

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

Herrera Mora, Ivan Mauricio; López Vicario, José (tut.). *Modelo de aprendizaje automático para la predicción de grado de resección quirúrgica de los meningiomas craneales utilizando datos clínicos y radiómicos*. (Universitat Autònoma de Barcelona), 2025

This version is available at <https://ddd.uab.cat/record/322556>

under the terms of the  license.

MODELO DE APRENDIZAJE AUTOMÁTICO PARA LA PREDICCIÓN DE GRADO DE RESECCIÓN QUIRÚRGICA DE LOS MENINGIOMAS CRANEALES UTILIZANDO DATOS CLÍNICOS Y RADIÓMICOS

Ivan Mauricio Herrera Mora

RESUMEN

Introducción: El meningioma es el tumor intracraneal primario más frecuente en adultos. La resecabilidad es un determinante del control tumoral en meningioma. La radiómica aplicada a IRM preoperatoria y la clínica puede aportar predicciones cuantitativas para planificar la cirugía. **Objetivos:** Desarrollar y evaluar un modelo multivariable que estime, en el preoperatorio del paciente con meningioma intracraneal, la probabilidad de alcanzar resección Simpson I frente a Simpson II. **Métodos:** Estudio retrospectivo (2018–2024). Se incluyeron adultos con meningioma y IRM T1 con contraste técnicamente adecuada. El desenlace fue binario (Simpson I vs II). La segmentación se realizó en 3D Slicer y la extracción radiómica con PyRadiomics bajo configuración estandarizada (IBSI). Los predictores se restringieron a variables numéricas preoperatorias (clínicas y radiómicas). Se comparó un banco selector×clasificador (ANOVA-KBest, LASSO, Elastic Net × LDA, Naive Bayes gaussiano, Regresión Logística; exploratorio: SVM lineal/RBF y Random Forest). **Resultados.** De ~70 casos cribados, 13 cumplieron elegibilidad; se analizaron 12 (Simpson I=4; Simpson II=8) con 126 predictores. La mejor configuración fue LASSO (C=0.2; ≤3 variables) + LDA: exactitud 0,833; exactitud balanceada 0,812; F1-macro 0,812; AUROC 0,844; Strength fue el predictor más estable (12/12 iteraciones). **Conclusiones:** El modelo preoperatorio basado en radiómica y aprendizaje automático mostró discriminación y calibración favorables para predecir resecabilidad (Simpson I vs II). Se requieren validación externa multicéntrica y evaluación de impacto clínico.

Palabras clave: meningioma; radiómica; aprendizaje automático; grado de Simpson; validación cruzada; calibración.

ABSTRACT

Background: Meningiomas are the most frequent primary central nervous system neoplasms. The extent of resection drives tumor control in meningioma. Preoperative MRI radiomics may provide quantitative predictions to support surgical planning. **Objective:** To develop and internally evaluate a multivariable model estimating the probability of achieving Simpson I versus Simpson II resection. **Methods:** Retrospective study (2018–2024). Adults with meningioma and adequate preoperative contrast-enhanced T1-weighted MRI were included. Outcome: binary (Simpson I vs II). 3D segmentation in 3D Slicer; radiomics extraction with PyRadiomics under IBSI-compliant settings. Predictors were numeric preoperative clinical and radiomic variables. An internal selector×classifier benchmark compared ANOVA-KBest, LASSO, Elastic Net × LDA, Gaussian Naive Bayes, Logistic Regression; exploratory SVM (linear/RBF) and Random Forest. **Results:** Of ~70 screened cases, 13 were eligible; 12 were modeled (Simpson I=4; Simpson II=8) with 126 predictors. Best configuration: LASSO (C=0.2; ≤3 features) + LDA: accuracy 0.833; balanced accuracy 0.812; macro-F1 0.812; AUROC 0.844; AUPRC 0.867. Strength was most stable (12/12). **Conclusions:** A preoperative radiomics-based model showed favorable discrimination and calibration for predicting resectability (Simpson I vs II). Multicenter external validation and clinical impact studies are warranted.

Keywords: meningioma; radiomics; machine learning; Simpson grade; cross-validation; calibration.

1 INTRODUCCIÓN

Los meningiomas constituyen el tumor intracraneal primario más frecuente en adultos y su incidencia anual se ha estimado en ~5,3 por 100.000 habitantes, con incremen-

to sostenido con la edad (1). La clasificación de la OMS 2021 reconoce tres grados histológicos con implicaciones de pronóstico; los grados II–III exhiben mayor agresividad biológica y tasas superiores de recurrencia en comparación con el grado I (2). El tratamiento se sustenta en la resección quirúrgica, con radioterapia adyuvante en escenarios seleccionados de alto riesgo (3). La extensión de

- *E-mail de contacto:* cronos7411@yahoo.com
- *Trabajo tutorizado por:* Jose Lopez Vicario
- *Curso:* 2025

resección, tradicionalmente cuantificada por la escala de Simpson, se asocia de forma consistente con control tumoral y riesgo de recaída a largo plazo (4); además, factores anatómicos y morfológicos, incluida la localización y el volumen condicionan la reseccabilidad y la planificación preoperatoria (5,6). Las limitaciones de la evaluación visual convencional de la imagen por resonancia magnética (IRM) han motivado el uso de radiómica y aprendizaje automático para extraer fenotipos cuantitativos de imagen (intensidad, textura, forma) no evidentes a simple vista, con potencial para mejorar la estratificación de riesgo (7). En meningiomas, se han desarrollado modelos que estiman el riesgo de recurrencia a partir de descriptores radiómicos (8) y enfoques que integran radiómica con variables clínicas para predecir el grado tumoral (9). Este marco se alinea con avances en biología tumoral como la clasificación basada en metilación del material genético que refinan la heterogeneidad del pronóstico y pueden servir de referencia para el diseño de modelos multimodales (10). En paralelo, en otras neoplasias del sistema nervioso y de cabeza y cuello, los perfiles radiómicos han superado a modelos puramente clínicos al predecir desenlaces, aportando validez externa conceptual a estas aproximaciones (11-13). A pesar de los avances descritos, los estudios publicados sobre radiómica en meningioma presentan limitaciones metodológicas relevantes. El metaanálisis más reciente (14) reportó una calidad promedio baja (Radiomics Quality Score 19%) y alta heterogeneidad entre protocolos de adquisición y segmentación, con tamaños muestrales pequeños. En términos de desempeño, los modelos existentes para clasificación de grado tumoral alcanzan valores promedio de área bajo la curva (AUC) entre 0,78 y 0,91 dependiendo de las secuencias empleadas y del algoritmo de aprendizaje (15,16).

