

ORIGINAL

Clasificadores de aprendizaje supervisado no lineales basados en radiómica de la TC cerebral sin contraste para predecir el pronóstico funcional en pacientes con hematoma intracerebral espontáneo

E. Serrano^a, J. Moreno^b, L. Llull^c, A. Rodríguez^c, C. Zwanzger^d, S. Amaro^c, L. Oleaga^e y A. López-Rueda^{e,f,*}

^a Departamento Radiología, Hospital Universitario de Bellvitge, Hospitalet de Llobregat, Barcelona, España

^b Clínica Iribas-IRM, Asunción, Paraguay

^c Departamento de Neurología, Hospital Clínic, Barcelona, España

^d Departamento Radiología, Hospital del Mar, Barcelona, España

^e Departamento Radiología, Hospital Clínic, Barcelona, España

^f Servicio de Informática Clínica, Hospital Clínic, Barcelona, España

Recibido el 21 de mayo de 2023; aceptado el 3 de agosto de 2023

Disponible en Internet el 4 de octubre de 2023

PALABRAS CLAVE

Accidente cerebrovascular agudo;
Hemorragia intracerebral;
CT Scanner;
X-ray;
Inteligencia artificial;
Biomarcadores

Resumen

Objetivo: Evaluar si clasificadores de aprendizaje supervisado no lineales basados en radiómica de la TC cerebral sin contraste (TCCSC) pueden predecir el pronóstico funcional al alta en pacientes con hematoma intracerebral espontáneo (HIE).

Material y método: Análisis observacional retrospectivo y unicéntrico de pacientes con diagnóstico de HIE confirmado por TCCSC entre enero de 2016 y abril de 2018. Se incluyeron pacientes con HIE > 18 años y con TCCSC realizado dentro de las primeras 24 horas del inicio de los síntomas. Se excluyeron los HIE secundarios y en los que no se disponía de las variables de radiómica. Se recogieron datos clínicos, demográficos y variables al ingreso. Los pacientes se clasificaron según la Escala Modificada de Rankin (mRS) al alta en buen (mRS 0-2) y mal pronóstico (mRS 3-6). Tras la segmentación manual de la TCCSC de cada HIE se obtuvieron las variables de radiómica. La muestra se dividió en una cohorte de entrenamiento y prueba y otra cohorte de validación (70-30%, respectivamente). Se usaron diferentes métodos de selección de variables y reducción de dimensionalidad, así como diferentes algoritmos para la construcción del modelo. Se realizaron 10 iteraciones de validación cruzada estratificada en la cohorte de entrenamiento y prueba y se calculó la media de los valores de área bajo la curva (AUC). Una vez entrenados los modelos, se calculó la sensibilidad de cada uno para predecir el pronóstico funcional al alta en la cohorte de validación.

* Autor para correspondencia.

Correo electrónico: alrueda81@hotmail.com (A. López-Rueda).

Resultados: Se analizaron 105 pacientes con HIE. Se evaluaron 105 variables de radiómica de cada paciente. Los algoritmos P-SVM, KNN-E y RF-10, en combinación con el método de selección de variables ANOVA, fueron los clasificadores con mejor rendimiento en la cohorte de entrenamiento y prueba (AUC: 0,798, 0,752 y 0,742, respectivamente). Las predicciones de estos modelos, en la cohorte de validación, tuvieron una sensibilidad de 0,897 (IC 95%: 0,778-1), con una tasa de falsos negativos del 0% para la predicción de mal pronóstico funcional al alta.

Conclusión: Los clasificadores de aprendizaje supervisado no lineales basados en radiómica son una herramienta de diagnóstico prometedora para predecir el resultado funcional al alta en pacientes con HIE, con una baja tasa de falsos negativos, aunque todavía son necesarios estudios con mayor tamaño muestral y balanceados para desarrollar y mejorar su rendimiento.

© 2023 SERAM. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

KEYWORDS

Acute cerebrovascular accident; Cerebral hemorrhage; CT Scanner; X-ray; AI (artificial intelligence); Biomarker

Radiomic-based nonlinear supervised learning classifiers on non-contrast CT to predict functional prognosis in patients with spontaneous intracerebral hematoma

Abstract

Purpose: To evaluate if nonlinear supervised learning classifiers based on non-contrast cerebral CT can predict functional prognosis at discharge in patients with spontaneous intracerebral hematoma (HIE).

Methods: Retrospective, single-center, observational analysis of patients with a diagnosis of spontaneous intracerebral hematoma confirmed by non-contrast CT between January 2016 and April 2018. Patients with HIE > 18 years and with non-contrast CT performed within the first 24 hours of symptom onset were included. Patients with secondary spontaneous intracerebral hematoma and in whom radiomic variables were not available were excluded. Clinical, demographic and admission variables were collected. Patients were classified according to the Modified Rankin Scale (mRS) at discharge into good (mRS 0-2) and poor prognosis (mRS 3-6). After manual segmentation of each spontaneous intracerebral hematoma, the radiomics variables were obtained. The sample was divided into a training and testing cohort and a validation cohort (70-30%, respectively). Different methods of variable selection and dimensionality reduction were used, and different algorithms were used for model construction. Stratified 10-fold cross-validation were performed on the training and testing cohort and the mean area under the curve (AUC) were calculated. Once the models were trained, the sensitivity of each was calculated to predict functional prognosis at discharge in the validation cohort.

Results: 105 patients with spontaneous intracerebral hematoma were analyzed. 105 radiomic variables were evaluated for each patient. P-SVM, KNN-E and RF-10 algorithms, in combination with the ANOVA variable selection method, were the best performing classifiers in the training and testing cohort (AUC: 0.798, 0.752 and 0.742, respectively). The predictions of these models, in the validation cohort, had a sensitivity of 0.897 (95% CI: 0.778-1), with a false-negative rate of 0% for predicting poor functional prognosis at discharge.

Conclusion: The use of radiomics-based nonlinear supervised learning classifiers are a promising diagnostic tool for predicting functional outcome at discharge in HIE patients, with a low false negative rate, although larger and balanced samples are still needed to develop and improve their performance.

© 2023 SERAM. Published by Elsevier España, S.L.U. All rights reserved.

Introducción

El accidente cerebrovascular (ACV) es la segunda causa de muerte en todo el mundo y una de las principales causas de incapacidad¹. El hematoma intracerebral espontáneo (HIE) es el segundo tipo más frecuente de ACV, después del ictus isquémico, y representa el 10-20% de todos los ACV².

Aunque el pronóstico es variable, el HIE sigue siendo una causa importante de mortalidad y de morbilidad en todo el mundo³, con una tasa de mortalidad de aproximadamente

el 40% al mes, 54% al año y 75% a los 5 años. Solo del 12% al 39% de los pacientes logran independencia funcional³. Identificar a los pacientes con riesgo de mal pronóstico funcional permite una mejor estratificación para ofrecer terapias intensivas adaptadas a estos pacientes⁴.

