

Más allá del radiólogo:

Rendimiento de la inteligencia artificial en

RM prostática

Adolfo Manuel Delgado Brito¹, José Marlon Rodríguez Ortega¹, Alberto Benito Boillos¹,
Manuel Rafael López De la Torre Carretero¹, Daniel Alfonso Zambrano Andrade¹, María
Soledad Muhn Valdemarin¹, César Pedro Caldas Valdez¹, David Cano Rafart¹.

Clínica Universidad de Navarra¹, Pamplona

Objetivo

Evaluar la concordancia y el rendimiento diagnóstico de la inteligencia artificial (IA) en la detección de cáncer de próstata clínicamente significativo (csPCa) mediante resonancia magnética multiparamétrica prostática (RMmp), comparando dos lecturas radiológicas independientes, un informe asistido por IA y la IA autónoma.

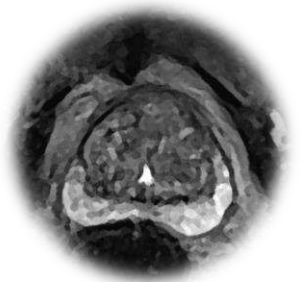
Se plantea que las estrategias basadas en IA ofrecen un rendimiento igual o superior al de la lectura radiológica independiente.

Introducción



La RM multiparametrica (mpMRI) se ha convertido en un pilar fundamental en el diagnóstico del cáncer de próstata, permitiendo una mejor detección del cáncer de próstata clínicamente significativo (csPCa) al tiempo que reduce biopsias innecesarias y el sobrediagnóstico de enfermedad indolente.

La estandarización mediante el PI-RADS v2.1 ha mejorado la consistencia en la elaboración de informes, aunque persiste una variabilidad sustancial entre lectores y la precisión diagnóstica continúa dependiendo en gran medida de la experiencia del radiólogo.



En este sentido, la inteligencia artificial (IA) ha emergido como una herramienta prometedora para apoyar la interpretación de la mpMRI mediante la detección de lesiones, la reducción de la variabilidad y la posible mitigación del sobrediagnóstico.

Introducción

No obstante, el papel clínico de la IA, ya sea como herramienta de apoyo independiente para la toma de decisiones o integrada en la elaboración rutinaria de informes radiológicos, sigue siendo objeto de investigación. Este estudio explora el equilibrio entre la detección de cáncer y el sobrediagnóstico mediante la comparación entre lectores humanos independientes, informes clínicos asistidos por IA y sistemas de IA utilizados de forma autónoma en una cohorte real de mpMRI.



Material y métodos



Estudio retrospectivo
unicéntrico

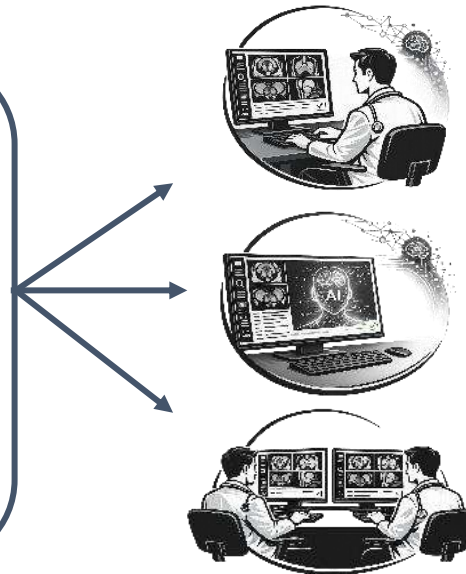


68 mpMRI
consecutivas



Informes según criterios
PI-RADS v2.1
Positividad para csPCa se
definió como PI-RADS ≥ 4

Los informes clínicos se generaron asistido por IA, en el cual los radiólogos revisaban las sugerencias de la IA y emitían la interpretación final. Los resultados de la IA en modo autónomo se archivaron de forma independiente. Posteriormente, dos radiólogos (R1 y R2) relevaron todas las exploraciones de forma independiente, ciegos a los datos clínicos, a los resultados de la IA y a la histopatología.



La concordancia entre lectores se evaluó mediante la κ de Cohen. El rendimiento diagnóstico (sensibilidad, especificidad, valor predictivo positivo, valor predictivo negativo y exactitud) se evaluó en el subgrupo de pacientes que se sometieron a **biopsia prostática**, utilizando la **histopatología como estándar de referencia**.

Resultados



Cohorte final: 67 estudios (se excluyó un paciente por prostatitis) (edad media $68,0 \pm 8,9$ años; PSA mediano 5,56 ng/mL).

La concordancia entre los dos radiólogos para PI-RADS ≥ 4 fue sustancial ($\kappa = 0,684$).

La concordancia entre los informes clínicos asistidos por IA y los lectores humanos independientes fue alta, con valores de κ de 0,720 frente a R1 y 0,745 frente a R2.

