

TRABAJO FIN DE GRADO



UCAM

UNIVERSIDAD CATÓLICA
DE MURCIA

FACULTAD DE MEDICINA

Grado en Medicina

Aplicación de la inteligencia artificial en el abordaje diagnóstico del cáncer de mama

Autor: Ricardo Cecilia Clement

Director: Dr. Jose Manuel Alarte Garvía

Murcia, mayo de 2024

TRABAJO FIN DE GRADO



UCAM

**UNIVERSIDAD CATÓLICA
DE MURCIA**

FACULTAD DE MEDICINA

Grado en Medicina

**Aplicación de la inteligencia artificial en el abordaje
diagnóstico del cáncer de mama**

Autor: Ricardo Cecilia Clement

Director: Dr. Jose Manuel Alarte Garví

Murcia, mayo de 2024

TRABAJO FIN DE GRADO



UCAM
UNIVERSIDAD CATÓLICA
SAN ANTONIO

DEFENSA TRABAJO FIN DE GRADO

DATOS DEL ALUMNO	
Apellidos: Cecilia Clement	Nombre: Ricardo
DNI: 20526647K	Grado de Medicina
Facultad de Medicina	
Título del trabajo: Aplicación de la inteligencia artificial en el abordaje diagnóstico del cáncer de mama	

El Dr. Jose Manuel Alarte Garvı, tutor del trabajo reseñado arriba, acreditó su idoneidad y otorgó el V. B.º a su contenido para ir a Tribunal de Trabajo fin de Grado.

En Murcia, a 15 de mayo de 2024.

A handwritten signature in blue ink, appearing to be 'J. Alarte Garvı', is written over a light blue circular stamp or watermark.

Fdo.: José Manuel Alarte Garvı

AGRADECIMIENTOS

En primer lugar, quiero dar las gracias a mi familia, por quererme y apoyarme siempre y por confiar en mí. Especialmente a mis padres, todo lo que consiga habrá sido gracias a ellos, y a mi hermano Rodrigo, por existir y ser tan importante en mi vida.

Agradecer a todos mis amigos, a quienes han estado en las buenas y, sobre todo, en las malas. Gracias a Murcia, por esta etapa en la que tanto he aprendido de la vida y por hacerme coincidir con una persona tan maravillosa como Marina, gracias por estar siempre y permitirme compartir la vida contigo. Gracias a Javier por ser un pilar fundamental durante la carrera.

Gracias también a mi tutor Jose Manuel por saber orientarme a la perfección con este trabajo, por su carisma y su profesionalidad.

Gracias a todos, espero que estéis orgullosos.

ABREVIATURAS:

AA – Aprendizaje Automático

AP – Aprendizaje Profundo

AUROC – Área bajo la curva ROC

CAD – Diseño asistido por computadora

CDIS – Carcinoma Ductal In Situ

CM – Cáncer de Mama

FP – Falsos positivos

IA – Inteligencia Artificial

RE – Receptores de Estrógenos

RNA – Redes Neuronales Artificiales

RP – Receptores de Progesterona

TFN – Tasa Falsos Negativos

TFP – Tasa Falsos Positivos

VPN – Valor Predictivo Negativo

VPP – Valor Predictivo Positivo

ÍNDICE

AGRADECIMIENTOS	9
ABREVIATURAS:	11
RESUMEN	15
ABSTRACT	17
INTRODUCCIÓN	19
EL CÁNCER DE MAMA	19
INTELIGENCIA ARTIFICIAL	20
JUSTIFICACIÓN Y OBJETIVOS	23
MATERIAL Y MÉTODOS	25
DISEÑO DE ESTUDIO	25
CRITERIOS DE SELECCIÓN	25
ESTRATEGIAS DE BÚSQUEDA	25
CRITERIOS DE EXCLUSIÓN	26
SELECCIÓN DE ESTUDIOS.....	26
EXTRACCIÓN DE DATOS.....	26
RESULTADOS Y DISCUSIÓN	27
CONCLUSIONES	33
PROYECTO DE INVESTIGACIÓN	35
BIBLIOGRAFÍA	37
ANEXOS	41

RESUMEN

Introducción: El cáncer de mama (CM) es el cáncer más común y la segunda causa principal de muerte relacionada con el cáncer en todo el mundo. Para el cribado de cáncer de mama se utilizan pruebas como la mamografía, la ecografía y el diagnóstico histopatológico, que guiarán el tratamiento y el pronóstico de los pacientes. Paralelamente, los avances en inteligencia artificial (IA) a la vez que la creciente digitalización del sistema sanitario supone un enfoque prometedor para una posible implantación de esta técnica en la práctica clínica. La detección del CM es un procedimiento susceptible de ser complementado con esta herramienta con el objetivo de mejorar la rentabilidad diagnóstica, así como el flujo de trabajo.