Al ser la tomografía computarizada cerebral sin contraste (TCCSC) el método diagnóstico de elección en el HIE agudo⁵, han surgido varios parámetros cualitativos (signos radiológicos) que se asocian a crecimiento del HIE y a mal pronóstico en estos pacientes. Los signos radiológicos de la TCCSC son

una forma de representar las características de morfología (irregularidad) y densidad (heterogeneidad) del HIE.

Sin embargo, estos parámetros cualitativos (signos radiológicos)^{6–12} pueden estar sujetos a cierta subjetividad en la interpretación, y algunos de ellos tienen definiciones superpuestas¹³, habiéndose publicado resultados de concordancia interobservador variables¹⁴. Además, se han descrito tasas bajas de sensibilidad en la predicción pronóstica de estos signos (entre el 14,3 y el 39,2% para pronóstico funcional según el grupo de Law)¹⁵, lo que puede dar lugar a estimaciones erróneas y eventualmente afectar a la toma de decisiones.

Estas limitaciones ponen de manifiesto la necesidad de desarrollar y utilizar parámetros cuantitativos reproducibles (biomarcadores) que minimicen el componente subjetivo de la evaluación radiológica cualitativa y que mejoren el rendimiento en la predicción del pronóstico de los signos radiológicos.

La radiómica es un enfoque cuantitativo en el que se extraen un gran número de características (biomarcadores) de las imágenes radiológicas^{16,17}.

Al aplicar algoritmos de aprendizaje automático supervisado¹⁸ sobre los parámetros cuantitativos (biomarcadores) se pueden crear modelos que puedan realizar la predicción de una variable resultado sobre datos no evaluados previamente. Nuestra hipótesis es que clasificadores de aprendizaje supervisado no lineales basados en radiómica pueden predecir el mal pronóstico funcional en pacientes con HIE, al objetivar características no apreciables por el ojo humano (intensidad, forma, textura...).

El objetivo de este estudio es evaluar si clasificadores de aprendizaje supervisado no lineales basados en radiómica de la TCCSC pueden predecir el pronóstico funcional al alta en pacientes con HIE.

Material y métodos

El manuscrito ha sido estructurado atendiendo a la iniciativa CLAIM (de sus siglas en inglés *Checklist for Artificial Intelligence in Medical Imaging*)¹⁹.

Para la segmentación de cada HIE se utilizó el Software 3D Slicer (versión 4.10.2). El software utilizado para el procesamiento de datos fue Orange data mining versión 3.31 (<https://orangedatamining.com/>).

Diseño del estudio

Análisis observacional retrospectivo unicéntrico de pacientes consecutivos con diagnóstico de HIE confirmado por TCCSC entre enero de 2016 y abril de 2018 en un centro terciario de ictus. El objetivo es crear un algoritmo de aprendizaje supervisado no lineal de cribado para predecir el pronóstico funcional al alta en pacientes con HIE. Dado el objetivo marcado, las métricas que se quieren optimizar son la sensibilidad y el valor predictivo negativo.

Datos

El protocolo del estudio fue aprobado por el Comité de Ética de Investigación Clínica local (número de registro

HCB/2020/0180) según lo establecido en las leyes y normas nacionales (Ley 14/2007, de 3 de julio, de Investigación Biomédica) e internacionales (Declaración de Helsinki, en su última actualización de Fortaleza, Brasil, 2013). Dada la naturaleza retrospectiva del estudio, no se requirió el consentimiento informado específico para la inclusión de los datos en el estudio. Para garantizar el anonimato de los participantes en el estudio se llevó a cabo un proceso de seudonimización. El conjunto de datos empleado en este estudio no se ha utilizado previamente. Los datos que respaldan las conclusiones de este estudio están disponibles mediante el autor de correspondencia, previa solicitud justificada de los mismos.

Inicialmente se valoraron un total de 128 pacientes mayores de 18 años con diagnóstico de HIE espontáneo y con TCCSC realizado dentro de las primeras 24 horas desde el inicio de los síntomas. De estos se excluyeron los pacientes con HIE secundario y los pacientes en los que no se disponía de todas las variables de radiómica. Se recogieron datos demográficos (edad y sexo), hábitos tóxicos (alcohol y tabaco), factores de riesgo cerebrovascular y cardiovascular (hipertensión, dislipidemia, diabetes mellitus, fibrilación auricular y cardiopatía isquémica), antecedentes médicos de HIE o ictus previo y tratamiento farmacológico antiagregante o anticoagulante concomitante.

Al ingreso, se registraron la presión arterial sistólica y diastólica (mmHG), la glucemia (mmol/l) y la evaluación neurológica inicial mediante la *National Institutes of Health Stroke Scale* (NIHSS). El pronóstico funcional al alta se determinó utilizando la escala de Rankin modificada (mRS) dicotomizada en buen pronóstico (mRS 0-2) o mal pronóstico (mRS 3-6).

Se realizó una TCCSC secuencial en dos equipos de TC multicorte (Somatom Definition Flash y Somatom Sensation 64, Siemens Healthcare, Erlangen, Alemania). Se obtuvieron imágenes secuenciales axiales paralelas a la línea orbitomeatal desde la base del cráneo hasta el vértece, utilizando parámetros estándar de 140 kV, 230 mAs y reconstrucciones axiales con un grosor de 5 mm.

Se analizaron la localización del HIE (ganglios basales, lobar, tronco encefálico y cerebelo), la presencia de hemorragia intraventricular y el volumen del hematoma (ml). El volumen del hematoma se calculó según el método validado A × B × C / 2²⁰.

Las imágenes de la TCCSC de cada paciente incluido en el estudio se importaron desde el *Picture Archiving and Communication System* (PACS) al Software 3D Slicer (versión 4.10.2), donde se utilizó el módulo «Segment Editor» para la segmentación.

El proceso de segmentación fue llevado a cabo por un radiólogo cualificado sin acceso a la información clínica. Los contornos de todos los HIE fueron dibujados manualmente corte a corte y se formaron los volúmenes de interés (VOI) tridimensionales de cada HIE.

Desde el módulo «Radiomics» del Software 3D Slicer se obtuvieron de forma automática un total de 105 variables de cada uno de los VOI, relacionadas con la intensidad, la forma y la textura del hematoma. El módulo «Radiomics» de 3D Slicer se basa en la librería *pyRadiomics*²¹, que cumple con el estándar *Image Biomarker Standardisation Initiative* (IBSI)²². En la figura 1 se resume el proceso de segmentación.

