



UNIVERSIDAD DE CHILE
FACULTAD DE MEDICINA
ESCUELA DE POSTGRADO

Cuantificación Volumétrica de la Carga Tumoral en Imágenes TC y su Impacto en la Clasificación de la Respuesta al Tratamiento según el Protocolo RECIST 1.1

Felipe Andrés Bravo Bahamóndez

Tesis para optar al grado de Magíster en Informática Médica.

Director de Tesis: **Prof. Dr. Steffen Härtel**
Codirectora: **MSc. Constanza Vásquez**

Santiago, Julio de 2025

Índice

1. Resumen	2
2. Abstract	3
3. Introducción	4
3.1. El cáncer en Chile y el mundo	4
3.2. Criterios de evaluación de respuesta al tratamiento en tumores sólidos	4
3.2.1. Criterios unidimensionales (1D) y bidimensionales (2D)	4
3.2.2. Criterios adaptados a la evaluación tridimensional (3D)	16
3.2.3. Problemática, brecha de investigación y aporte al conocimiento	19
4. Hipótesis	21
5. Objetivo General	22
6. Objetivos Específicos	22
7. Materiales y Métodos	22
7.1. Métricas	24
7.2. Estrategia de búsqueda y selección de artículos	26
7.3. Recolección de datos	29
7.4. Anonimización de datos	30
7.5. Anotación y etiquetado de datos	31
7.6. Medición de diámetros de tumores en imágenes de TC con OpenCV .	32
7.6.1. <code>cv2.fitEllipse</code> : Ajuste elíptico por mínimos cuadrados . . .	33
7.6.2. <code>cv2.minAreaRect()</code> : Ajuste de rectángulo mínimo rotado so-	
bre lesiones segmentadas	35
7.6.3. <code>cv2.convexHull()</code> : Envoltente convexa y diámetro de Feret .	40
7.6.4. Conversión de unidades y formato NIfTI	41
7.6.5. Generación y validación de lesiones tumorales sintéticas con	
OpenCV	43
7.6.6. Medición de diámetros tumorales en TC con OpenCV en pla-	
nos axial, sagital y coronal	55
7.6.7. Algoritmo de Cuantificación volumétrica de la carga tumoral .	57
8. Resultados	58
8.1. Definición de umbrales volumétricos mediante regresión de LOESS . .	60
8.2. Medición de segmentaciones en 1D, 2.5D y 3D	62
8.3. Concordancia entre métodos	63
8.4. Comparaciones de cambios relativos de carga tumoral	63
9. Discusión	64
10. Conclusiones	65
Referencias	65

1. Resumen

La clasificación de la respuesta al tratamiento en pacientes oncológicos es esencial para el manejo de la enfermedad y se basa en la cuantificación de la carga tumoral. El protocolo RECIST 1.1 (Response Evaluation Criteria In Solid Tumors), que utiliza mediciones manuales de diámetros máximos en el plano axial (1D) en imágenes de Tomografía Computarizada (TC), se considera el estándar actual. No obstante, esta aproximación presenta limitaciones relevantes, como la escasa precisión y sensibilidad para calcular la carga tumoral total y sus cambios en el tiempo, además de una alta variabilidad interobservador al realizarse de manera manual, lo que podría desencadenar en clasificaciones de respuesta al tratamiento y afectar la toma de decisiones terapéuticas.

Ante estas limitaciones, han surgido nuevas metodologías de medición, como la evaluación multiplanar o en planos ortogonales (axial, sagital y coronal) (2.5D), y la cuantificación volumétrica (3D), que ha sido estudiada de manera preliminar también en el seguimiento de pacientes oncológicos.

Esta tesis evalúa la concordancia en la clasificación de respuesta al tratamiento entre el protocolo RECIST 1.1 y métodos de medición automática en 1D, 2.5D y 3D, utilizando imágenes de TC de pacientes oncológicos del Hospital Clínico de la Universidad de Chile (HCUCH). Se realizó un análisis estadístico para determinar las diferencias en la medición de los cambios en la carga tumoral entre los distintos métodos, con el objetivo de evaluar las implicancias clínicas de estas mediciones.

Se implementó el algoritmo OBB (Oriented Bounding Box) para medición automática de diámetros 1D y se realizaron segmentaciones manuales de lesiones tumorales con 3D Slicer. Mediante regresión LOESS (Locally Estimated Scatterplot Smoothing) se establecieron umbrales volumétricos de clasificación de la respuesta al tratamiento adaptados a RECIST 1.1. Posteriormente se evaluó la concordancia en la clasificación de respuesta al tratamiento y la significancia estadística entre los métodos de medición.

Las mediciones 3D mostraron significancia estadística en la detección de cambios en la carga tumoral durante la evaluación inicial del paciente, en comparación con la medición 1D automática ($p < 0.05$). A pesar de que el protocolo RECIST 1.1 y la medición 1D presentaron una buena concordancia ($Kappa = 0.59$), las técnicas 2.5D y 3D mostraron concordancias pobres ($Kappa = 0.34$) y débiles ($Kappa = 0.14$), respectivamente. Esto sugiere que estos métodos difieren en sus clasificaciones en relación con la medición 1D, siendo similares entre sí, lo que podría influir en la toma de decisiones clínicas.

Esta tesis resalta el potencial del análisis 2.5D y 3D como complemento a los métodos tradicionales y enfatiza la necesidad de futuros estudios para evaluar su impacto en la práctica clínica.

2. Abstract

The classification of treatment response in oncology patients is essential for disease management and is based on the quantification of tumor burden. The RECIST 1.1 protocol (Response Evaluation Criteria in Solid Tumors), which relies on manual unidimensional (1D) measurements of the maximum diameter of lesions in axial computed tomography (CT) images, is considered the current clinical standard. However, this approach presents significant limitations, such as low precision and sensitivity in estimating total tumor burden and its variations over time. In addition, the manual nature of the measurements introduces high interobserver variability, which may lead to discrepancies in response classification and affect therapeutic decision-making.

To address these limitations, new measurement methodologies have emerged, including multiplanar evaluation or orthogonal plane measurements (axial, sagittal, and coronal) (2.5D), and volumetric quantification (3D), which has also been preliminarily explored in the follow-up of oncology patients.

This thesis evaluates the concordance in treatment response classification between the RECIST 1.1 protocol and automatic measurement methods in 1D, 2.5D, and 3D, using CT images from oncology patients at the Clinical Hospital of the University of Chile (HCUCH). A statistical analysis was performed to assess the differences in tumor burden variation measured by each method, with the aim of evaluating their clinical implications.

The Oriented Bounding Box (OBB) algorithm was implemented for automatic 1D diameter measurements, and manual tumor segmentations were performed using 3D Slicer. LOESS regression (Locally Estimated Scatterplot Smoothing) was used to define volumetric response thresholds adapted to RECIST 1.1. Subsequently, response classification concordance and statistical significance between measurement methods were analyzed.

3D measurements showed statistically significant sensitivity in detecting changes in tumor burden during the initial evaluation ($p < 0.05$). While RECIST 1.1 and 1D automatic measurements showed good concordance (Kappa = 0.59), the 2.5D and 3D techniques demonstrated poor (Kappa = 0.34) and slight (Kappa = 0.14) agreement, respectively. These results suggest that 2.5D and 3D methods yield response classifications that differ from 1D, although they show internal consistency.

This study highlights the potential of 2.5D and 3D analysis as complementary tools to traditional methods and emphasizes the need for future research to assess their impact on clinical practice.

3. Introducción

3.1. El cáncer en Chile y el mundo

El cáncer persiste como una de las principales causas de morbilidad y mortalidad a nivel mundial, representando un desafío significativo para la salud pública. Según el Observatorio Global del Cáncer, en 2022 se registraron cerca de 20 millones de nuevos casos y 9.7 millones de muertes, siendo los cánceres de pulmón, mama y colorrectal los más prevalentes (1). La Agencia Internacional para la Investigación en Cáncer (IARC) proyectó a nivel mundial un aumento de 47 % de los casos para 2040 en comparación a 2020(2).

En Chile, la situación refleja esta tendencia global. Durante 2020 se registraron 54,200 nuevos casos y 28,600 fallecimientos. Los cánceres de próstata, mama, colon, estómago y pulmón representaron el 48 % de la incidencia total, mientras que los cánceres de pulmón, estómago, próstata, colon y páncreas concentraron el 46 % de la mortalidad (2). En respuesta a esta carga de enfermedad, el país ha implementado políticas como el Plan Nacional de Cáncer (3) y la Ley Nacional del Cáncer (4). No obstante, persisten desafíos operacionales significativos, incluyendo demoras en el informe de exámenes imagenológicos, componente crítico del manejo oncológico (5).

Aunque los avances en detección temprana, tratamiento y prevención han mejorado los resultados, se anticipa que la carga sobre los sistemas de salud aumentará considerablemente en las próximas décadas (6). Este escenario exige optimizar el flujo de trabajo clínico en la atención oncológica, desarrollando intervenciones eficaces que gestionen la creciente demanda y aseguren acceso sostenido a atención médica oportuna y de calidad (7).

Sin embargo, la eficiencia operativa por sí sola no es suficiente. Para que esta optimización tenga un impacto real en los resultados clínicos, es fundamental que esté acompañada de precisión diagnóstica. En este contexto, evaluar la respuesta tumoral de manera precisa y estandarizada se vuelve crucial para guiar decisiones terapéuticas, determinar la eficacia de los tratamientos y predecir el pronóstico del paciente (8). Esta evaluación sistemática resulta esencial tanto en la práctica clínica diaria como en ensayos clínicos, donde se han desarrollado diversos métodos y criterios para la medición longitudinal de la carga tumoral (9).

3.2. Criterios de evaluación de respuesta al tratamiento en tumores sólidos

3.2.1. Criterios unidimensionales (1D) y bidimensionales (2D)

Evolución histórica y fundamentos metodológicos

Los criterios de la Organización Mundial de la Salud (OMS), establecidos en 1979, representaron el primer esfuerzo por estandarizar la evaluación objetiva de la respuesta tumoral. Basados en una metodología bidimensional, calculan el producto

de los dos diámetros perpendiculares más largos de cada lesión y definen categorías específicas de respuesta, resumidas en la **Tabla 1**(10).

Tabla 1. Criterios de Respuesta Tumoral de la OMS

Categoría	Respuesta tumoral
CR	Desaparición de toda la enfermedad conocida, confirmada a las 4 semanas ¹
PR	$\geq 50\%$ de disminución respecto a la basal, confirmada a las 4 semanas
SD	No se cumplen criterios de PR ni PD
PD	$\geq 25\%$ de incremento; sin CR, PR ni SD; nueva(s) lesión(es); $\geq 25\%$ de incremento en 1 lesión

¹Intervalo mínimo entre la evaluación basal y la primera evaluación de la respuesta. CR: Respuesta Completa, PR: Respuesta Parcial, SD: Enfermedad Estable, PD: Enfermedad Progresiva. **Fuente:** Aykan y Özatlı, «*Objective Response Rate Assessment in Oncology*» (10).

Aunque representaron un avance significativo frente a la evaluación visual subjetiva, los criterios OMS presentaban limitaciones importantes. La variabilidad interobservador era considerable, el tiempo requerido para medir todas las lesiones resultaba excesivo, y existía ambigüedad en la definición de progresión tumoral (10). Estas deficiencias motivaron el desarrollo de aproximaciones más simplificadas y reproducibles.

El modelo EASL en carcinoma hepatocelular

El reconocimiento temprano de que diferentes neoplasias podrían requerir enfoques evaluativos distintos llevó al desarrollo de criterios específicos por tipo tumoral. Los criterios de la Asociación Europea para el Estudio del Hígado (EASL) surgieron como una adaptación especializada de los principios OMS para el carcinoma hepatocelular, reconociendo las características únicas de las lesiones hepáticas y su patrón de respuesta particular (10).

Si bien los criterios EASL conservaron la lógica bidimensional de la OMS, es decir, calcular el producto de los dos diámetros perpendiculares máximos de cada lesión, introdujeron posteriormente una modificación decisiva para el carcinoma hepatocelular: medir únicamente el tejido tumoral viable que realza con contraste, ignorando el componente necrótico. Esta adaptación permite cuantificar con mayor precisión la necrosis inducida por terapias locorregionales (11).

La **Figura 1** contrasta visualmente esta estrategia con la de la OMS y los criterios RECIST en su versión original y su actualización posterior. La OMS evalúa el área total del tumor; RECIST resume la suma unidimensional de los diámetros largos, incorporando límites para el número de lesiones diana y ajustes para ganglios linfáticos; y EASL se centra en la viabilidad tumoral realzada (11).

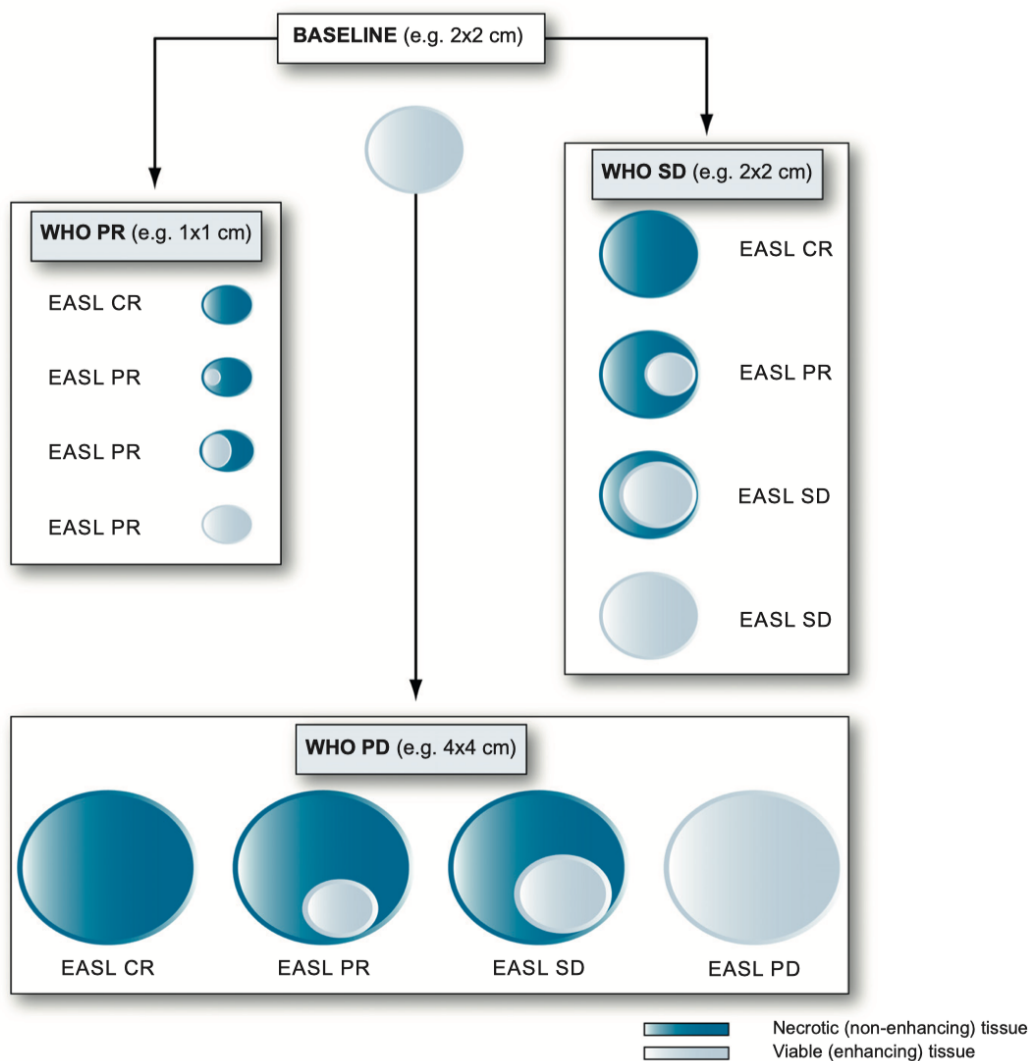


Figura 1. Representación esquemática de la discordancia entre los criterios de respuesta OMS y de la Asociación Europea para el Estudio del Hígado (EASL). A partir de una lesión basal de 2×2 cm, la OMS clasifica la respuesta según el cambio bidimensional total (círculos de igual contorno), mientras que EASL evalúa solo el tejido viable que realza con contraste (azul claro) y descarta el tejido necrótico (azul oscuro). Así, la misma lesión puede variar desde respuesta completa hasta progresión dependiendo de si se mide la masa total o únicamente la fracción viable. Fuente: Riaz et al., «Role of the EASL, RECIST, and OMS Response Guidelines Alone or in Combination for Hepatocellular Carcinoma» (12).

Aunque los cuatro estados de respuesta clínica se mantienen nominalmente iguales, cada sistema los aplica a métricas diferentes: superficie bidimensional total en la OMS, suma unidimensional en RECIST y tejido viable en EASL. Esta evolución metodológica demostró que la heterogeneidad biológica entre tumores exige criterios específicos y sentó las bases para el desarrollo posterior de guías adaptadas a terapias dirigidas y contextos oncológicos particulares (12).

Transición hacia la estandarización unidimensional: RECIST

Con el objetivo de simplificar la evaluación, los criterios RECIST 1.0, introducidos en 2000 por un grupo de trabajo internacional, revolucionaron la metodología mediante la transición a mediciones unidimensionales del diámetro mayor de hasta 10 lesiones diana (máximo 5 por órgano). La **Figura 2** ilustra claramente la diferencia metodológica fundamental: mientras que OMS requiere la obtención de dos medidas perpendiculares (largo y ancho) para calcular el área bidimensional de la lesión, RECIST 1.0 simplifica este proceso utilizando únicamente el diámetro más largo como medida unidimensional (13; 14).

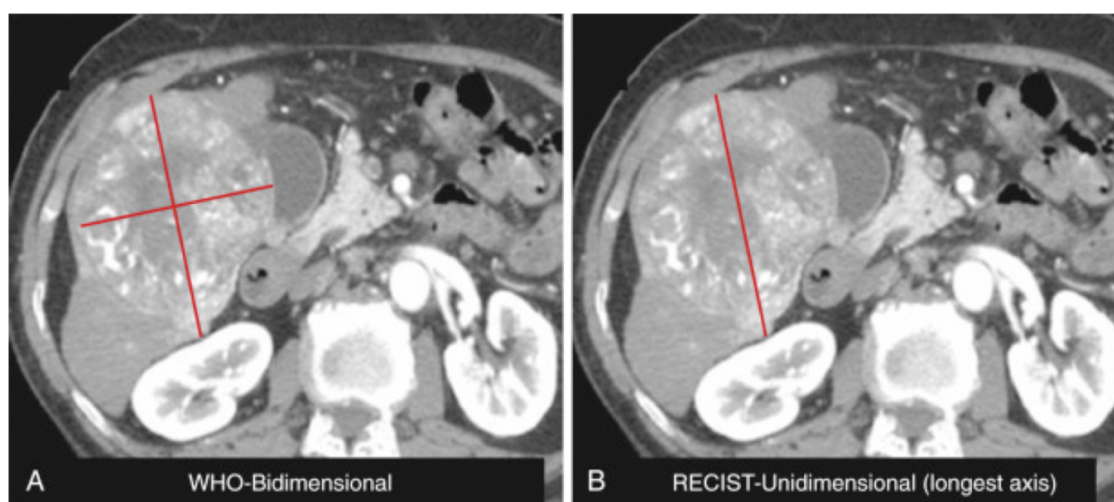


Figura 2. Medición de lesiones objetivo según los criterios OMS y RECIST. (A) Con los criterios OMS, se obtiene el diámetro tumoral más largo y una segunda diagonal que sea perpendicular y la más larga, y se multiplican (medición bidimensional). (B) Con RECIST 1.0, solo se obtiene el diámetro más largo. Fuente: Themes, «Response Evaluation Criteria in Solid Tumors, World Health Organization, and Other Response Criteria» (15).

Esta simplificación metodológica demostró menor variabilidad en los errores de medición comparado con OMS, estableciendo umbrales específicos: RP con disminución $\geq 30\%$ y EP con aumento $\geq 20\%$ del diámetro mayor, además de categorizar las lesiones como "lesiones diana" y "lesiones no diana", permitiendo un enfoque más sistemático y eficiente (10).

Optimización metodológica basada en evidencia: desarrollo de RECIST 1.1

La experiencia clínica acumulada con RECIST 1.0 reveló oportunidades significativas de refinamiento metodológico. El análisis exhaustivo de más de 6,500 pacientes y 18,000 lesiones diana proporcionó la base empírica para el desarrollo de RECIST 1.1, publicado en 2009 (20,22). Esta versión optimizada introdujo modificaciones sustanciales que mejoraron tanto la eficiencia como la precisión del sistema de evaluación (14; 16).

La **Tabla 2** resume las principales diferencias entre ambas versiones, destacando cuatro innovaciones fundamentales: (i) la reducción del número de lesiones diana de 10 a 5 totales (con un máximo de 2 por órgano), simplificando el proceso sin comprometer la representatividad; (ii) el establecimiento de criterios específicos para ganglios linfáticos patológicos, requiriendo medición del eje corto ≥ 15 mm; (iii) la incorporación formal de FDG-PET/CT (Tomografía por Emisión de Positrones/Tomografía Computarizada con Fluorodesoxiglucosa) para la detección de nuevas lesiones, reconociendo el valor de la imagen funcional; y (iv) una definición más rigurosa de progresión que exige no solo un incremento porcentual del 20% sino también un aumento absoluto mínimo de 5 mm, reduciendo así las clasificaciones erróneas por variaciones mínimas en lesiones pequeñas (16).

Tabla 2. Características principales y cambios importantes entre RECIST 1.1 y RECIST 1.0

Características	RECIST 1.1	RECIST 1.0
Método de medición	Diámetro tumoral más largo para lesiones extranodales	Diámetro tumoral más largo
Tamaño de lesión diana (lesión medible)	≥ 10 mm en TC	10 mm en TC espiral y 20 mm en TC convencional
Número de lesiones diana	Máximo de 5 (hasta 2 por órgano)	Máximo de 10 (hasta 5 por órgano)
Evaluación de ganglios linfáticos	Eje corto registrado; ganglio diana ≥ 15 mm; patológico entre 10–15 mm	No especificado
Clarificación sobre progresión	> 20 % de aumento en la suma de diámetros más largos (requiere incremento absoluto > 5 mm) y nuevas lesiones	> 20 % de aumento en la suma de diámetros (no requiere incremento absoluto) y nuevas lesiones
Modalidades de imagen	Radiografía de tórax, TC y RMN	TC, RMN y FDG-PET (PET solo para detección); radiografía de tórax no preferida

RECIST 1.1 permite monitorizar la carga tumoral durante todo el curso del tratamiento y cuantificar su evolución. Para cada lesión diana se mide el diámetro mayor en un corte axial; la única excepción son los ganglios linfáticos, donde se utiliza el eje corto. A fin de garantizar la reproducibilidad, se establecen umbrales mínimos de tamaño: tumores sólidos ≥ 10 mm en TC (≥ 20 mm en radiografía de tórax) y ganglios linfáticos con eje corto ≥ 15 mm. La **Figura 3** sintetiza estos criterios dimensionales que determinan si una lesión puede incluirse como diana (16).

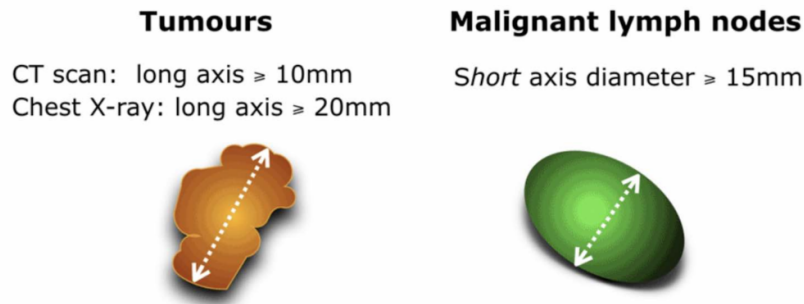


Figura 3. Criterios de tamaño mínimo para seleccionar lesiones diana según RECIST 1.1. El esquema destaca el umbral de 10 mm para tumores sólidos en TC (20 mm en radiografía de tórax) y el umbral de 15 mm en el eje corto para ganglios linfáticos. Fuente: «The Radiology Assistant: RECIST 1.1 - the basics» (17).

La carga tumoral total se calcula como la suma de los diámetros más largos de estas lesiones objetivo. La respuesta se determina comparando la suma de diámetros en los escaneos de seguimiento con la basal, clasificando la respuesta en categorías como Respuesta Completa (CR), Respuesta Parcial (PR), Enfermedad Estable (SD) o Enfermedad Progresiva (PD), basándose en cambios porcentuales predefinidos en la suma de diámetros (**Figura 4**) (17).

RECIST 1.1		Response
CR	Disappearance of all lesions and pathologic lymph nodes	
PR	$\geq 30\%$ decrease SLD no new lesions no progression of non-target lesions	
SD	no PR - no PD	
PD	$\geq 20\%$ increase SLD* compared to smallest SLD in study or progression of non-target lesions or new lesions	

Figura 4. Criterios de respuesta al tratamiento según RECIST 1.1. El diagrama presenta las cuatro categorías de respuesta al tratamiento: CR (Respuesta Completa): desaparición de todas las lesiones y ganglios linfáticos patológicos; PR (Respuesta Parcial): disminución del SLD $\geq 30\%$, sin nuevas lesiones ni progresión de lesiones no diana; SD (Enfermedad Estable): sin criterios para PR ni para PD; y PD (Enfermedad Progresiva): aumento del SLD $\geq 20\%$ comparado con el menor SLD registrado durante el estudio, progresión de lesiones no diana o aparición de nuevas lesiones. Fuente: *The Radiology Assistant: RECIST 1.1 - the basics* (17).