Objetivo: Realizar una revisión de la literatura sobre la utilidad de la aplicación de la IA en el abordaje diagnóstico del CM.

Material y métodos: La obtención de datos para la elaboración de este estudio se ha llevado a cabo mediante una revisión bibliográfica en diferentes bases de datos, incluyéndose finalmente un total de 23 artículos.

Resultados y discusión: En los artículos seleccionados se compara la rentabilidad diagnóstica del CM con el método convencional (de interpretación por dos radiólogos) con las nuevas estrategias que implican el uso de IA. Se analizan tanto las ventajas como los riesgos y limitaciones que pueden presentar. Se aprecia una mejora en el diagnóstico utilizando IA, como también en el flujo de trabajo, a falta de amplios estudios prospectivos que aumenten la validez externa de estos resultados antes de poder implantarlo en la práctica clínica.

Conclusiones: La IA aplicada al abordaje diagnóstico del CM tiene un gran potencial, pero hacen falta estudios con mayor validez externa para permitir su implantación en la práctica clínica.

Palabras clave/ Descriptores: Cáncer de mama, inteligencia artificial, mamografía, rentabilidad diagnóstica.

ABSTRACT

Background: Breast cancer (BC) is the most common cancer and the second leading cause of cancer-related death worldwide. Screening for breast cancer involves tests such as mammography, ultrasound, and histopathological diagnosis, which guide the treatment and prognosis of patients. Concurrently, advances in artificial intelligence (AI) along with the increasing digitization of the healthcare system represent a promising approach for potential implementation of this technique in clinical practice. The detection of BC is a procedure that can be complemented with this tool with the aim of improving diagnostic yield as well as workflow efficiency.

Objective: To conduct a literature review on the utility of applying AI in the diagnostic approach to breast cancer (BC).

Material and methods: The data for this study was obtained through a literature review in databases, ultimately including a total of 23 articles.

Results and discussion: The selected articles compare the diagnostic yield of BC using conventional methods (interpretation by two radiologists) versus new strategies involving AI. Both the advantages and the risks and limitations that may arise are analyzed. An improvement in diagnosis using AI is observed, as well as in workflow efficiency, pending extensive prospective studies to increase the external validity of these results before implementation in clinical practice can be considered.

Conclusions: AI applied to the diagnostic approach of BC holds great potential, but further studies with higher external validity are needed to enable its implementation in clinical practice.

Key words: Breast cancer, artificial intelligence, mammography, diagnostic yield.

INTRODUCCIÓN

EL CÁNCER DE MAMA

El cáncer de mama (CM) es el tipo de neoplasia más común y la segunda causa de muerte por cáncer en el mundo ((1). Consiste en la proliferación veloz y descontrolada de las células del epitelio glandular, con una capacidad reproductiva incrementada. Las células cancerígenas pueden diseminarse por vía linfática o sanguínea hacia otras partes del cuerpo y allí adherirse formando metástasis. El 99% se da en mujeres frente al 1% en varones (2). Los estudios realizados hasta la fecha destacan diferentes factores de riesgo asociados al desarrollo o progresión de la enfermedad (Tabla 1) (3).

La detección e identificación temprana es muy conveniente, ya que aumenta las opciones de tratamiento y disminuye la tasa de mortalidad. Por eso, muchos países han implementado sistemas de cribado para las mujeres a partir de los 40 años (4). El cribado consiste en una mamografía bianual analizada por dos radiólogos en busca de cualquier indicio de cáncer incipiente (5). Una vez hemos detectado el cáncer por mamografía, el siguiente paso será el estudio histopatológico e inmunohistoquímico en el que obtendremos el diagnóstico de confirmación y lo clasificaremos según su subtipo (Tabla 2). Posteriormente, se realiza un estudio de extensión, mediante técnicas como la RMN y el PET-TC para obtener su estadiaje en base a su TNM. Con todos estos datos podremos orientarnos hacia una terapia y un pronóstico determinado, con el que se determinará un plan de tratamiento individualizado para cada paciente (6).

