

#guardsymposium2026
X @GuardConsortium



GUARD INTERNATIONAL SYMPOSIUM · 16-18 SEPTIEMBRE 2026 · MADRID

INTELIGENCIA ARTIFICIAL EN EL DIAGNÓSTICO DEL CÁNCER DE PRÓSTATA

Dr. Mario Domínguez Esteban
Sº Urología - H.U. Marques de Valdecilla
Santander

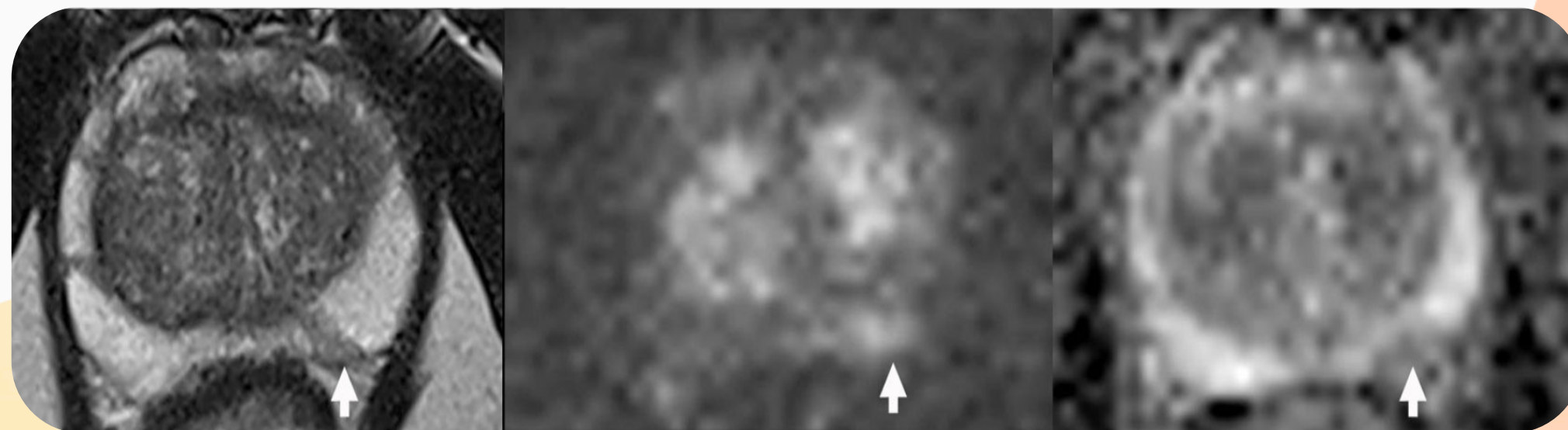
Conflictos de interés

- He proporcionado asesoramiento científico a Janssen, Astellas, Ipsen, Boston Scientific, Almirall, Bayer, AstraZeneca, Pfizer
- He participado en reuniones médicas organizadas por Janssen, Astellas, Ipsen, Boston Scientific, Almirall, Bayer, Roche, Novartis, AAA, Abex
- Proctor quirúrgico de Boston-Scientific y Abbex-Intuitive



50 años

- **PSA elevado:** 3,5 ng/dl ('24) – 5 ng/dl ('25) - 6,5 ng/dl ('26)
- **PSAI/t 8%**
- **RMmp PI-RADS 2, HBP 20 cc**
- **D-PSA 0,30 ng/ml**
- **TR Normal**



¿Biopsiar o vigilar?

Los cuatro cuellos de botella del diagnóstico

01 **Capacidad y variabilidad de lectura de la RM**

Las exploraciones aumentan más rápido que radiólogos con experiencia

02 **Reproducibilidad de la gradación histológica**



03 **IA - Especificidad: Descarta Bien – Asistencia Dx**



Metanálisis – Deep Learnig: 36 estudios · 9.411 pacientes

Mucha variabilidad

04 **Información pronóstica que Gleason no captura**

La misma categoría agrupa pacientes con evoluciones muy distintas.

IA: Tres funciones distintas bajo una misma palabra

COMBINA DATOS

Calcular riesgo

Modelos sobre edad, PSA, densidad de PSA, tacto y biomarcadores.

Salida: probabilidad de cáncer significativo. (De regresión a redes neuronales)

Stockholm3, MPS2, ClarityDX.
Evidencia: ensayos de cribado.

IDENTIFICA PATRONES

Interpretar imágenes

Redes convolucionales sobre RM, PET o portaobjetos.

Salida: mapa de sospecha, lesión marcada o grupo de grado.

PI-CAI, PANDA, CADe Evidencia: estudios de lectores.