La **Figura 5** esquematiza el flujo operativo de RECIST 1.1. En la fase basal —realizada idealmente dentro de las cuatro semanas previas al tratamiento y con cortes tomográficos de grosor ≤ 5 mm y contraste intravenoso— se eligen las lesiones diana (tumores con diámetro mayor ≥ 10 mm o ganglios linfáticos con eje corto ≥ 15 mm), con un máximo de cinco en total y dos por órgano, y se calcula la suma

de los diámetros mayores (SLD).

Las lesiones no diana (diámetro menor a 10 mm o no medibles, como derrame pleural, ascitis o linfangitis) se consignan aparte. Durante el seguimiento se actualiza el SLD, se revisa la evolución de las lesiones no diana y se comprueba la aparición de nuevas lesiones. Con esos datos se adjudica la respuesta: completa, parcial, estable o progresiva; la detección de cualquier lesión nueva se interpreta como progresión de la enfermedad (17).

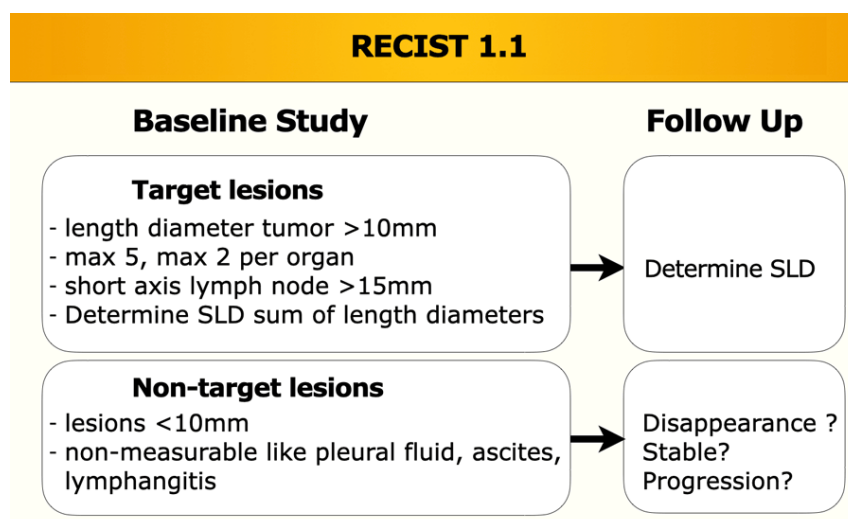


Figura 5. Diagrama de RECIST 1.1 que distingue las dos fases del protocolo. Estudio Basal (selección y medición de lesiones diana, registro de lesiones no diana) y Seguimiento (re-medición del SLD y reevaluación de todas las lesiones para clasificar la respuesta). Fuente: «The Radiology Assistant: RECIST 1.1 - the basics» (17).

Sin embargo, estos refinamientos de RECIST 1.1, aunque efectivos para quimioterapias citotóxicas convencionales, resultaron insuficientes para abordar los patrones de respuesta atípicos observados con las nuevas terapias inmunológicas (14). El reconocimiento de que los agentes inmunoterápicos pueden manifestar respuestas no convencionales, incluyendo incrementos iniciales en el tamaño tumoral debido a infiltración inflamatoria, motivó el desarrollo de criterios específicos para esta modalidad terapéutica que mantuvieran la metodología unidimensional de RECIST. Estas modificaciones consolidaron al protocolo RECIST como el estándar predominante, estableciendo la base metodológica sobre la cual se desarrollarían posteriormente adaptaciones más especializadas para terapias dirigidas e inmunoterapia (18).

Adaptaciones para terapias dirigidas: integración de criterios funcionales

El reconocimiento de las limitaciones de los criterios anatómicos convencionales en el contexto de agentes biológicos dirigidos motivó el desarrollo de criterios modificados que integraron parámetros funcionales. Los criterios de Choi (2007), validados prospectivamente para tumores del estroma gastrointestinal tratados con imatinib, demostraron superioridad significativa sobre RECIST al incorporar cam-

bios en la densidad tumoral medida en Unidades Hounsfield (disminución $\geq 15\%$ o reducción $\geq 10\%$ del diámetro) (10), alcanzando una sensibilidad del 97 % y una especificidad del 100 % (19). La **Figura 6** ilustra la aplicación práctica de dichos criterios, mostrando casos en los que la evaluación por densidad identifica respuesta (pseudoprogresión) cuando las mediciones dimensionales permanecen estables, correlacionándose mejor con el tiempo hasta progresión y la supervivencia específica de la enfermedad (18).



Figura 6. Pseudoprogresión de metástasis hepática en un hombre de 65 años con tumor estromal gastrointestinal avanzado tratado con imatinib. (a) La tomografía computarizada con contraste previa al tratamiento mostró una metástasis hepática con realce en anillo (flecha) (40 UH). A los 7 meses, las imágenes de tomografía computarizada revelaron que la masa (flecha) había aumentado de tamaño (b), pero las imágenes fusionadas correspondientes de tomografía por emisión de positrones–tomografía computarizada no mostraron captación apreciable de 2-fluoro-2-desoxi-D-glucosa (flecha) (c). Fuente: Choi, «Role of Imaging in Response Assessment and Individualised Treatment for Sarcomas» (18)

Desarrollo de criterios evaluativos ante patrones atípicos de respuesta inmunológica

La introducción de inmunoterapias en oncología ha evidenciado patrones de respuesta tumoral atípicos, como la *pseudoprogresión*, caracterizada por un aumento transitorio del volumen tumoral debido a infiltración linfocitaria e inflamación, lo que motivó el desarrollo de criterios evaluativos específicos. En este contexto, Wolchok *et al.* propusieron en 2009 los *immune-related response criteria* (irRC), que retomaron el enfoque bidimensional de la OMS para medir lesiones mediante el producto de diámetros perpendiculares, en contraste con el modelo unidimensional de RECIST 1.1 basado en el eje mayor de las lesiones (10). Esta decisión metodológica buscó mejorar la sensibilidad frente a patrones inmunológicos, aunque mantuvo limitaciones inherentes a la evaluación bidimensional, como la menor reproducibilidad.

Revalidación metodológica unidimensional y criterios de respuesta: sustento empírico y operacionalización

La evidencia aportada por Nishino *et al.* en 2013 constituyó un punto de inflexión metodológico clave en la evaluación de respuesta tumoral bajo inmunoterapia. Su análisis de 57 pacientes con melanoma avanzado demostró que las mediciones

unidimensionales presentaban una superioridad significativa en términos de concordancia interobservador y reproducibilidad, en comparación con los enfoques bidimensionales tradicionales (20). Este hallazgo cuestionó la reintroducción del modelo bidimensional propuesto por irRC, particularmente por sus limitaciones operativas en contextos clínicos de alta variabilidad. A partir de esta evidencia, Bohnsack et al. propusieron en 2014 el marco metodológico de irRECIST, el cual mantuvo plena compatibilidad con RECIST 1.1 en cuanto a la estructura de medición unidimensional, pero incorporó modificaciones específicas orientadas a capturar los patrones de respuesta propios de la inmunoterapia (21).

Una de las innovaciones centrales de *irRECIST* fue la incorporación del concepto de carga tumoral total medida (*Total Measured Tumor Burden*, TMTB), con umbrales cuantitativos definidos para cada categoría de respuesta: respuesta completa inmuno-relacionada (*irCR*), como la desaparición de todas las lesiones sin necesidad de confirmación temporal; respuesta parcial (*irPR*), como una reducción $\geq 30\%$ en el TMTB respecto al valor basal; y progresión inmuno-relacionada (*irPD*), definida como un incremento $\geq 20\%$ desde el nadir, con un aumento absoluto mínimo de 5 mm, requiriendo en este caso confirmación mediante estudios repetidos al menos 4 semanas después. Esta asimetría en los requisitos de confirmación, exigida para progresión pero no para respuesta, representa un avance metodológico orientado a evitar la clasificación errónea de pseudoprogresiones (18).

La **Tabla 3** sintetiza y compara los criterios de respuesta establecidos por las principales guías clínicas (RECIST 1.1, irRC, irRECIST e iRECIST), evidenciando las similitudes, diferencias y requisitos de confirmación específicos que rigen cada marco evaluativo frente a inmunoterapias y terapias convencionales.

Tabla 3. Criterios de respuesta según diferentes guías clínicas

Criterio de respuesta	RECIST 1.1	irRC	irRECIST	iRECIST
Respuesta completa (CR)	Desaparición de todas las lesiones	Desaparición de todas las lesiones (medibles o no, sin nuevas lesiones). Confirmación con evaluación repetida ≥ 4 semanas	Desaparición de todas las lesiones medibles y no medibles (ganglios < 10 mm eje corto). No requiere confirmación	Desaparición de todas las lesiones
Respuesta parcial (PR)	$\geq 30\%$ de disminución en SLD desde línea basal	$\geq 50\%$ de disminución en SPD desde línea basal. Confirmación con evaluación repetida ≥ 4 semanas	$\geq 30\%$ de disminución en SLD desde línea basal	$\geq 30\%$ de disminución en SLD desde línea basal
Enfermedad estable (SD)	No cumple criterios para CR, PR ni PD	No cumple criterios para CR, PR ni PD	No cumple criterios para CR, PR ni PD	No cumple criterios para CR, PR ni PD
Enfermedad progresiva (PD)	$\geq 20\%$ de aumento en SLD desde el nadir y aumento absoluto ≥ 5 mm, o progresión inequívoca de lesiones no objetivo	$\geq 25\%$ de aumento en SPD desde el nadir, o progresión inequívoca de lesiones no objetivo. Confirmación ≥ 4 semanas	$\geq 20\%$ de aumento en SLD y ≥ 5 mm absolutos desde el nadir, o progresión inequívoca de lesiones no objetivo o nuevas lesiones no medibles. Confirmación ≥ 4 semanas	<i>iUPD</i> : requiere confirmación en 4–8 semanas. Si no se confirma, se reclasifica como <i>iCR</i> , <i>iPR</i> o <i>iSD</i> . <i>iCPD</i> : progresión confirmada en seguimiento

iCR: Respuesta Completa Inmuno-Relacionada; **irPR**: Respuesta Parcial Inmuno-relacionada; **SLD**: Suma de los Diámetros más largos; **SPD**: Suma de Productos de los Diámetros Perpendiculares; **iUPD**: Progresión Inmunológica No Confirmada; **iCPD**: Progresión Inmunológica Confirmada. **Fuente**: Lalchandani *et al.*, «*A Radiologist's Guide to Response Evaluation Criteria in Solid Tumors*» (14).

Estandarización definitiva: iRECIST como consenso metodológico

La culminación de esta evolución metodológica se materializó con la publicación de iRECIST en 2017 por el RECIST Working Group, estableciendo criterios de consenso que preservan la metodología unidimensional de RECIST 1.1 mientras incorporan las categorías diferenciadas de enfermedad progresiva no confirmada (iUPD) y confirmada (iCPD). Estos criterios fueron específicamente diseñados para estandarizar la evaluación de respuesta en ensayos clínicos de inmunoterapia, asegurando consistencia en diseño experimental, recolección de datos, interpretación y análisis, sin intención de influir en decisiones clínicas individuales (19).

La **Figura 7** sintetiza las diferencias metodológicas clave entre RECIST 1.1, irRECIST/iRECIST, OMS e irRC. RECIST y sus variantes inmunológicas emplean mediciones unidimensionales del eje largo (≥ 10 mm), limitan la selección a un máximo de cinco lesiones diana (hasta dos por órgano) e incluyen ganglios linfáticos si el eje corto es ≥ 15 mm; la carga tumoral se expresa como la suma de los diámetros más largos (SLD). Por el contrario, OMS e irRC utilizan mediciones bidimensionales perpendiculares, admiten hasta quince lesiones (diez viscerales y cinco cutáneas) y

calculan la carga tumoral como la suma de los productos de dichos diámetros (SPD), sin un protocolo específico para ganglios linfáticos. Estas discrepancias repercuten directamente en la detección de pseudoprogresión y en la toma de decisiones clínicas (14).

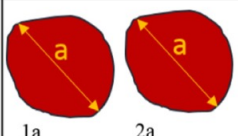
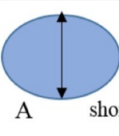
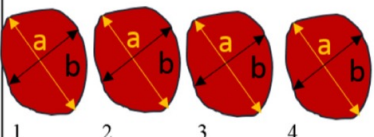
	Target/Index Lesion		Tumor Burden
RECIST 1.1, irRECIST, iRECIST	 Unidimensional Long axis $\geq 10\text{mm}$ ≤ 2 in any one organ Max. 5 lesions	 Lymph node short axis $\geq 15\text{mm}$	Sum of longest diameters (SLD) and short axis of lymph nodes $= 1a + 2a + A$
WHO irRC	 Perpendicular bidimensional ≤ 5 in any one organ Max. 15 lesions (10 visceral + 5 cutaneous) Lymph node assessment not further described		Sum of products of perpendicular diameters (SPD) = $1(a \times b) + 2(a \times b) + 3(a \times b) + 4(a \times b)$

Figura 7. Evaluación de lesiones medibles y carga tumoral según diferentes criterios. Evaluación de lesiones diana y carga tumoral según criterios unidimensionales (RECIST 1.1/irRECIST/iRECIST) frente a bidimensionales (OMS/irRC). El esquema contrasta la medición del eje largo $\pm$ eje corto ganglionar con la medición bidimensional y el cálculo de SPD. **Fuente:** Lalchandani *et al.*, «A Radiologist's Guide to Response Evaluation Criteria in Solid Tumors» (14).

Esta evolución representa la adaptación exitosa de la metodología unidimensional establecida por RECIST para abordar los patrones de respuesta únicos de la inmunoterapia, manteniendo la simplicidad y reproducibilidad de las mediciones unidimensionales mientras incorpora las consideraciones específicas requeridas para evaluar adecuadamente la eficacia de los agentes inmunoterápicos (22; 14).

Transición hacia la evaluación tridimensional (3D)

Si bien la evaluación de la respuesta tumoral ha evolucionado significativamente desde enfoques unidimensionales hacia métodos que consideran la viabilidad tumoral como un marcador más representativo de la eficacia terapéutica, aún persisten desafíos importantes. En el contexto del carcinoma hepatocelular, los criterios mRECIST (2010) marcaron un hito al priorizar el realce arterial como indicador de tejido viable, desplazando la simple medición de dimensiones totales. Posteriormente, los criterios qEASL (2012) expandieron este enfoque mediante la incorporación de mediciones volumétricas tridimensionales del volumen realzado, estableciendo umbrales específicos para respuesta parcial ($RP \geq 65\%$ de reducción) y progresión ($EP \geq 73\%$ de aumento)(10).

Esta progresión metodológica, ilustrada en la **Tabla 4**, refleja una transición desde criterios lineales hacia evaluaciones más integrales de la carga tumoral. Estudios clínicos han demostrado que qEASL ofrece una mejor correlación entre respuesta

radiológica y supervivencia global tras quimioembolización transarterial, superando a RECIST, mRECIST y EASL (10).

Tabla 4. Comparación de criterios modificados de evaluación de respuesta en tumores sólidos

	mRECIST (1D)	EASL (2D)	qEASL (3D)
CR	Desaparición de cualquier realce intratumoral en todas las lesiones diana	Desaparición de cualquier realce intratumoral en todas las lesiones diana	Desaparición de cualquier realce intratumoral en todas las lesiones diana
PR	Disminución de al menos un 30 % en la suma de los diámetros máximos del área tumoral realzada	Disminución de al menos un 50 % en el producto del diámetro máximo y su perpendicular del área tumoral realzada	Disminución de al menos un 65 % en el volumen tumoral realzado
SD	Ni PR ni PD	Ni PR ni PD	Ni PR ni PD
PD	Aumento de al menos un 20 % en la suma de los diámetros del área tumoral realzada	Aumento de al menos un 25 % en el producto del diámetro máximo y su perpendicular del área tumoral realzada	Aumento de al menos un 73 % en el volumen tumoral realzado

EASL: Asociación Europea para el Estudio del Hígado; **qEASL:** Asociación Europea Cuantitativa para el Estudio del Hígado; **HCC:** Carcinoma hepatocelular; **1D:** Unidimensional; **2D:** Bidimensional; **3D:** Tridimensional (volumétrico); **CR:** Respuesta completa; **PR:** Respuesta parcial; **SD:** Enfermedad estable; **PD:** Enfermedad progresiva. **Fuente:** Aykan y Özath *et al.*, «Objective Response Rate Assessment in Oncology» (10).

No obstante, las limitaciones persistentes de los enfoques uni y bidimensionales han estimulado el desarrollo de estrategias más precisas y reproducibles para cuantificar la carga tumoral. En este contexto, la volumetría tridimensional ha emergido como una alternativa metodológica robusta, al permitir la estimación del volumen total de las lesiones mediante la sumatoria de voxels delimitados por segmentación semiautomatizada. Esta técnica ofrece una representación más completa de la morfología tumoral, especialmente útil en lesiones irregulares o multifocales, y mejora la sensibilidad para detectar cambios sutiles en la respuesta terapéutica.

3.2.2. Criterios adaptados a la evaluación tridimensional (3D)

La medición volumétrica representa la aproximación más integral a la cuantificación tumoral, ya que se basa en la reconstrucción tridimensional completa de la lesión y la cuantificación directa del volumen total de la masa tumoral (23). Al utilizar todo el volumen del tumor en lugar de mediciones lineales selectivas, este enfoque proporciona una representación más completa de la morfología y una evaluación más precisa y detallada de la carga tumoral total. Además, esta metodología posee ventajas técnicas fundamentales: al emplear algoritmos automatizados o semiautomatizados de segmentación volumétrica, se logra reducir considerablemente la variabilidad inter e intraobservador típica de las mediciones manuales basadas en diámetros lineales (24; 25).

La transición desde mediciones lineales (RECIST 1.1) hacia enfoques tridimensionales ha estado motivada, ante todo, por la elevada variabilidad que muestran los diámetros manuales: estudios en metástasis pulmonares describen discordancias en hasta 28 % de los pacientes y un coeficiente kappa $\leq 0,66$ cuando se comparan lecturas entre observadores o contra volumetría automatizada (?). Este margen de error se traduce en decisiones terapéuticas potencialmente inconsistentes y ha impulsado la adopción de algoritmos de segmentación automática que reducen la variabilidad intra- e interobservador a valores residuales ($< 2\%$) (9).

Entre los criterios adaptados más extendidos figuran los vRECIST (volumetric Response Evaluation Criteria in Solid Tumors), que trasladan los puntos de corte clásicos de RECIST al espacio volumétrico: una reducción $\geq 65\%$ del volumen se considera respuesta parcial (equivalente al -30% de RECIST), mientras que un aumento $\geq 73\%$ indica progresión tumoral (equivalente al $+20\%$ de RECIST). Estos porcentajes proceden del análisis geométrico original y fueron ya recogidos por Vogel *et al.* al comparar diámetros y volúmenes de metástasis pulmonares (26).

Posteriormente, Winter *et al.* validaron empíricamente esos umbrales en metástasis hepáticas de cáncer colorrectal y demostraron que $\pm 65\%$ (-65% para respuesta, $+65\%$ para progresión) reproduce con alta concordancia los criterios lineales de RECIST. (27).

Otro ejemplo es qEASL (quantitative EASL), que supera la simple medición volumétrica al segmentar de forma tridimensional el volumen realizado arterialmente y, por tanto, el tejido tumoral viable. Esta estrategia resulta particularmente útil en tumores hepáticos tratados con terapias locorregionales como la quimioembolización transarterial (TACE), donde la necrosis ocupa un papel central; estudios recientes que aplican volumetría 3D tras TACE han mostrado cómo la cuantificación del componente viable mejora la clasificación de respondedores frente a no-respondedores respecto a RECIST bidimensional (28).

Además de los umbrales absolutos, la evidencia respalda la superioridad técnica de la volumetría: en un análisis comparativo, el coeficiente de variación interobservador pasó de 6,3 % con medidas manuales a 4,1 % con segmentación semi-automática, permitiendo imaginar cortes volumétricos más estrictos en el futuro (29).

Mientras vRECIST mide “todo” el tumor, qEASL (quantitative European Association for the Study of the Liver) añade la dimensión biológica al cuantificar únicamente el volumen realizado arterialmente, es decir, el tejido viable. Este matiz resulta crucial tras terapias locorregionales en hepatocarcinoma. Lin *et al.* (28) demostraron que la volumetría qEASL automatizada detecta respuesta en focos viables con mayor precisión que RECIST o mRECIST. De forma concordante, un estudio retrospectivo en 50 pacientes tratados con TACE reveló que vRECIST y qEASL reclasificaron a un tercio de los casos respecto a mRECIST, modificando la estrategia de retratamiento (28).

Otros marcos volumétricos amplían el alcance de la medición. El Whole-Body Tumor Volume (WBTV), evaluado en neoplasias metastásicas diversas, redujo la discordancia entre tres lectores de un 100 % con RECIST a solo 2 % al incorporar

todo el volumen tumoral, con un coeficiente kappa de Fleiss $\kappa \approx 0,92$ (30). Por su parte, el cálculo de tasas de crecimiento y regresión basado en volumetría mejoró el poder estadístico para detectar beneficio terapéutico en ensayos fase III de cáncer colorrectal, como demostró Maitland et al. (31).

La evidencia publicada confirma que la volumetría 3D ofrece mayor sensibilidad, porque permite detectar antes variaciones pequeñas de tamaño que se diluyen en la métrica lineal; por ejemplo, al aplicar umbrales volumétricos reducidos en metástasis pulmonares se adelantó la identificación de progresión y se modificó teóricamente la decisión terapéutica en un 12 %–13 % de los pacientes (26), y en ensayos fase III de cáncer colorrectal aumentó la potencia estadística para mostrar beneficios del tratamiento frente a las mediciones 1D (31).

Asimismo, la variabilidad inter- e intraobservador se reduce de forma sustancial: estudios comparativos en metástasis pulmonares y en cohortes multiórgano muestran coeficientes de variación que caen del 6 %–7 % con medidas manuales al 0,3 %–4 % con segmentación (semi)automática, con límites de concordancia clínicamente aceptables (31). En cuanto a la correlación con desenlaces, el cambio volumétrico del tumor se asocia de manera independiente con la supervivencia global en cáncer de pulmón avanzado, mientras que la variación en diámetros RECIST no alcanza significación (32).

Metodologías emergentes

Más allá de los enfoques volumétricos actualmente en uso, han comenzado a emerger nuevas metodologías que buscan capturar con mayor precisión los cambios biológicos y funcionales que acompañan la respuesta tumoral. En este contexto, los métodos de evaluación metabólica mediante PET-CT, como PERCIST e iPERCIST, se perfilan como alternativas prometedoras para detectar cambios funcionales tempranos relacionados con la actividad metabólica del tumor. No obstante, su uso clínico aún requiere validación adicional a través de ensayos clínicos aleatorizados que respalden su utilidad pronóstica y estandarización operativa (10).

En paralelo, el desarrollo de algoritmos de inteligencia artificial aplicados al procesamiento de imágenes médicas y el análisis radiómico han abierto una nueva vía de exploración en la evaluación de respuesta al tratamiento oncológico. Estas herramientas, al analizar patrones complejos dentro de volúmenes tridimensionales mediante técnicas de aprendizaje automático y extracción masiva de características, prometen mejorar la caracterización tumoral más allá de las capacidades humanas convencionales. Sin embargo, su implementación en la práctica clínica rutinaria aún se encuentra en fase experimental, sujeta a procesos rigurosos de validación, reproducibilidad y transparencia algorítmica (33).

Así, la evolución de los criterios de evaluación de respuesta tumoral ha sido impulsada por una necesidad constante de mayor estandarización, reproducibilidad y precisión. Esta progresión no solo refleja el avance paralelo de las tecnologías de imagen y las modalidades terapéuticas oncológicas, sino que también deja en evi-

dencia que ningún sistema actual captura plenamente la complejidad de la respuesta tumoral moderna. En consecuencia, la comunidad científica continúa impulsando el desarrollo de enfoques metodológicos más comprehensivos, que integren dimensiones anatómicas, funcionales y moleculares, y que puedan adaptarse a los desafíos de una oncología cada vez más personalizada (33).

3.2.3. Problemática, brecha de investigación y aporte al conocimiento

Criterios adaptados a métodos multiplanares (2.5D)

Aunque el impacto clínico de los criterios de evaluación unidimensionales y volumétricos ha sido ampliamente documentado, de la revisión de la literatura emerge el método 2.5D como una aproximación intermedia en diversas aplicaciones de procesamiento de imágenes médicas, pero aún no ha sido explorado en el contexto de la evaluación de respuesta tumoral según el protocolo RECIST. Este término tiene sus orígenes en el aprendizaje profundo y las redes neuronales convolucionales, en particular las arquitecturas U-Net empleadas para la segmentación de imágenes médicas (34; 35; 36; 27). El término 2.5D surge precisamente porque, aunque utiliza información de los tres planos ortogonales principales, no proporciona acceso completo a las estructuras del paciente en verdaderas tres dimensiones (37).

Si bien diversos trabajos han comparado la cuantificación tumoral obtenida en cortes axiales, vistas multiplanares y reconstrucciones 3D, ninguno ha evaluado de manera sistemática la utilidad del enfoque 2.5D para clasificar la respuesta tumoral conforme a los criterios RECIST. Los estudios disponibles se centran, ante todo, en mejorar la estadificación y la predicción pronóstica en tumores sólidos, lo que subraya la necesidad de investigar este método específicamente en la evaluación de la respuesta terapéutica en dicha población:

En cáncer de pulmón no microcítico, un análisis retrospectivo de 544 pacientes mostró que las mediciones axiales, multiplanares y con máscara 3D ofrecen un rendimiento pronóstico casi idéntico (índices de concordancia $\approx 0,66$ – $0,67$), pero la aproximación multiplanar alcanzó la mayor coincidencia con la estadificación patológica (67,5 % frente a 60,5 % axial y 50,9 % 3D) (38).