No obstante, la mayoría de estos trabajos se centra en la predicción histológica o de recurrencia, sin abordar de manera directa la reseccabilidad quirúrgica, a pesar de su impacto en la planificación operatoria y el pronóstico funcional. En este contexto, el presente trabajo propone un modelo preoperatorio basado en radiómica y aprendizaje automático orientado a predecir la probabilidad de alcanzar resecciones Simpson I frente a II, aportando un enfoque novedoso y clínicamente aplicable para apoyar la toma de decisiones quirúrgicas. En conjunto, la evidencia respalda que enfoques de radiómica y aprendizaje automático pueden complementar la valoración convencional en neurooncología y aportar señales cuantitativas útiles para estratificación y toma de decisiones preoperatorias (7-9,11,12). A partir de estas consideraciones, el objetivo de este trabajo es desarrollar y validar internamente un modelo preoperatorio, basado en datos radiómicos de segmentaciones semiautomatizadas de IRM con contraste e integración variables clínicas, que estime la probabilidad de alcanzar resecciones Simpson I frente a II con la finalidad de apoyar la planificación quirúrgica individualizada en meningiomas.

2 MATERIALES Y MÉTODOS

2.1 ASPECTOS ÉTICOS Y GOBERNANZA DE DATOS

Aspectos éticos y gobernanza de datos. El estudio se condujo de acuerdo con la Declaración de Helsinki y fue aprobado por el Comité de Ética de la Universidad Tecnológica de Pereira (aval 67-040725). Los datos clínicos e imagenológicos se anonimizaron antes del análisis, sin conservar identificadores ni llaves de re-identificación. La información se almacenó en Google Cloud Platform con cifrado en reposo (AES-256) y en tránsito (TLS), acceso restringido y registros de auditoría. El estudio no implicó intervenciones adicionales ni modificación de la atención clínica.

2.2. DISEÑO

Se condujo un estudio observacional, retrospectivo, que incluyó adultos con diagnóstico histopatológico de meningioma intracraneal intervenidos entre 2018 y 2024 en el Hospital Universitario San Jorge de Pereira. El carácter retrospectivo del diseño permitió aprovechar la base de datos institucional consolidada, garantizando uniformidad en los criterios diagnósticos, quirúrgicos y de imagen. Este enfoque resulta congruente con las recomendaciones de TRIPOD+AI para etapas iniciales de desarrollo y validación de modelos predictivos, donde la trazabilidad y la transparencia del flujo de información priman sobre el volumen muestral.

Los criterios de inclusión (IRM preoperatoria técnicamente adecuada y disponibilidad de variables clínicas esenciales) aseguraron la integridad entre el dominio radiológico y el contexto clínico, preservando la correspondencia entre las variables de entrada y el desenlace quirúrgico real. Se excluyeron casos con imágenes de baja calidad o inconsistencias histopatológicas para preservar la estabilidad analítica de las características radiómicas y evitar desviaciones en la distribución de los predictores.

El modelo se estructuró como una tarea binaria que distingue entre resecciones Simpson I y Simpson II, dado que estas dos categorías concentraron la mayoría de los casos disponibles y representan el punto clínico más relevante en términos de extensión quirúrgica y riesgo de recurrencia. Las demás clases (Simpson III-V) presentaron frecuencias marginales que habrían generado un desequilibrio entre categorías y posible sobreajuste comprometiendo la estabilidad del aprendizaje.

2.3. RADIÓMICA

La segmentación volumétrica del tumor se realizó sobre la IRM preoperatoria en secuencia T1 con contraste, utilizando 3D Slicer con herramientas semiautomáticas y revisión

en consenso por observadores entrenados (Figura 1) (Figura 2) con el fin de reducir la variabilidad interobservador conservando precisión anatómica.

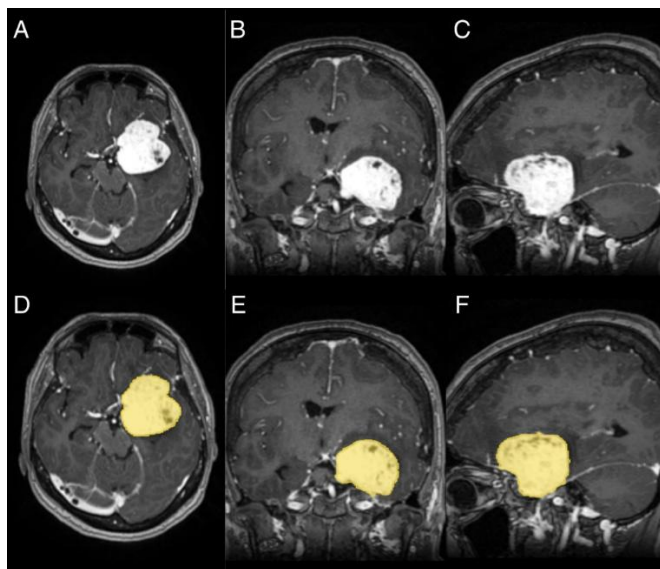


Fig. 1 Segmentación semiautomática de meningioma en IRM T1 con contraste. Ejemplo representativo del proceso de segmentación semiautomática realizado en 3D Slicer. (A–C) Cortes axial, coronal y sagital de una IRM T1 con contraste, donde se observa un meningioma de base craneal con realce homogéneo. (D–F) Máscara tumoral correspondiente (en amarillo) generada mediante herramientas semiautomáticas y revisada en consenso por observadores entrenados.