GENERA LENGUAJE

Procesar texto

Modelos de lenguaje sobre informes de RM o patología.

Salida: dato estructurado, estadificación o recomendación.

Validado para extraer datos;
no validado para decidir.

Las cifras de esta charla

1.368 identificados en PubMed

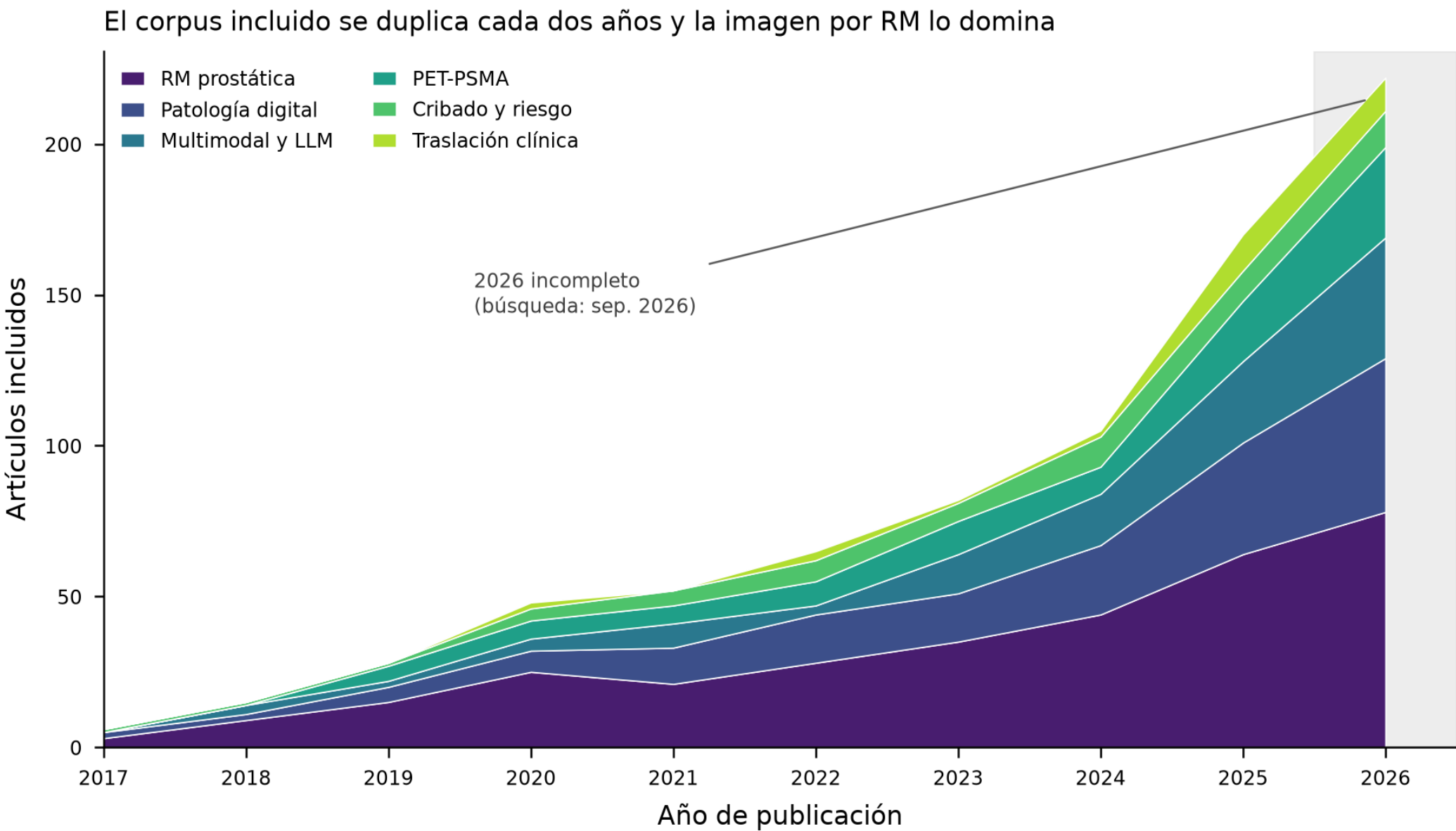
1.345 tras deduplicar

802 incluidos tras cribado

279 estudios con extracción

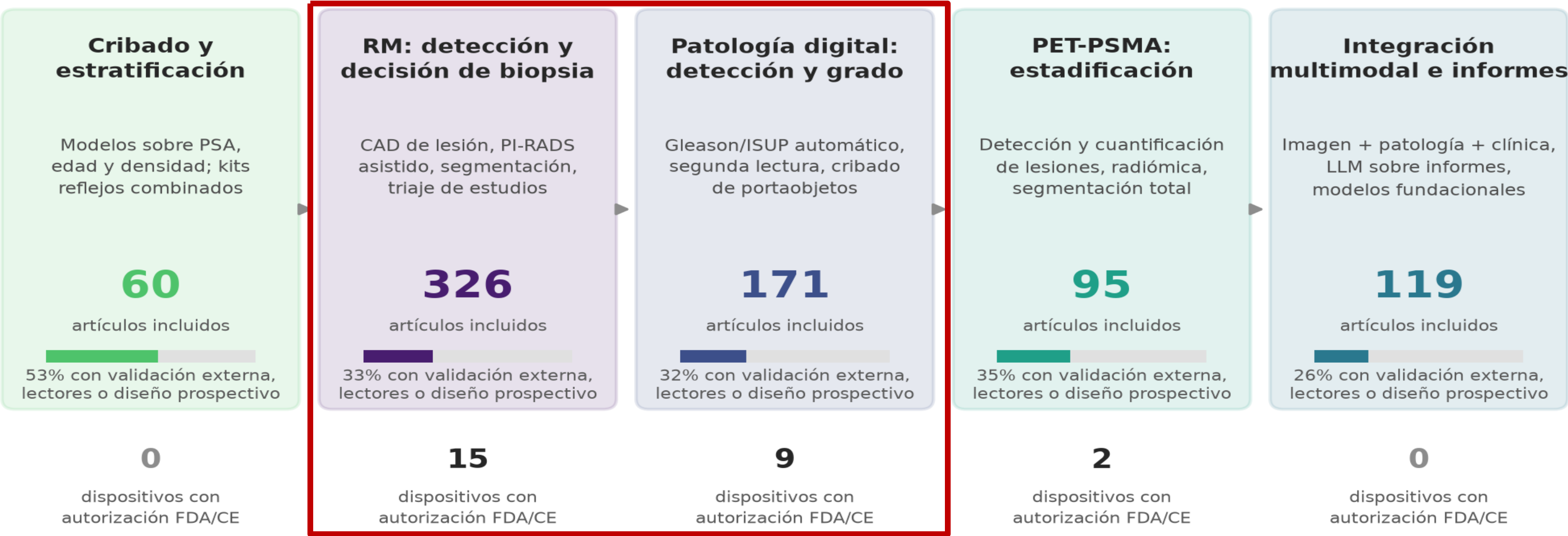
797/802 DOI verificados en Crossref · 2 retractados excluidos

49,7 % de las filas proceden del resumen · cribado asistido por modelo de lenguaje, no doble revisión



Panorama de la IA a lo largo de la ruta diagnóstica del cáncer de próstata

La literatura y las autorizaciones regulatorias se concentran en RM y patología digital; el cribado algorítmico y la integración multimodal no tienen todavía ningún dispositivo autorizado con esa indicación.



Recuentos propios sobre el corpus incluido (n=802 artículos, búsqueda de septiembre de 2026) y sobre el inventario de 28 dispositivos (24 autorizaciones FDA verificadas en openFDA, 3 marcados CE no verificables en fuente primaria, 1 sin verificar). Dos dispositivos autorizados —ecografía prostática y gammagrafía ósea— no corresponden a ninguno de los cinco eslabones y no se cuentan. ArteraAI Prostate se cuenta en patología, aunque su indicación autorizada es pronóstica y no de detección.