En adenocarcinoma de páncreas, al evaluar 183 tumores, los diámetros medidos en los planos axial (2,9 cm), coronal (3,2 cm) y sagital (3,2 cm) fueron menores que el diámetro patológico (3,4 cm), salvo que el mayor de los tres planos se aproximó al valor patológico. Además, el volumen tumoral en 3D mostró mejor capacidad de predicción de muerte temprana ($AUC = 0,714$) y correlacionó más estrechamente con la supervivencia que las medidas unidimensionales (39).

En un estudio de 89 tumores pulmonares, la medición multiplanar incrementó el tamaño máximo y provocó un cambio de estadificación T (“upstaging”) en el 18 %–20 % de los casos en comparación con la medición axial. La repetibilidad entre lectores fue elevada para axial ($ICC = 0,900$) y sagital ($ICC = 0,874$), y moderada para coronal ($ICC = 0,754$) (40).

A pesar de las aplicaciones de los métodos 2.5D multiplanares en diversas aplicaciones de procesamiento de imágenes médicas, estos enfoques no han sido evaluados específicamente en el contexto de los criterios RECIST ni en la evaluación sistemática de tumores sólidos. Además, no se ha explorado exhaustivamente su impacto clínico en comparación directa con los métodos unidimensionales y tridimensionales tradicionalmente utilizados en la práctica oncológica. Esta brecha en el conocimiento representa una oportunidad para explorar el potencial de los métodos 2.5D como una herramienta que podría ofrecer ventajas en términos de precisión diagnóstica, eficiencia computacional y aplicabilidad clínica. Actualmente no existen estudios que permitan comparar el potencial de los métodos volumétricos versus el método multiplanar 2.5D y el estándar unidimensional en la evaluación de la respuesta terapéutica en tumores sólidos.

Motivación

La adopción de RECIST 1.1 ha representado un avance fundamental en la estandarización de la evaluación de la respuesta tumoral. No obstante, su efectividad y eficiencia continúan viéndose limitadas por diversas debilidades operativas y metodológicas. En primer lugar, su implementación presenta una reproducibilidad restringida debido a la subjetividad en la selección de lesiones objetivo y a la dependencia de mediciones manuales del diámetro mayor (41). Esta variabilidad interobservador ha sido ampliamente documentada, especialmente en contextos clínicos con alta carga tumoral, donde la precisión es crítica (8; 42; 13).

Adicionalmente, el uso exclusivo de mediciones unidimensionales puede conducir a estimaciones inexactas de la carga tumoral total, subestimando la magnitud real del tumor (43). Esta simplificación geométrica compromete la sensibilidad de los criterios RECIST para detectar cambios terapéuticos sutiles, pero clínicamente relevantes (8). A ello se suma el carácter manual del procedimiento, que requiere en promedio más de 11 minutos por evaluación, incluso por radiólogos experimentados, lo cual restringe significativamente su aplicabilidad en entornos clínicos de alta demanda (41).

Más allá de estas limitaciones técnicas, RECIST se basa en la evaluación de un número reducido de lesiones observadas en cortes axiales bidimensionales de tomografía computada (TC), con el propósito de representar una carga tumoral inherentemente tridimensional. Esta aproximación ha sido objeto de cuestionamientos, ya que múltiples estudios han demostrado que la selección de un subconjunto limitado de lesiones no necesariamente refleja de manera fidedigna la progresión real ni la magnitud completa del volumen tumoral. El número y tipo de lesiones seleccionadas en cortes bidimensionales influyen directamente en la clasificación de la respuesta terapéutica, generando inconsistencias e incluso subestimaciones relevantes de la carga tumoral total (44; 45; 46).

Se ha descrito que la clasificación de respuesta puede variar considerablemente dependiendo de las lesiones seleccionadas, sin que ello se relacione necesariamente con la verdadera carga tumoral. Asimismo, se ha documentado una elevada tasa de desacuerdo en la selección y medición de lesiones entre observadores, especialmente

en pacientes con múltiples metástasis (47; 34; 35), así como dificultades metodológicas en la evaluación de ganglios linfáticos, particularmente cuando estos se fusionan o dividen entre estudios de seguimiento (48).

Pese al reconocimiento generalizado de la superioridad técnica de la volumetría tridimensional frente a los métodos unidimensionales y bidimensionales, su adopción clínica se ha visto restringida por limitaciones operativas y metodológicas. En este contexto, han surgido enfoques intermedios, como los métodos multiplanares (también denominados 2.5D), que permiten incorporar información tridimensional parcial mediante mediciones en los planos ortogonales axial, sagital y coronal, sin requerir segmentaciones volumétricas completas. Sin embargo, el impacto real de estos enfoques sobre la clasificación clínica de la respuesta tumoral según los criterios RECIST sigue siendo incierto. Esta brecha de conocimiento resulta particularmente relevante si se considera que los métodos multiplanares podrían representar una alternativa viable para mejorar la precisión evaluativa sin incurrir en los elevados requerimientos técnicos y temporales de la volumetría completa.

Esta tesis tiene como objetivo evaluar la concordancia en la clasificación de la respuesta al tratamiento oncológico, según los criterios RECIST, utilizando tres enfoques de medición: el tradicional unidimensional (1D), una aproximación multiplanar (2.5D) y la medición tridimensional (3D). Esta comparación permite analizar el impacto clínico de los métodos volumétricos en contextos locales, así como explorar el potencial del enfoque multiplanar como una herramienta intermedia en la evaluación de la respuesta terapéutica en oncología. En este sentido, el presente trabajo busca aportar evidencia sobre la relevancia de los métodos geométricos en la interpretación de la efectividad del tratamiento y en la toma de decisiones terapéuticas para los pacientes del Hospital Clínico de la Universidad de Chile (HCUCH).

En síntesis, esta tesis aborda la variabilidad e inconsistencias del protocolo RECIST 1.1, especialmente al compararlo con las mediciones 2.5D y 3D, en términos de la concordancia en la clasificación de respuesta. Se plantea la hipótesis de que RECIST 1.1 podría estar subestimando la carga tumoral total, lo que podría derivar en categorizaciones divergentes con consecuencias clínicas relevantes, al influir en la evaluación de la efectividad del tratamiento y en la selección de las estrategias terapéuticas más adecuadas para los pacientes oncológicos del HCUCH.

4. Hipótesis

Las diferencias en la clasificación de la respuesta al tratamiento obtenidas mediante mediciones 1D, en comparación con aquellas obtenidas a través de métodos 2.5D y 3D, en imágenes de TC TAP de pacientes sometidos al protocolo RECIST 1.1 del Hospital Clínico de la Universidad de Chile (HCUCH), son estadísticamente significativas.

5. Objetivo General

Evaluar la concordancia en la clasificación de respuesta al tratamiento al comparar el método de medición 1D con los métodos de medición 2.5D y 3D en imágenes de TC TAP de pacientes sometidos al protocolo RECIST 1.1 del HCUCH.

6. Objetivos Específicos

- OE1.** Cuantificar la carga tumoral total de las lesiones objetivo, observadas en imágenes de TC TAP de pacientes bajo el protocolo RECIST 1.1 del HCUCH, utilizando los métodos de medición 1D, 2.5D y 3D.
- OE2.** Establecer umbrales volumétricos según los criterios RECIST 1.1 para comparar la clasificación de respuesta al tratamiento utilizando diferentes métodos de medición.
- OE3.** Analizar la concordancia entre la clasificación de respuesta al tratamiento obtenida mediante los distintos métodos de medición.
- OE4.** Analizar la significancia estadística de las diferencias en la clasificación de respuesta al tratamiento entre los métodos de medición 1D, 2.5D y 3D para validar la hipótesis planteada.

7. Materiales y Métodos

Las **Figuras 8 y 9** ofrecen una síntesis del flujo de materiales y métodos: la primera detalla la fase de preparación y análisis, mientras que la segunda describe la fase de aplicación. En ambas, las columnas especifican las actividades efectuadas, las fuentes de medición empleadas y el tipo de análisis desarrollado. Un panorama consolidado de cada etapa metodológica se presenta en la **Tabla 5**.

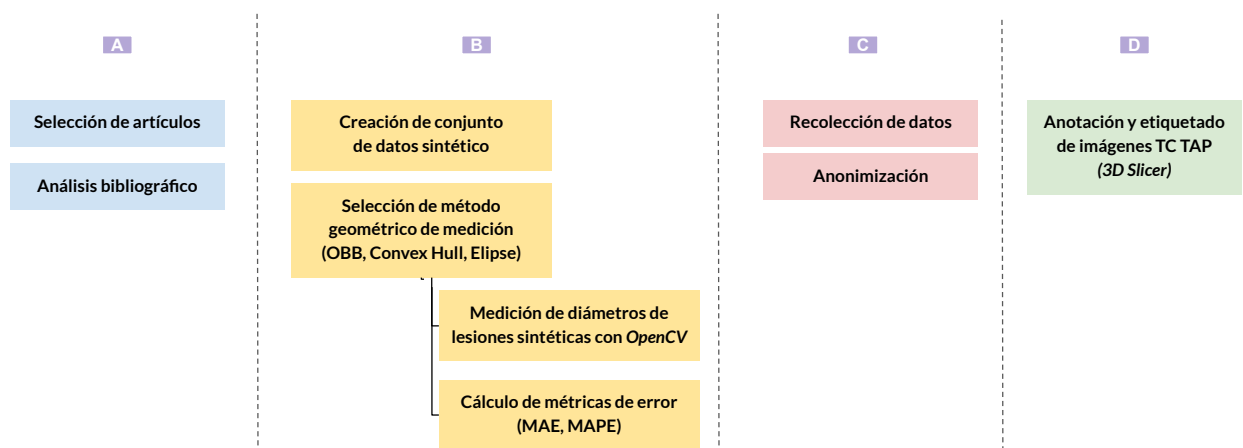


Figura 8. Fase de Preparación. Las columnas A-D representan la etapa de análisis bibliográfico, implementación de métodos geométricos y medición de rendimiento, tratamiento de datos y segmentaciones manuales, respectivamente.

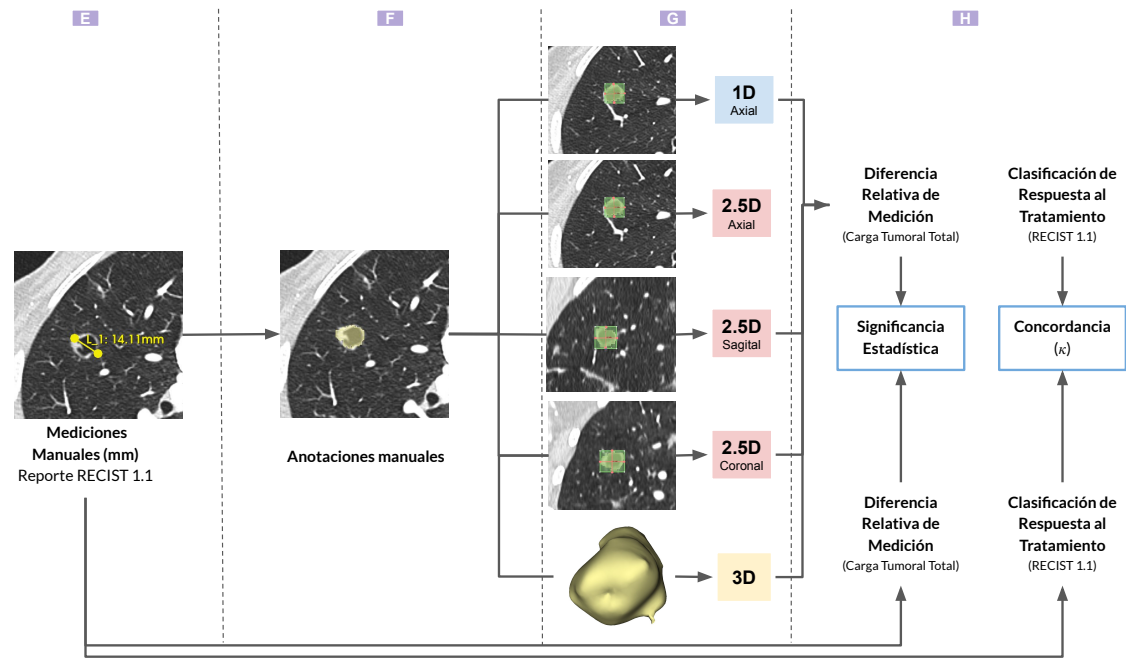


Figura 9. Fase de Análisis. Las columnas E–H muestran, respectivamente, el reporte experto, la segmentación semiautomática, las mediciones automáticas (1D, 2.5D, 3D) y las métricas de concordancia con significancia estadística y clasificación RECIST 1.1.

Tabla 5. Resumen de etapas metodológicas

Col.	Descripción
A	Revisión sistemática y análisis crítico de la literatura sobre los criterios de respuesta tumoral fundamentados en mediciones geométricas.
B	Generación de lesiones sintéticas con geometrías controladas (circulares, elipsoidales e irregulares) y diámetros conocidos; comparación de algoritmos geométricos en OpenCV mediante los errores absoluto medio (MAE) y porcentual medio (MAPE).
C	Cohorte retrospectiva: 45 estudios de TC de tórax–abdomen/pelvis que abarcan 78 lesiones en 15 pacientes oncológicos (edad media 60 años), seleccionados según RECIST 1.1 y con aprobación ética. Los archivos DICOM fueron anonimizados y seudonimizados en un servidor seguro.
D	Segmentación manual de 78 lesiones por dos radiólogos con 10 y 5 años de experiencia; las discrepancias se resolvieron por consenso.
E	Informe inicial de 78 lesiones diana conforme al protocolo RECIST 1.1, emitido por radiólogos con más de 10 años de experiencia.
F	Ajuste manual de las máscaras por los mismos radiólogos expertos (10 y 5 años de experiencia), sin herramientas semiautomáticas, para garantizar la fidelidad anatómica.
G	Obtención automática de los diámetros mayor y menor a partir de las máscaras validadas, empleando los enfoques de medición 1D, 2.5D y 3D.
H	Evaluación de concordancia: cálculo de las diferencias relativas entre los tres métodos y las mediciones expertas, aplicación de pruebas de significancia y categorización de la respuesta al tratamiento conforme a RECIST 1.1.

7.1. Métricas

Error Absoluto Mediano (MAE) y Error Porcentual Absoluto Mediano (MAPE)

Las métricas usadas para validar la precisión de estos métodos frente a los diámetros de referencia fueron MAE (Error Absoluto Mediano) y MAPE (Error Porcentual Absoluto Mediano) las cuales son métricas estadísticas ampliamente utilizadas para cuantificar la precisión de estimaciones o predicciones. MAE mide la magnitud promedio del error en las mismas unidades de los datos (por ejemplo, milímetros). Se calcula como el valor medio de las diferencias absolutas entre los valores reales y los predichos:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (1)$$

donde y_i son los valores reales y $\hat{y}_i$ los valores estimados. Esta métrica es fácil de interpretar (un MAE bajo indica alta precisión) y trata todos los errores por igual sin penalizar excesivamente los errores grandes.

MAPE (Error Porcentual Absoluto Mediano) expresa el error en términos porcentuales respecto al valor real. Se define matemáticamente como el promedio de los errores absolutos relativos:

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad (2)$$

donde y_i es el valor real y $\hat{y}_i$ el valor predicho. Al multiplicar por 100 se obtiene un porcentaje mediano de error. Esto hace que MAPE sea independiente de la escala original de los datos y fácilmente interpretable (por ejemplo, un MAPE del 10 % indica un error mediano del 10 % respecto al valor real).

Ambas expresiones usan la suma de errores absolutos, pero MAPE agrega una normalización por el valor real, lo que proporciona un criterio relativo.

En el contexto del procesamiento de imágenes médicas, MAE y MAPE sirven para validar funciones geométricas que estiman características como diámetros o áreas de lesiones. Por ejemplo, si se tiene un conjunto sintético de imágenes tumorales con diámetros conocidos (circulares, elipsoides, irregulares), los algoritmos geométricos pueden calcular diámetros estimados que se comparan con los valores de referencia.

Calcular el MAE entrega el error mediano absoluto en unidades físicas (p.ej. milímetros), facilitando ver qué tan cerca, en promedio, están las estimaciones de la realidad. En cambio, el MAPE convierte cada error en un porcentaje relativo del tamaño real del tumor, de modo que se puede comparar el rendimiento entre tumores grandes y pequeños.

Por ejemplo, si un método subestima el diámetro de un tumor pequeño en 1 mm (error grande en proporción) y el de un tumor grande en 2 mm (error pequeño en proporción), MAE promediaría estos errores absolutos, pero MAPE reflejará que la primera estimación fue mucho menos precisa porcentualmente que la segunda.

Cálculo del coeficiente Kappa de Cohen

El coeficiente de concordancia κ de Cohen es una medida estadística que cuantifica el grado de acuerdo entre dos evaluadores categóricos, corrigiendo el acuerdo observado por el que podría ocurrir al azar. Fue introducido por Jacob Cohen en 1960 como una alternativa más robusta al simple porcentaje de coincidencia. Formalmente, si P_o es la proporción de casos en que ambos evaluadores coinciden (el acuerdo observado) y P_e es la proporción de acuerdo esperada por azar (calculada

a partir de las frecuencias marginales), el coeficiente κ se define por la fórmula:

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

donde:

$$P_o = \frac{\text{número de coincidencias observadas}}{N}$$

$$P_e = \sum_{i=1}^C \left(\frac{n_{i\cdot}}{N} \cdot \frac{n_{\cdot i}}{N} \right)$$

Aquí N es el total de casos evaluados, $n_{i\cdot}$ y $n_{\cdot i}$ son las sumas marginales de la categoría i , y C es el número de categorías. Por construcción, $\kappa = 1$ indica acuerdo perfecto (todos los casos coincidentes), $\kappa = 0$ indica que el nivel de acuerdo observado equivale al esperado por azar, y valores negativos señalan que el acuerdo es peor que el azar. En la práctica se interpreta como una especie de correlación estandarizada: valores más cercanos a 1 implican mayor fiabilidad interevaluador, mientras que valores próximos a 0 o negativos denotan poco o ningún acuerdo efectivo.

Interpretación del valor de κ

Aunque κ varía en teoría entre -1 y 1 , en aplicaciones suele caer entre 0 y 1 . Cohen propuso una escala interpretativa: valores $\kappa \leq 0$ indican ausencia de acuerdo, 0.01 – 0.20 un acuerdo mínimo, 0.21 – 0.40 escaso, 0.41 – 0.60 moderado, 0.61 – 0.80 sustancial y 0.81 – 1.00 casi perfecto. Por ejemplo, $\kappa = 0.5$ se consideraría “moderado” según esta regla empírica. Sin embargo, estos umbrales son orientativos y dependen del contexto; en áreas como la salud, a menudo se exige un κ elevado para confiar en la coincidencia de diagnósticos. En todo caso, la ventaja de κ sobre la simple concordancia porcentual es que descuenta el efecto del azar, lo que lo convierte en una medida más informativa de la fiabilidad interobservador.

7.2. Estrategia de búsqueda y selección de artículos

La búsqueda y selección de artículos se realizó en las bases de datos PubMed, Scopus y Google Scholar (**Figura 10**). El objetivo fue identificar y analizar los distintos métodos geométricos y sus criterios utilizados en la evaluación de la respuesta al tratamiento en tumores sólidos mediante imágenes de tomografía computada (TC), con énfasis en aquellos comparados los criterios RECIST 1.1. La estrategia se centró en caracterizar las limitaciones del enfoque unidimensional tradicional, así como en revisar el desempeño de métodos volumétricos tridimensionales, reconocidos por su mayor sensibilidad pero limitados por barreras operativas.

Asimismo, se puso especial atención en identificar una brecha metodológica relevante: la escasa exploración de enfoques multiplanares ortogonales o “2.5D”, que podrían representar una alternativa intermedia entre los métodos unidimensionales y volumétricos. En consecuencia, esta revisión no solo buscó estudios comparativos entre RECIST 1.1 y métodos 3D, sino también evidencia sobre métodos 2.5D,

evaluando su aplicabilidad clínica y concordancia diagnóstica. Las preguntas clave y objetivos específicos que guiaron esta revisión se presentan en la **Tabla 6**. Los criterios de inclusión y exclusión aplicados a los artículos revisados se describen en la **Tabla 7**.

Tabla 6. Objetivos y preguntas clave de la búsqueda sistemática

Objetivos	Preguntas clave
Identificar los métodos geométricos utilizados para evaluar la respuesta al tratamiento de tumores sólidos en imágenes TC, incluyendo métodos unidimensionales, volumétricos y multiplanares	¿Qué métodos geométricos han sido aplicados en la evaluación de respuesta tumoral en imágenes TC, y qué características los diferencian en términos de reproducibilidad y aplicabilidad clínica?
Analizar comparaciones entre RECIST y otros criterios (volumétricos o multiplanares), evaluando diferencias en la clasificación de respuesta y su correlación clínica	¿Qué diferencias metodológicas y resultados clínicos han reportado los estudios que comparan RECIST con métodos tridimensionales o multiplanares?
Explorar la aplicabilidad clínica de los métodos volumétricos y multiplanares, considerando sus limitaciones operativas y su posible valor como alternativa o complemento a RECIST	¿Qué evidencia existe sobre el impacto clínico y la viabilidad práctica de los métodos volumétricos y multiplanares como alternativa o complemento a los criterios RECIST?

Los parámetros descritos fueron aplicados a un total de 20 artículos publicados entre los años 2008–2024.

Tabla 7. Criterios de inclusión y exclusión de artículos

Criterios de inclusión	Criterios de exclusión
CI1. Artículos publicados hasta el 13 de julio del 2024.	CE1. Opiniones, editoriales, cartas al editor, tesis no publicadas y resúmenes de congresos sin texto completo disponible.
CI2. Artículos en idioma español e inglés.	CE2. Artículos que no proporcionen resultados clínicos entre los distintos métodos y criterios de medición descritos.
CI3. Estudios que comparen RECIST con métodos geométricos bidimensionales, volumétricos (3D) o multiplanares (2.5D) en la evaluación de tumores sólidos.	CE3. Artículos que no realicen un análisis comparativo entre RECIST o RECIST 1.1 y otros métodos de evaluación geométrica o volumétrica en imágenes de TC.
	CE4. Artículos que no se enfoquen en tumores sólidos evaluados mediante tomografía computada.

CI: Criterios de inclusión. **CE:** Criterios de exclusión.

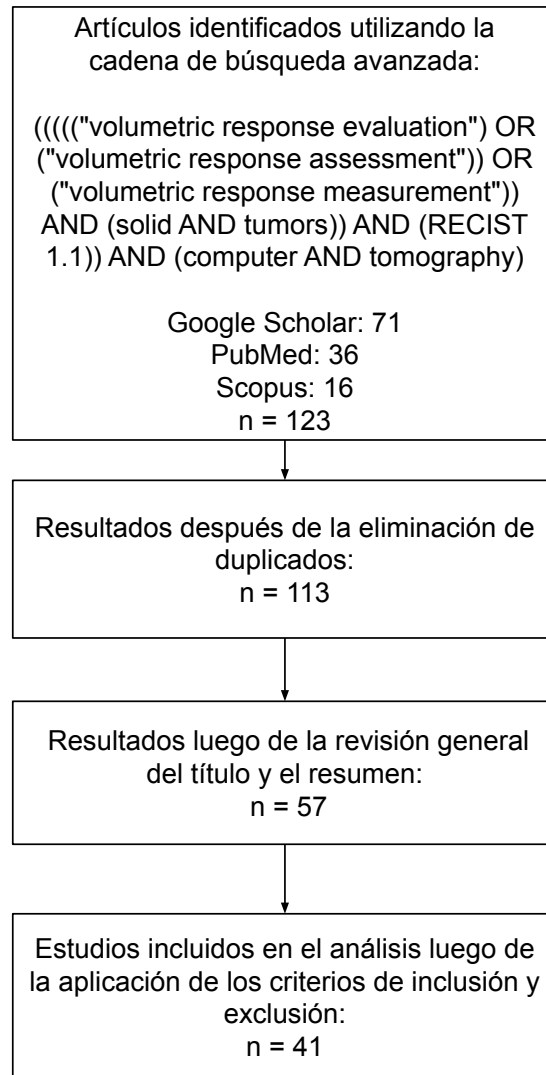


Figura 10. Diagrama de estrategia de búsqueda bibliográfica. Se detalla el proceso sistemático empleado para la identificación y selección de artículos relevantes. Se identificaron 123 artículos en tres bases de datos (Google Scholar, PubMed y Scopus) mediante una cadena de búsqueda específica. La eliminación de duplicados, dejó un total de 113 artículos únicos. A continuación, se revisaron los títulos y resúmenes, reduciendo la muestra a 40 artículos. Finalmente, tras aplicar criterios de inclusión y exclusión, se seleccionaron 20 estudios para el análisis final.

7.3. Recolección de datos

La publicación de tomografías computarizadas anonimizadas recolectadas de forma retrospectiva fue aprobada por el Comité de Ética Científica del HCUCH, bajo el certificado N°45. Se incluyeron pacientes mayores de 18 años, con diagnóstico confirmado de cáncer, evaluados según los criterios RECIST 1.1. Los criterios de exclusión contemplaron a aquellos pacientes con tipos de cáncer no evaluables mediante RE-

CIST 1.1 o que no contaban con estudios de seguimiento además del examen basal. Asimismo, solo se consideraron pacientes que disponían de tomografías computarizadas de tórax y abdomen/pelvis.