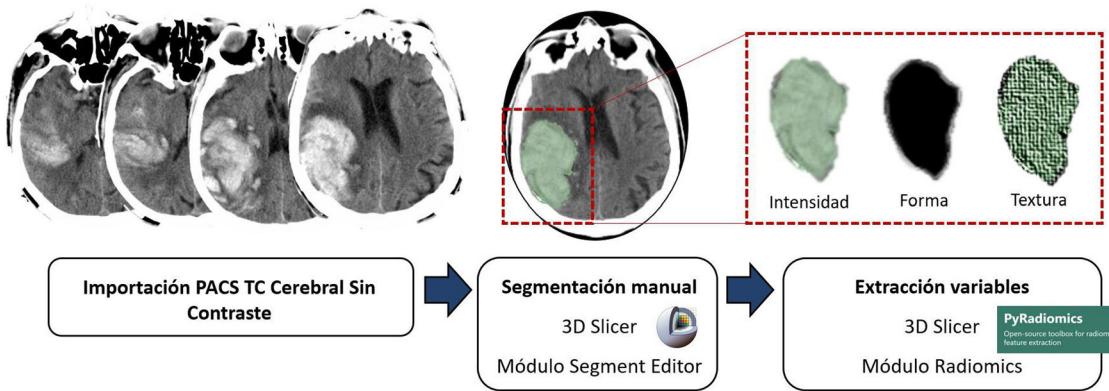


Figura 1 Proceso de segmentación. Las imágenes de la TCCSC se importaron desde el *Picture Archiving and Communication System* (PACS) al Software 3D Slicer (versión 4.10.2), donde se utilizó el módulo «Segment Editor» para la segmentación. Los contornos de todos los HIE fueron dibujados manualmente corte a corte y se formaron los volúmenes de interés tridimensionales (VOI) de cada HIE. Desde el módulo «Radiomics» del Software 3D Slicer se obtuvieron de forma automática un total 105 variables de cada uno de los VOI, relacionadas con la intensidad, la forma y la textura del hematoma.

Verdad fundamental

El objetivo principal de este trabajo fue evaluar si los clasificadores de aprendizaje supervisado basados en radiómica de la TCCSC son capaces de predecir el mal pronóstico funcional al alta en pacientes con HIE. Se definió mal pronóstico como aquellos pacientes con dependencia para realizar las actividades diarias al alta o defunción tras el HIE (mRS 3-6). La variable resultado (mRS al alta) fue dicotomizada en buen pronóstico (independencia funcional al alta: mRS 0-2) y mal pronóstico (dependencia funcional o defunción al alta: mRS 3-6). La escala mRS evalúa la discapacidad global tras un ACV y es la medida de resultado funcional más completa y utilizada en los ensayos de ictus. Sin embargo, se ha descrito variabilidad interobservador que puede implicar una clasificación errónea y limitar la validez de los resultados²³. Por ello, el mRS al alta ha sido determinado por seis neurólogos vasculares con más de cinco años de experiencia y con certificación en su determinación. Además, se ha dicotomizado el resultado del mRS en lugar de emplearlo en su escala original, ya que se ha demostrado que las tasas de error de clasificación son menores con la dicotomización del mRS, en comparación con la determinación ordinal²⁴.

Distribución de datos

La muestra se dividió en dos cohortes estratificadas de pacientes, una cohorte de entrenamiento y prueba (70%, n=70) y otra de validación (30%, n=29). La estratificación de la muestra se refiere a que cada cohorte tiene la misma proporción de pacientes con buen y mal pronóstico funcional que la muestra global.

Modelo

El software utilizado para el procesamiento de datos fue Orange data mining versión 3.31 (<https://orangedatamining.com/>). Todas las variables de radiómica obtenidas (tanto en la cohorte de entrenamiento y prueba como en la cohorte de validación) se sometieron a un procesamiento inicial para asegurar el correcto funcionamiento de los clasificadores. Este procesamiento consiste,

en primer lugar, en excluir los pacientes que tenían valores perdidos (ausentes) en alguna de las 105 variables. En segundo lugar, mediante el algoritmo «Isolation Forest» se eliminaron el 5% de los pacientes con valores de radiómica extremos. Por último, se llevó a cabo un proceso de estandarización y normalización para convertir las variables de radiómica en variables con una distribución normal con valores comprendidos entre 0 y 1.

Reducción de la dimensionalidad - Selección de variables

Una vez realizado el procesamiento inicial de las variables de radiómica, para seleccionar las variables a estudio se utilizaron, por un lado, todas las variables de radiómica sin seleccionar ni reducir su dimensionalidad (evaluando las 105 variables de cada uno de los pacientes), y, por otro, se aplicaron técnicas de reducción de la dimensionalidad para identificar y retirar información irrelevante y redundante²⁵. Dado que contamos con una muestra de 70 pacientes para entrenar el modelo, se redujo la dimensionalidad en 7 variables para minimizar el efecto del sobreajuste (*overfitting*). Los métodos empleados de selección de variables y de reducción de la dimensionalidad fueron:

- Método ANOVA: seleccionamos 7 variables según tengan mayor diferencia entre los valores promedio de la característica en diferentes clases, siempre y cuando no exista correlación entre ellas (coeficiente de correlación de Spearman < 0,5).
- Algoritmos de reducción de la dimensionalidad lineal:
- o *Principal Component Analysis* (PCA)²⁶: es un procedimiento estadístico que transforma ortogonalmente las n dimensiones numéricas originales de un conjunto de datos en un nuevo conjunto de n dimensiones llamadas componentes principales. En este caso, el nuevo conjunto tendrá tantas dimensiones como sean necesarias para conservar el 80% de la variación de los datos.
- o PCA-90: en este caso, el nuevo conjunto tendrá tantas dimensiones como sean necesarias para conservar el 90% de la variación de los datos.



Figura 2 Resumen del método de procesamiento de datos, selección de variables y construcción del modelo.

- Algoritmos de reducción de la dimensionalidad no lineales con los que seleccionaremos 7 nuevas dimensiones (variables) para representar nuestras variables originales:
- o tSNE-7 (incrustación de vecinos estocásticos distribuidos en t)²⁷: es un algoritmo que calcula la probabilidad de que pares de puntos de datos en el espacio de alta dimensión estén relacionados y luego elige una inserción de baja dimensión que produce una distribución similar.
- o Isomap-7 (mapeo de características isométricas)²⁸: es un algoritmo que proyecta los datos a una dimensión inferior, preservando la distancia más corta entre dos puntos de una curva.
- o LLE-7 (incrustación localmente lineal)²⁹: es un algoritmo que mantiene las características lineales locales de las muestras, por lo que cada punto puede estar representado como una suma lineal y ponderada de sus vecinos.

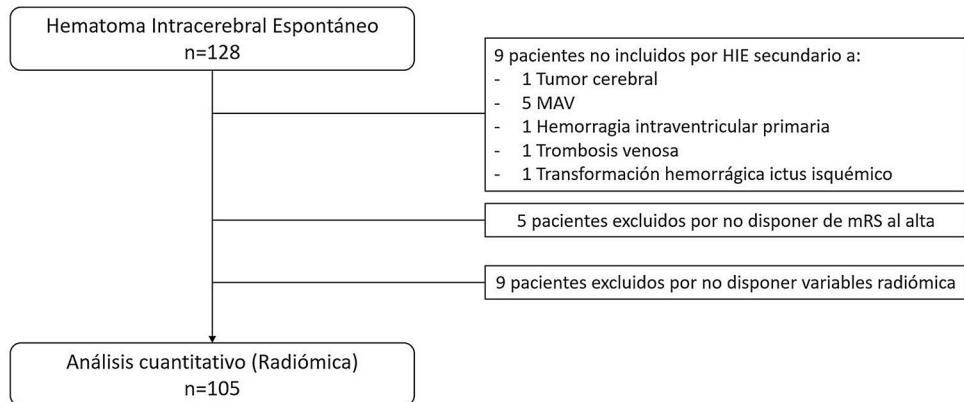
Construcción del modelo

Para la construcción del modelo se emplearon algoritmos utilizados previamente en la literatura con buenos resultados en la evaluación del crecimiento y el pronóstico del HIE³⁰⁻³². Se aplicaron los siguientes algoritmos:

- K vecinos más cercanos (KNN): es uno de los algoritmos de clasificación más básicos. Es un algoritmo de aprendizaje no paramétrico, es decir, que no hace suposiciones sobre la forma funcional de los datos. Por el contrario, es un algoritmo basado en instancias, es decir, el algoritmo no aprende un modelo, sino que memoriza las instancias de entrenamiento que son usadas como «base de conocimiento» para realizar las predicciones^{33,34}.
- o Distancia euclídea (KNN-E): es la distancia en línea recta o la trayectoria más corta posible entre dos puntos.
- o Distancia Manhattan (KNN-M): la distancia Manhattan entre 2 puntos es la suma de las diferencias absolutas de sus coordenadas. Es decir, es la suma de las longitudes de los dos catetos del triángulo rectángulo. Es la distancia entre dos puntos en una cuadrícula de calles tipo ciudad, en la que es no es posible recorrer los dos puntos en línea recta.

- Máquinas de vectores de soporte (*Support Vector Machine [SVM]*): es un algoritmo que representa las instancias de la muestra en el espacio, separando las clases a dos espacios mediante un hiperplano de separación. Cuando las nuevas instancias se introducen en el modelo, en función de los espacios a los que pertenezcan, pueden ser clasificadas a una u otra clase. Una SVM construye un hiperplano o conjunto de hiperplanos en un espacio de alta dimensionalidad que puede ser utilizado en problemas de clasificación o regresión. La manera más simple de realizar la separación es mediante una línea recta, un plano recto, pero los problemas a estudiar no se suelen representar en casos de dos dimensiones. Habitualmente, un algoritmo de SVM debe tratar con curvas no lineales de separación y más de dos variables predictoras. La representación por medio de funciones Kernel ofrece una solución a este problema, ya que toma un espacio de entrada de baja dimensión y lo transforma en un espacio dimensional más alto, es decir, convierte el problema no separable en un problema separable^{33,35,36}. Se utilizaron Kernel Polinomial (P-SVM); Kernel Radial (R-SVM) y Kernel Sigmoideo (S-SVM).
- Random Forest (RF): en este algoritmo se ejecutan varios algoritmos de árboles de decisiones en *bagging*, es decir, los distintos árboles ven distintas proporciones de datos, por lo que cada árbol se entrenará con distintas muestras de datos para un mismo problema. Para clasificar una nueva instancia, cada árbol de decisión da una clasificación, y finalmente la decisión con mayor número de «votos» es la predicción del algoritmo^{33,37}. Se utilizaron algoritmos sobre la decisión de 10 árboles (RF-10) y 50 árboles (RF-50).
- Gradient boosting CatBoost (GB): algoritmo construido con árboles de decisión individuales entrenados de forma secuencial, de forma que cada nuevo árbol trata de mejorar los errores del anterior (*boosting*). La predicción de una nueva instancia se obtiene agregando las predicciones de todos los árboles individuales que forman el modelo^{33,38}.

En la figura 2 se resume el método de procesamiento de datos, de selección de variables y de construcción del modelo.

**Figura 3** Diagrama de flujo de la selección de pacientes.

Entrenamiento y prueba

En la cohorte de entrenamiento y prueba se realizaron 10 iteraciones de validación cruzada estratificada, es decir, se entrenó al modelo con el 90% de los casos de la cohorte de entrenamiento y prueba y se realizó la predicción con el 10% restante de la muestra. Esto se realizó 10 veces y se calculó la media de los valores de área bajo la curva ROC (AUC) de las 10 predicciones. No se realizaron estrategias de *data augmentation* en nuestra muestra.

Una vez entrenados los algoritmos en la cohorte de entrenamiento y prueba se realizaron las predicciones con la cohorte de validación. Se calcularon la sensibilidad y el valor predictivo negativo de los clasificadores en la cohorte de validación. Se definió un IC del 95%.

Resultados

Datos

Un total de 105 pacientes cumplieron finalmente los criterios de inclusión y exclusión, y fueron analizados; en la figura 3 se representa el diagrama de flujo de los pacientes. Las principales características demográficas, clínicas y de imagen se recogen en las tablas 1 y 2.

Procesamiento de los datos y cohorte de entrenamiento y prueba (fig. 4)

Tras excluir los pacientes con valores perdidos y aplicar el algoritmo «Isolation Forest», con en el que se eliminaron el 5% de los pacientes con valores extremos, las variables de radiómica de 99 pacientes con HIE fueron analizadas. La muestra se dividió en una cohorte de entrenamiento y prueba (70%, n = 70) y otra de validación (30%, n = 29). En la cohorte de entrenamiento y prueba se realizaron 10 iteraciones de validación cruzada estratificada (se entrenó al modelo con el 90% de los casos de la cohorte y se realizó la predicción con el 10% restante de la muestra). Esto se llevó a cabo 10 veces y se calculó la media de los valores de área bajo la curva (AUC) de las 10 predicciones. Una vez entrenados los algoritmos en la cohorte de entrenamiento y prueba se realizaron las predicciones con la cohorte de validación.

Tabla 1 Características demográficas

<i>Edad (años), mediana (RIQ)</i>	78 (66-84)
<i>Sexo (masculino)</i>	59 (56,2)
<i>Antecedentes, n (%)</i>	
Alcohol	7 (6,7)
Tabaco	7 (6,7)
HTA	66 (62,9)
Dislipemia	41 (39)
DM	23 (21,9)
FA	22 (21)
Cardiopatía isquémica	8 (7,6)
HIE previo	5 (4,8)
Ictus previo	14 (13,3)
Antiagregantes	26 (24,8)
Anticoagulantes	25 (23,8)
<i>Variables al ingreso</i>	
Glucemia, mediana (RIQ), Mmol/l	138 (114-173)
PAS, mediana (RIQ), mmHG	159 (141-188)
PAD, mediana (RIQ), mmHG	81 (67-100)
NIHSS inicial, mediana (RIQ)	15 (5-21)
<i>mRS al alta, n (%)</i>	
mRS 0-2	16 (15,2)

Tabla 2 Características de imagen

	Total, n (%)
<i>Localización</i>	
Lobar	47 (44,8)
Profunda	47 (44,8)
Cerebelo	8 (7,6)
Tronco encefálico	3 (2,9)
<i>Volumen (ml), mediana (RIQ)</i>	17 (8-47)
<i>Extensión ventricular</i>	47 (44,8)

Se calcularon la sensibilidad y el valor predictivo negativo de los clasificadores en la cohorte de validación.