La extracción de características radiómicas se efectuó con PyRadiomics bajo un único archivo de configuración documentado, garantizando la consistencia paramétrica. El preprocesamiento incluyó resampleo isotrópico (111 mm), normalización z-score y bin width de 31, siguiendo el estándar de la Image Biomarker Standardization Initiative (IBSI), que establece criterios de homogeneización espacial e intensidad estadística para minimizar la dependencia de hardware y parámetros de adquisición (17). De este modo, las características derivadas reflejan propiedades intrínsecas del tejido tumoral y no artefactos de la adquisición. Se extrajeron variables de primer orden, forma y textura, integrando dominios complementarios que describen tanto la distribución de intensidades como la geometría y la heterogeneidad intratumoral, atributos frecuentemente asociados a invasividad, vascularización y consistencia mecánica. La estandarización de todo el flujo de trabajo bajo las directrices IBSI permite la reproducibilidad cuantitativa en imagen, permitiendo su eventual extrapolación a entornos multicéntricos.

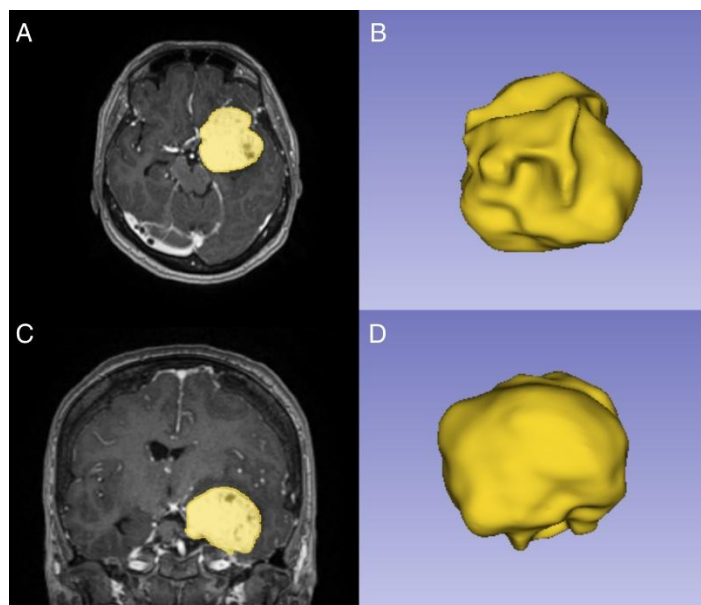


Fig. 2 Reconstrucción tridimensional del volumen tumoral segmentado. (A, C) Cortes axial y coronal de IRM T1 con contraste mostrando el meningioma y su máscara segmentada (en amarillo). (B, D) Reconstrucción tridimensional del volumen tumoral obtenida a partir de la segmentación.

3. ANÁLISIS ESTADÍSTICO Y APRENDIZAJE AUTOMÁTICO

El análisis estadístico y de aprendizaje automático se implementó en Python 3.10 utilizando librerías abiertas (pandas, scikit-learn, matplotlib y seaborn) dentro de un entorno reproducible (Kaggle Notebooks), lo que permitió garantizar control de versiones, trazabilidad y replicabilidad. Dado el tamaño muestral y la elevada dimensionalidad del conjunto de variables, el flujo metodológico se diseñó con el objetivo de preservar la parsimonia estadística y controlar el sobreajuste, maximizando la estabilidad del aprendizaje. En este contexto, la relación de pacientes y de predictores significó un riesgo de colinealidad y varianza elevada en los coeficientes estimados, por lo que se incorporaron etapas de preprocesamiento, selección de variables y validación cruzada anidada. Durante el preprocesamiento, se restringieron los predictores a variables numéricas estrictamente preoperatorias y se excluyeron sistemáticamente identificadores, metacampos técnicos y descriptores sin interpretación clínica directa. Para garantizar homogeneidad estadística, se aplicó imputación por mediana y estandarización mediante transformación z-score, procedimientos ejecutados de manera independiente en cada iteración del esquema de validación, con el fin de evitar fugas de información entre entrenamiento y prueba.

La selección de variables se abordó mediante tres estrategias complementarias: ANOVA-KBest, LASSO (penalización L1) y Elastic Net (L1+L2). La reducción final a un subconjunto de 3–5 variables fue con el objetivo de conservar una proporción de sujetos/predicotes adecuada y

reducir la varianza de estimación, preservando interpretabilidad. Posteriormente, se exploraron distintos clasificadores supervisados incluyendo Análisis Discriminante Lineal (LDA), Naive Bayes Gaussiano, Regresión Logística, Máquinas de Vectores de Soporte y Random Forest con el fin de comparar enfoques lineales, probabilísticos y de ensamblado bajo condiciones controladas. Esta elección metodológica permitió evaluar la interacción entre técnicas de selección y clasificadores. La comparación se realizó mediante un esquema anidado de validación cruzada leave-one-out (nested LOOCV), estrategia reconocida por su respuesta en muestras limitadas. En cada iteración del esquema leave-one-out, se emplearon 11 casos para entrenamiento y 1 para validación externa. Este enfoque evalúa simultáneamente la selección de hiperparámetros en un bucle interno y la generalización del modelo en un bucle externo, asegurando independencia entre entrenamiento y validación. esta metodología ofrece una estimación menos sesgada del error y reduce el riesgo de sobreajuste, motivo por el cual se privilegió sobre esquemas de partición aleatoria.

El criterio de selección del modelo óptimo fue la exactitud balanceada, esto debido a la distribución de clases pues pondera equitativamente la sensibilidad de ambas categorías y evita sesgo hacia la clase dominante. Otras métricas evaluadas fueron AUROC y la AUPRC, las cuales evaluaron su capacidad de discriminación entre clases; la F1-macro valoró también el equilibrio global y el Brier score permitió medir la calidad de la calibración. Para estimar la variabilidad de cada métrica y expresar el grado de confianza de los resultados, se aplicó un procedimiento de bootstrap con reemplazo sobre las predicciones obtenidas fuera de las muestras de entrenamiento (out-of-fold), generando intervalos de confianza al 95%. El conjunto de parámetros empleados para cada componente del flujo metodológico se resume en la Tabla 1.