Posterior a esta selección inicial, las imágenes fueron revisadas por radiólogos, excluyéndose pacientes cuyas imágenes no presentaran tumores sólidos o cuyas lesiones identificadas fueran menores a 10 mm en su diámetro mayor. El conjunto de datos final comprende 117 lesiones analizadas obtenidas de 45 imágenes de TC obtenidas entre 2017 y 2023 de 15 pacientes con diversos tipos de cáncer (**Figura 11**). Las edades oscilaron entre 53 y 74 años, con mediana de 60 años. El conjunto presenta distribución equilibrada por sexo: 7 pacientes femeninos y 8 masculinos.

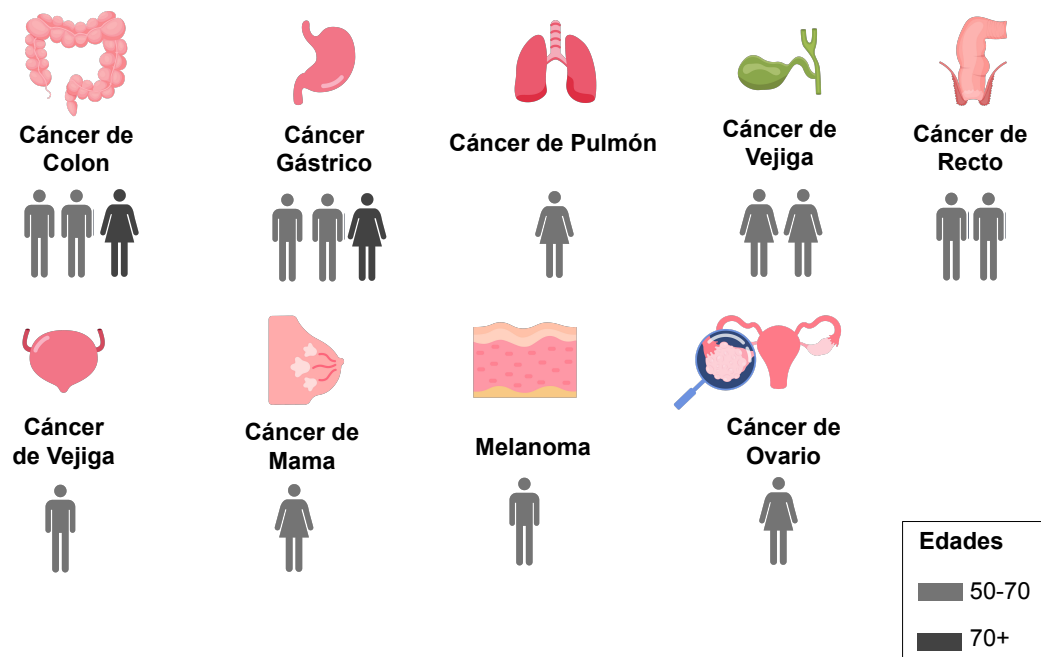


Figura 11. Distribución del número de pacientes, sexo y edad en el conjunto de datos según los distintos tipos de cáncer.

7.4. Anonimización de datos

La información sensible de los pacientes fue anonimizada y seudonimizada en un servidor local inmediatamente después de abandonar la infraestructura del HCUCH. Para este propósito, las imágenes DICOM se transmitieron desde el Sistema de Archivo y Comunicación de Imágenes (PACS) del departamento de radiología del HCUCH hacia un endpoint de MirthConnect ubicado en el entorno local.

Un servicio encargado de la curación de datos procedió a eliminar metadatos sensibles, tales como la dirección y la fecha de nacimiento del paciente. Simultáneamente, tanto el nombre como el identificador del paciente fueron seudonimizados, reemplazando sus valores originales con identificadores numéricos únicos que garantizan la distinción entre diferentes pacientes.

7.5. Anotación y etiquetado de datos

Todas las tomografías computarizadas fueron revisadas y segmentadas manualmente por dos radiólogos con 5 y 10 años de experiencia, respectivamente; las discrepancias se resolvieron por consenso. El flujo completo de anotación se muestra en la **Figura 12**.

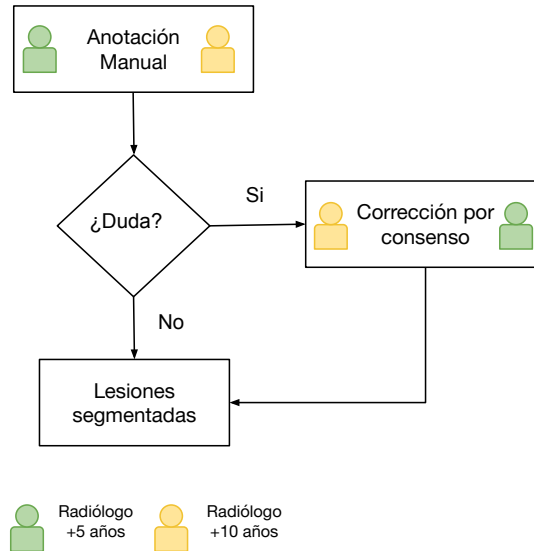


Figura 12. Diagrama del flujo de anotación de lesiones en estudios de tomografía computarizada. Dos radiólogos lectores independientes participaron en la segmentación manual: R1, con más de 10 años de experiencia, anotó las imágenes de 8 pacientes (24 TC TAP), y R2, con más de 5 años de experiencia, anotó las de 7 pacientes (21 TC TAP). Los exámenes se distribuyeron de manera simple-aleatoria, sin solapamiento entre lectores. Ante cualquier duda diagnóstica, ambos revisores resolvieron la discrepancia mediante consenso, generándose así la máscara definitiva de cada lesión.

Este proceso resultó en un total de 78 lesiones delineadas, que incluyen metástasis, 73 ganglios linfáticos (adenopatías) y 5 tumores primarios. La **Figura 13** presenta imágenes representativas de TC torácica y abdominal con ejemplos de lesiones delineadas.

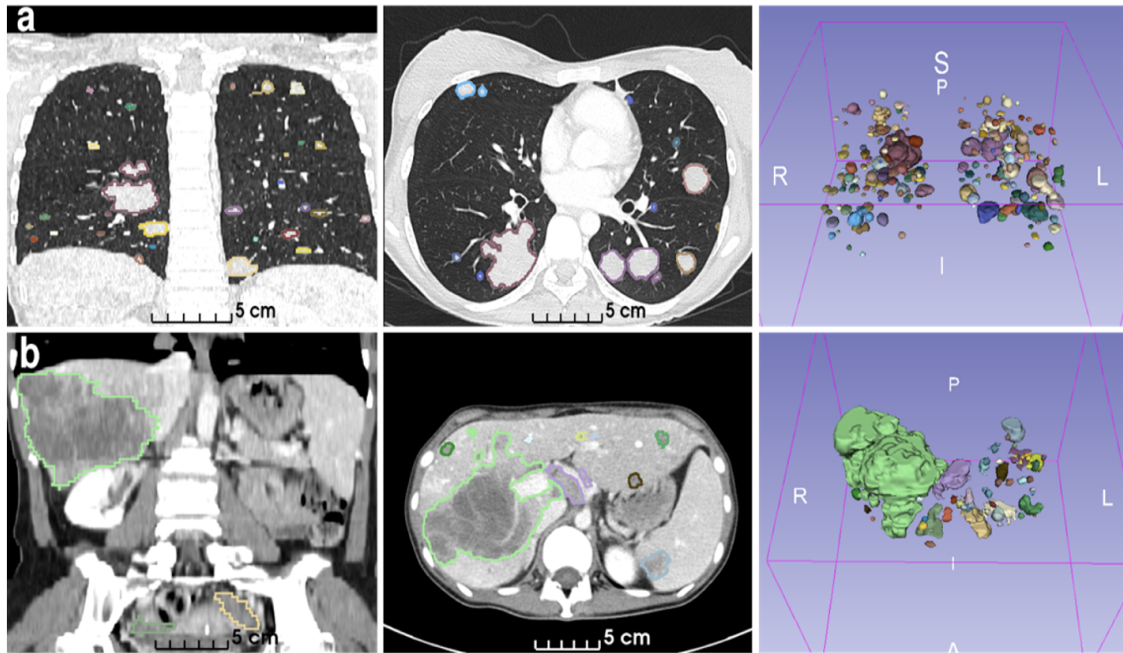


Figura 13. Ejemplos de lesiones anotadas en imágenes de TC. Cada fila corresponde a una serie obtenida del mismo paciente y estudio de tomografía computarizada. (a) Tórax. (b) Abdomen. Cada columna muestra una vista específica de la misma imagen de TC, con las lesiones anotadas representadas en color. De izquierda a derecha: plano coronal, plano axial y reconstrucción tridimensional en el *software 3D Slicer*.

7.6. Medición de diámetros de tumores en imágenes de TC con OpenCV

OpenCV (*Open Source Computer Vision Library*) es una biblioteca ampliamente utilizada para procesamiento digital de imágenes. Ofrece funciones para filtrado, detección de bordes, extracción de características y segmentación, siendo muy versátil en diversos campos, incluido el análisis de imágenes médicas.

En el contexto biomédico, OpenCV facilita la manipulación y análisis de máscaras binarias de segmentación de lesiones (imágenes resultantes de delinear tumores u otras estructuras en estudios radiológicos). Por ejemplo, permite limpiar artefactos mediante operaciones morfológicas, detectar contornos de una lesión segmentada y calcular propiedades geométricas de dicha región de interés. Estas capacidades permiten automatizar la cuantificación de características tumorales (como tamaños y formas) a partir de imágenes médicas de forma reproducible y precisa.

Una tarea clínica habitual en oncología radiológica es medir los diámetros mayores y menores de un tumor en las imágenes de tomografía computarizada (TC) para evaluar su evolución (criterios RECIST). Tradicionalmente, un radiólogo mide manualmente el mayor diámetro en el plano axial de la TC. Mediante Python y OpenCV es posible automatizar esta medición: existen funciones geométricas (`cv2.fitEllipse`, `cv2.minAreaRect`, `cv2.convexHull`) que permiten estimar la mayor extensión de una lesión segmentada en una imagen axial.

7.6.1. cv2.fitEllipse: Ajuste elíptico por mínimos cuadrados

Recolección de puntos (x_i, y_i) del contorno

El primer paso en el ajuste de una elipse sobre una lesión tumoral segmentada es la recolección de los puntos que forman su contorno. Estos puntos se extraen a partir de una máscara binaria generada en 3D Slicer, donde los píxeles de valor 1 indican la presencia de la lesión. En este contexto, se utiliza la función `cv2.findContours()` de OpenCV.

En la máscara, los vóxeles con intensidad 1 representan tejido tumoral y los de intensidad 0 corresponden al fondo. Esta función analiza la imagen binaria y localiza los bordes de las regiones blancas (valor 1), devolviendo como resultado una lista de contornos. Cada contorno se representa como una secuencia ordenada de coordenadas que siguen el perímetro de una región conectada en la imagen, formando así una curva cerrada en dos dimensiones.

En términos simples, cada contorno es un conjunto de puntos que, al unirse secuencialmente, rodean completamente una lesión segmentada. Esta representación es clave para calcular propiedades geométricas de la lesión.

$$\text{Puntos extraídos} = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \quad (3)$$

La **Figura 14.** muestra un ejemplo de cómo los puntos son extraídos del borde de una lesión segmentada.

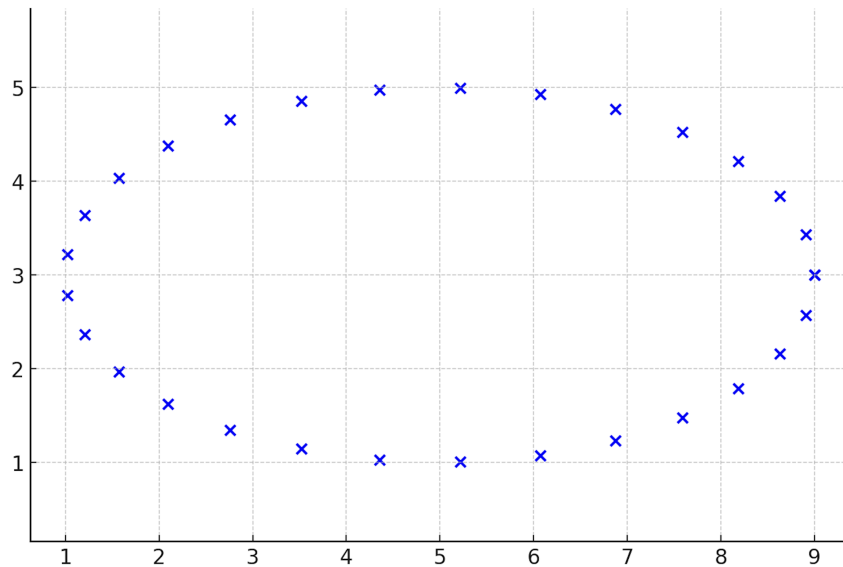


Figura 14. Puntos x_i, y_i del contorno de una lesión.

Estos datos son la entrada del proceso de ajuste, representando la forma bidimensional de la región de interés. Una vez definidos los puntos del contorno, se aplica el método de mínimos cuadrados para ajustar una elipse que represente de forma aproximada la forma de la lesión. Esto se realiza con la función `fitEllipse`.

Este algoritmo ajusta una curva elíptica que minimiza el error cuadrático respecto a los puntos del contorno. La elipse se modela mediante una ecuación cuadrática general en dos variables:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \quad (4)$$

El algoritmo encuentra los coeficientes A, B, C, D, E, F que mejor se ajustan a los puntos del contorno minimizando:

$$\min \sum_{i=1}^n (Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F)^2 \quad (5)$$

La solución se restringe a los casos en que la ecuación resultante define efectivamente una elipse válida, descartando aquellos en los que la curva resultante sea una parábola, hipérbola u otra forma geométrica que no sea adecuada para representar lesiones segmentadas. La **Figura 15** representa gráficamente el resultado del algoritmo `fitEllipse`, que encuentra los parámetros de la elipse que mejor se ajusta (en el sentido de mínimos cuadrados) a los puntos anteriores.

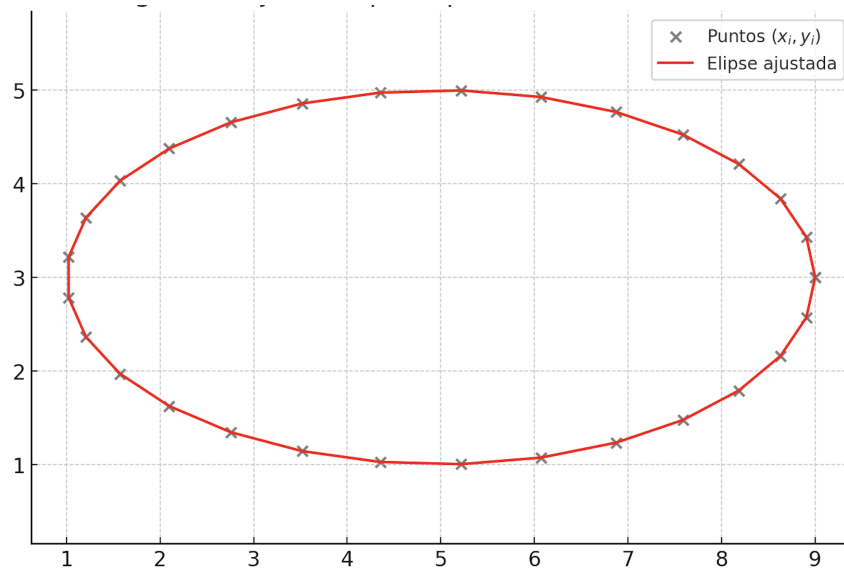


Figura 15. Ajuste elíptico por mínimos cuadrados

Extracción de parámetros geométricos de la elipse

Del resultado del ajuste, se obtienen directamente los siguientes parámetros de la elipse:

- **Centro:** (x_0, y_0) , punto central de la elipse.
- **Ejes:** Longitudes del eje mayor a y eje menor b , definidos como los diámetros de mayor y menor longitud.
- **Ángulo de rotación:** θ , que indica la orientación de la elipse respecto al eje horizontal de la imagen.

Estos parámetros simplifican la forma irregular de una lesión en una representación regular y analíticamente manejable. La **Figura 16** representa un ejemplo de la extracción de estos parámetros.

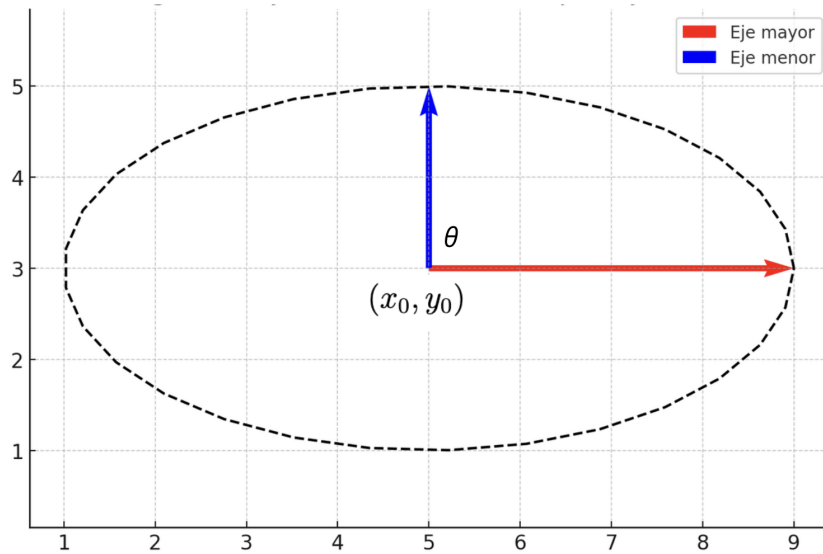


Figura 16. Parámetros geométricos obtenidos del ajuste elíptico

Interpretación clínica: diámetros tumorales

Los ejes de la elipse ajustada se interpretan como **estimaciones geométricas** del diámetro mayor y menor de la lesión. En aplicaciones clínicas, estas medidas son esenciales para:

- Clasificar el tamaño tumoral según criterios como RECIST o mRECIST.
- Evaluar la respuesta a tratamientos antitumorales.
- Comparar dimensiones en distintos momentos temporales.

Para convertir estos valores de píxeles a milímetros, se multiplica cada eje por el espaciado físico del píxel:

$$\text{Diámetro real} = \text{Diámetro en píxeles} \times \text{Tamaño de píxel (mm)} \quad (6)$$

Esto permite expresar los diámetros en unidades clínicas estándar (mm), lo que facilita la interpretación del impacto clínico.

7.6.2. `cv2.minAreaRect()`: Ajuste de rectángulo mínimo rotado sobre lesiones segmentadas

Concepto del Rectángulo Delimitador Orientado (OBB)

Para caracterizar geométricamente una lesión segmentada en una imagen médica, una técnica ampliamente utilizada es el cálculo del **rectángulo de área mínima**

rotado, que permite estimar las dimensiones máximas de la región, adaptándose a su orientación. En esta tesis, se utiliza la función `cv2.minAreaRect()` de OpenCV, que calcula el rectángulo de menor área posible que encierra todos los puntos del contorno de la lesión, sin estar restringido a la alineación con los ejes de la imagen.

Este rectángulo puede estar rotado, lo que permite adaptarse mejor a la forma y orientación real de la lesión. A diferencia del rectángulo alineado con los ejes (`boundingRect`), el rectángulo mínimo rotado minimiza el área contenida, proporcionando una envolvente más ajustada y precisa.

El resultado de `cv2.minAreaRect()` es una estructura denominada `RotatedRect`, que contiene:

- Centro del rectángulo: (x_0, y_0)
- Dimensiones del rectángulo: ancho w y alto h (en píxeles)
- Ángulo de rotación θ respecto al eje horizontal de la imagen

El contorno (línea fina) se encierra, primero, con un rectángulo alineado a los ejes (`cv2.boundingRect`, línea discontinua), que incluye área vacía adicional, y luego con un rectángulo de mínima área rotado (`cv2.minAreaRect`, línea continua), que se ajusta a la orientación real de la lesión y minimiza el área envolvente.

En la **Figura 17** se indican el centro (x_0, y_0) , el ancho w , el alto h y el ángulo de rotación θ devueltos por `cv2.minAreaRect()`, ilustrando cómo este método proporciona una envolvente geoméricamente más precisa para caracterizar la lesión.

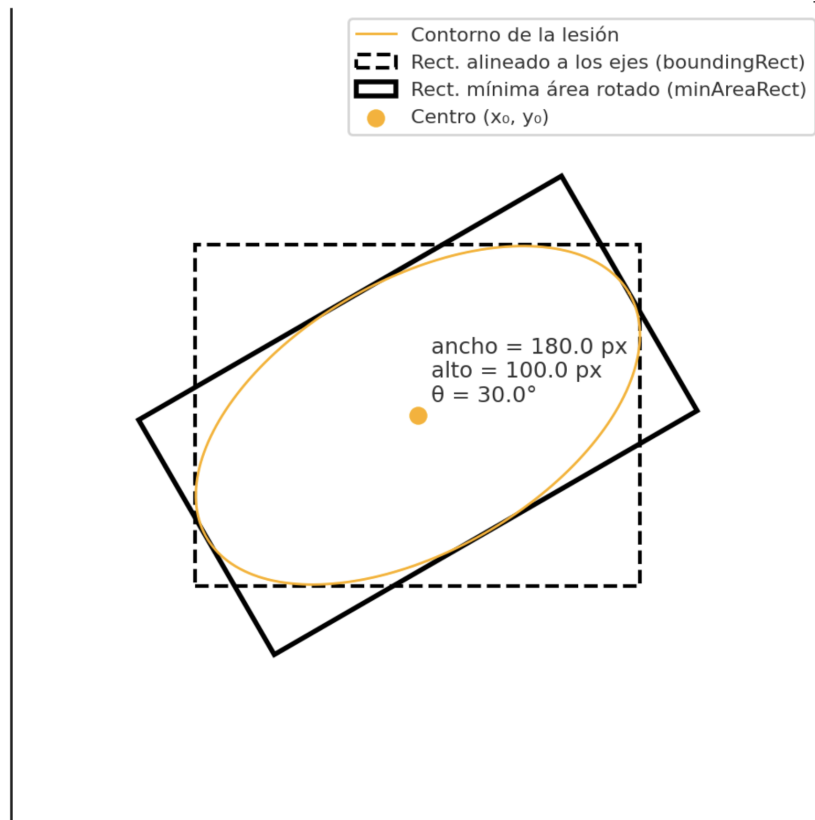


Figura 17. Rectángulo de mínima área rotado frente a rectángulo alineado a los ejes sobre el contorno de una lesión simulada.

Fundamento geométrico y algoritmo

Internamente, `cv2.minAreaRect()` implementa el algoritmo conocido como *rotating calipers* o calibradores rotantes. Este procedimiento analiza múltiples orientaciones candidatas y selecciona aquella que produce el rectángulo de área mínima.

El algoritmo funciona de la siguiente manera:

1. Se calcula primero el **casco convexo** del conjunto de puntos del contorno, ya que los puntos extremos determinan la envolvente mínima.
2. Para cada arista del casco convexo, se calcula el rectángulo que la contiene y cuya base está alineada con dicha arista.
3. Se calcula el área de cada rectángulo propuesto y se conserva el de menor área total.

Las **Figura 18 y 19** ilustran visualmente estos pasos, mostrando el casco convexo, los rectángulos candidatos y el rectángulo final de mínima área que envuelve la lesión, tanto en lesiones simples (elipsoidales) como en complejas (irregulares).

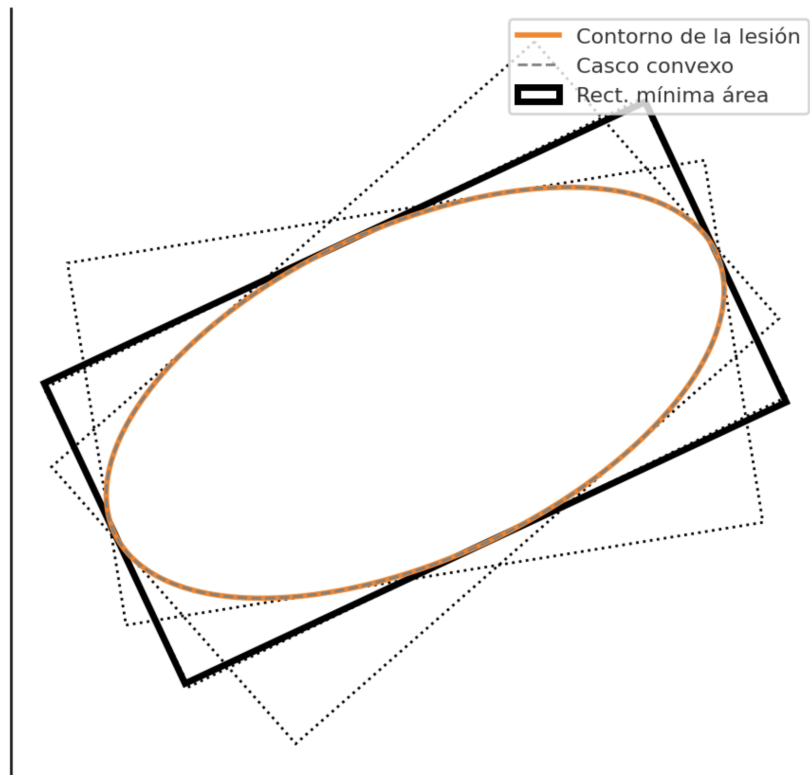


Figura 18. Ejemplo del algoritmo de calibradores rotantes (`cv2.minAreaRect()`). Sobre el contorno segmentado (naranja), se dibuja el casco convexo (gris) y los rectángulos candidatos generados al rotar *calipers* sobre cada arista (líneas punteadas). El rectángulo negro sólido corresponde al rectángulo de mínima área finalmente seleccionado, que envuelve la lesión con el menor área posible y refleja su orientación real.