La AUC media de los diferentes métodos de selección de variables y clasificadores con la cohorte de entrenamiento y prueba se resume en la tabla 3. KNN-E, P-SVM y RF-10, en combinación con el método de selección de características de correlación ANOVA, fueron los clasificadores de

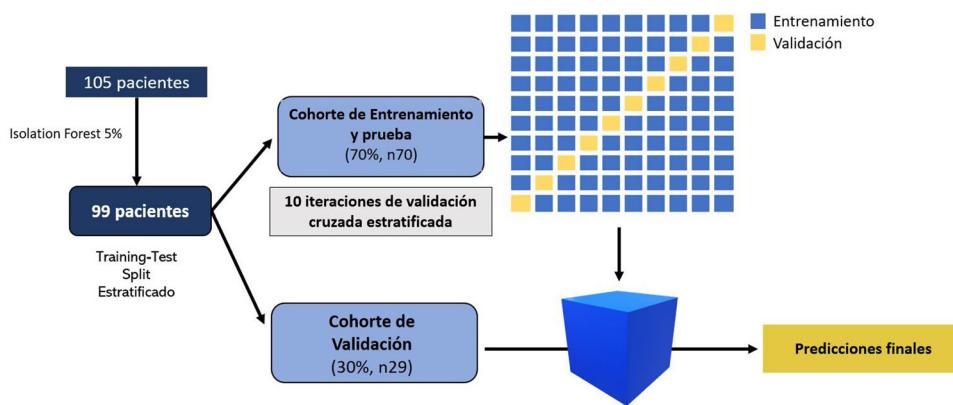


Figura 4 Procedimiento de entrenamiento y validación. Un total de 105 pacientes cumplieron los criterios de inclusión. Se excluyeron pacientes con valores perdidos y se aplicó el algoritmo «Isolation Forest» con el que se eliminaron el 5% de los pacientes con valores extremos. Tras el procesamiento inicial, se analizaron finalmente 99 pacientes. La muestra se dividió en dos cohortes estratificadas de pacientes: una cohorte de entrenamiento y prueba (70%, n=70) y otra de validación (30%, n=29). En la cohorte de entrenamiento y prueba se realizaron 10 iteraciones de validación cruzada estratificada. Una vez entrenados los algoritmos en la cohorte de entrenamiento y prueba, se realizaron las predicciones con la cohorte de validación.

Tabla 3 Media AUC de clasificadores tras 10 iteraciones de validación cruzada estratificada en cohorte entrenamiento y prueba

AUC media	Clasificadores							
	KNN-E	KNN-M	P-SVM	R-SVM	S-SVM	RF-10	RF-50	GB
Métodos de selección de variables								
Sin seleccionar	0,554	0,472	0,693	0,575	0,330	0,514	0,607	0,709
ANOVA + SPERAMAN (7)	0,752	0,629	0,798	0,636	0,522	0,742	0,715	0,690
PCA-80 (23)	0,558	0,466	0,487	0,462	0,425	0,572	0,536	0,601
PCA-90 (34)	0,494	0,528	0,384	0,419	0,512	0,548	0,581	0,505
tSNE-7	0,600	0,590	0,462	0,288	0,328	0,574	0,638	0,596
Isomap-7	0,585	0,475	0,530	0,399	0,411	0,355	0,311	0,268
LLE-7	0,462	0,491	0,433	0,468	0,482	0,486	0,482	0,513

GB: gradient boosting CatBoost; Isomap-7: mapeo de características isométricas; KNN: K vecinos más cercanos; KNN-E: distancia euclídea; KNN-M: distancia Manhattan; LLE-7: incrustación localmente lineal; PCA: principal component analysis; P-SVM: Kernel polynomial; RF: Random forest; R-SVM: Kernel radial; S-SVM: Kernel sigmoidal; SVM: Support vector machine; tSNE: incrustación de vecinos estocásticos distribuidos en t.

mejor rendimiento en la cohorte de entrenamiento y prueba (AUC de 0,752, 0,798 y 0,742, respectivamente). Las siete variables de radiómica seleccionadas con el método ANOVA fueron: «Run Length Non-Uniformity», «Gray Level Non-Uniformity», «High Gray Level Emphasis», «Run Entropy», «Busyness», «Long Run Emphasis», «Interquartile Range».

Rendimiento del modelo

Después del entrenamiento de los diferentes modelos, los clasificadores se evaluaron con los datos de la cohorte de validación (n=29) para predecir mal pronóstico funcional del paciente al alta, comparando la predicción de los clasificadores con el pronóstico funcional que tuvieron dichos pacientes. La sensibilidad de los clasificadores en la cohorte de validación se detalla en la tabla 4. KNN-E, KNN-M, P-SVM y RF-10 en combinación con el método de selección de características de correlación ANOVA fueron los clasificadores con mejor rendimiento en la cohorte de validación, la sensibilidad fue de 0,897 (IC 95%: 0,778-1), con 0 falsos negativos,

un valor predictivo positivo del 89% y un valor predictivo negativo del 100%. El clasificador RF-10 en combinación con el método de selección de variables Isomap-7 logró la misma tasa de sensibilidad. En la tabla 5 se muestra la matriz de confusión de los cinco modelos con mejores resultados de sensibilidad.

Discusión

En este estudio retrospectivo se han desarrollado diferentes modelos de aprendizaje supervisado no lineales basados en radiómica para predecir el pronóstico funcional al alta en pacientes con HIE. Los algoritmos P-SVM, KNN-E y RF-10, en combinación con el método de selección de variables ANOVA, fueron los clasificadores con mejor rendimiento en la cohorte de entrenamiento y prueba (AUC de 0,798, 0,752 y 0,742, respectivamente). Las predicciones de estos modelos, en la cohorte de validación, tuvieron una sensibilidad de 0,897 (IC 95%: 0,778-1), con una tasa de falsos negativos del 0% para la predicción de mal pronóstico

Tabla 4 Sensibilidad de los clasificadores en la cohorte de validación

AUC media	Clasificadores							
	KNN-E	KNN-M	P-SVM	R-SVM	S-SVM	RF-10	RF-50	GB
Métodos de selección de variables								
Sin seleccionar	0,862	0,862	0,862	0,862	0,862	0,862	0,828	0,862
ANOVA + SPERAMAN (7)	0,897	0,897	0,897	0,862	0,828	0,897	0,862	0,793
PCA-80 (23)	0,862	0,862	0,828	0,862	0,828	0,862	0,828	0,862
PCA-90 (34)	0,862	0,862	0,828	0,862	0,862	0,828	0,862	0,828
tSNE-7	0,828	0,793	0,793	0,862	0,828	0,862	0,862	0,828
Isomap-7	0,862	0,862	0,862	0,862	0,862	0,897	0,862	0,828
LLE-7	0,862	0,862	0,793	0,862	0,862	0,862	0,862	0,828

GB: *gradient boosting CatBoost*; Isomap-7: mapeo de características isométricas; KNN: K vecinos más cercanos; KNN-E: distancia euclídea; KNN-M: distancia Manhattan; LLE-7: incrustación localmente lineal; PCA: *principal component analysis*; P-SVM: *Kernel polynomial*; RF: *Random forest*; R-SVM: *Kernel radial*; S-SVM: *Kernel sigmoidal*; SVM: *Support vector machine*; tSNE: incrustación de vecinos estocásticos distribuidos en t.