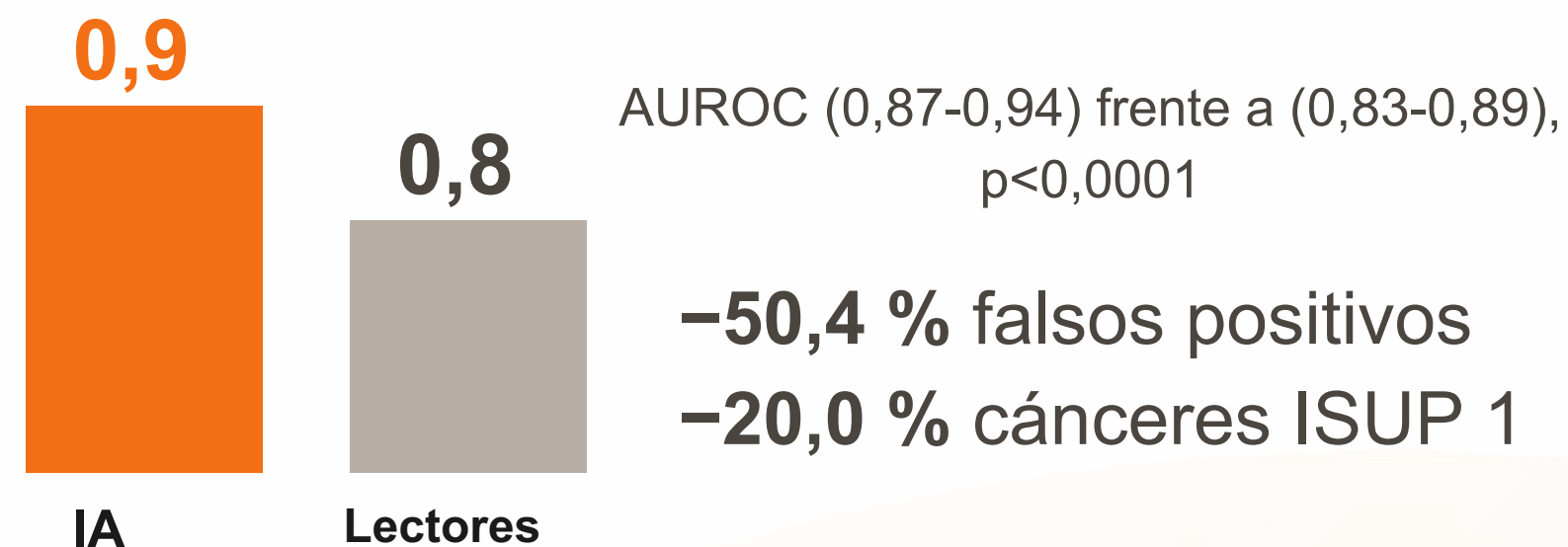


PI-CAI: el estudio de referencia

Confirmatorio internacional pareado

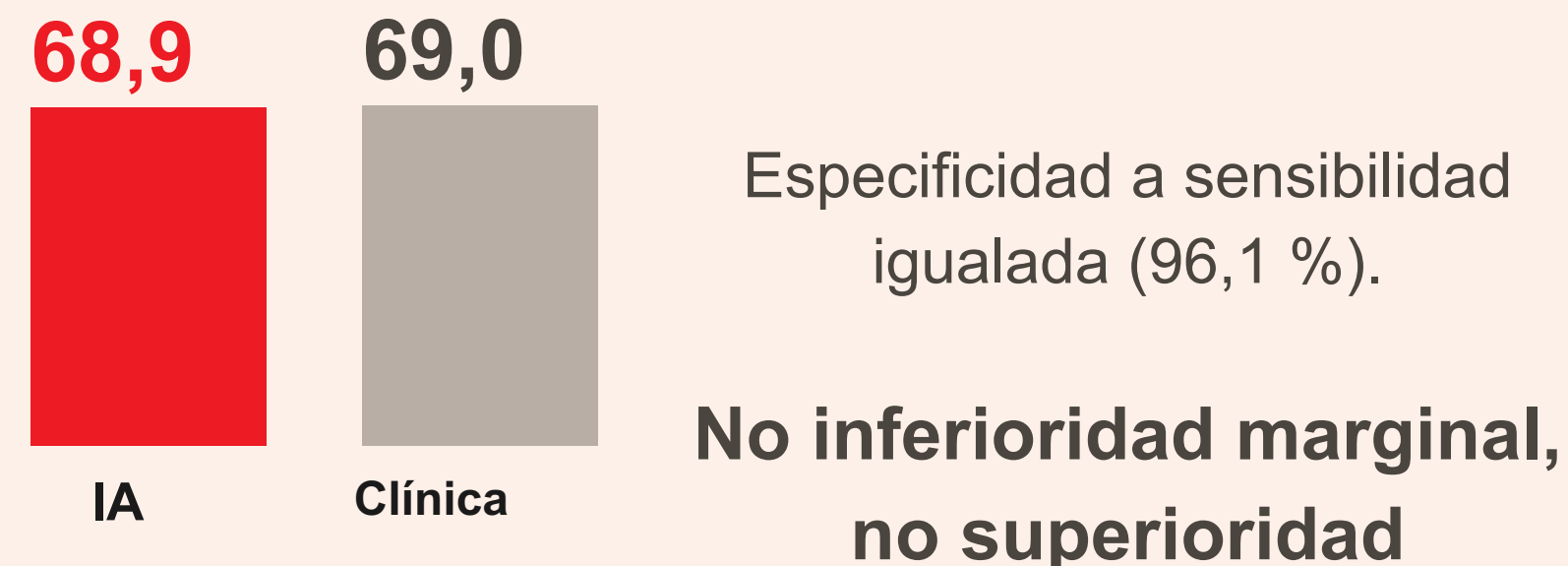
- 1.- 10.207 RM recopiladas, lectura retrospectiva (histología disponible o 3 años de seguimiento)
- 2.- 400 casos frente a 62 radiólogos de 20 países con PI-RADS v2.1
- 3.- 1.000 casos frente a la práctica multidisciplinar real

400 CASOS · 62 LECTORES 20 PAISES



Barras a escala relativa sobre el rango 0,5-1,0 de AUROC.