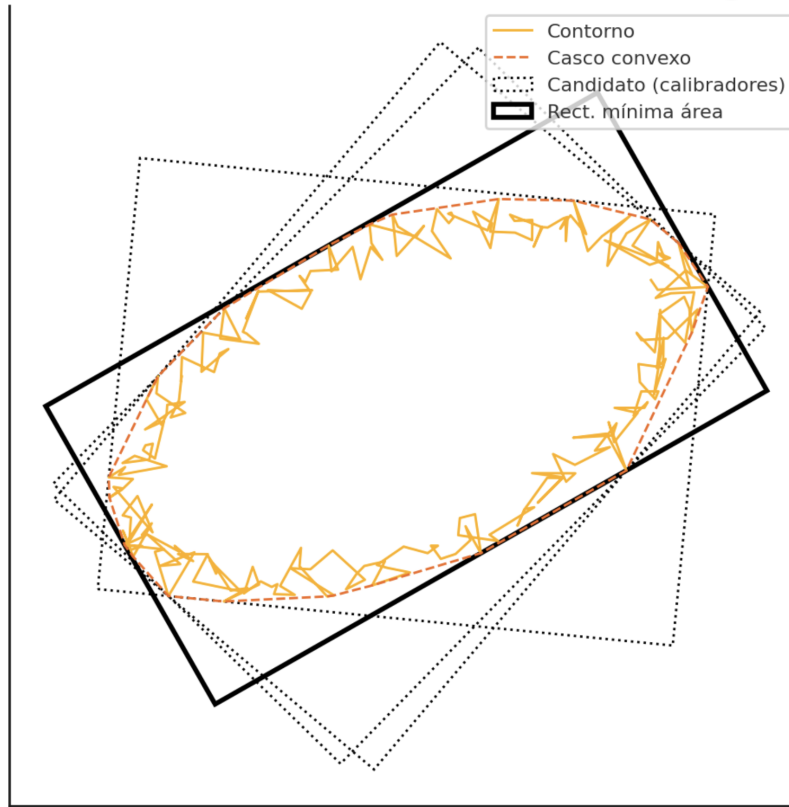


Figura 19. Demostración gráfica del algoritmo de *calibradores rotantes* (*rotating calipers*) empleado por `cv2.minAreaRect()` en una lesión irregular. El contorno de la lesión (línea fina) se reduce primero a su *casco convexo* (línea discontinua). Sobre cada arista del casco se generan *rectángulos candidatos* (líneas punteadas) alineando su base con la arista correspondiente. El algoritmo evalúa el área de cada candidato y selecciona el de *mínima área* (rectángulo negro sólido), constituyendo así la envolvente rotada óptima de la región.

Parámetros geométricos extraídos

Del `RotatedRect` resultante se obtienen directamente dos parámetros clave:

- **Diámetro mayor estimado:** el lado más largo del rectángulo ($\max(w, h)$)
- **Diámetro menor estimado:** el lado más corto del rectángulo ($\min(w, h)$)

Ambos valores se expresan inicialmente en píxeles. Para convertirlos a milímetros (unidad estándar en imágenes médicas), se multiplican por el tamaño físico del píxel, extraído del encabezado del archivo NIfTI:

$$\text{Diámetro real} = \text{Diámetro en píxeles} \times \text{Tamaño de píxel (mm)} \quad (7)$$

Utilidad clínica y robustez del método

Este método es especialmente robusto para lesiones de forma irregular, ya que no requiere que la forma sea convexa ni simétrica. Detecta automáticamente la orientación predominante y la extensión máxima de la lesión. Su aplicación clínica es relevante para:

- Medición reproducible de diámetros tumorales.
- Evaluación de la respuesta terapéutica.
- Comparación de tamaños entre distintos cortes o momentos.

Una ventaja adicional es que, en muchos casos, la orientación del rectángulo mínimo rotado coincide con la de la elipse de mejor ajuste, ya que ambos métodos tienden a alinearse con la dirección de máxima extensión de la lesión. Por lo tanto, `cv2.minAreaRect()` es una herramienta útil, rápida y precisa para la estimación de dimensiones tumorales a partir de imágenes segmentadas, especialmente cuando la morfología de la lesión es compleja o no elipsoidal.

7.6.3. `cv2.convexHull()`: Envoltente convexa y diámetro de Feret

La función `cv2.convexHull()` calcula el **envoltente convexa o polígono convexo mínimo** de un conjunto de puntos que definen el contorno de la lesión. Se puede imaginar como "estirar una banda elástica" alrededor del contorno: el polígono resultante es el menor convexo que contiene todos los puntos, eliminando irregularidades internas. OpenCV implementa `convexHull` usando el método de la cadena monotónica (algoritmo de Sklansky). Primero ordena los puntos por coordenada, luego recorre esta lista agregando vértices y descartando aquellos que crearían giros hacia dentro.

Cálculo del diámetro de Feret

A partir del casco convexo, el *diámetro de Feret máximo* se define como la distancia euclídea más grande entre dos vértices del polígono convexo:

$$D_{\text{Feret}} = \max_{i,j} \|\mathbf{v}_i - \mathbf{v}_j\|, \quad (8)$$

donde $\mathbf{v}_i$ y $\mathbf{v}_j$ son vértices del casco.

La **Figura 20** ilustra ambos conceptos: muestra el contorno original, la envoltente convexa obtenida con `cv2.convexHull()` y la línea del diámetro de Feret con su longitud en píxeles, destacando cómo esta estrategia simplifica la geometría de la lesión y permite una medición consistente.

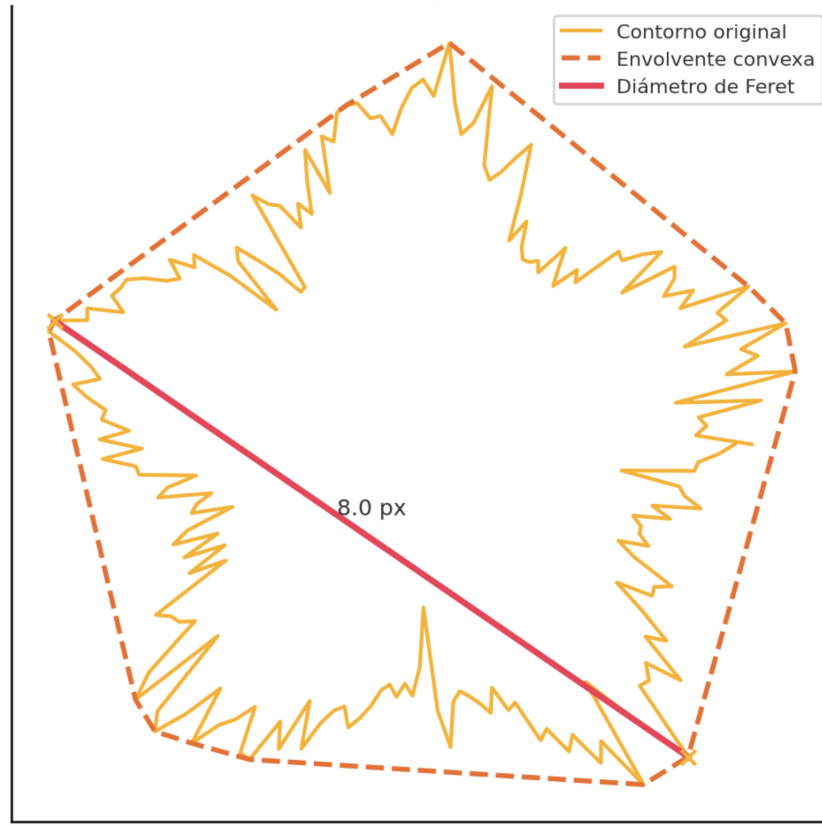


Figura 20. Ajuste de envoltura convexa en lesión tumoral irregular. Contorno original de la lesión (línea amarilla), su envoltura convexa calculada mediante `cv2.convexHull()` (línea naranja discontinua) y el diámetro de Feret (línea roja), que corresponde a la máxima distancia entre dos puntos del casco convexo. El valor anotado indica la longitud en píxeles, ilustrando cómo la envoltura convexa “suaviza” irregularidades internas y facilita la medición de dimensiones críticas de la lesión.

Medición del diámetro mayor en milímetros

El valor D_{Feret} en píxeles se convierte a milímetros multiplicando por el espacio físico del píxel (extraído del encabezado NIfTI):

$$\text{Diámetro real (mm)} = D_{\text{Feret}} \times \text{Tamaño de píxel (mm)}. \quad (9)$$

7.6.4. Conversión de unidades y formato NIfTI

Conversión de píxeles a milímetros

En imágenes médicas tridimensionales (como las obtenidas mediante tomografía computarizada, TC), cada volumen se almacena junto con metadatos espaciales que describen el tamaño real de cada píxel o vóxel. En este trabajo se utilizaron volúmenes en formato *NIfTI* (.nii o .nii.gz), ampliamente empleado en radiología e investigación biomédica, ya que permite integrar los datos de imagen (matriz 3D de intensidades) con información espacial en el encabezado del archivo.

Este encabezado contiene parámetros como `pixdim[1]`, `pixdim[2]` y `pixdim[3]`, que corresponden a las dimensiones físicas del vóxel en los ejes X, Y y Z, respectivamente. Por ejemplo, un volumen con espaciado de $0.5 \times 0.5 \times 3.0$ mm indica que cada píxel en el plano axial representa 0.5 mm en las direcciones X e Y, mientras que la separación entre cortes (espesor de slice) es de 3.0 mm en el eje Z.

Toda medición inicial en imágenes segmentadas suele darse en unidades de píxeles. Para convertir a unidades reales (milímetros) se multiplica el número de píxeles por el tamaño físico de cada píxel. Por ejemplo, en un corte axial con espaciado isotrópico de 0.5 mm/píxel, una distancia de D píxeles se convierte directamente a $D \cdot 0.5$ mm.

En una imagen de tomografía computarizada (TC), cada píxel representa la proyección de un **vóxel** (volumetric pixel) del cuerpo. Si el vóxel tiene dimensiones iguales en los ejes X y Y ($s_x = s_y$), se considera **isotrópico**, y el píxel correspondiente es un cuadrado perfecto. Sin embargo, con mayor frecuencia el espaciado difiere entre ambos ejes ($s_x \neq s_y$), generando **vóxeles anisotrópicos**: el píxel pasa a ser un rectángulo cuyo lado horizontal mide s_x mm y el vertical s_y mm. Esta asimetría obliga a ponderar cada eje con su propio factor de escala al convertir distancias de píxeles a milímetros.

Así, si la línea que se desea medir recorre Δx píxeles en el eje X y Δy píxeles en el eje Y, la longitud física L se calcula mediante la distancia euclídea ponderada:

$$L = \sqrt{(\Delta x \cdot s_x)^2 + (\Delta y \cdot s_y)^2}. \quad (10)$$

Cuando la diferencia entre s_x y s_y es mínima ($s_x \approx s_y$), puede utilizarse un único factor de escala—por ejemplo, el promedio $(s_x + s_y)/2$ —sin introducir un error clínicamente relevante. Una vez hecha la conversión, los diámetros mayor y menor medidos en píxeles se expresan en milímetros, de modo que las dimensiones reportadas reflejen la anatomía real y sean comparables con las mediciones convencionales.

Implementación con NiBabel y OpenCV

El siguiente procedimiento describe cómo, a partir de un volumen *NIfTI* que contiene la máscara segmentada de un estudio de TC, se extraen los diámetros axial mayor y menor de cada lesión en unidades físicas: primero se leen la imagen y sus metadatos espaciales, después se identifican los contornos de cada lesión en cada corte 2D y, finalmente, se convierten las dimensiones obtenidas de píxeles a milímetros según el espaciado real de vóxel. El proceso garantiza que las mediciones geométricas reflejen la anatomía con precisión y sean comparables con los estándares clínicos (p. ej., RECIST).

1. **Lectura del volumen y metadatos.** Cargar el archivo *NIfTI* y extraer, del *header*, la tupla de espaciados de vóxel $\text{zoom} = (s_x, s_y, s_z)$ en milímetros.
2. **Recorrido de cortes axiales.** Iterar sobre cada índice axial k ; cada iteración proporciona la imagen 2-D `mask_slice = mask_data[:, :, k]`, convertida a binaria (`uint8`) con valores 255 (lesión) y 0 (fondo).

3. **Identificación de etiquetas.** Obtener los valores únicos distintos de 0 en `mask_slice`; cada valor corresponde a una lesión. Para cada etiqueta generar la máscara booleana `segment_mask`.
4. **Detección del contorno externo.** Aplicar un detector de contornos y seleccionar, si hubiese varios, el de mayor área.
5. **Rectángulo mínimo y dimensiones en píxeles.** Ajustar el rectángulo de mínima área que circunscribe el contorno; anotar width_{px} y $\text{height}_{\text{px}}$.
6. **Conversión a milímetros.**

$$\text{width}_{\text{mm}} = \text{width}_{\text{px}} \times s_x, \quad \text{height}_{\text{mm}} = \text{height}_{\text{px}} \times s_y. \quad (11)$$

7. **Cálculo de diámetros.**

$$D_{\text{mayor}} = \text{máx}(\text{width}_{\text{mm}}, \text{height}_{\text{mm}}), \quad (12)$$

$$D_{\text{menor}} = \text{mín}(\text{width}_{\text{mm}}, \text{height}_{\text{mm}}). \quad (13)$$

8. **Actualización del registro por lesión.** Si el diámetro actual D_{mayor} supera el almacenado para la etiqueta, actualizar el valor y el índice de corte k .
9. **Resultado final.** Tras procesar todos los cortes, un diccionario recoge, para cada lesión, su diámetro axial máximo (y opcionalmente el menor) en milímetros y el corte donde se midió. El mismo algoritmo se aplica en los planos coronal y sagital usando los espaciados correspondientes.

7.6.5. Generación y validación de lesiones tumorales sintéticas con OpenCV

Generación de Datos Sintéticos

Para disponer de un banco de pruebas controlado, se diseñaron imágenes binarias que representan tres morfologías tumorales ideales—circular, elipsoidal e irregular—utilizando primitivas geométricas de `OpenCV`. El diámetro real de cada lesión sintética se fijó a priori con los parámetros de las primitivas, lo que permite comparar directamente la medida “verdadera” con la estimación del algoritmo. El procedimiento se resume a continuación:

- 1: **Definición del lienzo.** Se crea una imagen binaria de tamaño $N \times N$ píxeles inicializada a cero.
- 2: **Círculo.** Con centro (x_c, y_c) y radio R , los píxeles internos se fijan a 255; el diámetro de referencia vale $2R$.
- 3: **Elipse.** Con centro (x_c, y_c) , semiejes (a, b) y ángulo α , se asigna valor 255 a toda elipse completa, quedando diámetros mayores $2a$ y $2b$.
- 4: **Forma irregular.** Se definen vértices $\{\mathbf{p}_i\}$ bajo restricciones de conectividad, se traza y rellena el polígono. Su diámetro de referencia se obtiene con el rectángulo mínimo o el casco convexo.

Evaluación geométrica

A partir de las máscaras sintéticas se validan los métodos geométricos del algoritmo:

5: Contorno externo. Se detecta el contorno de cada máscara y se conserva el de mayor área.

6: Rectángulo mínimo. Sobre el contorno se obtiene el rectángulo de área mínima que lo circunscribe, con dimensiones width_{px} y $\text{height}_{\text{px}}$.

7: Conversión a milímetros. Sea (s_x, s_y) el espaciado de vóxel; entonces

$$\text{width}_{\text{mm}} = \text{width}_{\text{px}} \times s_x, \quad \text{height}_{\text{mm}} = \text{height}_{\text{px}} \times s_y. \quad (14)$$

8: Diámetros mayor y menor. Definidos como

$$D_{\text{mayor}} = \max(\text{width}_{\text{mm}}, \text{height}_{\text{mm}}), \quad D_{\text{menor}} = \min(\text{width}_{\text{mm}}, \text{height}_{\text{mm}}). \quad (15)$$

9: Errores de medición. Para n ejemplos sintéticos se calcula

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{D}_i - D_{\text{ref},i}|, \quad \text{MAPE} = \frac{100}{n} \sum_{i=1}^n \frac{|\hat{D}_i - D_{\text{ref},i}|}{D_{\text{ref},i}}. \quad (16)$$

10: Selección del método. El algoritmo con menor MAE y MAPE se adopta como referencia para las mediciones automáticas en los volúmenes clínicos.

Las ecuaciones (14)–(16) siguen la numeración global del documento y permiten referenciar con consistencia los resultados presentados en secciones posteriores.

La **Figura 21** ilustra visualmente cómo se generaron estas morfologías mediante OpenCV.

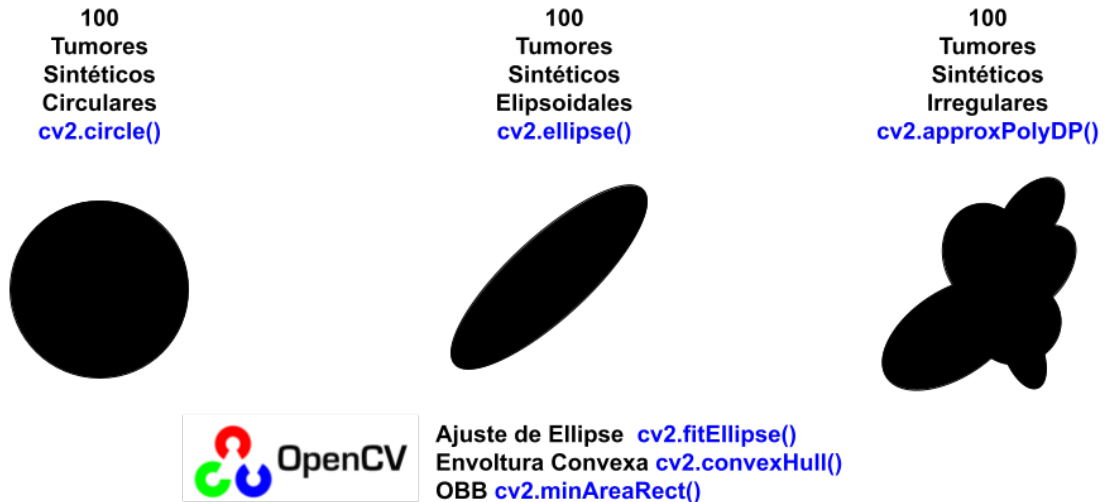


Figura 21. Generación del conjunto sintético de 300 lesiones empleadas para la validación. 100 tumores *circulares* (`cv2.circle()`), 100 *elipsoidales* (`cv2.ellipse()`) y 100 *irregulares* (`cv2.approxPolyDP()`). En la parte inferior se listan los métodos evaluados: ajuste de **Elipse** (`cv2.fitEllipse()`), **OBB** (`cv2.minAreaRect()`) y **Convex Hull** (`cv2.convexHull()`).

Cálculo de Ejes y Validación Geométrica

Una vez generados los contornos sintéticos, se calcularon los diámetros mayores y menores utilizando funciones geométricas de OpenCV (**Figura 22**):

- **Ajuste Elíptico** (`cv2.fitEllipse`): ajusta una elipse óptima, minimizando distancias cuadráticas a los puntos del contorno.
- **Rectángulo de Área Mínima** (`cv2.minAreaRect`): obtiene el rectángulo rotado de menor área que encierra el contorno.
- **Casco Convexo** (`cv2.convexHull`): determina la envolvente convexa del contorno; luego se aplica `minAreaRect` sobre esta envolvente para obtener diámetros.

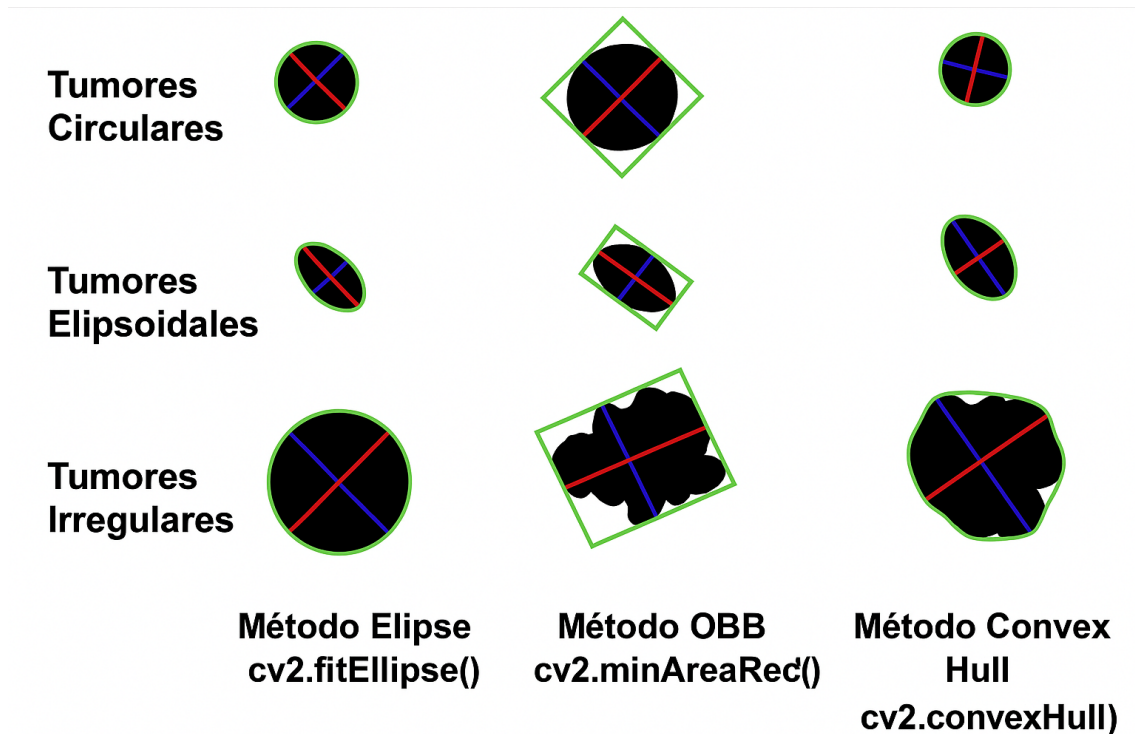


Figura 22. Comparación de tres métodos geométricos para estimar los diámetros mayor (rojo) y menor (azul) en tumores de distinta morfología. Las columnas corresponden a: (i) **Elipse** (`cv2.fitEllipse()`), (ii) **OBB** (`cv2.minAreaRect()`) y (iii) **Convex Hull** (`cv2.convexHull()`); las filas muestran tumores *circulares*, *elipsoidales* e *irregulares*. La máscara binaria de la lesión aparece en blanco, la envolvente calculada en verde y los ejes ortogonales en rojo y azul. Obsérvese cómo cada método ajusta la envolvente según la complejidad del contorno.

Resultados de Validación

En general, para formas simples (círculo y elipse), los métodos `fitEllipse` y `minAreaRect` recuperan con alta precisión los diámetros verdaderos, obteniendo

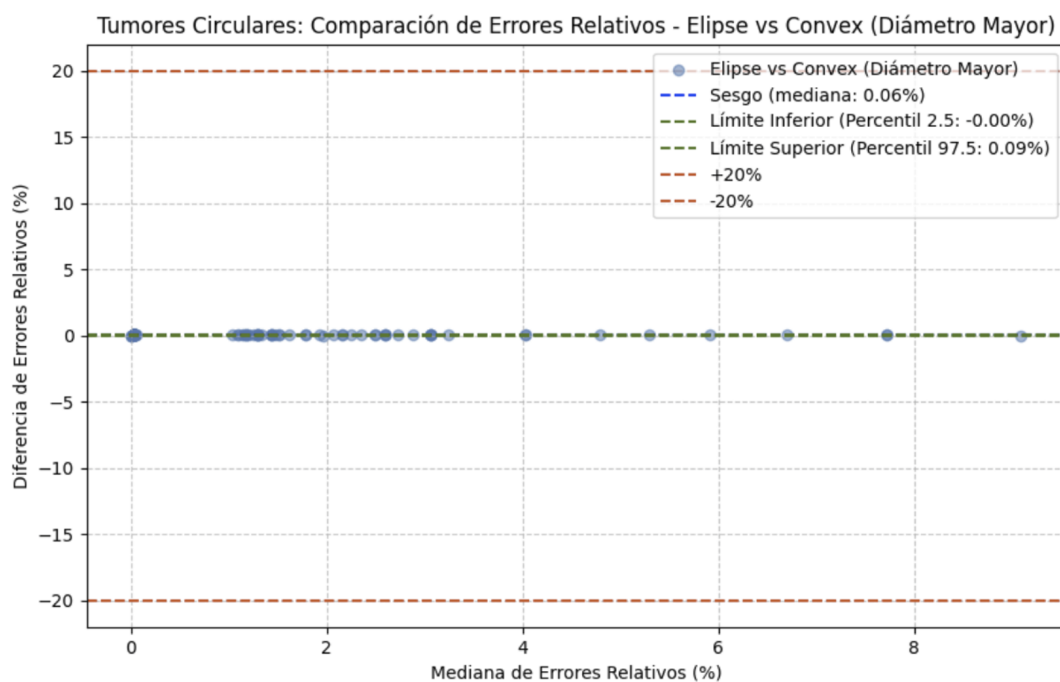
errores mínimos (MAE y MAPE cercanos a cero). Sin embargo, para formas irregulares, se observan mayores discrepancias debido a la dificultad inherente de ajustar estas geometrías a modelos elípticos o rectangulares. La **Tabla 8** ilustra estos resultados.

