Gac Sanit. 2026;40(S4):14

Gaceta Sanitaria

https://www.gacetasanitaria.org

675 - VALIDACIÓN Y CALIBRACIÓN DE UN SISTEMA DE INTELIGENCIA ARTIFICIAL PARA EL CRIBADO DE MAMA

A. Muñoz de la Parte, N. López Mintegi, E. Arizaga Batiz, J. González Ocio, K. López Guridi, I. Portillo Villares, I. Moratinos Moreno, P. Fisteos Mazaira, N. González Nieto

Dirección General, Osakidetza; OSI Donostialdea, Osakidetza; OSI Araba, Osakidetza; OSI Barrualde-Galdakao, Osakidetza.

Resumen

Antecedentes/Objetivos: La implementación de sistemas de inteligencia artificial (IA) en programas de cribado mamográfico requiere una validación rigurosa adaptada al contexto local. El Reglamento Europeo de IA (AI Act) establece que estos sistemas no pueden actuar como lectores únicos, lo que obliga a definir estrategias de integración en el flujo de trabajo. El objetivo de este estudio fue validar el rendimiento del sistema Transmural en el Programa de Detección Precoz de Cáncer de Mama del País Vasco y desarrollar un marco metodológico reproducible para la evaluación de futuros algoritmos comerciales.

Métodos: Estudio retrospectivo con diseño de casos y controles (ratio 1:50) utilizando una muestra enriquecida de 51.000 estudios de tomosíntesis realizados entre 2020 y 2023. Se incluyeron 1.000 casos positivos (cánceres detectados por el programa y cánceres de intervalo) confirmados mediante los comités de tumores de Osakidetza. El algoritmo evaluó un único estudio por paciente, sin acceso a imágenes previas ni datos clínicos, considerándose positivo cualquier resultado de riesgo elevado para microcalcificaciones o nódulos. Se analizó el rendimiento diagnóstico (sensibilidad, especificidad, VPP, VPN) y se realizó análisis cualitativo de los falsos negativos. El estudio fue aprobado por el Comité de Ética, que eximió la obtención de consentimiento informado.

0213-9111 © 2026 SESPAS. Publicado por Elsevier España, S.L.U. Todos los derechos reservados

Resultados: Con el punto de corte original del fabricante, la especificidad fue del 6,8%. La recalibración permitió optimizar el rendimiento para diferentes escenarios de estratificación por riesgo, evaluando el impacto en sensibilidad y valores predictivos en cada umbral. El análisis de falsos negativos proporcionó información relevante sobre las limitaciones del algoritmo en nuestro contexto poblacional. Ajustando el punto de corte a una especificidad del 40%, el VPP del riesgo estándar es de 9,33 positivos por 1.000 mujeres cribadas, obteniendo un cáncer por cada 1.000 mujeres cribadas en el grupo bajo riesgo.

Conclusiones/Recomendaciones: Se ha diseñado un marco metodológico robusto para evaluar algoritmos de IA en cribado mamográfico, aplicable a futuros sistemas comerciales. Es esencial evaluar y ajustar estas herramientas al contexto local. El modelo propuesto permite la estratificación de casos en grupos de riesgo (bajo riesgo vs. riesgo estándar), cumpliendo con la normativa europea y optimizando el flujo de trabajo del programa poblacional que realiza 3.000 mamografías semanales.

0213-9111 © 2026 SESPAS. Publicado por Elsevier España, S.L.U. Todos los derechos reservados