1.000 EXPLORACIONES · LECTURA CLÍNICA REAL



La IA democratiza y reduce diferencias



Mejora sobre todo al lector menos experto y duplica la reproducibilidad de PI-RADS.

Estudio de lectores

RM biparamétricas de PI-CAI · 61 lectores (34 expertos, 27 no expertos), cada uno con y sin el mapa de la IA · retrospectivo: no es un ensayo de decisiones asistenciales

SIN ASISTENCIA → CON ASISTENCIA · 61 LECTORES

AUROC  0,882 → 0,916

Sensibilidad PIRADS ≥ 3 DEL 94,3 AL 96,8

Kappa interlector PI-RADS  0,336 → 0,748

Ganancia · no expertos  +0,053

Ganancia · expertos  +0,018

+3,3 puntos de AUROC (1,8-4,9; $p < 0,001$)

LA CARA B

Un estudio sin mejora alguna: AUC 0,87 con asistencia frente a 0,86 sin ella ($p=0,734$)

VPP 67,2 % → 77,2 % a costa de
Sensibilidad lesional 48,0 % → 44,4 %

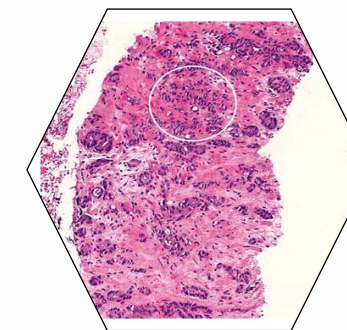


IA autónoma frente a radiólogos

En validaciones independientes, la IA autónoma empata por paciente y pierde por lesión.

Estudio	Unidad paciente	Unidad lesión	Consecuencia
Validación multicéntrica y multiescáner	IA - AUC 0,91 (0,87-0,95) frente a 0,95 (0,92-0,97)	13 de 96 lesiones de grupo ≥ 2 no detectadas, frente a 7 de 96	Pierde frente al comité multidisciplinar
Prospectivo bicéntrico	IA - AUC 0,79 frente a 0,77 (p=0,73)	IA - Sensibilidad lesional 80 % frente a 93 % (p=0,02)	Mejor valor predictivo positivo (58 % frente a 48 %)
Validación externa, 658 participantes	96 % de pacientes identificados frente al 98 % (p=0,23)	Sensibilidad lesional 55 % Dice 0,29	Sirve para decidir dónde mirar, no para delimitar

PANDA: el techo es la variabilidad del patólogo



Reto internacional y validación ciega

- 1.- PANDA: 10.616 biopsias digitalizadas: Entrenamiento
- 2.- Validación externa en EE. UU. (741 biopsias, panel de 6 uropatólogos) y Europa (330, otro laboratorio)
- 3.- Comparación «en condiciones reales»: 113 portaobjetos y 10 patólogos

KAPPA CUADRÁTICA · PANDA

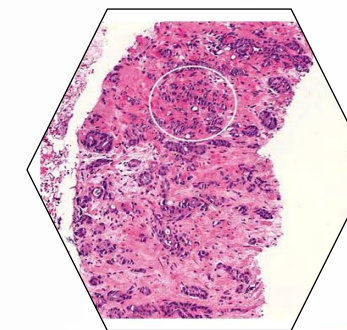
Validación interna		0,931
Externa EE. UU.		0,862
Externa Europa		0,868

- **La caída se concentra en la especificidad:**
75,2 % Sensibilidad conservada al 98,6 %.

EN CONDICIONES REALES (sobre 113 preparaciones)



Escala de 0,70 a 0,95 de kappa. El rango humano abarca 0,14: dos algoritmos separados por 0,04 son indistinguibles.