Tabla 5 Matriz de confusión de los cinco modelos con mejores resultados de sensibilidad en la cohorte de validación

		PREDICCIÓN		Total
		Buen pronóstico	Mal pronóstico	
Realidad	Buen pronóstico	1	3	4
	Mal pronóstico	0	25	25
Total		1	28	29

funcional al alta. Identificar a los pacientes con riesgo de mal pronóstico funcional permite una mejor estratificación para ofrecer terapias intensivas adaptadas a estos pacientes.

La principal fortaleza de este análisis es que es el primer estudio basado en algoritmos de aprendizaje supervisado no lineales en el que el objetivo principal es predecir el pronóstico funcional al alta en pacientes con HIE.

Existen dos estudios previamente publicados basados en algoritmos de aprendizaje supervisado no lineales, en los que el objetivo fue la predicción de crecimiento del HIE. El primero fue el publicado en el año 2019 por el equipo de Hui Li et al.³⁹. En él se investigó si los valores de radiómica de la TCCSC podrían predecir el crecimiento del HIE. Tras el proceso de selección de datos, analizaron 4 variables de radiómica para construir el modelo y aplicaron 23 algoritmos de aprendizaje supervisado. El que obtuvo mejor rendimiento para predecir crecimiento del HIE fue «Linear Support Vector Classifier» ([tabla 6](#)).

El segundo estudio, desarrollado por el equipo de Song³⁰, tuvo como objetivo determinar si los modelos de TCCSC basados en valores de radiómica y algoritmos de aprendizaje supervisado podrían mejorar la predicción de la expansión temprana del hematoma en pacientes con HIE. Construyeron varios modelos para predecir el crecimiento del HIE: modelo radiológico, modelo radiómico, modelo clínico-radiológico, modelo radiológico-radiómico y modelo combinado. En sus resultados observaron que el modelo radiómico (en concreto, el algoritmo de regresión logística) demostró un mejor rendimiento y mayor sensibilidad que el modelo clínico-radiológico y que el modelo radiológico ([tabla 6](#)).

Otra de las fortalezas de este estudio es que, pese a su naturaleza retrospectiva, el protocolo de adquisición y de reconstrucción de las imágenes está estandarizado y no existe variabilidad. Debido a la naturaleza retrospectiva de la mayoría de los estudios de radiómica, los protocolos de imagen, incluyendo la adquisición, y los ajustes de reconstrucción a menudo no están controlados o estandarizados.

Múltiples investigaciones han evaluado el impacto de estos parámetros (voltaje, miliamperaje, *pitch*, *field of view*, grosor de corte, adquisición, marca, movimiento...) y su influencia sobre las variables de radiómica. En 2016 Lu et al.⁴⁰ evaluaron la concordancia en los valores de radiómica al variar parámetros de grosor de corte y el algoritmo de reconstrucción de la TCCSC y llegaron a la conclusión de que existe variabilidad entre las características de radiómica cuando se calculan a partir de imágenes de TCCSC reconstruidas con diferentes algoritmos y grosores de corte, destacando la importancia de estandarizar la adquisición de las imágenes. En el presente análisis se han utilizado dos equipos diferentes (Somatom Definition Flash, Siemens Healthcare y Somatom Sensation 64, Siemens Healthcare) de la misma marca; en investigaciones futuras habría que analizar si al utilizar marcas diferentes existe variabilidad en los resultados de las características de radiómica.

Una de las limitaciones es que, tal y como se ha mencionado en la introducción, el HIE es una de las causas principales de incapacidad, por lo que un alto porcentaje de pacientes incluidos en este estudio (84,8%) tuvieron un mal pronóstico funcional al alta (mRS 3-6). Esto hace que

Tabla 6 Resumen de los resultados de las cohortes de validación interna de los trabajos de radiómica, cuyo objetivo fue la predicción de crecimiento y pronóstico del hematoma intracerebral espontáneo (HIE), que han sido mencionados en la discusión del manuscrito

Variable resultado	Estudio	Método	Modelo	Resultados
Crecimiento HIE	Hui Li et al. ³⁹	Algoritmos de aprendizaje supervisado no lineales	Radiómica: <i>Linear Support Vector Classifier</i>	S 0,726; E 0,717; AUC 0,729
	Song et al. ³⁰	Algoritmos de aprendizaje supervisado no lineales	Radiológico: <i>Black Hole</i> Clínico-radiológico	S 0,367; E 0,853; AUC 0,610 S 0,645; E 0,775; AUC 0,766
			Radiómica: <i>Logistic regression</i> Radiómica + radiológico	S 0,761; E 0,818; AUC 0,850 S 0,795; E 0,879; AUC 0,867
			Combinado (radiómica + radiológico + clínico)	S 0,804; E 0,881; AUC 0,867
Pronóstico + crecimiento HIE	Pszczolkowski et al. ⁴¹	Modelos lineales generalizados	Radiómica	Pronóstico S 0,698; E 0,741; AUC 0,783
			Radiológico	S 0,318; E 0,880; AUC 0,621
			Radiómica + radiológico	S 0,698; E 0,741; AUC 0,783
			Clínico	S 0,620; E 0,815; AUC 0,789
			Radiómica + clínico	S 0,694; E 0,826; AUC 0,818
Pronóstico HIE	Huang et al. ⁴²	Modelos lineales generalizados	Radiómica	S 0,705; E 0,725; AUC 0,773
			Clínico	S 0,767; E 0,725; AUC 0,828
			Radiómica + clínico	S 0,775; E 0,739; AUC 0,844
	Nuestros resultados	Algoritmos de aprendizaje supervisado no lineales	Radiómica: P-SVM, KNN-E y RF-10	S 0,897 (IC 95%: 0,778-1)

AUC: área bajo la curva; E: especificidad; KNN-E: distancia euclíadiana; P-SVM: Kernel polinomial; RF: *Random forest*; S: sensibilidad.

los grupos de la cohorte de entrenamiento y prueba y de la cohorte de validación no estén balanceados.

El HIE es una patología con elevada morbimortalidad, lo que hace que la probabilidad pretest de mal pronóstico funcional al alta en estos pacientes sea elevada. La probabilidad pretest de mal pronóstico funcional tras un HIE es del 0,86 en la muestra global, y tras aplicar los clasificadores de aprendizaje supervisado basados en radiómica, obtenemos una sensibilidad del 0,89 en la cohorte de validación. Esta es la principal limitación del estudio, ya que la probabilidad de predecir el pronóstico no aumenta de manera sustancial una vez aplicado el modelo. El rendimiento del modelo podría mejorar aumentando la muestra para conseguir cohortes mejor balanceadas.

Según los resultados publicados por los equipos de Pszczolkowski⁴¹ y Huang⁴², otra de las medidas que se podría aplicar para mejorar el rendimiento de nuestro modelo sería crear modelos combinados en los que incorporar, al modelo de radiómica, información sobre los factores demográficos y clínicos.