El creciente conocimiento de la complejidad del cáncer y la disponibilidad de tratamientos individualizados sumado a los continuos avances en IA y la digitalización de las imágenes para el diagnóstico primario impulsan la necesidad de una detección, clasificación y predicción más precisas del comportamiento de los tumores de mama (1). Paralelamente coexiste un incremento en la demanda de radiólogos especializados en mama y una preocupante tendencia hacia el agotamiento profesional.

Por consiguiente, la mamografía supone un entorno idóneo para la implementación de la IA, que ofrece una solución prometedora para el desafío de gestionar eficientemente procesos de trabajo repetitivos de grandes volúmenes de datos y el análisis de patrones complejos. El objetivo es la evaluación de las imágenes de manera más precisa y cuantitativa, además de una considerable mejora de la eficiencia del flujo de trabajo de diagnóstico (1).

INTELIGENCIA ARTIFICIAL

La IA se define como "la rama de la informática dedicada al desarrollo de algoritmos informáticos para llevar a cabo tareas tradicionalmente asociadas con la inteligencia humana, como la capacidad de aprender y resolver problemas"(7). Este amplio término abarca una serie de elementos y métodos de entrenamiento distintos, que incluyen las redes neuronales artificiales (RNA), el diseño asistido por computadora (CAD), el aprendizaje automático (AA) y el aprendizaje profundo (AP). Esta tecnología tiene un énfasis creciente en radiología porque los píxeles de las imágenes suponen datos primarios que pueden alimentar y enseñar a diferentes algoritmos de IA (8).

Desde la década de 1980, los investigadores han estado desarrollando técnicas de AA para el CAD con el fin de distinguir entre lesiones mamarias malignas y benignas. A partir de la imagen de una lesión, el algoritmo de IA extrae características radiómicas mediante ingeniería humana o integra automáticamente características predictivas en el caso del AP, para luego generar una probabilidad de malignidad de la lesión. (6). El análisis radiómico implica la extracción y correlación de múltiples parámetros de imagen con diversas variables de interés, que incluyen características del paciente y datos histopatológicos, genómicos y moleculares. Estos datos se utilizan para crear modelos de apoyo en la toma de decisiones, que pueden emplearse en varios contextos, como la planificación del tratamiento, la evaluación del riesgo y la predicción de resultados. (9).

El sistema CAD no puede utilizarse como herramienta única de diagnóstico, los radiólogos deberán examinar y evaluar las imágenes bajo su criterio y usar el CAD como un sistema de apoyo verificador. Esta diseñado y entrenado para identificar características específicas, como masas o calcificaciones y presenta una elevada sensibilidad, lo que conlleva una alta tasa de falsos positivos (FP). Por ello, supone un riesgo de sobrediagnóstico y las consecuencias que ello implica, como en el carcinoma ductal in situ (CDIS) de bajo grado, perturbando el equilibrio entre beneficios y riesgos. Esta herramienta ha evolucionado con el tiempo, pasando de enfoques tradicionales a métodos modernos de AP, capaces de procesar grandes volúmenes de datos, aprender características tanto de las imágenes como de los errores y mejorar su rendimiento con el tiempo (4).

Por otro lado, el AA representa un subcampo cuya función es crear algoritmos capaces de adaptarse a nuevos problemas a partir de datos, aplicando métodos estadísticos, sin necesidad de ser reprogramados (1). Mediante algoritmos matemáticos, se adaptan y resuelven problemas relacionando variables que pueden ser simples, como talla, peso, edad, etc ; o complejas, como imágenes histopatológicas. La complejidad puede aumentar en base a la precisión que se desee obtener, con el objetivo de asistir a los patólogos en el diagnóstico (1).

El AP es una modificación extrema del AA en la que la imagen entra a una red neuronal que modifica y reduce su tamaño hasta convertirla en una secuencia numérica (10). Estos números representan una serie de características que conforman el algoritmo de AA, cuyo objetivo es detectarlas en las imágenes, como líneas, bordes, texturas e intensidad, así como otras más complejas como formas, lesiones u órganos (8).

Con los avances continuos en el análisis radiológico y el AA, como el AP, se está progresando hacia la meta de proporcionar una atención más efectiva, eficiente y centrada en el paciente para el manejo del CM ((9)

JUSTIFICACIÓN Y OBJETIVOS

El objetivo principal de este trabajo es realizar una revisión de la literatura sobre la utilidad de la aplicación de la inteligencia artificial en el diagnóstico del cáncer de mama.