Tabla 1. Configuración metodológica y parámetros principales del flujo analítico

| Etapas | Método / Técnica | Parámetros principales |
|--------------------------|---|--|
| Preprocesamiento | Imputación de valores faltantes y estandarización | Imputación por mediana; de estandarización tipo <i>z-score</i> aplicada de forma independiente en cada iteración de validación |
| Selección de variables | ANOVA-KBest (<i>Analysis of Variance - K Best Features</i>) | Selección de las 3–5 variables con mayor varianza entre clases (criterio F-value) |
| | LASSO (<i>Least Absolute Shrinkage and Selection Operator</i>) | Parámetro de regularización C = 0.2 , equivalente a un nivel alto de penalización L1 |
| | Elastic Net | Parámetro de regularización C = 0.5–1.0 ; proporción L1/L2 = 0.5–0.7 |
| Modelos de clasificación | Análisis Discriminante Lineal (LDA) | Solución SVD (<i>Singular Value Decomposition</i>) |
| | Naive Bayes Gaussiano | Supone distribución normal de los predictores |
| | Regresión Logística | Regularización L2; C = 1.0 |
| | Máquina de Vectores de Soporte (SVM) | <i>Kernels</i> lineal y RBF (<i>Radial Basis Function</i>); parámetro gamma = "scale" |
| | Random Forest | 100 árboles; <i>bootstrap</i> = True |
| Validación cruzada | <i>Nested</i> Leave-One-Out Cross-Validation (LOOCV) | Bucle interno: ajuste de hiperparámetros; bucle externo: prueba independiente |
| Evaluación del desempeño | Métricas: Exactitud balanceada, F1-macro, AUROC, AUPRC, Brier score | 1.000 iteraciones de <i>bootstrap</i> sobre predicciones <i>out-of-fold</i> |

4. RESULTADOS

Se identificaron aproximadamente 70 pacientes con diagnóstico de meningioma. Tras control de calidad de imagen (DICOM), verificación de integridad clínica e histopatológica y depuración de duplicados/inconsistencias, 13 cumplieron criterios de elegibilidad. Para el desarrollo y validación del modelo preoperatorio de reseccabilidad, se restringió la variable objetivo al grado de resección de Simpson I vs II; se excluyó un caso Simpson IV (n=1) por inviabilidad estadística en un esquema multiclase con tamaño muestral reducido. De este modo, la cohorte analítica final fue de 12 pacientes (Simpson I = 4; Simpson II = 8), con 126 predictores numéricos preoperatorios (clínicos y radiómicos).

Se realizó una evaluación comparativa mediante validación cruzada "dejar-uno-fuera" (LOOCV) que combinó tres estrategias de selección de variables (KBest-ANOVA, LASSO y Elastic Net), todas con filtro de colinealidad y límite de 3 predictores, con siete algoritmos de clasificación. El criterio de selección interna fue la exactitud balanceada para controlar el desbalance entre clases. La configuración con mejor desempeño correspondió a LASSO (C=0.2; 3–5 variables) + Análisis Discriminante Lineal, con exactitud=0.833, exactitud balanceada=0.812 y F1-macro=0.812 en predicciones out-of-fold (Tabla 2).

Tabla 2. Evaluación comparativa de configuraciones de selección de variables y clasificadores durante la validación cruzada

| Método de selección de variables | Clasificador supervisado | Parámetro de regularización (C) | Número máximo de variables seleccionadas | Exactitud | Exactitud balanceada | Puntaje F1-macro |
|--|---|---------------------------------|--|-----------|----------------------|------------------|
| Regresión LASSO (Least Absolute Shrinkage and Selection Operator) | Análisis Discriminante Lineal | 0.2 | 3 | 0.833 | 0.812 | 0.812 |
| Regresión LASSO | Análisis Discriminante Lineal | 0.2 | 5 | 0.833 | 0.812 | 0.812 |
| Regresión LASSO | Clasificador Bayesiano Gaussiano | 0.2 | 3 | 0.833 | 0.812 | 0.812 |
| Regresión LASSO | Clasificador de Vecinos más Cercanos (distancia de Mahalanobis) | 0.2 | 3 | 0.833 | 0.812 | 0.812 |
| Regresión LASSO | Clasificador de Vecinos más Cercanos (distancia de Mahalanobis) | 0.2 | 5 | 0.833 | 0.812 | 0.812 |
| Regresión Elastic Net (combinación L1/L2) | Máquina de Vectores de Soporte con Función de Base Radial | 0.5 | 5 | 0.583 | 0.556 | 0.556 |
| Regresión LASSO | Regresión Logística | 0.7 | 3 | 0.667 | 0.562 | 0.562 |
| Regresión LASSO | Máquina de Vectores de Soporte con Núcleo Lineal | 0.2 | 3 | 0.667 | 0.562 | 0.562 |
| Selección ANOVA-KBest (Análisis de Varianza) | Análisis Discriminante Lineal | — | 3 | 0.583 | 0.496 | 0.496 |

Configuraciones con LASSO+Naive Bayes Gaussiano y LASSO+Clasificador por Centroide mostraron métricas equivalentes en exactitud pero con menor área bajo la curva. Los enfoques con Máquinas de Vectores de Soporte y Bosques Aleatorios demostraron desempeño inferior de manera consistente en este tamaño muestra. El flujo de trabajo LASSO (C=0.2; 3 variables) + Análisis Discriminante Lineal se evaluó bajo validación cruzada anidada, obteniendo exactitud 0.833, exactitud balanceada 0.812, F1-macro 0.812, área bajo la curva ROC 0.844, área bajo la curva de precisión-exhaustividad 0.867 y score de Brier 0.129 sobre las probabilidades out-of-fold (Tabla 3) (Figura 3).