Resultados

La correlación histopatológica estuvo disponible en 29 pacientes, de los cuales 11 presentaban csPCa. En este subgrupo de pacientes con resultado anatomopatológico:

- La IA autónoma alcanzó una sensibilidad del 81,8%, especificidad del 50,0% y una exactitud global del 62,1%
- El informe asistido por IA priorizó la sensibilidad (90,9%) a expensas de la especificidad (33,3%), con una exactitud del 55,2%.
- La unión de los dos radiólogos (R1 U R2) mantuvo la sensibilidad (81,8%), pero redujo la especificidad (27,8%), resultando en menor exactitud (48,3%).

Resultados

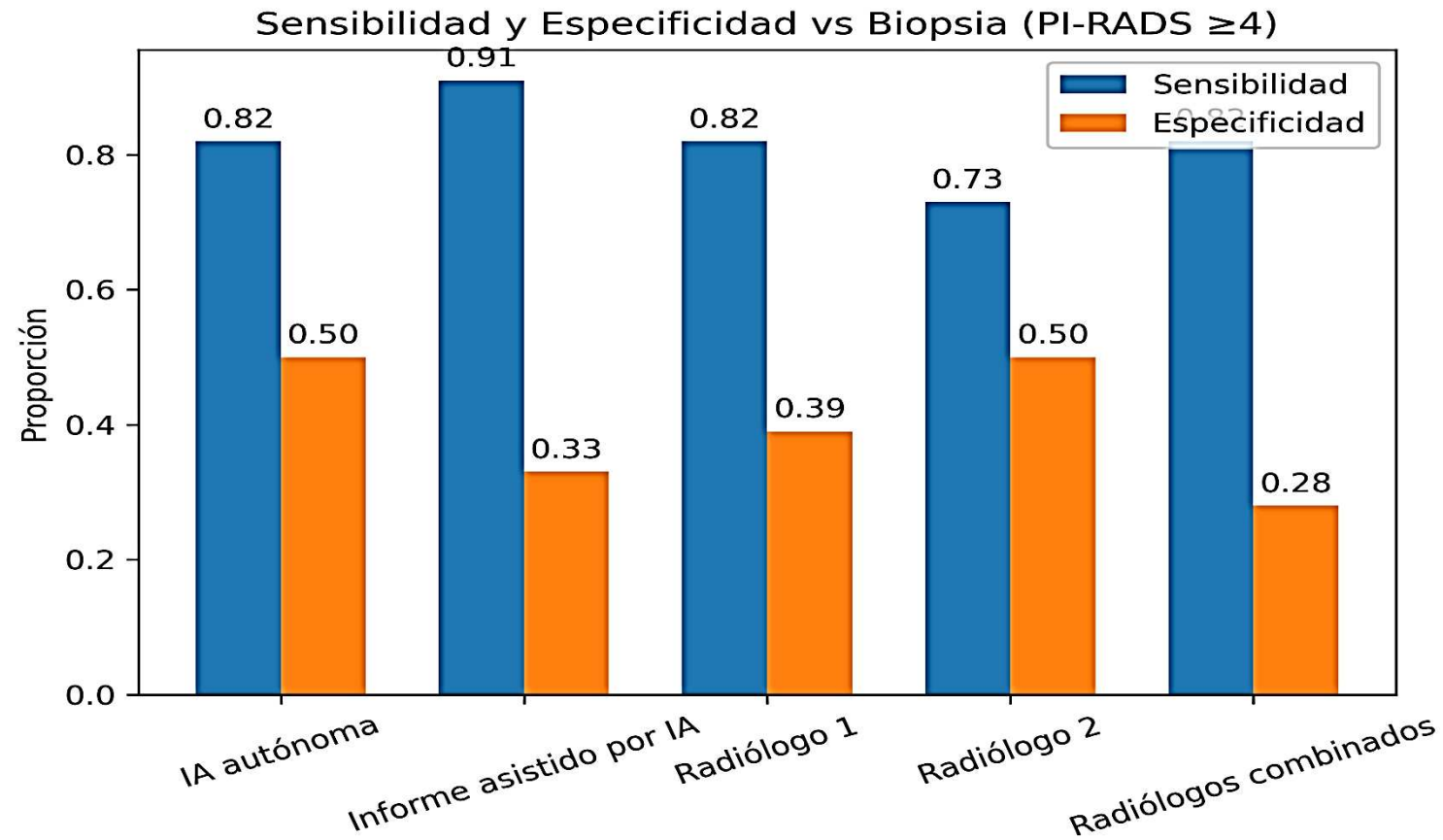


Figura 1. Gráfico de barras muestra la sensibilidad y especificidad para la detección de csCPa (PI-RADS ≥ 4) frente a la biopsia, comparando IA autónoma, informes asistidos por IA y dos radiólogos independientes.