Como objetivos secundarios nos hemos propuesto:

1. Valorar el grado de rentabilidad diagnóstica que tendría su implantación en la práctica clínica.
2. Determinar las ventajas e inconvenientes de la aplicación de técnicas de IA en las técnicas de imagen diagnóstica de patología mamaria
3. Analizar los riesgos que puede conllevar su utilización y las limitaciones de su uso.

MATERIAL Y MÉTODOS

DISEÑO DE ESTUDIO

El estudio se ha llevado a cabo mediante una búsqueda de artículos en diferentes bases de datos con el fin de realizar una revisión bibliográfica de la literatura referente a la aplicación de la inteligencia artificial en el abordaje diagnóstico del cáncer de mama según la evidencia científica publicada en la actualidad.

CRITERIOS DE SELECCIÓN

De los criterios que se obtuvieron al realizar la búsqueda sistemática en las bases de datos, se seleccionaron los artículos incluidos en el estudio teniendo en cuenta los siguientes criterios de selección:

- Artículos redactados en castellano o en inglés, independientemente de la nacionalidad de los investigadores.
- Artículos publicados en los últimos 10 años.
- Artículos que incluyesen las palabras claves.
- Incluir revisiones sistemáticas, metaanálisis, estudios prospectivos, retrospectivos y guías de las sociedades científicas.

ESTRATEGIAS DE BÚSQUEDA

De acuerdo al modelo PICO, se establecieron el objetivo y las características de la revisión. A la hora de realizar la búsqueda bibliográfica, se utilizó la terminología MeSH (Medical Subject Headings) mediante la estrategia de búsqueda: [(breast cancer) AND (ai artificial intelligence)].

Se realizó una primera búsqueda dirigida en la base de datos de PubMed. Después se realizó una segunda búsqueda, dirigida y complementaria, en esta misma base de datos para la búsqueda de guías clínicas de la inteligencia artificial en el campo del cáncer de mama. Se consultaron además las páginas web de la Organización Mundial de la Salud (OMS), Sociedad Española de Oncología Médica y el Ministerio de Sanidad.

CRITERIOS DE EXCLUSIÓN

- Artículos duplicados
- No accesibilidad o disponibilidad del texto completo

SELECCIÓN DE ESTUDIOS

Tras comenzar la búsqueda, los artículos se seleccionaron en base a la lectura tanto del título como del “abstract”, de acuerdo con los criterios de selección mencionados anteriormente.

EXTRACCIÓN DE DATOS

En la primera búsqueda en la base de datos PubMed con las palabras [(breast cancer) AND (ai artificial intelligence)] obtuve un total de 304 artículos.

Ante los exponenciales avances de la IA durante los últimos años y a su vez con el objetivo de acotar el resultado, se ajustaron ciertos criterios de búsqueda, ajustándose a los criterios de inclusión previamente definidos:

- Incluir artículos en los últimos 10 años lo que reduce el número de artículos a 207, de los cuales eran gratuitos 116.
- Del resultado anterior seleccionamos con la lectura del título y del “Abstract” un total de 23 artículos.

Para proceder a la selección de los 23 artículos definitivos, se revisaron los artículos completos, con el fin de decidir si la información que contenían estaba o no relacionada con nuestro objetivo (Figura 1).

RESULTADOS Y DISCUSIÓN

Los estudios analizados tienen diversos países de origen, siendo EE. UU. el país que más ha contribuido presentando 5 artículos, en segundo lugar, Suecia con 2 publicaciones. Destacar la presencia de una publicación local española, y finalmente Países Bajos, China, UK con 1 estudio. En cuanto al diseño, se analizan 8 estudios retrospectivos, 2 estudios prospectivos de cohortes y una revisión bibliográfica.

Los objetivos propuestos en la mayoría de los estudios se basan en la evaluación de la utilidad y del rendimiento diagnóstico de la IA en la detección del CM, comparándolo con el método convencional de detección por dos radiólogos e incluso creando un modelo híbrido compuesto por radiólogos ayudados por IA (Tabla 3). También se quiso analizar la tasa de falsos positivos (TFP) obtenidos (11,12), así como las limitaciones (13) y una posible mejora de la carga de trabajo sin reducir la precisión (14).

Respecto al material y métodos utilizados (Tabla 4), se han determinado como criterios de inclusión a mujeres que se habían realizado alguna prueba de imagen en un período concreto de tiempo (12,15–20), mientras que otras publicaciones no especificaban sobre éstos (13,14,21). En cuanto a los criterios de exclusión, el más comunmente observado ha sido el antecedente de CM (15–18); seguido de la presencia de implantes mamarios (15,17,18).