En el estudio del grupo de Pszczolkowski⁴¹ se evaluó el rendimiento predictivo de las variables basadas en radiómica de la TCCSC para predecir no solo la expansión del HIE, sino también el mal pronóstico funcional utilizando modelos lineales generalizados. Investigaron también el rendimiento predictivo de los signos radiológicos y los factores clínicos de forma independiente y en combinación con las variables basadas en radiómica. Llegaron a la conclusión de que los modelos que utilizan variables basadas en radiómica de la TCCSC superan a los modelos individuales que utilizan signos radiológicos o factores clínicos aislados. Además, observaron que los modelos combinados, en los que se incorporaban factores demográficos y clínicos al modelo de radiómica, mejoraban la predicción de mal pronóstico en los pacientes con HIE (**tabla 6**).

De igual forma, en 2022, el equipo de Huang⁴² evaluó el rendimiento predictivo de las variables basadas en radiómica de la TCCSC del HIE y del edema perihematoma, y desarrollaron varios modelos basados en las características de radiómica y clínicas para predecir el pronóstico funcional a los tres meses, utilizando modelos lineales generalizados. Demostraron que el modelo conjunto de radiómica y clínico obtuvo un mejor rendimiento y una mayor sensibilidad en la predicción de mal pronóstico tanto en la cohorte de entrenamiento y prueba como en la de validación interna y de validación externa) (**tabla 6**).

Por último, otra de las limitaciones del estudio radica en que, aunque la segmentación del HIE en la TCCSC ha sido realizada por un radiólogo especialista, no se ha analizado la robustez de los valores de radiómica con varias segmentaciones. La segmentación manual y semiautomática introduce el sesgo de observador, y los estudios han demostrado que muchas variables de radiómica no son robustas frente a las variaciones intra e interobservador relativas a la delineación del ROI/VOI⁴³. En consecuencia, habría que realizar evaluaciones de la reproducibilidad intra e interobservador de las variables de radiómica derivadas y excluir las variables no reproducibles. A pesar de ello, es el método de segmentación utilizado previamente en la literatura⁴⁴, y además consideramos que el gran contraste que existe en la TCCSC entre el HIE (hiperdenso) y el resto de las estructuras adyacentes facilita la identificación y la segmentación

de este, por lo que no creemos que invalide nuestros resultados.

En conclusión, el uso de clasificadores de aprendizaje supervisado no lineales basados en radiómica con métodos de aprendizaje automático son una herramienta de diagnóstico prometedora para predecir el resultado funcional al alta en pacientes con HIE, con una baja tasa de falsos negativos, aunque todavía son necesarios estudios con mayor tamaño muestral, balanceados y en los que se combinen características de radiómica y clínicas para desarrollar y mejorar su rendimiento.

Financiación

Este trabajo no ha recibido ninguna financiación externa.

Autoría

1. Responsable de la integridad del estudio: ES, JM, LL, AR, CZ, SA, LO, ALR.
2. Concepción del estudio: ES, ALR, SA, LO.
3. Diseño del estudio: ES, ALR, SA, LO.
4. Obtención de los datos: ES, JM, LL, AR, CZ, SA, LO, ALR.
5. Análisis e interpretación de los datos: ES, JM, LL, AR, CZ, SA, LO, ALR.
6. Tratamiento estadístico: ALR.
7. Búsqueda bibliográfica: ES, ALR.
8. Redacción del trabajo: ES, ALR.
9. Revisión crítica del manuscrito con aportaciones intelectualmente relevantes: ES, JM, LL, AR, CZ, SA, LO, ALR.
10. Aprobación de la versión final: ES, JM, LL, AR, CZ, SA, LO, ALR.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Bibliografía

1. Collaborators GS. Global, regional, and national burden of stroke and its risk factors, 1990-2019: A systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol. 2021;20:795-820, [http://dx.doi.org/10.1016/S1474-4422\(21\)00252-0](http://dx.doi.org/10.1016/S1474-4422(21)00252-0).
2. Ikram MA, Wieberdink RG, Koudstaal PJ. International epidemiology of intracerebral hemorrhage. Curr Atheroscler Rep. 2012;14:300-6, <http://dx.doi.org/10.1007/s11883-012-0252-1>.
3. An SJ, Kim TJ, Yoon BW. Epidemiology, risk factors, and clinical features of intracerebral hemorrhage: An update. J Stroke. 2017;19:3-10, <http://dx.doi.org/10.5853/jos.2016.00864>.
4. Chen S, Zhao B, Wang W, Shi L, Reis C, Zhang J. Predictors of hematoma expansion predictors after intracerebral hemorrhage. Oncotarget. 2017;8:89348-63, <http://dx.doi.org/10.18632/oncotarget.19366>.
5. Hemphill JC, Greenberg SM, Anderson CS, Becker K, Bendok BR, Cushman M, et al. Guidelines for the management of spontaneous intracerebral hemorrhage: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2015;46:2032-60, <http://dx.doi.org/10.1161/STR.0000000000000069>.