Discusión

- Nuestro estudio muestra un **equilibrio sensibilidad–especificidad** entre las distintas estrategias de interpretación de la mpMRI, en línea con la evidencia publicada [1,2].
- La **concordancia entre radiólogos fue sustancial** ($\kappa = 0,684$), aunque la variabilidad observada se encuentra asociada a la subjetividad del sistema PI-RADS.
- El **informe asistido por IA** alcanzó la mayor sensibilidad, reforzando su posible **papel en la minimización de cánceres no detectados**, aunque con menor especificidad y más falsos positivos. Estos hallazgos respaldan la idea de que **la interacción radiólogo–IA puede favorecer interpretaciones más conservadoras** [4].
- Por el contrario, **la IA autónoma** mostró el perfil diagnóstico más equilibrado en nuestra cohorte. Esto sugiere que los sistemas de IA independientes pueden ser **especialmente útiles en flujos de trabajo de triaje, segunda lectura o como confirmación**, orientados a reducir el sobrediagnóstico y las biopsias innecesarias, más que como sustitutos directos de los radiólogos [5,1].

Discusión

- La unión de ambos radiólogos preservó la sensibilidad pero produjo la menor especificidad, reforzando la idea de que **la doble lectura tienden a aumentar la detección de cáncer a costa de incrementar los falsos positivos**.
 - Cabe destacar que, en nuestros resultados observamos una menor especificidad en la lectura asistida por IA. Este hallazgo probablemente se encuentre influido por el tamaño muestral reducido y la verificación histopatológica parcial.
 - El **informe asistido por IA resulta especialmente útil cuando interesa evitar pasar por alto csPCa**, mientras que la IA autónoma podría estar mejor posicionada para limitar el sobrediagnóstico mediante una toma de decisiones más equilibrada [5].
 - Finalmente, en cuanto a las **limitaciones de este estudio** podemos incluir su diseño retrospectivo, cohorte unicéntrica y verificación parcial mediante biopsia. Por tanto, **estos resultados requieren confirmación en estudios prospectivos multicéntricos de mayor tamaño**.
-

Conclusiones

- La **IA autónoma** alcanzó la mayor exactitud global en nuestra cohorte, manteniendo la sensibilidad y mejorando la especificidad, lo que respalda su papel como **herramienta de triaje o para limitar el sobrediagnóstico, más que como sustituto directo de los radiólogos**.
- En línea con la literatura publicada, nuestros resultados confirman la alta sensibilidad de la IA para la detección de cáncer de próstata clínicamente significativo (csPCa), especialmente en la lectura asistida por IA, aunque con una especificidad menor en condiciones clínicas reales.
- La **lectura asistida por IA maximizó la sensibilidad** pero redujo sustancialmente la especificidad, lo que respalda la idea de que la **interacción humano-IA puede favorecer interpretaciones más conservadoras**, destacando su papel en la minimización de cánceres no detectados.

Referencias

1. Saha A, Bosma JS, Twilt JJ, van Ginneken B, Bjartell A, Padhani AR, Bonekamp D, Villeirs G, Salomon G, Giannarini G, Kalpathy-Cramer J, Barentsz J, Maier-Hein KH, Rusu M, Rouvière O, van den Bergh R, Panebianco V, Kasivisvanathan V, Obuchowski NA, Yakar D, Elschot M, Veltman J, Fütterer JJ, de Rooij M, Huisman H; PI-CAI consortium. Artificial intelligence and radiologists in prostate cancer detection on MRI (PI-CAI): an international, paired, non-inferiority, confirmatory study. *Lancet Oncol*. 2024 Jul;25(7):879-887.
2. Bhattacharya I, Khandwala YS, Vesal S, Shao W, Yang Q, Soerensen SJC, Fan RE, Ghanouni P, Kunder CA, Brooks JD, Hu Y, Rusu M, Sonn GA. A review of artificial intelligence in prostate cancer detection on imaging. *Ther Adv Urol*. 2022 Oct 10;14:17562872221128791.
3. Rosenkrantz AB, Ginocchio LA, Cornfeld D, Froemming AT, Gupta RT, Turkbey B, Westphalen AC, Babb JS, Margolis DJ. Interobserver Reproducibility of the PI-RADS Version 2 Lexicon: A Multicenter Study of Six Experienced Prostate Radiologists. *Radiology*. 2016 Sep;280(3):793-804.
4. van Leeuwen KG, de Rooij M, Schalekamp S, van Ginneken B, Rutten MJCM. Clinical use of artificial intelligence products for radiology in the Netherlands between 2020 and 2022. *Eur Radiol*. 2024 Jan;34(1):348-354.
5. Görtz M, Baumgärtner K, Schmid T, et al. An artificial intelligence-based chatbot for prostate cancer education: Design and patient evaluation study. *DIGITAL HEALTH*. 2023;9.