Tabla 3 Desempeño del modelo final (LASSO + Análisis Discriminante Lineal) tras la validación cruzada anidada.

| Métrica | Resultado |
|---|--|
| Exactitud | 0.833 |
| Exactitud balanceada | 0.812 |
| Puntaje F1 promedio | 0.812 |
| Área bajo la curva ROC (AUROC) | 0.844 |
| Área bajo la curva Precisión–Recall (AUPRC) | 0.867 |
| Brier Score | 0.129 |
| Calibración global | Intercepto = 0.000; Pendiente = 0.973 |

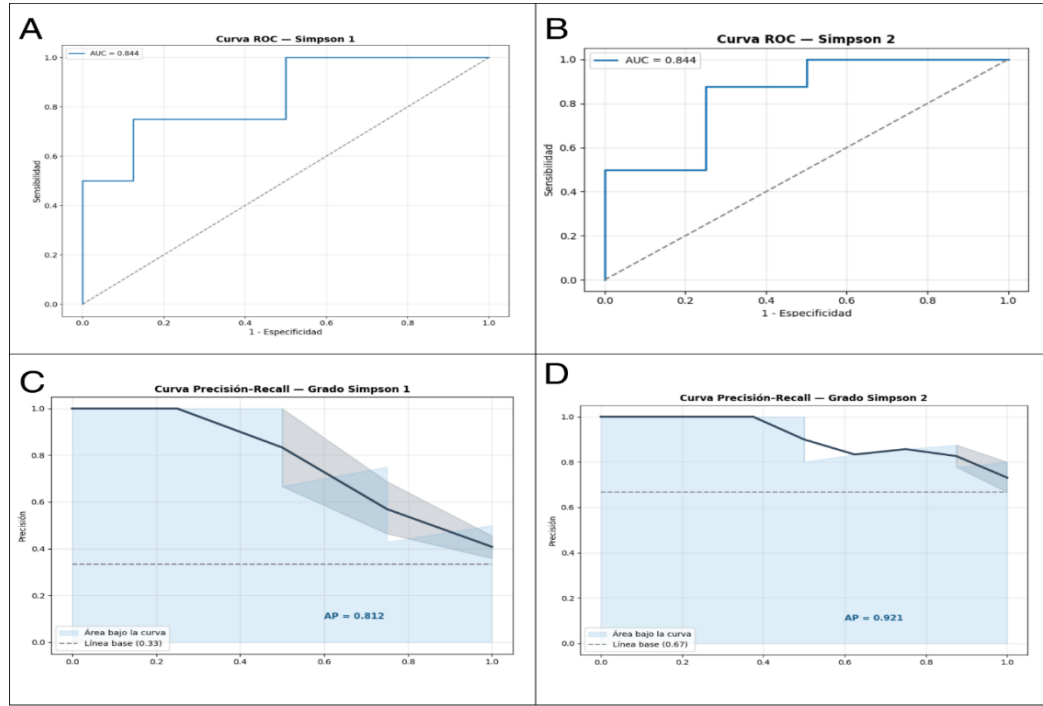


Fig. 3 Desempeño del modelo final (LASSO + LDA) bajo validación cruzada anidada. Los paneles A y B muestran las curvas ROC para las clases Simpson I y Simpson II, respectivamente. Los paneles C y D presentan las curvas de Precisión–Recall (PR) para cada clase.

Desempeño del modelo final obtenido mediante la combinación de Regresión LASSO y LDA tras la validación cruzada anidada. En comparación con la validación simple inicial, el uso del esquema anidado permitió recalibrar los parámetros en cada iteración, evitando la fuga de información y proporcionando una estimación más realista del comportamiento del modelo frente a datos no vistos.

La matriz de confusión mostró 3/4 aciertos para Simpson I (sensibilidad 0.75) y 7/8 para Simpson II (sensibilidad 0.875) (Figura 3). El análisis de calibración reveló intercepto promedio 0.00 y pendiente 0.97 (Figura 4). La evaluación de incertidumbre mediante bootstrap sobre las

predicciones out-of-fold arrojó intervalos de confianza al 95%: exactitud 0.583–1.000; exactitud balanceada 0.500–1.000; F1-macro 0.478–1.000; área bajo la curva de precisión-exhaustividad 0.583–1.000; score de Brier 0.026–0.272

El análisis de asociación mediante correlación punto-biserial identificó que características de forma e intensidad concentraron la mayor señal discriminativa entre Simpson I y II, con Mean.1, Flatness, Strength, Elongation y Sphericity encabezando el ranking de asociación (Figura 5).

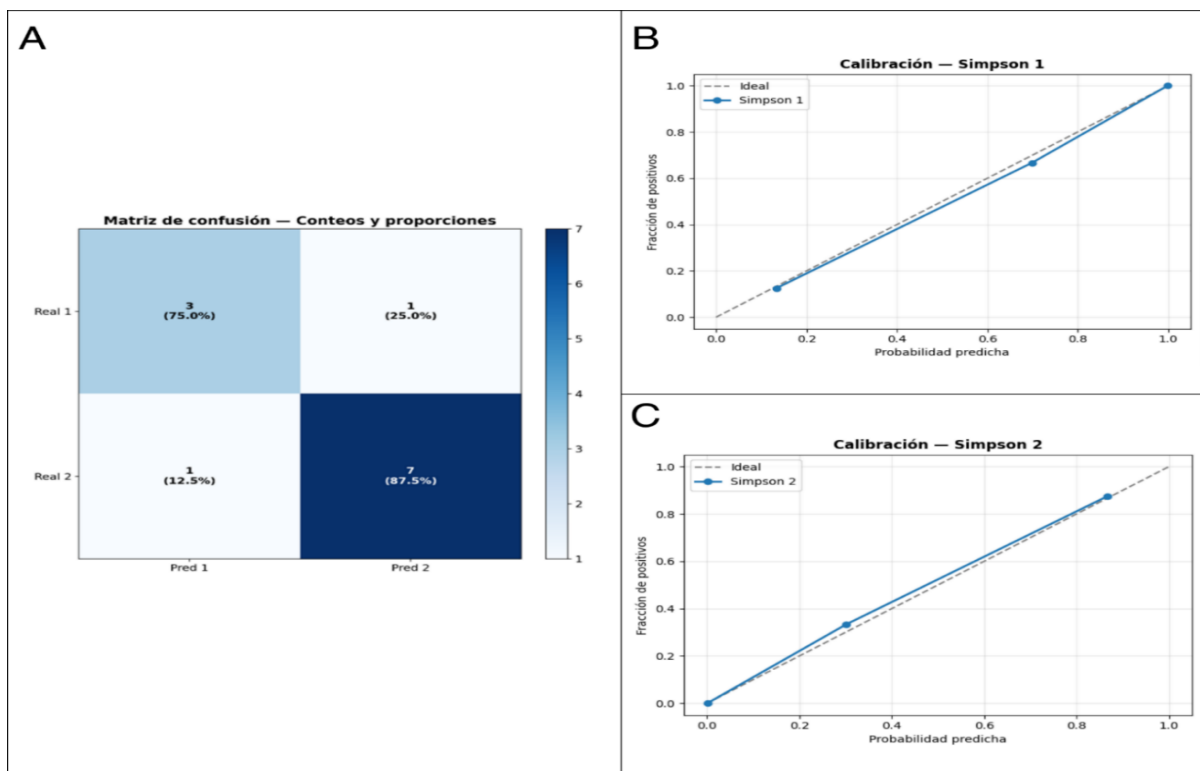


Fig. 3 Matriz de confusión y curvas de calibración del modelo final (LASSO + LDA) tras validación cruzada anidada. (A) Matriz de confusión mostrando el rendimiento de clasificación del modelo final para los grados de resección Simpson I y Simpson II. (B–C) Curvas de calibración por clase que comparan las probabilidades predichas frente a las proporciones reales observadas.