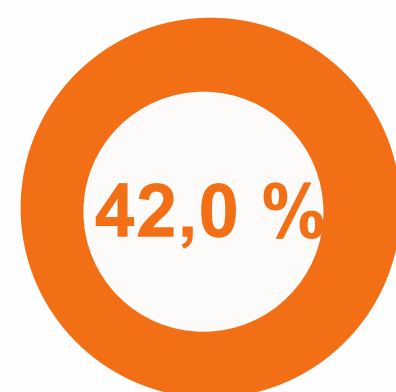


El giro pronóstico: donde la IA supera a Gleason

Análisis de ensayos fase III

Patología digital + variables clínicas · 8 ensayos de NRG Oncology, **9.787 pacientes**, frente a los grupos de riesgo NCCN · Componente predictivo en RTOG 9202: 1.192 pacientes y **17,2 años de seguimiento**

PRONÓSTICO · 9.787 PACIENTES, 8 ENSAYOS FASE III



de los pacientes **reclasificados** respecto a los grupos NCCN

Metástasis a 10 años en los subgrupos de alto riesgo



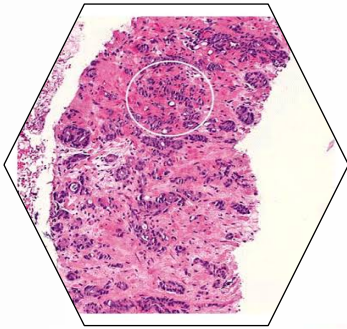
PREDICCIÓN DE BENEFICIO · RTOG 9202, 1.192 PACIENTES, 17,2 AÑOS

Biomarcador positivo
sHR 0,55 (0,41-0,73)

Biomarcador negativo
sHR 1,06 (0,61-1,84)



Interacción con el tratamiento $p=0,04$. La hormonoterapia larga beneficia a los positivos y no a los negativos: diferencia absoluta a 15 años de 14 puntos frente a 0



Asistencia en patología: beneficio real y sesgo de automatización

Estudios de lectores

La asistencia mejora a quien parte de más abajo, y cuando la IA se equivoca arrastra al lector.

Paneles de 3, 14 y 18 patólogos sobre biopsias digitalizadas, cada uno con y sin asistencia · más un prospectivo unicéntrico en entorno especializado · ninguno mide desenlaces del paciente

Medida	Sin IA	Con IA
Sensibilidad de patólogos generales	74 %	90 % Casos Difíciles
Kappa mediana de un panel de 14 patólogos	0,799	0,872
Concordancia con subespecialistas	referencia	+5,6 pts <28% sobregraduacion
Exactitud en entorno ya especializado	95,00 %	93,81 %

SESGO DE AUTOMATIZACIÓN

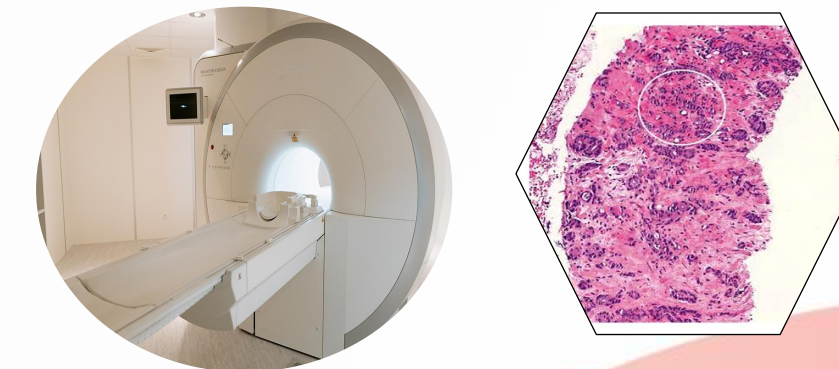
45,1% a 38,0 %

concordancia asistida en las 61 biopsias donde la interpretación de la IA era incorrecta

En uropatología, el beneficio es de proceso

- 21,94 % de tiempo de lectura e informe
- 9,5 puntos de peticiones de inmunohistoquímica

Triage: la arquitectura más plausible y la menos probada



El reparto de casos entre IA y humano rinde más que la competición entre ambos, pero solo está probado en simulación.

Simulación retrospectiva

400 exploraciones de PI-CAI, 4 centros europeos · en patología, cohorte de Yale (1.876 cilindros) y vigilancia activa PRIAS (4.744 portaobjetos)

RESONANCIA · 400 EXPLORACIONES SIMULADAS

49 % IA autónoma

51 % radiólogo

Sensibilidad y especificidad del 94,7 % en el tramo autónomo.