6. Li Q, Liu QJ, Yang WS, Wang XC, Zhao LB, Xiong X, et al. Island Sign: An imaging predictor for early hematoma expansion and poor outcome in patients with intracerebral hemorrhage. *Stroke.* 2017;48:3019–25, <http://dx.doi.org/10.1161/STROKEAHA.117.017985>.
7. Selariu E, Zia E, Brizzi M, Abul-Kasim K. Swirl sign in intracerebral haemorrhage: Definition, prevalence, reliability and prognostic value. *BMC Neurol.* 2012;12:109, <http://dx.doi.org/10.1186/1471-2377-12-109>.
8. Li Q, Zhang G, Xiong X, Wang XC, Yang WS, Li KW, et al. Black Hole Sign: Novel imaging marker that predicts hematoma growth in patients with intracerebral hemorrhage. *Stroke.* 2016;47:1777–81, <http://dx.doi.org/10.1161/STROKEAHA.116.013186>.
9. Li Q, Yang WS, Wang XC, Cao D, Zhu D, LV FJ, et al. Blend sign predicts poor outcome in patients with intracerebral hemorrhage. *PLoS One.* 2017;12:e0183082, <http://dx.doi.org/10.1371/journal.pone.0183082>.
10. Boulouis G, Morotti A, Brouwers HB, Charidimou A, Jessel MJ, Auriel E, et al. Noncontrast computed tomography hypodensities predict poor outcome in intracerebral hemorrhage patients. *Stroke.* 2016;47:2511–6, <http://dx.doi.org/10.1161/STROKEAHA.116.014425>.
11. Boulouis G, Morotti A, Charidimou A, Dowlatshahi D, Goldstein JN. Noncontrast computed tomography markers of intracerebral hemorrhage expansion. *Stroke.* 2017;48:1120–5, <http://dx.doi.org/10.1161/STROKEAHA.116.015062>.
12. Serrano E, López-Rueda A, Moreno J, Rodriguez A, Llull L, Zwanzger C, et al. The new Hematoma Maturity Score is highly associated with poor clinical outcome in spontaneous intracerebral hemorrhage. *Eur Radiol.* 2022;32:290–9, <http://dx.doi.org/10.1007/s00330-021-08085-4>.
13. Morotti A, Boulouis G, Dowlatshahi D, Li Q, Barras CD, Delcourt C, et al. Standards for detecting, interpreting, and reporting noncontrast computed tomographic markers of intracerebral hemorrhage expansion. *Ann Neurol.* 2019;86:480–92, <http://dx.doi.org/10.1002/ana.25563>.
14. Quintas-Neves M, Marques L, Silva L, Amorim JM, Ferreira C, Pinho J. Noncontrast computed tomography markers of outcome in intracerebral hemorrhage patients. *Neurology Res.* 2019;41:1083–9, <http://dx.doi.org/10.1080/01616412.2019.1673279>.
15. Law ZK, Ali A, Krishnan K, Bischoff A, Appleton JP, Scutt P, et al. Noncontrast computed tomography signs as predictors of hematoma expansion, clinical outcome, and response to tranexamic acid in acute intracerebral hemorrhage. *Stroke.* 2020;51:121–8, <http://dx.doi.org/10.1161/STROKEAHA.119.026128>.
16. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* 2012;48:441–6, <http://dx.doi.org/10.1016/j.ejca.2011.11.036>.
17. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. *Radiology.* 2016;278:563–77, <http://dx.doi.org/10.1148/radiol.2015151169>.
18. Brink JA, Arenson RL, Grist TM, Lewin JS, Enzmann D. Bits and bytes: The future of radiology lies in informatics and information technology. *Eur Radiol.* 2017;27:3647–51, <http://dx.doi.org/10.1007/s00330-016-4688-5>.
19. Mongan J, Moy L, Kahn CE. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A guide for authors and reviewers. *Radiol Artif Intell.* 2020;2:e200029, <http://dx.doi.org/10.1148/ryai.2020200029>.
20. Kothari RU, Brott T, Broderick JP, Barsan WG, Sauerbeck LR, Zuccarello M, et al. The ABCs of measuring intracerebral hemorrhage volumes. *Stroke.* 1996;27:1304–5, <http://dx.doi.org/10.1161/01.str.27.8.1304>.
21. Van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 2017;77:e104–7, <http://dx.doi.org/10.1158/0008-5472.CAN-17-0339>.
22. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andreanczyk V, Apté A, et al. The Image Biomarker Standardization Initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology.* 2020;295:328–38, <http://dx.doi.org/10.1148/radiol.2020191145>.
23. Quinn TJ, Dawson J, Walters MR, Lees KR. Reliability of the modified Rankin Scale: A systematic review. *Stroke.* 2009;40:3393–5, <http://dx.doi.org/10.1161/STROKEAHA.109.557256>.
24. Mandava P, Krumpelman CS, Shah JN, White DL, Kent TA. Quantification of errors in ordinal outcome scales using Shannon entropy: Effect on sample size calculations. *PLoS One.* 2013;8:e67754, <http://dx.doi.org/10.1371/journal.pone.0067754>.
25. Hall MA, Smith LA. Feature selection for machine learning: Comparing a correlation-based filter approach to the wrapper. 1999.
26. Sarker IH. Machine learning: Algorithms, real-world applications and research directions. *SN Comput Sci.* 2021;2:160, <http://dx.doi.org/10.1007/s42979-021-00592-x>.
27. Van der Maaten L, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res.* 2008;9:2579–605.
28. Tenenbaum JB, de Silva V, Langford JC. A global geometric framework for nonlinear dimensionality reduction. *Science.* 2000;290:2319–23, <http://dx.doi.org/10.1126/science.290.5500.2319>.
29. Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. *Science.* 2000;290:2323–6, <http://dx.doi.org/10.1126/science.290.5500.2323>.
30. Song Z, Guo D, Tang Z, Liu H, Li X, Luo S, et al. Non-contrast computed tomography-based radiomics analysis in discriminating early hematoma expansion after spontaneous intracerebral hemorrhage. *Korean J Radiol.* 2021;22:415–24, <http://dx.doi.org/10.3348/kjr.2020.0254>.
31. Mayerhoefer ME, Materka A, Langs G, Läggström I, Szczypinski P, Gibbs P, et al. Introduction to radiomics. *J Nucl Med.* 2020;61:488–95, <http://dx.doi.org/10.2967/jnumed.118.222893>.
32. Al-Mufti F, Thabet AM, Singh T, el-Ghanem M, Amuluru K, Gandhi CD. Clinical and radiographic predictors of intracerebral hemorrhage outcome. *Interv Neurol.* 2018;7:118–36, <http://dx.doi.org/10.1159/000484571>.
33. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. *Radiographics.* 2017;37:505–15, <http://dx.doi.org/10.1148/rg.2017160130>.
34. Zhou CY, Chen YQ. Improving nearest neighbor classification with cam weighted distance. *Pattern Recognition.* 2006;39:635–45 <https://doi.org/10.1016/j.patcog.2005.09.004>
35. Burges CJ. A tutorial on support vector machines for pattern recognition. *Data Min Knowl Discov.* 1998;2:121–67, <http://dx.doi.org/10.1023/A:1009715923555>.
36. Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other Kernel-based learning methods. Cambridge University Press; 2000, <http://dx.doi.org/10.1017/CBO9780511801389>.
37. Breiman L. Random Forests. *Machine Learning.* 2001;45:5–32, <http://dx.doi.org/10.1023/A:1010933404324>.
38. Dorogush AV, Ershov V, Gulin A. CatBoost: Gradient boosting with categorical features support. *ArXiv.* 2018, <abs/1810.11363>.
39. Li H, Xie Y, Wang X, Chen F, Sun J, Jiang X. Radiomics features on non-contrast computed tomography predict early enlargement of spontaneous intracerebral hemorrhage.

- Clin Neurol Neurosurg. 2019;185:105491, <http://dx.doi.org/10.1016/j.clineuro.2019.105491>.
40. Lu L, Ehmke RC, Schwartz LH, Zhao B. Assessing agreement between radiomic features computed for multiple CT imaging settings. PLoS One. 2016;11:e0166550, <http://dx.doi.org/10.1371/journal.pone.0166550>.
41. Pszczolkowski S, Manzano-Patrón JP, Law ZK, Krishnan K, Ali A, Bath PM, et al. Quantitative CT radiomics-based models for prediction of haematoma expansion and poor functional outcome in primary intracerebral haemorrhage. Eur Radiol. 2021;31:7945–59, <http://dx.doi.org/10.1007/s00330-021-07826-9>.
42. Huang X, Wang D, Zhang Q, Ma Y, Zhao H, Li S, et al. Radiomics for prediction of intracerebral hemorrhage outcomes: A retrospective multicenter study. Neuroimage Clin. 2022;36:103242, <http://dx.doi.org/10.1016/j.nicl.2022.103242>.
43. Van Timmeren JE, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B. Radiomics in medical imaging – «how-to» guide and critical reflection. Insights Imaging. 2020;11:91, <http://dx.doi.org/10.1186/s13244-020-00887-2>.
44. Haider SP, Qureshi AI, Jain A, Tharmaseelan H, Berson ER, Zeevi T, et al. Admission computed tomography radiomic signatures outperform hematoma volume in predicting baseline clinical severity and functional outcome in the ATACH-2 trial intracerebral hemorrhage population. Eur J Neurol. 2021;28:2989–3000, <http://dx.doi.org/10.1111/ene.15000>.