1
32. Lao J, Chen Y, Li ZC, et al. A Deep Learning-Based Radiomics Model for Prediction of Survival in Glioblastoma Multiforme. *Sci Rep*. 2017;7(1):10353. DOI: 10.1038/s41598-017-10649-8
33. Akbari H, Bakas S, Pisapia JM, et al. In vivo evaluation of EGFRvIII mutation in primary glioblastoma patients via deep learning-based analysis of MRI. *Neuro Oncol*. 2020;22(6):855–864. DOI: 10.1093/neuonc/noaa001
34. Chang K, Bai HX, Zhou H, et al. Residual convolutional neural network for the determination of IDH status in low- and high-grade gliomas from MR imaging. *Clin Cancer Res*. 2018;24(5):1073–1081. DOI: 10.1158/1078-0432.CCR-17-2236
35. Senders J, Staples P, Karhade AV, et al. Machine Learning and Neurosurgical Outcome Prediction: A Systematic Review. *World Neurosurg*. 2018; 109:476–486. DOI: 10.1016/j.wneu.2017.09.149
36. Liu Z, et al. Comparison of deep learning and machine learning for brain tumor segmentation. *Medical Physics*. 2021;48(9):4870–4881. DOI: 10.1002/mp.15143
37. Rathore S, Akbari H, Rozycki M, et al. Radiomic signature of infiltration in peritumoral edema predicts subsequent recurrence in glioblastoma. *Neurosurgery*. 2018;82(6):766–775. DOI: 10.1093/neuros/nyx316
38. Fathi Kazerooni A, Bakas S, Saligheh Rad H, Davatzikos C. Imaging signatures of glioblastoma molecular characteristics: A radiogenomics review. *J Magn Reson Imaging*. 2020;52(1):54–69. DOI: 10.1002/jmri.26907
39. Pessina F, Ghezzi F, Ceccon G, et al. A Pilot Study on Radiomic Analysis of Peritumoral Edema in Glioblastoma: Prognostic Value of Textural Features in Relation to Patient Survival. *World Neurosurg*. 2020;141: e55–e63. DOI: 10.1016/j.wneu.2020.06.012
40. Elshafeey N, El-Ghanam A, Helal S, et al. Multiregional quantitative texture analysis of T1-weighted MRI: a potential biomarker for IDH1 mutation status in glioblastoma, a pilot study. *J Neurooncol*. 2019;142(2):331–338. DOI: 10.1007/s11060-019-03125-9
41. Navarria P, Pessina F, Clerici E, et al. Radiomic Analysis in Glioblastoma: A Pilot Study for Correlation of Magnetic Resonance Imaging Features with Molecular Markers. *Cancers (Basel)*. 2019;11(10):1485. DOI: 10.3390/cancers11101485
42. Ghoneem A, Akbari H, Davatzikos C, et al. Investigating the Prognostic Value of Pre-Operative MRI-Based Radiomic Features in Glioblastoma: A Pilot Study. *Tomography*. 2022;8(4):1753–1768. DOI: 10.3390/tomography8040156
43. Larroza A, Moratal D, Arribas L, et al. A radiomics model based on Gd-EOB-DTPA-enhanced MRI for the prediction of early recurrence in glioblastoma. *Sci Rep*. 2021;11(1):1272. DOI: 10.1038/s41598-021-81516-w
44. Menze B, Jakab A, Bauer S, et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Trans Med Imaging*. 2015;34(10):1993–2024. DOI: 10.1109/TMI.2014.2377694
45. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. *Wiley Interdiscip Rev Data Min Knowl Discov*. 2019;9(4): e1312. DOI: 10.1002/widm.1312
46. Reyes M, Meier R, Pereira S, et al. On the interpretability of artificial intelligence in radiology: challenges and opportunities. *Radiol Artif Intell*. 2020;2(3): e190043. DOI: 10.1148/ryai.2020190043