El flujo completo mantiene la sensibilidad (89,0 % frente a 89,4 %; $p=0,36$) y gana 11,5 puntos de especificidad

PATOLOGÍA · CILINDROS

68,6 % cribado automático

31,4 % manual

Proporción de portaobjetos que quedarían para revisión humana en Yale

En vigilancia activa evitaría la lectura del 63 % de preparaciones

0,55 %

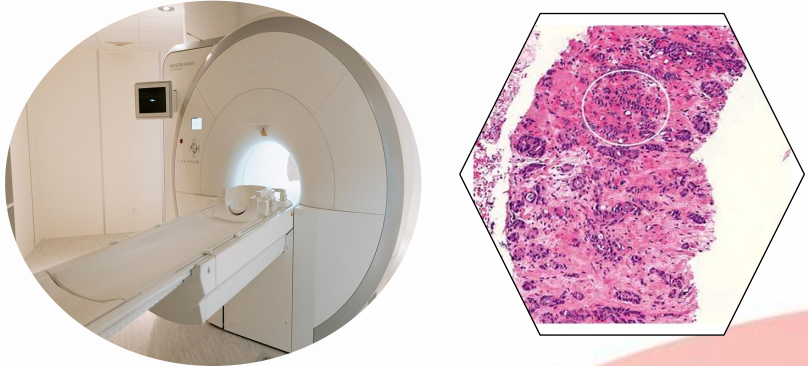
de preparaciones con cáncer clasificados como benignos en vigilancia activa.

Sensibilidad 0,96 y especificidad 0,73 sobre 4.744 portaobjetos.

Ninguna guía fija la pérdida aceptable

Multimodalidad y modelos de lenguaje:

El modelo multimodal iguala al radiólogo; los modelos de lenguaje solo están validados para extraer datos, no para decidir.



Diseños Fusión temprana en validación externa frente al radiólogo con PI-RADS · LLM sobre 912 informes PI-RADS 3-4

CON ESTÁNDAR DE REFERENCIA

(comparado contra histología o desenlace real del paciente)

Evaluacion de imagen + clinica en validación externa · AUC

Multimodal		0,77
Radiólogo con PI-RADS		0,75
Solo imagen		0,70
Solo clínica		0,67

Sin diferencia significativa frente al radiólogo
Extracción de informes: 98,0 % de exactitud en más de 130 centros

SIN VALIDACIÓN CLÍNICA

(comparado solo contra el texto del informe , nunca contra el resultado real)

Modelo de lenguaje general decidiendo Bx.P sobre el informe

AUC decisión de biopsia		0,65
Kappa PI-RADS global		0,538
Kappa zona de transición		0,077

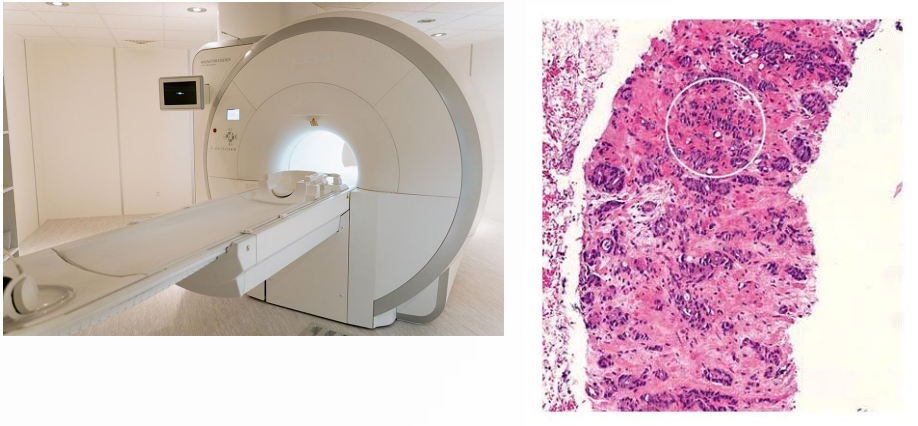
Sensibilidad 0,84 y especificidad 0,44 al decidir biopsia
No infiere las reglas zonales

Segundos lectores autorizados, no decisores

Ninguno está autorizado para indicar una biopsia y ninguno tiene un ensayo propio del circuito.



2022 · ProstatID	2023 · Lucida Pi (CE)	2024-25 · QP-Prostate	2025 · Siemens Prostate MR AI	2026 · DeepHealth ProstateAI
Producto	Autorización	Validación presentada		AUC del lector sin → con IA
ProstatID · Bot Image	FDA CE 2022 · 2025	Multilector · 150 casos · 9 lectores		0,629 → 0,671
QP-Prostate CAD · Quibim	CE FDA 2024 · 2025	Test independiente · 228 casos · 9 lectores		0,849 → 0,868
Prostate MR AI · Siemens	FDA 2025	Multilector · 340 casos · 12 radiólogos		0,676 → 0,701
DeepHealth ProstateAI	FDA CE 2026 · 2026	Multicéntrico · 250 casos · 6 instituciones		0,626 → 0,724
Lucida Pi	CE 2023	Externa independiente · 252 casos · 6 centros		IA 0,91 frente a 0,95