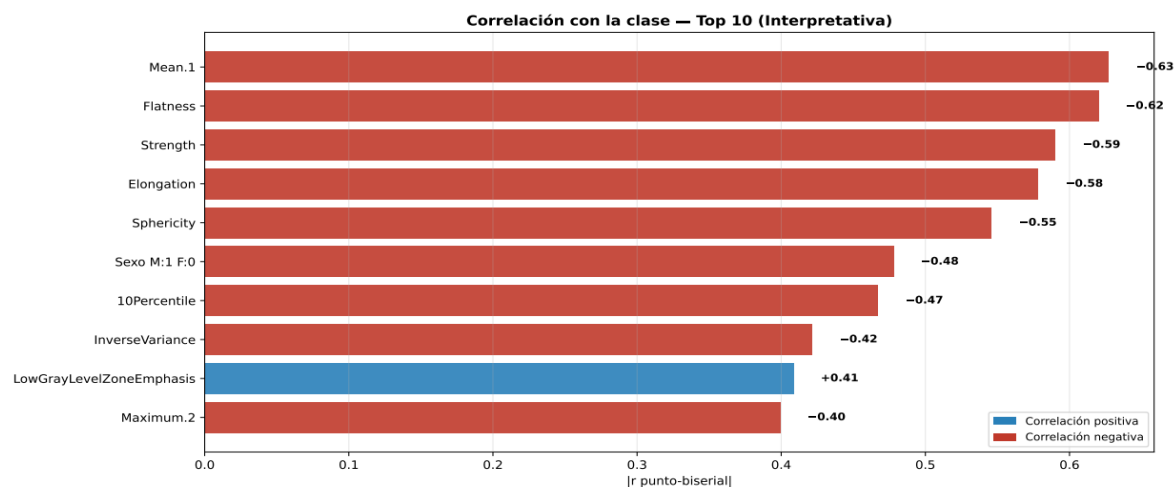


Fig. 4 Variables radiómicas y clínicas con mayor asociación con el grado de resección (Simpson I vs II). Se muestran las diez variables con mayor correlación punto-biserial respecto al desenlace quirúrgico. Correlaciones negativas asociados a Simpson II, mientras que correlaciones positivas asociados a Simpson I.

En el proceso de validación anidada, el análisis de estabilidad en la selección de variables mostró que Strength fue seleccionada en los 12/12 folds, seguido de Flatness y Elongation. Con el fin de visualizar las relaciones lineales entre todos los predictores, se generó un mapa de correlación completo de las variables preoperatorias (Figura X). Este análisis reveló patrones distintivos de asociación, particularmente entre las variables de forma (Figura 6).