Tabla 8. Comparación de diámetros verdaderos y estimados (en píxeles).

Forma	Referencia	fitEllipse	minAreaRect
Círculo (radio=100)	200, 200	199.17, 199.17	199.40, 199.40
Elipse (160×80)	160, 80	159.99, 79.95	160.50, 80.62
Irregular	215, 200	213.08, 193.71	213.00, 187.48

Los resultados de *MAE*, *MAPE*, sesgo y límites de acuerdo para los tres conjuntos de tumores sintéticos —circulares, elipsoidales e irregulares— se presentan en las **Figuras 23–31**. Cada una contrasta dos algoritmos mediante un diagrama de Bland–Altman y muestra, bajo la gráfica, la tabla con las métricas numéricas correspondientes.

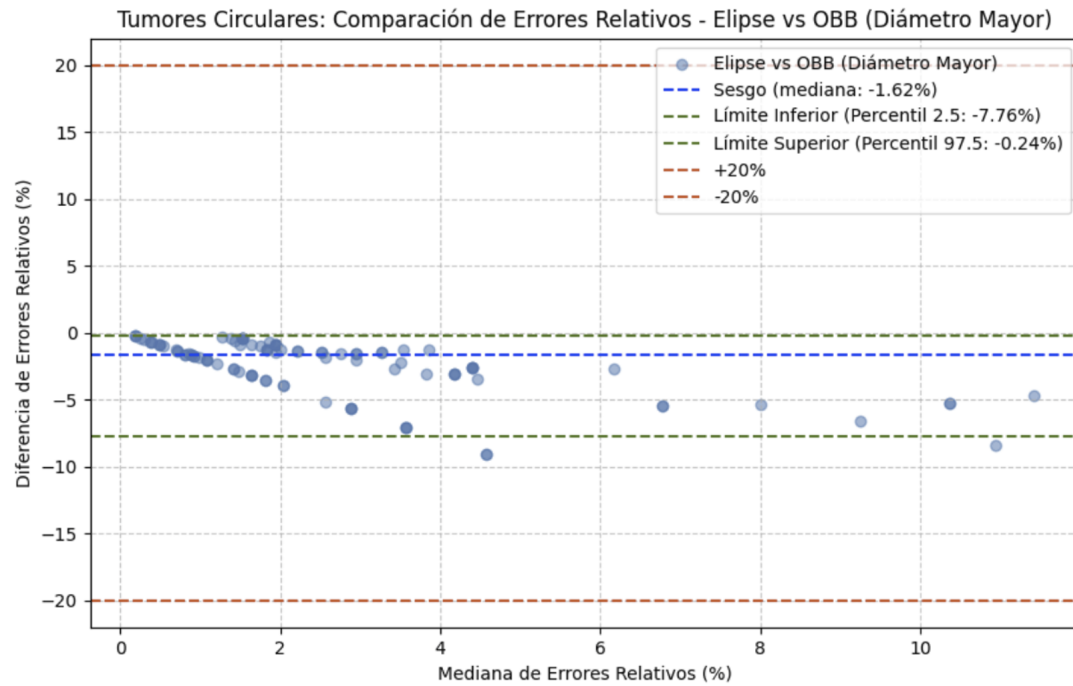
En primer lugar, la **Figura 23** compara Elipse y Convex Hull sobre tumores circulares. Los puntos se agrupan alrededor del cero (sesgo mediano 0,06 %); los límites de acuerdo (−0,09 % a 0,09 %) quedan muy por debajo del umbral clínico de ± 20 %. Las métricas (MAE 0,76 mm; MAPE 1,17 %) confirman que ambos métodos son prácticamente equivalentes en contornos perfectamente redondos.



MAE M1	MAPE M1	MAE M2	MAPE M2	Sesgo ($\mu_{1/2}$)	IQR	IC95% Inf	IC95% Sup	<i>p</i> -valor
0.76 mm	1.17%	0.76 mm	1.11%	0.06%	0.002%	0.053%	0.059%	<0.05

Figura 23. Comparación entre métodos *elipse* y *convex hull* en tumores sintéticos circulares.

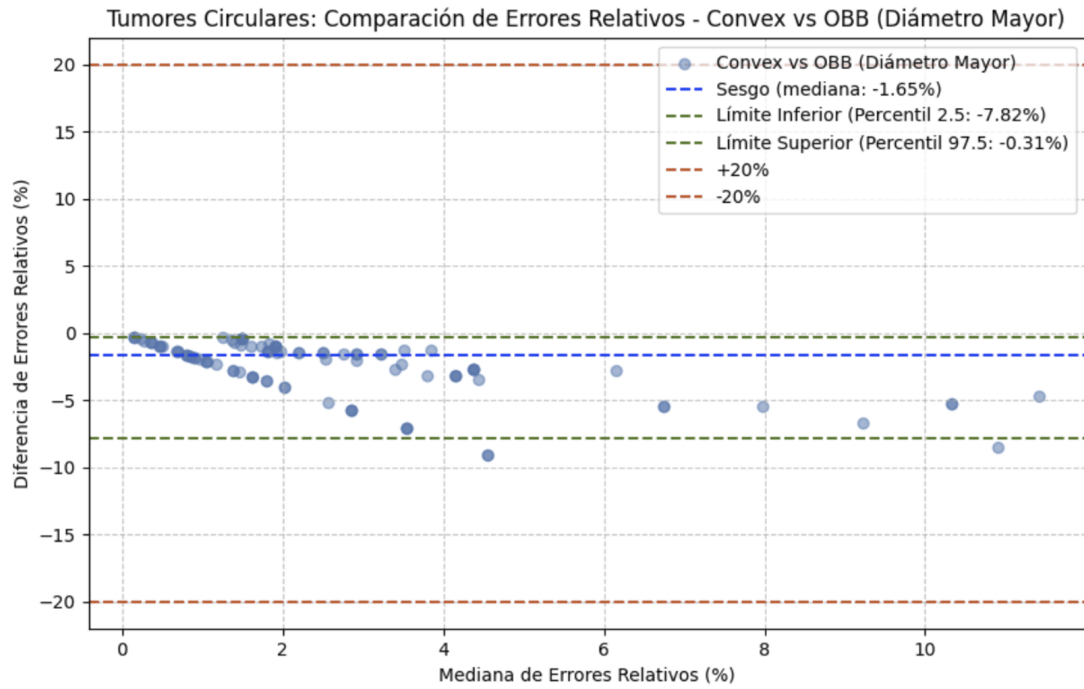
La **Figura 24** mantiene el escenario circular, pero enfrenta Elipse con el rectángulo de mínima área (*OBB*). Aparece una ligera tendencia negativa (sesgo $-1,62\%$; límites $-7,76\%$ a $-0,24\%$), señal de que Elipse sobrestima el diámetro; aun así, todas las diferencias permanecen en el rango clínico y son estadísticamente significativas.



MAE M1	MAPE M1	MAE M2	MAPE M2	Sesgo ($\mu_{1/2}$)	IQR	IC95% Inf	IC95% Sup	p-valor
0.76 mm	1.17%	1.05 mm	2.83 %	-1.62%	2.19%	-2.84%	-2.02%	<0.05

Figura 24. Comparación entre métodos elipse y OBB en tumores sintéticos circulares.

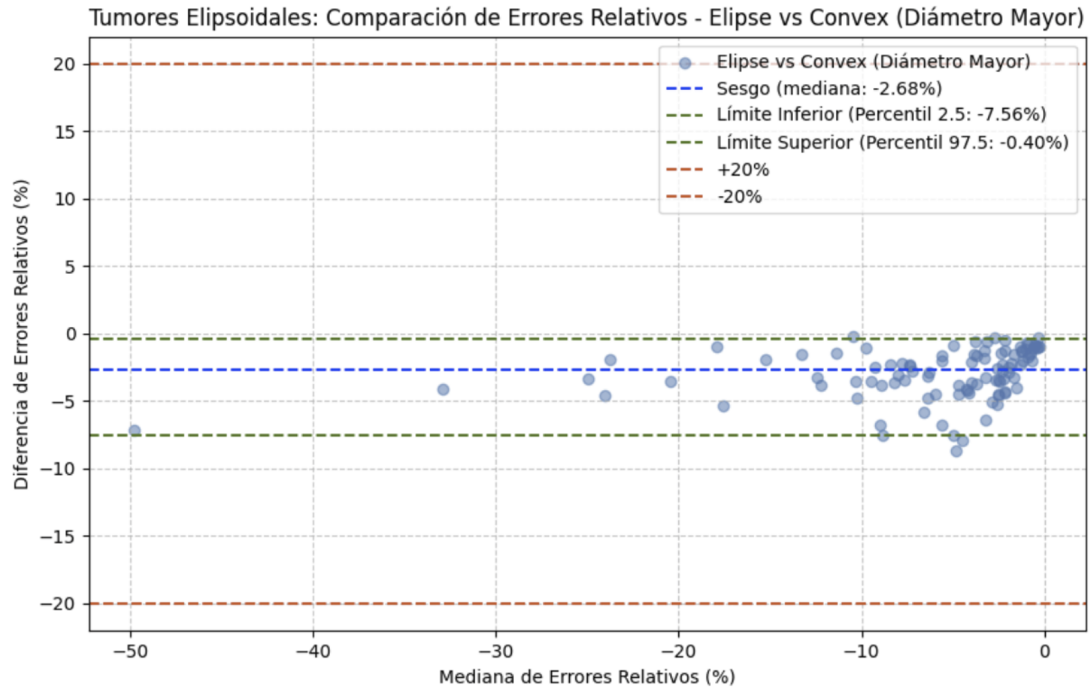
Para cerrar el bloque circular, la **Figura 25** contrasta Convex Hull y OBB. El sesgo (-1,65 %) y los límites (-7,82 % a -0,31 %) son similares, pero el MAE vuelve a favorecer a OBB (0,71 mm frente a 0,76 mm), subrayando su mejor rendimiento.



MAE M1	MAPE M1	MAE M2	MAPE M2	Sesgo ($\mu_{1/2}$)	IQR	IC95% Inf	IC95% Sup	p-valor
0.76 mm	1.11 %	1.05 mm	2.83 %	-1.65%	2.18%	-2.90%	-2.08%	<0.05

Figura 25. Comparación entre métodos convex hull y OBB en tumores sintéticos circulares.

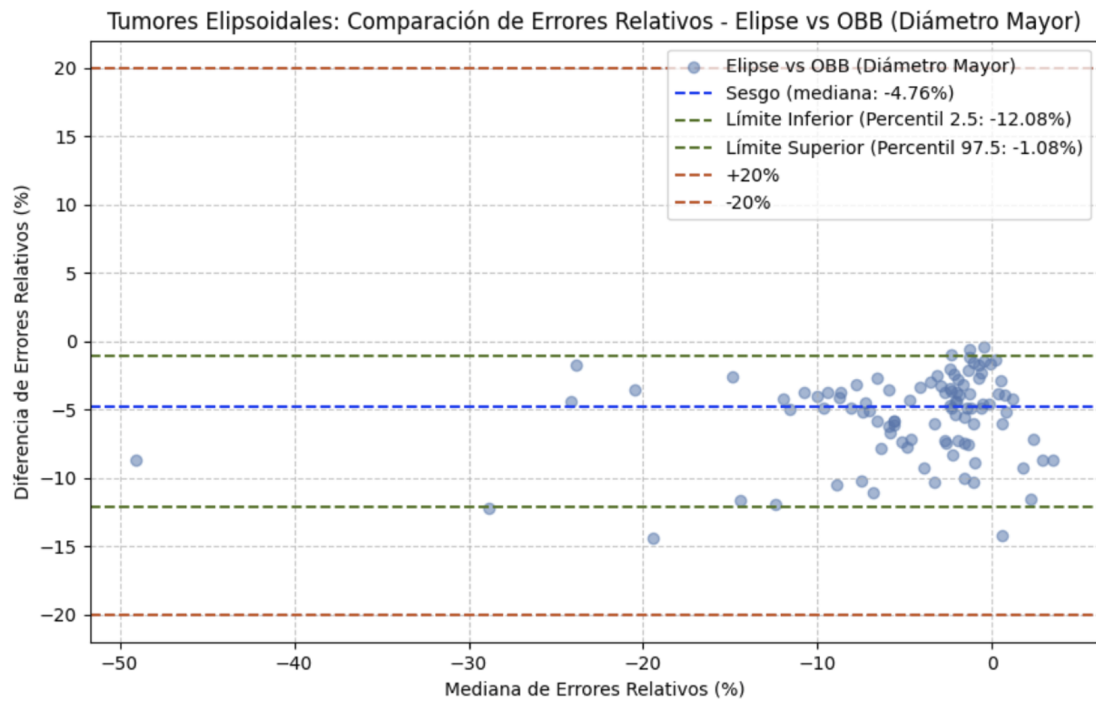
Al pasar al conjunto *elipsoidal*, la **Figura 26** muestra que Elipse subestima frente a Convex Hull (sesgo $-2,68\%$; límites $-7,56\%$ a $-0,40\%$), aunque ambos métodos mantienen errores pequeños y dentro de $\pm 20\%$.



MAE M1	MAPE M1	MAE M2	MAPE M2	Sesgo ($\mu_{1/2}$)	IQR	IC95% Inf	IC95% Sup	p-valor
0.50 mm	1.19%	0.49 mm	1.12%	-1.34%	1.71%	-2.41%	-1.57%	<0.05

Figura 26. Comparación entre métodos elipse y convex hull en tumores sintéticos elipsoidales.

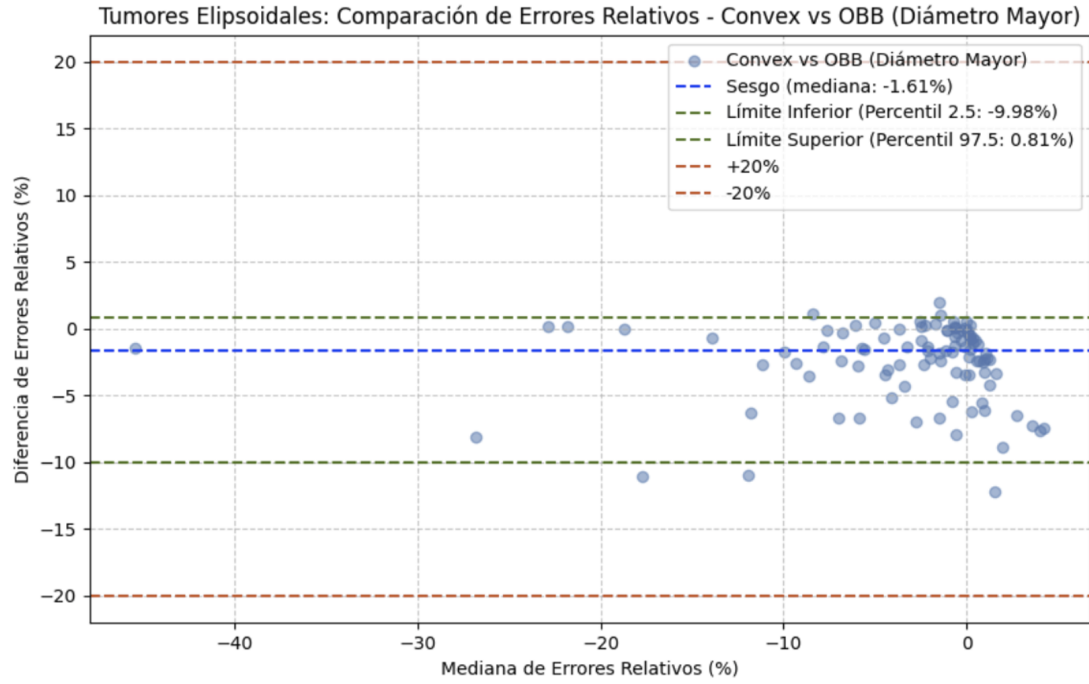
La **Figura 27** revela que la subestimación se acentúa al comparar Elipse con OBB (sesgo $-4,76\%$; límites $-12,08\%$ a $-1,08\%$). OBB reduce el MAE a $0,71$ mm (vs. $1,03$ mm), confirmando su mayor precisión.



MAE M1	MAPE M1	MAE M2	MAPE M2	Sesgo ($\mu_{1/2}$)	IQR	IC95% Inf	IC95% Sup	p-valor
0.71 mm	2.92%	1.03 mm	5.91%	-8.8%	10.75%	-14.26%	-10.54%	<0.05

Figura 27. Comparación entre métodos elipse y OBB en tumores sintéticos elipsoidales.

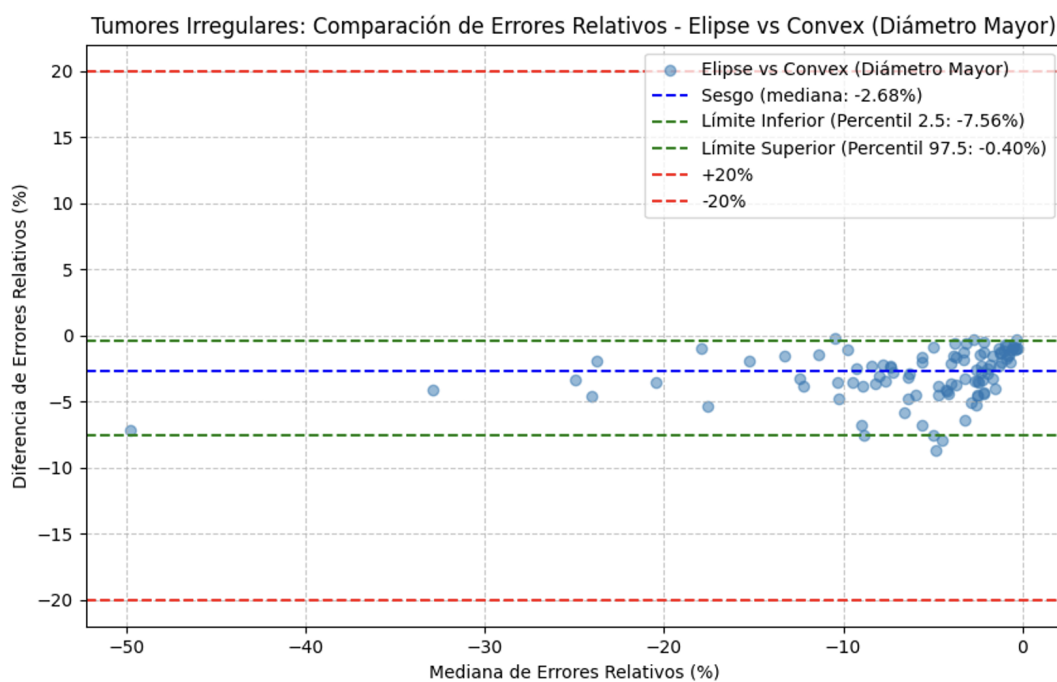
La **Figura 28** demuestra la concordancia entre Convex Hull y OBB (sesgo -1,61 %; límites -9,98 % a 0,81 %), con MAE muy similares (0,49 mm vs. 0,41 mm) y dispersión mínima.



MAE M1	MAPE M1	MAE M2	MAPE M2	Sesgo ($\mu_{1/2}$)	IQR	IC95% Inf	IC95% Sup	p-valor
0.49 mm	1.12%	0.41 mm	0.73%	0.18 %	1.52%	0.01%	0.76%	<0.05

Figura 28. Comparación entre métodos convex hull y OBB en tumores sintéticos elipsoidales.

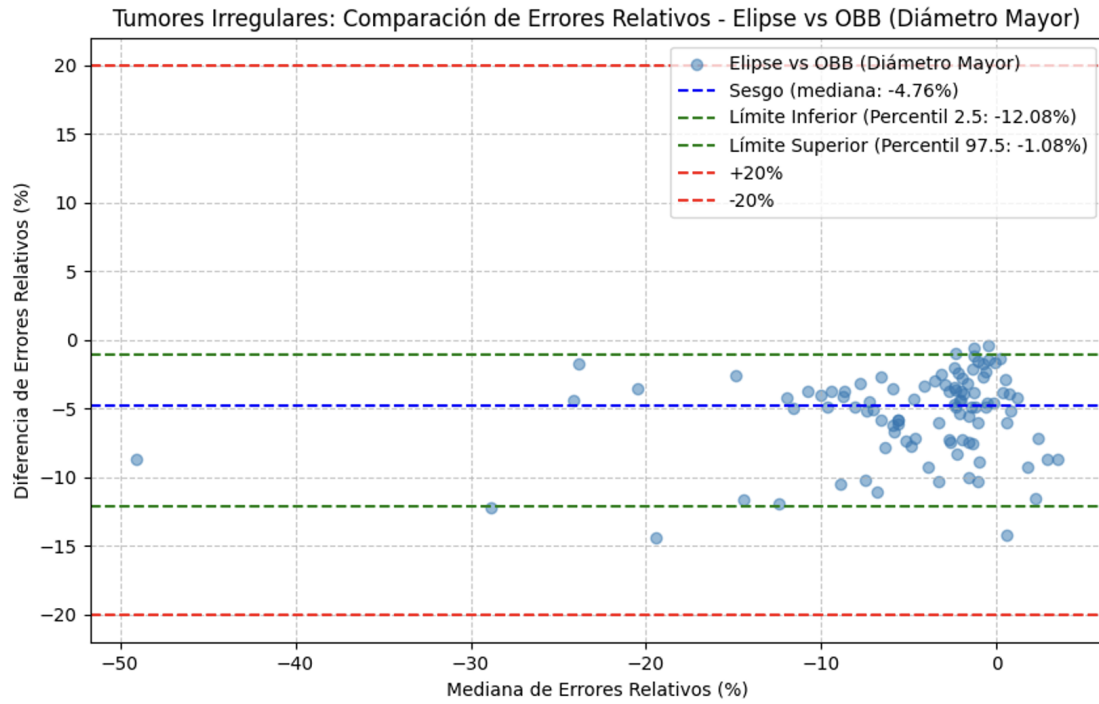
El análisis concluye con las lesiones sintéticas *irregulares*. En la **Figura 29**, Elipse arroja errores altos (MAE 7,54 mm; MAPE 5,38 %) frente a Convex Hull (3,08 mm; 2,16 %), pese a que el sesgo (-2,68 %) sigue dentro del margen clínico.



MAE M1	MAPE M1	MAE M2	MAPE M2	Sesgo ($\mu_{1/2}$)	IC95% Inf.	IC95% Sup.	p-valor
7.54 mm	5.38%	3.08 mm	2.16%	2.55%	2.17%	3.31%	<0.05

Figura 29. Comparación entre métodos elipse y convex hull en tumores sintéticos irregulares

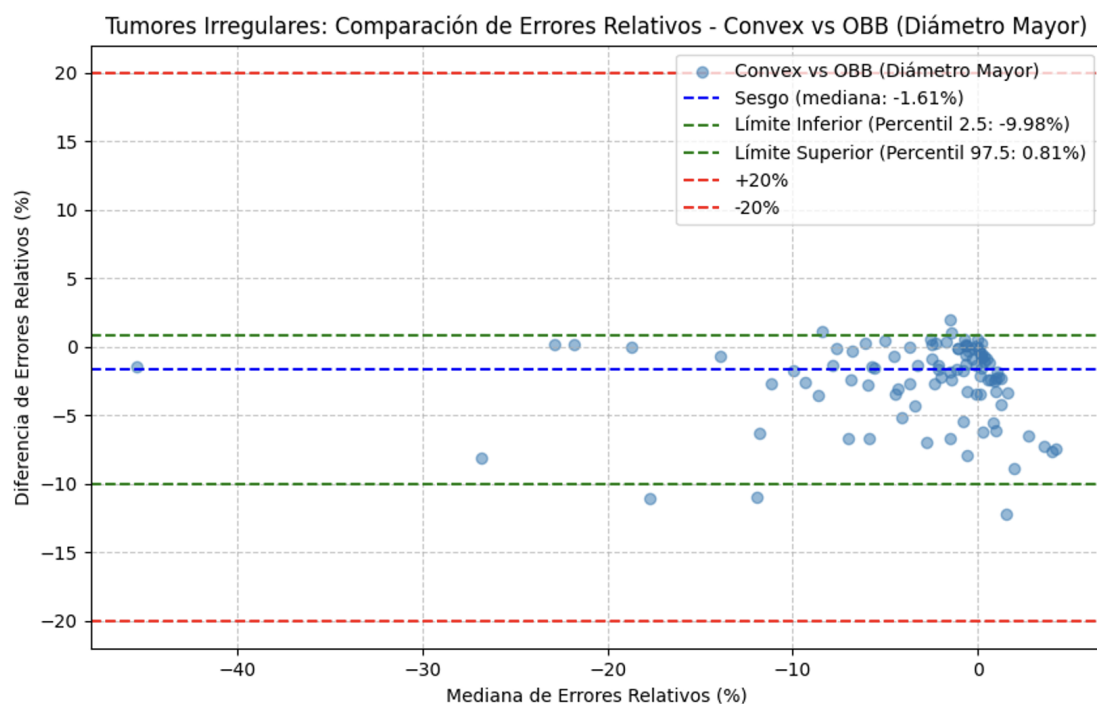
La **Figura 30** confirma la ventaja de OBB: MAE 3,20 mm y MAPE 2,20 % frente a 7,54 mm y 5,38 % de Elipse, con un sesgo de -4,76 % en límites aceptables.