Qué hay autorizado y para qué sirve

Etapa				
Etapa	Producto	Autorización	Indicación autorizada	Utilidad demostrada
Antes de la RM	ClarityDX Prostate	CE-IVD 2022 · sin FDA	Estimar riesgo antes de pedir más pruebas	Sensibilidad 95 %: hasta 35 % de biopsias evitables
Lectura de RM	ProstatID · QP-Prostate · Siemens · DeepHealth · Pi	FDA 510(k) 2022-2026 · Pi con CE MDR	Detección y caracterización · lectura concurrente	+2 a +10 puntos de AUC del lector; sin ensayo del circuito
Biopsia digital	Paige Prostate Detect	FDA De Novo 2021 · CE-IVD	Señalar focos sospechosos tras la revisión del patólogo	Sensibilidad de patólogos generales 74 → 90 %
Segunda lectura	Ibex · Galen Prostate	Marcado CE · uso rutinario	Cribado de portaobjetos ya informados	Cáncer omitido detectado · 122.000 láminas en 3 años
Pronóstico	ArteraAI Prostate	FDA De Novo · pronóstica	Pronóstico sobre histología digital, no detección	Reclasifica al 42,0 % frente a NCCN

Europa: Pi comunicó marcado CE en 2023 y está en piloto en cuatro NHS Trusts de Greater Manchester desde agosto de 2026.

No existe «aprobación EMA» de estos dispositivos: verificar certificado, versión y alcance antes de implantar.

«¿Se puede usar ya para evitar biopsias?»

Biopsias evitadas	Condición	Pérdida declarada
38-52 %	Modelo de riesgo a sensibilidad fijada del 90-95 %	Valor predictivo negativo del 97,5-98,0 %
22,7 %	Sistema multimodal validado prospectivamente en 1.629 pacientes	AUC 0,91 frente a 0,85 de PI-RADS
18-32 %	Sistema automatizado sobre datos multimodales	4,1-4,88 % de los cánceres significativos

¿Y los modelos de lenguaje, ChatGPT y similares?

Extracción estructurada

Excelente y con referencia clara: 98,0 % de exactitud para número de cilindros y 95% para cilindros afectados, en más de 130 centros

Estadificación desde el informe

Aceptable, pero sin estimación específica de próstata

Decisión de biopsia

No validado: AUC 0,65 sobre el informe de resonancia y kappa 0,077 en zona de transición

Generación de informes

Sin estándar de referencia: ningún estudio de lectores con desenlace diagnóstico

Semáforo de madurez: qué se puede usar y qué hay que medir

VERDE · YA REAL

Segmentación glandular y zonal.

Detección concurrente y segundo lector.

Estandarización de PI-RADS y de la gradación.

Respaldado por estudios de lectores y validación externa.

ÁMBAR · PROMETEDOR

Triaje semiautónomo: solo probado en simulación.

Integrar IA con densidad de PSA y biomarcadores.

Reducción de carga de lectura y de pruebas auxiliares.

Necesita evaluación prospectiva en el centro que lo implanta.

ROJO · NO DEMOSTRADO

Que un algoritmo descarte una RM por sí solo.

Decidir no biopsiar sin red de seguridad ni seguimiento.

Que un CADe comercial mejore el circuito completo.

Ningún producto de RM tiene ensayo propio del circuito.

ESTABLECIDO

- 1 La asistencia con IA mejora concordancia y eficiencia, más en lectores menos expertos
- 2 La gradación automatizada alcanza el rango interpatólogo en cohortes externas
- 3 El triaje reduce sustancialmente el material que requiere revisión humana

NO ESTABLECIDO

Que el uso de IA reduzca biopsias innecesarias o mejore la detección en una población de cribado, medido de forma prospectiva y aleatorizada con la IA como intervención asignada.

Artículos del corpus que declaran aleatorización



2,3 %

1 · Datos del paciente

Edad (años)

50

Ascendencia africana / Black-African American

No

Esperanza de vida

≥15 años

Comorbilidad relevante (texto libre)

NO

Ver criterios SIOG si hay duda

2 · Antecedentes familiares / genéticos

Familiar de 1er grado con CaP

No

Ese familiar, diagnosticado <60 años o fallecido por CaP

No

Mutación germinal conocida

No testado

Requiere consejo genético previo

3 · Síntomas y medicación

☒ Síntomas del tracto urinario inferior (LUTS) ☐ Hematuria / hematospermia

☐ Dolor óseo / signos de enfermedad avanzada ☐ Asintomático

Inhibidor de 5α-reductasa (finasteride/dutasteride)

No

Reduce el PSA ~50%: se corrige automáticamente

4 · Exploración y PSA

Tacto rectal

Normal

Histórico de PSA

17/09/20266,5

17/09/20255

Generar valoración