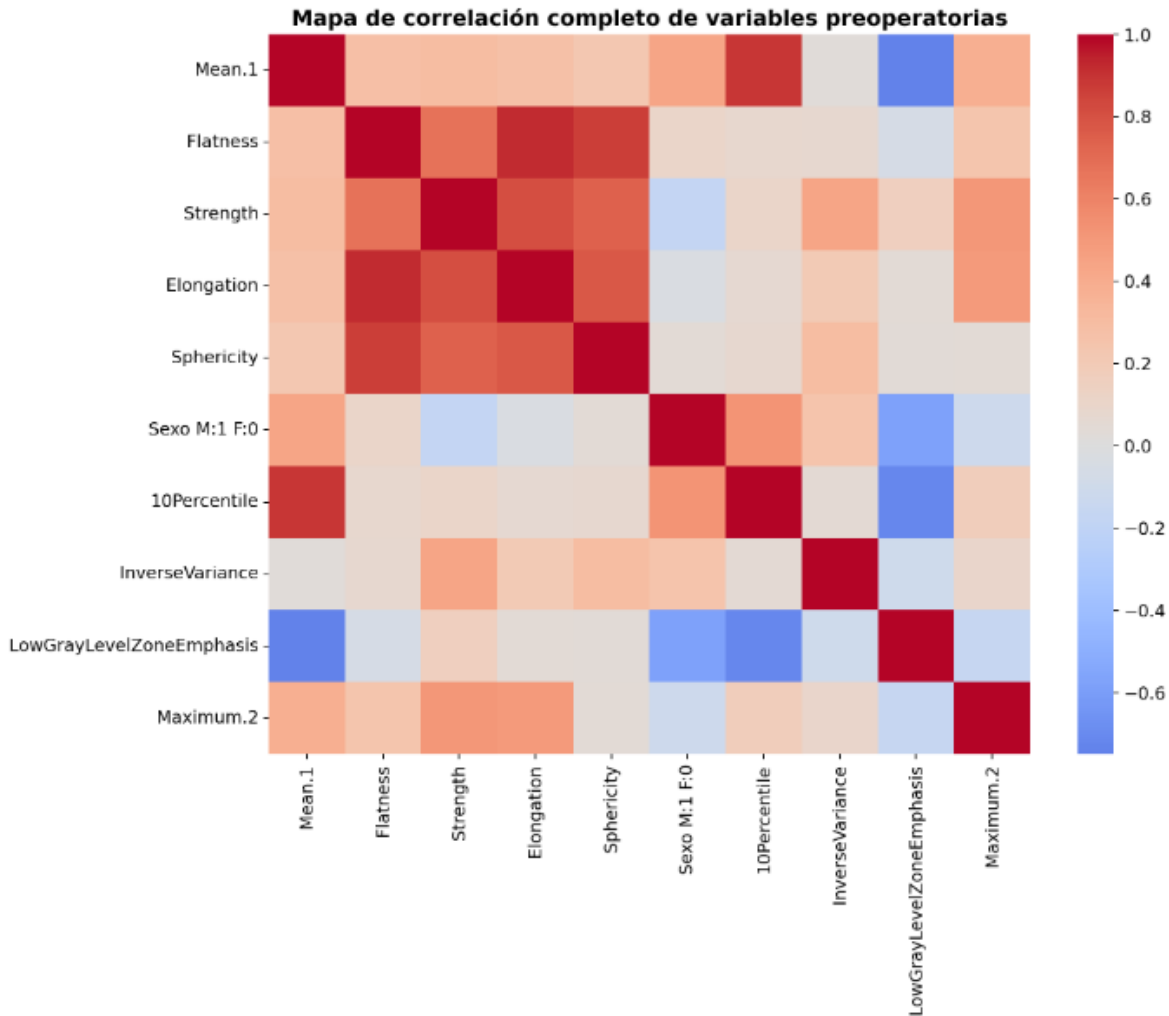


Fig 5 . Mapa de correlación completo entre las variables preoperatorias seleccionadas. Mapa de correlación de Pearson entre las diez variables con mayor asociación al grado de resección (Simpson I vs II) según el análisis punto-biserial.

5. DISCUSION

Este proyecto desarrolló y validó internamente un modelo de aprendizaje automático para predicción de la resecabilidad de meningiomas según la escala de Simpson utilizando características radiómicas y clínicas preoperatorias. El modelo combinó selección LASSO con Análisis Discriminante Lineal, alcanzando AUROC 0,844 y calibración adecuada (pendiente 0,97) en una cohorte de 12 pacientes. Esta capacidad predictiva podría transformar la planificación quirúrgica hacia estrategias más individualizadas (4,5,19).

El análisis identificó Strength como predictor más estable (12/12 folds), característica que cuantifica el contraste de intensidades intratumorales. Desde la perspectiva histopatológica, esta heterogeneidad textural podría reflejar variaciones en celularidad, contenido de colágeno o áreas de necrosis que afectan la consistencia tumoral y adherencia a estructuras adyacentes (20,21). Tumores con alta heterogeneidad textural frecuentemente presentan interfaces tumor-duramadre menos definidas, dificultando las resecciones Simpson I (5).

Características radiómicas de forma Elongation y Flatness mostraron asociación inicial fuerte pero menor estabilidad, explicada por su correlación mutua evidenciada en el mapa de correlación. Morfológicamente, meningiomas elongados o aplanados representan patrones de crecimiento adaptados a superficies durales curvadas o extensiones alongadas en la base craneal (6,22). Estos patrones suelen asociarse con mayor complejidad quirúrgica por proximidad a estructuras neurovasculares críticas (5,6).

La característica Mean.1 (intensidad promedio post-contraste) demostró alto valor discriminativo, sugiriendo que las propiedades de realce reflejan diferencias en vascularización tumoral o integridad de la barrera hematoencefálica (23,24). Estudios previos correlacionan realce homogéneo con meningiomas más encapsulados, mientras patrones heterogéneos indicarían mayor proliferación o transformación anaplásica (24,25).

En este contexto, la combinación de LASSO con Análisis Discriminante Lineal resultó ser la más adecuada porque logró mantener un equilibrio entre rendimiento y estabilidad sin perder interpretabilidad. En escenarios como el de este estudio, con un número alto de variables en relación con el tamaño de la muestra, los métodos lineales penalizados como LASSO controlan el sobreajuste al limitar la influencia de variables poco informativas y conservar solo aquellas con verdadero peso predictivo (23,26). Este mecanismo hace que el modelo sea más estable y fácil de entender desde el punto de vista clínico, a diferencia de enfoques más complejos como Random Forest o SVC, que requieren grandes volúmenes de datos para aprender relaciones no lineales y tienden a perder capacidad de generalización cuando las muestras son pequeñas (27,28). Además, los modelos lineales permiten visualizar de forma directa la contribución de cada variable, lo que aporta

trazabilidad y coherencia con los principios de transparencia y reproducibilidad recomendados (18). El desempeño obtenido confirma que un modelo simple, bien calibrado y explicable puede ser más robusto y clínicamente útil que aproximaciones de mayor complejidad en ciertos escenarios (14,27).

En cuanto a la validación leave-one-out (LOOCV) anidada, esta representó la estrategia óptima para el contexto del estudio, maximizando información disponible mientras proporciona estimación rigurosa del error (18,28). Estudios con cohortes similares han empleado exitosamente esta aproximación (15,16).

Respondiendo a un rendimiento del modelo (AUROC 0,844, exactitud balanceada 0,812) vemos que este se sitúa favorablemente dentro del espectro reportado en literatura para clasificación binaria en neuroimágenes (14,28). Es importante considerar la sensibilidad para Simpson II (0,875), ya que identificar correctamente casos con riesgo de resección subóptima es clínicamente útil para optimizar la planificación quirúrgica (4,5).

Durante el desarrollo del estudio se abordó proactivamente las limitaciones metodológicas identificadas en revisiones críticas del campo. El metaanálisis de Ugga y colaboradores reportó calidad metodológica promedio de solo 19% según el Radiomics Quality Score en estudios de meningiomas (14). En contraste, el presente trabajo implementa rigurosamente las guías IBSI para estandarización de características radiómicas (17) y sigue recomendaciones TRIPOD+AI para reporte de modelos de predicción (18).

La reciente iniciativa multicéntrica enfatiza la importancia crítica de protocolos estandarizados de preprocesamiento (17). El flujo de trabajo para la segmentación 3D, resampling isotrópico y normalización z-score responde directamente a estas recomendaciones, estableciendo bases sólidas para futuras validaciones externas.

La orientación del estudio, su desarrollo y resultados puede contribuir a la incorporación de estos métodos y modelos, considerando que predecir la extensión de resección podría influir en decisiones quirúrgicas como selección de abordaje o consideración de estrategias multimodales para que el paciente tenga el mejor desenlace clínico esperable (4,5,13).