MAE M1	MAPE M1	MAE M2	MAPE M2	Sesgo ($\mu_{1/2}$)	IC95% Inf.	IC95% Sup.	p-valor
7.54 mm	5.38 %	3.20 mm	2.20 %	-4.76%	0.76%	3.87%	<0.05

Figura 30. Comparación entre métodos elipse y OBB en tumores sintéticos irregulares

Por último, la **Figura 31** compara Convex Hull y OBB; el sesgo (-1,61 %) y los límites (-9,98 % a 0,81 %) evidencian elevada concordancia y MAE similares (3 mm), aunque OBB muestra intervalos de confianza más estrechos.



MAE M1	MAPE M1	MAE M2	MAPE M2	Sesgo ($\mu_{1/2}$)	IC95% Inf.	IC95% Sup.	p-valor
3.08 mm	2.16 %	3.20 mm	2.20 mm	-1.61%	-3.06%	-1.89%	<0.05

Figura 31. Comparación entre métodos convex hull y OBB en tumores sintéticos irregulares.

En síntesis, *OBB* exhibe los *MAE* y *MAPE* más bajos en siete de las nueve comparaciones, mantiene sesgos reducidos y presenta la dispersión más compacta de errores. Además, **todas las discrepancias cuantificadas permanecen dentro de los márgenes de variación ($\pm 20\%$ del diámetro mayor) que, según RECIST 1.1, serían necesarios para cambiar la categoría de respuesta tumoral**; por tanto, se consideran clínicamente no significativas. Estas ventajas sustentan la elección de *OBB* como **método de referencia** para el resto del estudio.

7.6.6. Medición de diámetros tumorales en TC con OpenCV en planos axial, sagital y coronal

Fundamentos y planos de análisis

El presente apartado describe los procedimientos implementados para cuantificar el tamaño y la carga tumoral a partir de las segmentaciones obtenidas en TC. Se adoptaron tres aproximaciones complementarias—1D, 2.5D y 3D—con el fin de contrastar métricas lineales tradicionales con mediciones geométricas derivadas por visión computacional y con la volumetría completa de la lesión. La **Figura 32** ofrece una visión general del flujo de trabajo y sirve de marco para los subapartados que siguen:

- **Medición 1D (RECIST 1.1).** Se emplea el diámetro máximo en el plano axial, siguiendo los criterios convencionales de respuesta en oncología.
- **Medición 2.5D.** Se calcula el rectángulo mínimo rotado (`cv2.minAreaRect`) sobre la proyección de la segmentación en los planos axial, coronal y sagital; de él se extraen los diámetros mayor y menor.
- **Cuantificación 3D.** Se determina el volumen total de la lesión a partir de la segmentación manual en cada *slice*.

En los subapartados siguientes se explican en detalle los algoritmos, las transformaciones espaciales y los factores de escala empleados para cada método, así como su papel en el análisis comparativo de respuesta terapéutica.

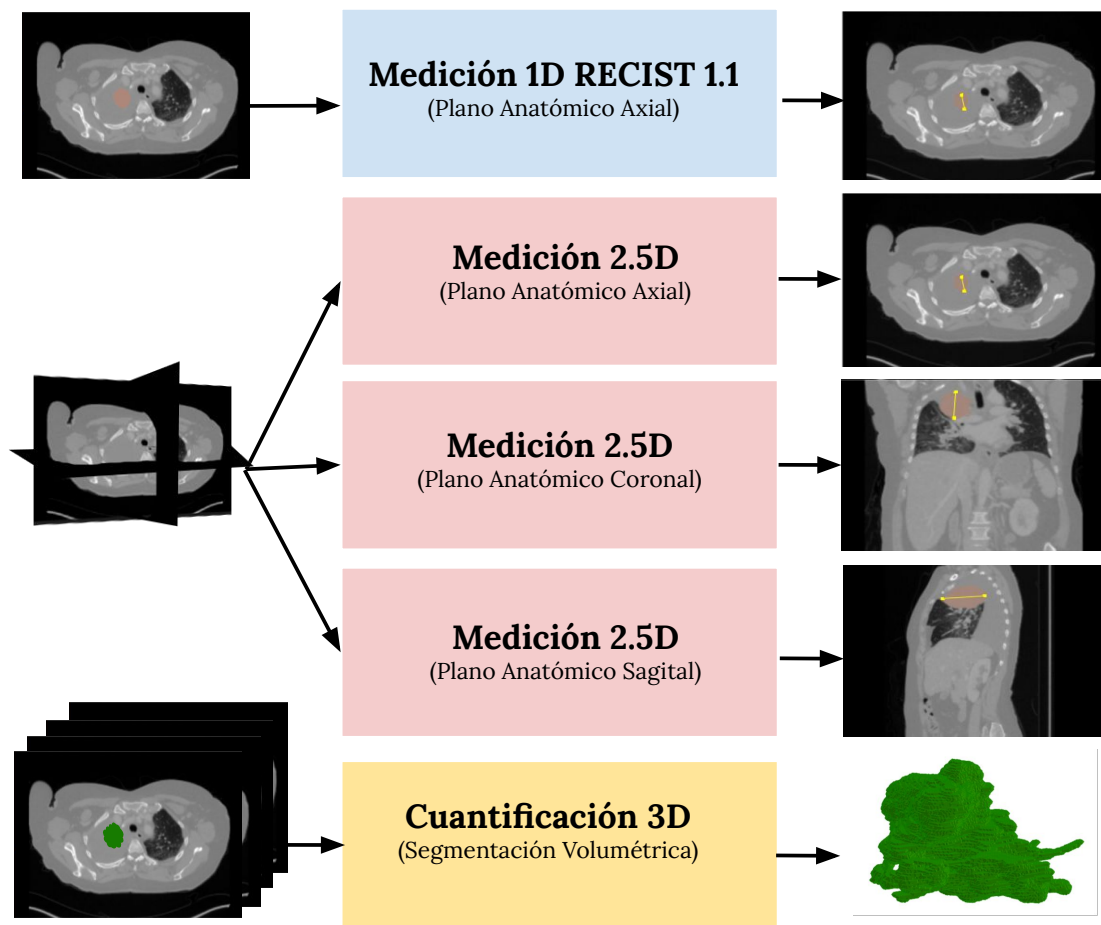


Figura 32. Métodos de medición de lesiones tumorales. Se presentan los métodos 1D, 2.5D y 3D mediante cortes de TC. La medición 1D se realiza según el diámetro mayor en el plano anatómico axial, conforme al protocolo RECIST 1.1 (color celeste). La medición en 2.5D emplea diámetros mayores en los planos axial, coronal y sagital (color rosado). La cuantificación 3D se efectúa a partir de la segmentación manual volumétrica en cada slice (color amarillo).

Flujo general del algoritmo en 2D

A continuación se presenta la descripción paso a paso del flujo general que el algoritmo sigue en un plano bidimensional determinado; este mismo esquema se

aplica de manera idéntica y autónoma a los tres planos ortogonales (axial, coronal y sagital):

1. **Selección del plano y extracción de cortes 2D:** Se fija el índice del volumen 3D correspondiente al plano deseado (Z para axial, X para sagital, Y para coronal) y se recorren todos los *slices* en ese eje. Cada *slice* es una imagen 2D de la máscara segmentada del tumor en ese plano.
2. **Preprocesamiento de la máscara:** La imagen de máscara del corte se convierte a formato binario 8-bit (valores 255 para el tumor, 0 fondo) y se puede aplicar una limpieza si es necesario (p.ej., eliminar pequeños artefactos). Esto garantiza que los contornos detectados correspondan a las lesiones relevantes.
3. **Detección de contornos:** Se aplica `cv2.findContours` para obtener los contornos del tumor en la máscara binaria. Usamos `RETR.EXTERNAL` para extraer solo el contorno externo (asumiendo que cada lesión es una región conectada) y `CHAIN_APPROX_SIMPLE` para simplificar la representación del contorno. Si hubiera múltiples contornos (p. ej. por fragmentación de la segmentación), se selecciona el contorno de mayor área como representativo de la lesión.
4. **Cálculo del rectángulo mínimo rotado:** Con el contorno principal, se invoca `cv2.minAreaRect`, obteniendo el rectángulo rotado de mínima área que lo envuelve. De este se extraen las dimensiones en píxeles (ancho y alto). El diámetro mayor en píxeles se toma como $\max(\text{ancho}, \text{alto})$ y el menor como $\min(\text{ancho}, \text{alto})$. A continuación, se convierte cada dimensión a milímetros multiplicando: el ancho por el tamaño de píxel correspondiente al eje horizontal del plano, y el alto por el tamaño de píxel del eje vertical. Así obtenemos el diámetro mayor y menor en mm para ese contorno en el corte actual.
5. **Registro del valor máximo:** Se compara el diámetro mayor obtenido en ese *slice* con los ya calculados en *slices* anteriores de la misma lesión. Si es mayor que el previamente guardado, se actualiza el valor máximo y se registra el índice de corte donde ocurrió. De esta manera, tras recorrer todos los cortes, se identifica para cada lesión cuál fue su diámetro mayor global en ese plano y en qué corte se dio.
6. **Repetición para otros planos:** El procedimiento se repite independientemente en los tres ejes ortogonales. Al final, se cuenta con tres posibles diámetros mayores globales (uno por plano). El más grande de entre los planos axial, sagital y coronal representará la mayor extensión lineal detectada del tumor en cualquier dirección cardinal, la cual suele aproximarse bastante a la máxima extensión 3D real. Asimismo, para cada plano se pueden obtener los respectivos diámetros menores asociados al corte de diámetro mayor, aportando información sobre el espesor de la lesión en el plano.

7.6.7. Algoritmo de Cuantificación volumétrica de la carga tumoral

El proceso completo para estimar el volumen tumoral y obtener la reconstrucción tridimensional de cada lesión —a partir de las segmentaciones exportadas desde 3D Slicer— puede expresarse en los pasos que se resumen en la **Tabla 9**.

Tabla 9. Flujo de cuantificación volumétrica de la carga tumoral

Paso	Operación	Resultado principal
1	Localización de archivos: recorrer ordenadamente el directorio de trabajo y seleccionar todos los ficheros NIfTI con extensión <i>.nii.gz</i> .	Lista trazable de máscaras segmentadas.
2	Carga de la máscara: abrir cada archivo y extraer (i) la matriz volumétrica de etiquetas (<i>mask data</i>) y (ii) el espaciado de vóxel (d_x, d_y, d_z) en milímetros.	Acceso simultáneo a la segmentación y a las dimensiones físicas de cada vóxel.
3	Generación de máscara binaria: para la etiqueta de interés, crear una matriz booleana donde los vóxeles de la lesión valgan <i>True</i> .	Delimitación exacta de la lesión dentro del volumen.
4	Cálculo del volumen: contar los vóxeles <i>True</i> y multiplicar por el volumen de un vóxel ($d_x \times d_y \times d_z$).	Volumen tumoral en milímetros cúbicos.
5	Reconstrucción de superficie 3-D: aplicar <i>Marching Cubes</i> sobre la máscara binaria con el espaciado real.	Malla poligonal que representa la geometría externa de la lesión.
6	Visualización: renderizar la malla en un visor 3-D con mapa de color continuo y proporciones isométricas.	Imagen tridimensional para inspección y validación.

Al finalizar estos pasos, se dispone tanto de la estimación cuantitativa del volumen tumoral como de su representación tridimensional, listas para análisis morfológico o para su incorporación en modelos posteriores.

8. Resultados

Siguiendo la ruta metodológica resumida en la **Figura 32**, se **aplicaron** los tres enfoques geométricos —1D manual, 2.5D automática mediante *Oriented Bounding Box* (OBB) y 3D volumétrica— a las segmentaciones reales de las 39 lesiones *target* en los tres tiempos de seguimiento (T_0 basal, T_1 control 1 y T_2 control 2). Como ejemplo, la **Figura 33** muestra una metástasis mediastínica con su OBB en los tres planos ortogonales, mientras que la **Figura 34** ilustra el modelo tridimensional derivado, útil para valorar orientación, tamaño y ubicación espacial antes de la planificación terapéutica.

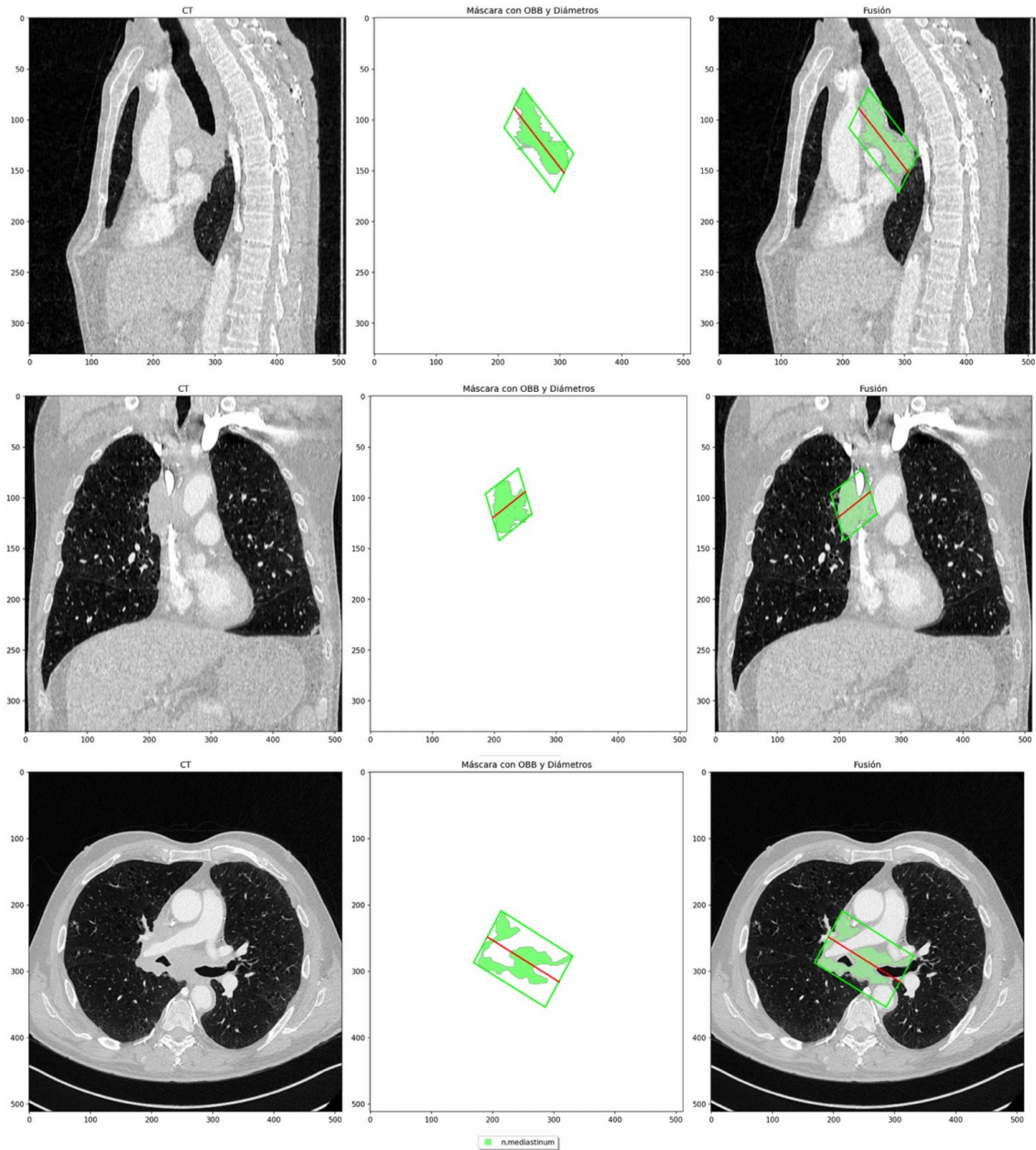


Figura 33. Implementación del método 2.5D de caja delimitadora orientada (Oriented Bounding Box, OBB) en una metástasis mediastínica. La lesión, resaltada en verde (n, mediastinum), aparece en cortes de tomografía computarizada con ventana mediastínica en tres planos anatómicos: sagital (fila superior), axial (fila media) y coronal (fila inferior).



Figura 34. Metástasis mediastínica (n,mediastinum). Este modelo facilita el análisis de características anatómicas, como la orientación, el tamaño y la localización espacial de la lesión, aspectos cruciales para la planificación quirúrgica y el desarrollo de estrategias terapéuticas personalizadas.

8.1. Definición de umbrales volumétricos mediante regresión de LOESS

La regresión LOESS (*Locally Estimated Scatterplot Smoothing*) es un método de ajuste no paramétrico que traza una curva suave a través de un diagrama de dispersión. Para cada punto se selecciona un subconjunto local de observaciones —determinado por un parámetro de ancho de banda— y se ajusta un polinomio de bajo grado mediante mínimos cuadrados ponderados, otorgando mayor peso a los puntos cercanos al punto de interés. Así, LOESS construye una función continua que captura la tendencia general de los datos sin requerir una forma funcional global predefinida.

Se construyeron curvas LOESS para tres intervalos de tiempo: Basal–Control 1, Basal–Control 2 y Control 1–Control 2, cuyos resultados se presentan en las **Figura 34**. La interpretación es la siguiente:

- **Umbrales volumétricos por intervalo (Figura 35)** Las curvas muestran la relación entre el cambio porcentual en diámetro (ΔD) y volumen (ΔV). El umbral de respuesta parcial se define en $\Delta D = -30\%$. En los tres intervalos se obtuvo $\Delta V = -75\%$, indicando que se requiere una reducción del 75 % en volumen para equivaler al -30% de diámetro en RECIST. El umbral de progresión se evaluó en $\Delta D = +20\%$, obteniéndose $\Delta V = +74\%$ (Basal–Control 1), $+71\%$ (Basal–Control 2) y $+73\%$ (Control 1–Control 2) (**Tabla 10**).

Tabla 10. Umbrales volumétricos adaptados de RECIST 1.1

Intervalo	Respuesta Parcial (RP)	Progresión de Enfermedad (PE)
Basal – Control 1	–75 %	+74 %
Basal – Control 2	–75 %	+71 %
Control 1 – Control 2	–75 %	+73 %

- **Porcentaje de concordancia (Tabla 11):** Se compararon las clasificaciones obtenidas con los umbrales LOESS y con RECIST. Los porcentajes de concordancia fueron 72.5 %, 77.5 % y 86.25 % para los intervalos Basal–Control 1, Basal–Control 2 y Control 1–Control 2, respectivamente. Esto indica que en el intervalo más tardío la clasificación entre ambos métodos fue más consistente.

Tabla 11. Concordancia de umbrales volumétricos

Intervalo	Porcentaje de Concordancia
Basal – Control 1	72.5 %
Basal – Control 2	77.5 %
Control 1 – Control 2	86.25 %

- **Elección de umbrales de referencia (Control 1–Control 2):** Dado que este intervalo presentó la mayor concordancia (86.25 %), se seleccionaron como referencia los umbrales derivados de este periodo: –75 % para respuesta parcial y +73 % para progresión. Al tratarse de un tramo pos-tratamiento, se considera que refleja más fielmente la dinámica de respuesta tumoral, con menor variabilidad inicial.

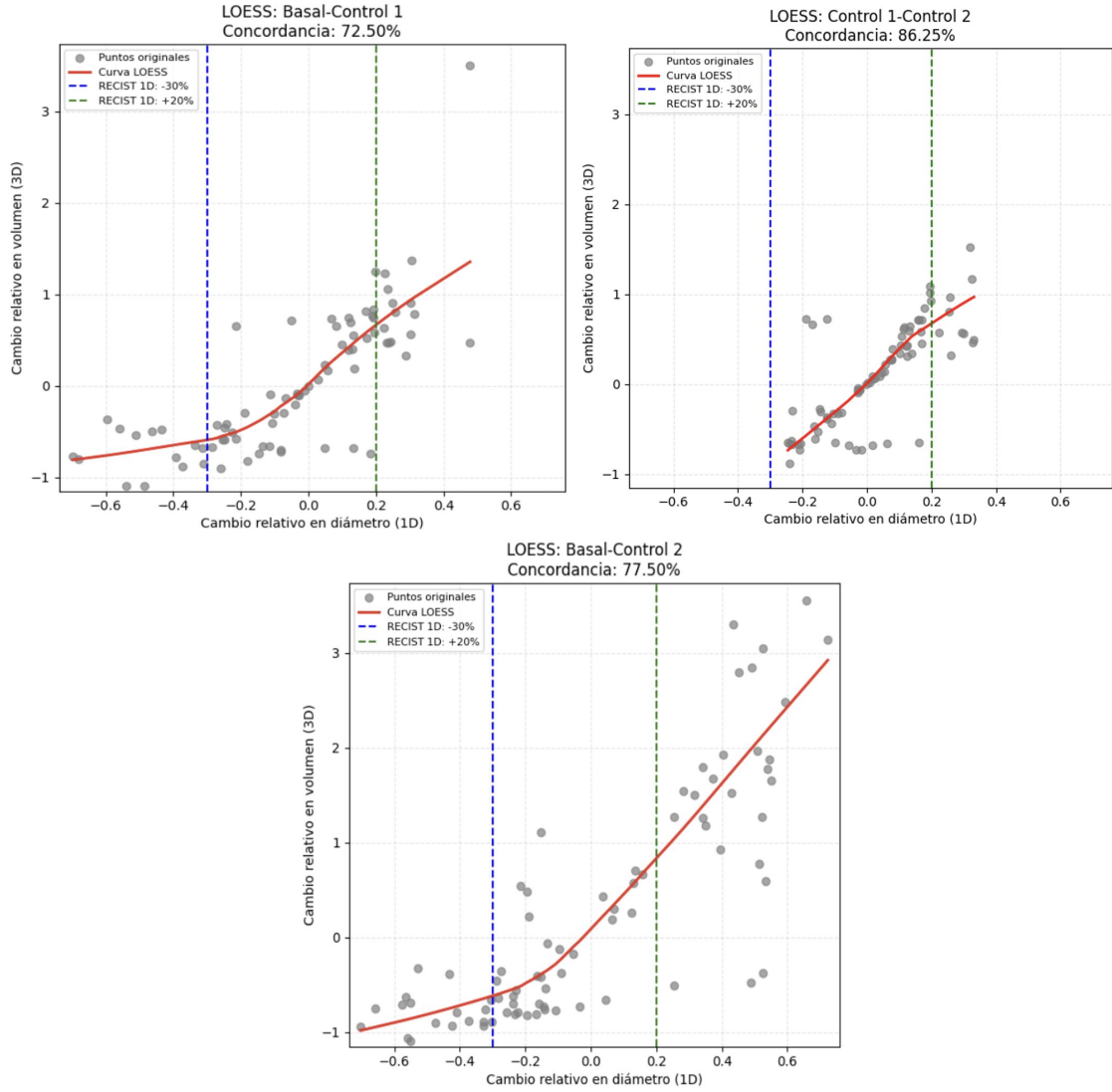


Figura 35. Análisis comparativo de la relación entre cambios relativos en diámetro (1D) y volumen (3D) mediante regresión LOESS.

8.2. Medición de segmentaciones en 1D, 2.5D y 3D

Se evaluó la concordancia inter-métodos y la significancia estadística de las diferencias en la carga tumoral estimada mediante cuatro enfoques geométricos (1D manual, 1D automático, 2.5D y 3D). Para cada comparación se calcularon dos indicadores complementarios:

1. **Coefficiente Kappa de Cohen (κ)**, que cuantifica el grado de acuerdo clasificatorio entre pares de métodos.
2. **Prueba de Mann-Whitney U**, aplicada a los cambios relativos de carga tumoral en dos intervalos temporales, a fin de determinar si las distribuciones de dichos cambios difieren de forma estadísticamente significativa.

8.3. Concordancia entre métodos

El coeficiente κ para cada par de métodos se resume en la **Tabla 12**. Los valores se interpretan de acuerdo con las categorías habituales (*débil*: $\kappa < 0.20$; *pobre*: $0.21 \leq \kappa \leq 0.40$; *moderada*: $0.41 \leq \kappa \leq 0.60$; *buen*a: $0.61 \leq \kappa \leq 0.80$).

Tabla 12. Coeficiente Kappa de Cohen para la concordancia entre métodos.

Medición 1	Medición 2	κ	Nivel de concordancia
1D manual	1D automático	0.59	Buena
1D manual	2.5D	0.34	Pobre
1D manual	3D	0.14	Débil
2.5D	3D	0.59	Buena

El acuerdo es más alto entre técnicas que comparten la misma dimensionalidad (1D manual–1D automático y 2.5D–3D), mientras que disminuye al comparar métodos lineales con volumétricos. Estos hallazgos motivaron un análisis adicional de las diferencias numéricas entre métodos, presentado a continuación.

8.4. Comparaciones de cambios relativos de carga tumoral

Para investigar diferencias en los cambios relativos de la carga tumoral, se aplicó la prueba de Mann–Whitney U en dos intervalos temporales. El primero (Basal–Control1) refleja la respuesta temprana pos-tratamiento; el segundo (Control1–Control 2) corresponde a un periodo de seguimiento posterior.

Intervalo Basal–Control 1. Los resultados se muestran en la **Tabla 13**. Un valor $p < 0.05$ indica diferencias significativas entre las medianas de los métodos comparados.

Tabla 13. Prueba Mann–Whitney U en el intervalo Basal–Control 1.

Comparación	U	p	r	Poder	Medianas (M1 vs. M2)	Diferencia medianas (IC 95 %)
1D manual vs. 1D automático	177	0.542	0.098	0.09	−1.28 vs. 4.26	−5.54 (−13.46, 7.96)
1D manual vs. 2.5D	197	0.946	0.013	0.05	−1.28 vs. 0.96	−2.24 (−10.20, 8.03)
1D manual vs. 3D	273	0.048	0.314	0.49	−1.28 vs. −10.51	9.22 (−0.12, 19.70)
1D automático vs. 2.5D	223	0.542	0.098	0.09	4.26 vs. 0.96	3.30 (−10.92, 12.62)
1D automático vs. 3D	303	0.005	0.441	0.78	4.26 vs. −10.51	14.76 (0.02, 24.61)
2.5D vs. 3D	285	0.022	0.364	0.61	0.96 vs. −10.51	11.47 (1.39, 21.66)

Los valores p significativos ($p < 0.05$) se observan en las comparaciones que incluyen la medición 3D, lo que sugiere que este método detecta cambios no captados por los enfoques lineales durante la fase temprana.

Intervalo Control 1–Control 2. La **Tabla 14** resume los resultados para el periodo de seguimiento tardío.

Tabla 14. Prueba Mann–Whitney U en el intervalo Control 1–Control 2.

Comparación	U	p	r	Medianas (M1 vs. M2)	Diferencia medianas (IC 95 %)
1D manual vs. 1D automático	203	0.946	0.013	−1.67 vs. −5.57	3.91 (−11.42, 12.92)
1D manual vs. 2.5D	157.5	0.255	0.182	−1.67 vs. 2.78	−4.44 (−22.95, 6.54)
1D manual vs. 3D	168.5	0.400	0.135	−1.67 vs. 0.00	−1.67 (−13.35, 8.16)
1D automático vs. 2.5D	157.5	0.255	0.182	−5.57 vs. 2.78	−8.35 (−24.12, 5.92)
1D automático vs. 3D	157.5	0.255	0.182	−5.57 vs. 0.00	−5.57 (−14.69, 7.14)
2.5D vs. 3D	205.5	0.892	0.024	2.78 vs. 0.00	2.78 (−8.43, 19.28)

En este segundo intervalo no se detectaron diferencias estadísticamente significativas, lo que indica que los métodos convergen en sus estimaciones una vez estabilizada la progresión tumoral.

9. Discusión

La presente tesis, desarrollada con pacientes oncológicos del Hospital Clínico de la Universidad de Chile y evaluada bajo RECIST 1.1, comparó medición manual 1D, 1D automática, aproximación 2.5D y volumetría 3D. La literatura corrobora el valor de este contraste: en cáncer pulmonar, la volumetría adelanta hasta 34 días la detección de respuesta frente a RECIST 1D y se asocia con mejor supervivencia global (8); algo similar se ha descrito en metástasis hepáticas colorrectales y otras neoplasias sólidas, donde el volumen inicial o su reducción temprana predicen la evolución clínica mejor que los diámetros lineales (49; 50).

En nuestro **Tiempo 1** (basal–primer control), el método 3D mostró diferencias significativas ($p = 0.005$ – 0.048) con tamaños del efecto moderados ($r \approx 0.32$ – 0.44) y poder ≥ 0.49 , confirmando su mayor sensibilidad diagnóstica. Estos hallazgos reflejan estudios donde la disminución volumétrica inicial se correlaciona con mayor supervivencia libre de progresión en pulmón y riñón (51; 52; 53). Durante el **Tiempo 2** (controles tardíos) no emergieron diferencias significativas; sin embargo, el poder ≤ 0.25 exige cautela. La convergencia observada respalda la intercambiabilidad de 1D, 2.5D y 3D cuando la enfermedad está estable, concordante con estudios fantoma QIBA sobre reproducibilidad volumétrica (54).

La **morfología tumoral** modula la utilidad de cada estrategia. Lesiones irregulares, necróticas o con cavidades—frecuentes en cáncer colorrectal, gástrico y pulmonar—se subestiman con un único diámetro; la volumetría 3D, al capturar la heterogeneidad espacial, ofrece mejor capacidad pronóstica. Por el contrario, en tumores de geometría regular (mama, ovario) las mediciones 1D automáticas y 2.5D podrían resultar adecuadas (55; 50).

Un hallazgo relevante es la **validación clínica del método 2.5D**: su concordancia con el 3D fue moderada ($\kappa \approx 0.59$), lo cual no fue significativamente diferente del desempeño de la volumetría 3D completa. En línea con investigaciones en segmentación multiprocesal, se confirma que los análisis multiplanares pueden monitorizar cambios volumétricos globales con eficiencia, constituyendo una opción viable cuando el tiempo o los recursos limitan el uso del método 3D (53; 56).

La **variabilidad entre observadores**, incluso con doble lectura, representa un desafío. Estudios como el de Wong et al. (2021) muestran DSC de 0.81 frente a 0.69 entre especialistas sénior y residentes, demostrando que la experiencia introduce inconsistencias. En nuestro caso, la medición 1D automática ($\kappa = 0.59$ respecto a la manual) redujo esta variabilidad, lo que coincide con evidencias en cáncer de próstata que demuestran que las herramientas asistidas igualan el rendimiento entre observadores con diferente nivel de experiencia (57; 58).

Desde la perspectiva de *outcomes* clínicos, numerosos trabajos vinculan una reducción volumétrica temprana con mayor supervivencia global y progresión libre de enfermedad en pulmón, hígado, riñón y tumores pediátricos (51; 52; 49).

Estos resultados evidencian la necesidad de estudios prospectivos multicéntricos que confirmen cómo la adopción rutinaria de mediciones 2.5D y 3D influye en las estrategias de tratamiento y en la supervivencia global. Además, será indispensable un análisis estratificado por subtipo tumoral para identificar en qué escenarios cada método ofrece ventaja, considerando la heterogeneidad intrínseca de las neoplasias y optimizando la oncología de precisión.

10. Conclusiones

En esta tesis, se observan indicios de que la volumetría 3D puede detectar variaciones tempranas en la carga tumoral que no son evidentes con una sola medida axial tradicional, especialmente en etapas tempranas de tratamiento. Esto sugiere que el 3D podría aportar valor adicional al protocolo RECIST 1.1.

Asimismo, la aproximación 2.5D surgió como una alternativa factible, mostrando resultados compatibles con la volumetría completa, lo cual favorece su uso en escenarios donde el acceso a segmentación volumétrica es limitado. Paralelamente, la medición automática en 1D demostró buena concordancia con el método manual, disminuyendo la variabilidad entre observadores y optimizando la eficiencia.

Estos hallazgos posibilitan la propuesta de un esquema escalonado para el seguimiento oncológico: volumetría 3D en fases iniciales o casos complejos, seguida de 2.5D o 1D automatizado durante la vigilancia, cuando la enfermedad está más estable. Esta estrategia equilibraría la precisión diagnóstica con la factibilidad operativa.

Desde una perspectiva clínica, se considera necesario diseñar estudios prospectivos y multicéntricos que evalúen el impacto de implementar rutinariamente métricas volumétricas (2.5D y 3D) en las decisiones de tratamiento y en desenlaces clínicos como supervivencia global y tiempo hasta la progresión. También es importante realizar un análisis por subtipo tumoral, para identificar en qué contextos cada método aporta la mayor ventaja, en función de la heterogeneidad tumoral.

Referencias

- [1] Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. 2024 May;74(3):229-63. Available from: <https://acsjournals.onlinelibrary.wiley.com/doi/10.3322/caac.21834>.
- [2] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: A Cancer Journal for Clinicians. 2021;71(3):209-49. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.3322/caac.21660>. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.3322/caac.21660>.
- [3] Plan Nacional de Cáncer 2018-2028;. Available from: https://cdn.digital.gob.cl/filer_public/d3/0a/d30a1f5e-53d9-4a31-a4fe-e90d8d9a2348/documento_plan_nacional_de_cancer.pdf.
- [4] Ley 21.258 CREA LA LEY NACIONAL DEL CÁNCER, QUE RINDE HOMENAJE PÓSTUMO AL DOCTOR CLAUDIO MORA;. Available from: <https://www.bcn.cl/leychile/navegar?idNorma=1149004>.
- [5] División de Gestión de la Red Asistencial. Modelo de Gestión para el Funcionamiento de la Red Oncológica de Chile 2018. MIN-SAL;. Available from: <https://farmaceuticosoncologia.cl/wp-content/uploads/2024/01/Modelo-de-Gestion-de-la-Red-Oncologica.pdf>.
- [6] Lewandowska A, Rudzki M, Rudzki S, Lewandowski T, Laszkowska B. Environmental risk factors for cancer – review paper. Annals of Agricultural and Environmental Medicine. 2019 Mar;26(1):1-7. Available from: <http://www.journalssystem.com/aaem/Environmental-risk-factors-for-cancer-review-paper,94299,0,2.html>.
- [7] Beaumont H, Bertrand AS, Klifa C, Patriti S, Cippolini S, Lovera C, et al. Radiology workflow for RECIST assessment in clinical trials: Can we reconcile time-efficiency and quality? European Journal of Radiology. 2019 Sep;118:257-63. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0720048X19302724>.
- [8] Hayes SA, Pietanza MC, O'Driscoll D, Zheng J, Moskowitz CS, Kris MG, et al. Comparison of CT volumetric measurement with RECIST response in patients with lung cancer. European Journal of Radiology. 2016 Mar;85(3):524-33. Publisher: Elsevier Ireland Ltd.
- [9] Marten K, Auer F, Schmidt S, Kohl G, Rummeny EJ, Engelke C. Inadequacy of manual measurements compared to automated CT volumetry in assessment of treatment response of pulmonary metastases using RECIST criteria. European Radiology. 2006 Apr;16(4):781-90. Available from: <http://link.springer.com/10.1007/s00330-005-0036-x>.

- [10] Aykan NF, Özath T. Objective response rate assessment in oncology: Current situation and future expectations. *World Journal of Clinical Oncology*. 2020 Feb;11(2):53-73. Available from: <https://www.wjgnet.com/2218-4333/full/v11/i2/53.htm>.
- [11] Riaz A, Memon K, Miller FH, Nikolaidis P, Kulik LM, Lewandowski RJ, et al. Role of the EASL, RECIST, and WHO response guidelines alone or in combination for hepatocellular carcinoma: Radiologic–pathologic correlation. *Journal of Hepatology*. 2011 Apr;54(4):695-704.
- [12] Riaz A, Memon K, Miller FH, Nikolaidis P, Kulik LM, Lewandowski RJ, et al. Role of the EASL, RECIST, and WHO response guidelines alone or in combination for hepatocellular carcinoma: Radiologic–pathologic correlation. *Journal of Hepatology*. 2011 Apr;54(4):695-704. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0168827810009104>.
- [13] Tran LN, Brown MS, Goldin JG, Yan X, Pais RC, McNitt-Gray MF, et al. Comparison of treatment response classifications between unidimensional, bi-dimensional, and volumetric measurements of metastatic lung lesions on chest computed tomography1. *Academic Radiology*. 2004 Dec;11(12):1355-60. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1076633204005513>.
- [14] Lalchandani UR, Sahai V, Hersberger K, Francis IR, Wasnik AP. A Radiologist’s Guide to Response Evaluation Criteria in Solid Tumors. *Current Problems in Diagnostic Radiology*. 2019 Nov;48(6):576-85. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0363018818301397>.
- [15] Themes UFO. Response Evaluation Criteria in Solid Tumors, World Health Organization, and Other Response Criteria; 2019. Available from: <https://radiologykey.com/response-evaluation-criteria-in-solid-tumors-world-health-organization-and-oth>
- [16] Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *European Journal of Cancer*. 2009 Jan;45(2):228-47. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0959804908008733>.
- [17] «The Radiology Assistant: RECIST 1.1 - the basics»; Available from: <https://radiologyassistant.nl/more/recist-1-1/recist-1-1-1>.
- [18] Choi H. Role of Imaging in Response Assessment and Individualised Treatment for Sarcomas. *Clinical Oncology*. 2017 Aug;29(8):481-8. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0936655517301577>.
- [19] Choi H, Charnsangavej C, Faria SC, Macapinlac HA, Burgess MA, Patel SR, et al. Correlation of Computed Tomography and Positron Emission Tomography in Patients With Metastatic Gastrointestinal Stromal Tumor Treated at a Single Institution With Imatinib Mesylate: Proposal of New Computed Tomography Response Criteria. *Journal of Clinical Oncology*. 2007 May;25(13):1753-9. Available from: <https://ascopubs.org/doi/10.1200/JCO.2006.07.3049>.

- [20] Nishino M, Giobbie-Hurder A, Gargano M, Suda M, Ramaiya NH, Hodi FS. Developing a Common Language for Tumor Response to Immunotherapy: Immune-Related Response Criteria Using Unidimensional Measurements. *Clinical Cancer Research*. 2013 Jul;19(14):3936-43. Available from: <https://aacrjournals.org/clincancerres/article/19/14/3936/77809/Developing-a-Common-Language-for-Tumor-Response-to>.
- [21] Bohnsack O, Hoos A, Ludajic K. Adaptation of the Immune Related Response Criteria: Irrecist. *Annals of Oncology*. 2014 Sep;25:iv369. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0923753419524802>.
- [22] Owen B, Gandara D, Kelly K, Moore E, Shelton D, Knollmann F. CT Volumetry and Basic Texture Analysis as Surrogate Markers in Advanced Non-small-cell Lung Cancer. *Clinical Lung Cancer*. 2020 May;21(3):225-31. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1525730419302530>.
- [23] Prasad SR, Jhaveri KS, Saini S, Hahn PF, Halpern EF, Sumner JE. CT Tumor Measurement for Therapeutic Response Assessment: Comparison of Unidimensional, Bidimensional, and Volumetric Techniques—Initial Observations. *Radiology*. 2002 Nov;225(2):416-9.
- [24] Yu SH, Choi SJ, Noh H, Lee IS, Park SH, Kim SJ. Comparison of CT Volumetry and RECIST to Predict the Treatment Response and Overall Survival in Gastric Cancer Liver Metastases. *Journal of the Korean Society of Radiology*. 2021;82(4):876. Available from: <https://jksronline.org/D0Ix.php?id=10.3348/jksr.2020.0085>.
- [25] Lubner MG, Stabo N, Lubner SJ, Rio AMd, Song C, Pickhardt PJ. Volumetric Versus Unidimensional Measures of Metastatic Colorectal Cancer in Assessing Disease Response. *Clinical Colorectal Cancer*. 2017 Dec;16(4):324-33.e1. Publisher: Elsevier Inc.
- [26] Vogel MN, Schmücker S, Maksimovic O, Hartmann J, Claussen CD, Horger M. Reduction in growth threshold for pulmonary metastases: An opportunity for volumetry and its impact on treatment decisions. *British Journal of Radiology*. 2012 Jul;85(1015):959-64.
- [27] Winter KS, Hofmann FO, Thierfelder KM, Holch JW, Hesse N, Baumann AB, et al. Towards volumetric thresholds in RECIST 1.1: Therapeutic response assessment in hepatic metastases. *European Radiology*. 2018 Nov;28(11):4839-48. Publisher: Springer Verlag.
- [28] Osman MF, Farag ASA, Samy HA, El-Baz TM, Elkholy SF. Role of multislice computed tomography 3D volumetric analysis in the assessment of the therapeutic response of hepatocellular carcinoma after transarterial chemoembolization. *Egyptian Journal of Radiology and Nuclear Medicine*. 2021 Dec;52(1). Publisher: Springer Science and Business Media Deutschland GmbH.
- [29] Wulff AM, Fabel M, Freitag-Wolf S, Tepper M, Knabe HM, Schäfer JP, et al. Volumetric response classification in metastatic solid tumors on MSCT: Initial

results in a whole-body setting. *European Journal of Radiology*. 2013;82(10). Publisher: Elsevier Ireland Ltd.

- [30] Zimmermann M, Kuhl CK, Engelke H, Bettermann G, Keil S. CT-based whole-body tumor volumetry versus RECIST 1.1: Feasibility and implications for inter-reader variability. *European Journal of Radiology*. 2021 Feb;135:109514.
- [31] Maitland ML, Wilkerson J, Karovic S, Zhao B, Flynn J, Zhou M, et al. Enhanced Detection of Treatment Effects on Metastatic Colorectal Cancer with Volumetric CT Measurements for Tumor Burden Growth Rate Evaluation. *Clinical Cancer Research*. 2020 Dec;26(24):6464-74. Available from: <https://aacrjournals.org/clincancerres/article/26/24/6464/82879/Enhanced-Detection-of-Treatment-Effects-on>.
- [32] Yu SH, Choi SJ, Noh H, Lee IS, Park SH, Kim SJ. Comparison of CT Volumetry and RECIST to Predict the Treatment Response and Overall Survival in Gastric Cancer Liver Metastases. *Journal of the Korean Society of Radiology*. 2021;82(4):876.
- [33] Napel S, Mu W, Jardim-Perassi BV, Aerts HJWL, Gillies RJ. Quantitative imaging of cancer in the postgenomic era: Radio(geno)mics, deep learning, and habitats. *Cancer*. 2018 Dec;124(24):4633-49. Available from: <https://acsjournals.onlinelibrary.wiley.com/doi/10.1002/cncr.31630>.
- [34] Bellomi M, De Piano F, Ancona E, Lodigiani AF, Curigliano G, Raimondi S, et al. Evaluation of inter-observer variability according to RECIST 1.1 and its influence on response classification in CT measurement of liver metastases. *European Journal of Radiology*. 2017 Oct;95:96-101. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0720048X17303133>.
- [35] Ghobrial FEI, Eldin MS, Razek AAKA, Atwan NI, Shamaa SSA. Computed Tomography Assessment of Hepatic Metastases of Breast Cancer with Revised Response Evaluation Criteria in Solid Tumors (RECIST) Criteria (Version 1.1): Inter-Observer Agreement. *Polish Journal of Radiology*. 2018 Feb;82:593-7. Available from: <https://www.polradiol.com/Computed-Tomography-Assessment-of-Hepatic-Metastases-of-Breast-Cancer-with-Revised-Response-Evaluation-Criteria-in-Solid-Tumors-RECIST-Criteria-Version-1.1-Inter-Observer-Agreement>.
- [36] Schiavon G, Ruggiero A, Schöffski P, Van Der Holt B, Bekers DJ, Eechoute K, et al. Tumor Volume as an Alternative Response Measurement for Imatinib Treated GIST Patients. *PLoS ONE*. 2012 Nov;7(11):e48372. Available from: <https://dx.plos.org/10.1371/journal.pone.0048372>.
- [37] Leuven K, Onderzoeksgroep OI. The use of 2D and 3D imaging modalities and its influence on diagnosis and treatment planning in orthodontics. 2014.
- [38] Jeong B, Lee SM, Park S, Choe J, Choi S, Do KH, et al. Prognostic performance in lung cancer according to tumor size: Comparison of axial, multiplanar, and 3-dimensional CT measurement to pathological size. *European Journal of Radiology*. 2021 Nov;144:109976. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0720048X21004575>.

- [39] Choi MH, Yoon SB, Song M, Lee IS, Hong TH, Lee MA, et al. Benefits of the multiplanar and volumetric analyses of pancreatic cancer using computed tomography. *PLOS ONE*. 2020 Oct;15(10):e0240318. Available from: <https://dx.plos.org/10.1371/journal.pone.0240318>.
- [40] Ridge CA, Huang J, Cardoza S, Zabor EC, Moskowitz CS, Zakowski MF, et al. Comparison of Multiplanar Reformatted CT Lung Tumor Measurements to Axial Tumor Measurement Alone: Impact on Maximal Tumor Dimension and T Stage. *American Journal of Roentgenology*. 2013 Nov;201(5):959-63. Available from: <https://www.ajronline.org/doi/10.2214/AJR.12.10033>.
- [41] Zimmermann M, Kuhl CK, Engelke H, Bettermann G, Keil S. CT-based whole-body tumor volumetry versus RECIST 1.1: Feasibility and implications for inter-reader variability. *European Journal of Radiology*. 2021 Feb;135. Publisher: Elsevier Ireland Ltd.
- [42] Tacher V, Lin M, Duran R, Yarmohammadi H, Lee H, Chapiro J, et al. Comparison of Existing Response Criteria in Patients with Hepatocellular Carcinoma Treated with Transarterial Chemoembolization Using a 3D Quantitative Approach. *Radiology*. 2016 Jan;278(1):275-84. Available from: <http://pubs.rsna.org/doi/10.1148/radiol.2015142951>.
- [43] Laubender RP, Lynghjem J, D'Anastasi M, Heinemann V, Modest DP, Mansmann UR, et al. Evaluating the agreement between tumour volumetry and the estimated volumes of tumour lesions using an algorithm. *European Radiology*. 2014 Jul;24(7):1521-8. Available from: <http://link.springer.com/10.1007/s00330-014-3195-9>.
- [44] Armato SG, Oxnard GR, MacMahon H, Vogelzang NJ, Kindler HL, Kocherginsky M, et al. Measurement of mesothelioma on thoracic CT scans: A comparison of manual and computer-assisted techniques. *Medical Physics*. 2004 May;31(5):1105-15. Available from: <https://aapm.onlinelibrary.wiley.com/doi/10.1118/1.1688211>.
- [45] Kuhl CK, Alparslan Y, Schmoe J, Sequeira B, Keulers A, Brümmendorf TH, et al. Validity of RECIST Version 1.1 for Response Assessment in Metastatic Cancer: A Prospective, Multireader Study. *Radiology*. 2019 Feb;290(2):349-56. Available from: <http://pubs.rsna.org/doi/10.1148/radiol.2018180648>.
- [46] Bucho TMT, Tissier RLM, Bodalal Z, Pizzi AD, Nguyen-Kim TDL, Beets-Tan RGH, et al. How Does Target Lesion Selection Affect RECIST? A Computer Simulation Study. *Investigative Radiology*. 2024;59(6).
- [47] Beaumont H, Evans TL, Klifa C, Guermazi A, Hong SR, Chadja M, et al. Discrepancies of assessments in a RECIST 1.1 phase II clinical trial – association between adjudication rate and variability in images and tumors selection. *Cancer Imaging*. 2018 Dec;18(1):50. Available from: <https://cancerimagingjournal.biomedcentral.com/articles/10.1186/s40644-018-0186-0>.

- [48] Shafiei A, Bagheri M, Farhadi F, Apolo AB, Biassou NM, Folio LR, et al. Ct evaluation of lymph nodes that merge or split during the course of a clinical trial: Limitations of recist 1.1. *Radiology: Imaging Cancer*. 2021 May;3(3). Publisher: Radiological Society of North America Inc.
- [49] Tai K, Komatsu S, Sofue K, Kido M, Tanaka M, Kuramitsu K, et al. Total tumour volume as a prognostic factor in patients with resectable colorectal cancer liver metastases. *BJS Open*. 2020 May;4(3):456-66. Available from: <https://academic.oup.com/bjsopen/article/4/3/456/6061435>.
- [50] Michiel Zeeuw J, Wesdorp NJ, Ali M, Bakker AJJJ, Voigt KR, Starmans MPA, et al. Prognostic value of total tumor volume in patients with colorectal liver metastases: A secondary analysis of the randomized CAIRO5 trial with external cohort validation. *European Journal of Cancer*. 2024 Aug;207:114185. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0959804924008414>.
- [51] Owen B, Gandara D, Kelly K, Moore E, Shelton D, Knollmann F. CT Volumetry and Basic Texture Analysis as Surrogate Markers in Advanced Non-small-cell Lung Cancer. *Clinical Lung Cancer*. 2020 May;21(3):225-31. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1525730419302530>.
- [52] Aktepe OH, Erdemir AG, Caliskan Yildirim E, Demirciler E, Ulasli T, Guven DC, et al. Early Changes in Volumetric Body Composition Parameters Predict Survival Outcomes in Patients with Metastatic Renal Cell Carcinoma Treated with Targeted Therapy. *Journal of Clinical Medicine*. 2025 May;14(9):3140. Available from: <https://www.mdpi.com/2077-0383/14/9/3140>.
- [53] Spencer D, Bonner ER, Tor-Díez C, Liu X, Bougher K, Prasad R, et al. Tumor volume features predict survival outcomes for patients diagnosed with diffuse intrinsic pontine glioma. *Neuro-Oncology Advances*. 2024 Jan;6(1). Available from: <https://academic.oup.com/noa/article/doi/10.1093/noajnl/vdae151/7745609>.
- [54] Athellogou M, Kim HJ, Dima A, Obuchowski N, Peskin A, Gavrielides MA, et al. Algorithm Variability in the Estimation of Lung Nodule Volume From Phantom CT Scans: Results of the QIBA 3A Public Challenge. *Academic Radiology*. 2016 Aug;23(8):940-52. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1076633216300101>.
- [55] Fournier L, De Geus-Oei LF, Regge D, Oprea-Lager DE, D'Anastasi M, Bidaut L, et al. Twenty Years On: RECIST as a Biomarker of Response in Solid Tumours an EORTC Imaging Group – ESOI Joint Paper. *Frontiers in Oncology*. 2022 Jan;11. Available from: <https://www.frontiersin.org/articles/10.3389/fonc.2021.800547/full>.
- [56] Sarkar S, Teo PT, Abazeed ME. Deep learning for automated, motion-resolved tumor segmentation in radiotherapy. *npj Precision Oncology*. 2025 Jun;9(1). Available from: <https://www.nature.com/articles/s41698-025-00970-1>.

- [57] Wong J, Baine M, Wisnoskie S, Bennion N, Zheng D, Yu L, et al. Effects of interobserver and interdisciplinary segmentation variabilities on CT-based radiomics for pancreatic cancer. *Scientific Reports*. 2021 Aug;11(1):16328. Available from: <https://www.nature.com/articles/s41598-021-95152-x>.
- [58] Scher HI, Morris MJ, Kelly WK, Schwartz LH, Heller G. Prostate Cancer Clinical Trial End Points: “RECIST”ing a Step Backwards. *Clinical Cancer Research*. 2005 Jul;11(14):5223-32. Available from: <https://aacrjournals.org/clincancerres/article/11/14/5223/290677/Prostate-Cancer-Clinical-Trial-End-Points-RECIST>.