6. LIMITACIONES Y RECOMENDACIONES

Este estudio contiene limitaciones. Una de las más importantes es el tamaño muestral reducido, aunque es pertinente contextualizar esta limitación dentro del panorama del desarrollo de modelos radiómicos, donde cohortes de tamaño similar han demostrado viabilidad para desarrollo inicial (14,15,16). La naturaleza retrospectiva y no multicéntrica restringe la posibilidad de extrapolarse de forma inmediata a una población definida.

Las direcciones futuras deben priorizar validación externa multicéntrica para evaluar el rendimiento del modelo con diferentes protocolos de adquisición (17,18). La integración de secuencias multiparamétricas adicionales representa otra condición a tener en cuenta, considerando hay reportes de mejoría significativas en rendimiento al combinar múltiples secuencias de imagen. Finalmente, estudios de implementación que evalúen el impacto real en decisiones quirúrgicas representan el siguiente escalón crítico a una aplicación clínica (13,18).

7. CONCLUSIÓN

En conclusión, el presente demuestra viabilidad de predecir resecabilidad en meningiomas utilizando características radiómicas preoperatorias y establece un referente metodológico con adecuado rendimiento para el desarrollo de herramientas de decisión quirúrgica cuantitativas. La identificación de características específicas de textura y forma como predictores claves proporcionan ideas y explicaciones sobre los sustratos biológicos que están relacionados a la complejidad quirúrgica, contribuyendo a avanzar hacia estándares más rigurosos de desarrollo y validación.

REFERENCIAS

- [1] Ostrom QT, Cioffi G, Gittleman H, et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2012–2016. *Neuro Oncol.* 2019;21(Suppl 5):v1–v100.
- [2] Louis DN, Perry A, Wesseling P, et al. The 2021 WHO classification of tumors of the central nervous system: a summary. *Neuro Oncol.* 2021;23(8):1231–1251.
- [3] Apra C, Peyre M, Kalamarides M. Current treatment options for meningioma. *Expert Rev Neurother.* 2018;18(3):241–249.
- [4] Simpson D. The recurrence of intracranial meningiomas after surgical treatment. *J Neurol Neurosurg Psychiatry.* 1957;20(1):22–39.
- [5] Lemée JM, Corniola MV, Da Broi M, et al. Extent of resection in meningioma: Predictive factors and clinical implications. *Sci Rep.* 2019;9:5944.
- [6] Voß KM, Spille DC, Sauerland C, et al. The Simpson grading in meningioma surgery: does the tumor location influence the prognostic value? *J Neurooncol.* 2017;133(3):641–651.
- [7] Aerts HJWL, Velazquez ER, Leijenaar RTH, et al. Decoding tumor phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014;5:4006.
- [8] Park YW, Han K, Ahn SS, et al. Prediction of recurrence in meningiomas using radiomics and machine learning-based approaches. *Sci Rep.* 2020;10(1):16816.
- [9] Wang YY, Fan XS, Lin P, et al. Integrating radiomics and clinical features for prediction of meningioma grades. *Cancer Imaging.* 2021;21(1):30.
- [10] Sahm F, Schrimpf D, Stichel D, et al. DNA methylation-based classification and grading system for meningioma: a multicenter, retrospective analysis. *Lancet Oncol.* 2020;21(11):1580–1589.
- [11] 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.
- [12] Zhang L, Liu Y, Zhang X, et al. Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer. *Front Oncol.* 2020;10:567736.
- [13] Cahill KS, Claus EB. Treatment and survival of patients with nonmalignant intracranial meningioma: results from the SEER Program of the National Cancer Institute. *J Neurosurg.* 2011;115(2):259–267.
- [14] Ugga L, Perillo T, Cuocolo R, Stanzione A, Romeo V, Green R, Aoki S, Brunetti A. Meningioma MRI radiomics and machine learning: systematic review, quality score assessment, and meta-analysis. *Neuroradiology.* 2021;63:1293–1304.
- [15] Laukamp KR, Shakirin G, Baeßler B, Thiele F, Zopfs D, Große Hokamp N, Timmer M, Kabbasch C, Perkuhn M, Borggrefe J. Accuracy of radiomics-based feature analysis on multiparametric magnetic resonance images for noninvasive meningioma grading. *World Neurosurg.* 2019;132:e366–e390.
- [16] Zhu Y, Man C, Gong L, Dong D, Yu X, Wang S, Fang M, Wang S, Fang X, Chen X, Tian J. A deep learning radiomics model for preoperative grading in meningioma. *Eur J Radiol.* 2019;116:128–134.
- [17] Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology.* 2020 May;295(2):328–38.

- [18] Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024 Apr 16;385:e078378.
- [19] Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278(2):563–577.
- [20] Zhang J, Yao K, Liu P, et al. A radiomics model for preoperative prediction of brain invasion in meningioma. *NeuroImage: Clinical*. 2020;28:102375.
- [21] Chu H, Lin X, He J, Pang P, Fan B, Lei P, Guo D, Ye C. Value of MRI Radiomics Based on Enhanced T1WI Images in Prediction of Meningiomas Grade. *Academic Radiology*. 2020;27(8):e227-e237.
- [22] Ke C, Chen H, Lv X, Li H, Zhang Y, Chen M, Hu D, Ruan G, Zhang Y, Zhang Y, Liu L, Feng Y. Differentiation between benign and nonbenign meningiomas by using texture analysis from multiparametric MRI. *Journal of Magnetic Resonance Imaging*. 2020;51(6):1810–1820.
- [23] Tibshirani R. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1996;58(1):267–288.
- [24] Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, et al. Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View. *Journal of Medical Internet Research*. 2016;18(12):e323.
- [25] Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. *JAMA*. 2017;318(14):1377–1384.
- [26] Chen C, Guo X, Wang J, Guo W, Ma X, Xu J. The diagnostic value of radiomics-based machine learning in predicting the grade of meningiomas using conventional magnetic resonance imaging: a preliminary study. *Frontiers in Oncology*. 2019;9:1338.
- [27] Morin O, Chen WC, Nassiri F, Susko M, Magill ST, Vasudevan HN, et al. Integrated models incorporating radiologic and radiomic features predict meningioma grade, local failure, and overall survival. *Neuro-Oncology Advances*. 2019;1(1):vdz011.
- [28] Park SH, Han K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. *Radiology*. 2018;286(3):800-809.