Se aprecian diferencias importantes entre los estudios encontrados en cuanto al tamaño muestral, que oscila desde 240 (17), hasta un máximo de 153.588 participantes (13). La edad de los participantes se sitúa entre los 39 y 89 años y la gran mayoría son mujeres.

En cuanto al método de valoración de los estudios, se llevaron a cabo diferentes estrategias. La mayoría fueron doble ciego, analizándolo de dos maneras: una lectura de las imágenes por un algoritmo de IA y otra lectura paralela por uno o dos radiólogos según el método tradicional. En algunos de ellos, se lleva a cabo una tercera lectura por un modelo híbrido conformado por

la unión de un radiólogo e IA (12,13,15,17,18,20,21). Los criterios evaluados fueron principalmente el área bajo la curva ROC (AUROC), la sensibilidad y la especificidad, en busca de analizar la precisión diagnóstica de cada estrategia. También se midieron otros parámetros como el valor predictivo positivo (VPP), el valor predictivo negativo (VPN), la tasa de falsos negativos (TFN), la tasa de falsos positivos (TFP) (15) y el tiempo de lectura con cada estrategia (16).

Dembrower et al. (15) realizaron un estudio prospectivo, doble ciego, en el que de 58.344 pacientes se detectan 250 CM (0,45%) con la lectura por dos radiólogos, 261 CM en la lectura híbrida por un radiólogo + IA (0,47%), 246 CM detectados por la lectura única por IA y 269 por la triple lectura de dos radiólogos + IA. Por lo tanto, respecto a la estrategia convencional de lectura por dos radiólogos, se consideró que: la doble lectura por un radiólogo + IA es superior, dando una proporción relativa de 1.04 (intervalo de confianza (IC) del 95% 1.00-1.09; $p=0.017$); la lectura única por IA es no inferior ($p<0.0001$), pero no superior ($p=0.73$), dando una proporción relativa de 0.98 (IC del 95% 0.93-1.04); la triple lectura por parte de dos radiólogos + IA es superior ($p<0.0001$) dando una proporción relativa de 1.08 (IC del 95% 1.04-1.11). La IA y el ojo humano perciben características diferentes de las imágenes, por lo tanto, crean sinergia para aumentar la sensibilidad en la detección del CM en la mamografía. Reemplazar a un radiólogo por IA en una población disminuiría tiempo de lectura a la vez que aumentaría las discusiones de consenso entre radiólogos. Aun así, la reducción de la carga de trabajo sería considerable; al igual que con la lectura única con IA. Todavía es necesario perfeccionar la técnica, que a su vez abre un debate sobre la responsabilidad médica, la aceptabilidad pública, la capacitación de los radiólogos y la certificación de los sistemas de IA.

Wang et al. (16) llevaron a cabo un estudio prospectivo de cohortes, doble ciego, en el que se analizaron a 2202 pacientes. Comparó la precisión diagnóstica de la IA con la estrategia convencional de dos radiólogos obteniendo respectivamente los siguientes datos: Sensibilidad 84,5% - 89,2%; especificidad 83,8% - 63%, VPP 92,3% - 84,7%, VPN 70% - 73%, TFN 15,5% - 10,8%, TFP 16,1% - 37%. La mamografía basada en

IA obtuvo un AUROC de 0.9. De acuerdo con estos parámetros, se halló una mayor precisión y rendimiento diagnóstico de la IA que de los radiólogos. Se determinó que la precisión y rentabilidad diagnóstica de la IA fue mayor que la de los radiólogos, lo que indica el potencial de esta herramienta de lectura de forma independiente en la detección preclínica del CM. Además, se calculó que el uso de la IA sería más rentable económicamente que el de la mamografía y ecografía, con un gasto de 21.266\$ frente a 83.346\$, respectivamente.

Rodríguez- Ruiz et al. (17) desarrollaron un estudio retrospectivo, doble ciego, en el que estudiaron a 240 mujeres. Se obtuvo un mejor rendimiento de detección con IA, obteniendo un AUROC de 0.89 frente a 0.87 sin ayuda de IA. La sensibilidad fue 3 puntos porcentuales y la especificidad 2 puntos porcentuales más alta con el uso de la IA. El tiempo de lectura sin IA fue de 146 segundos (IC del 95 %: 143 segundos - 149 segundos) y con IA de 149 segundos (IC del 95 %: 146 segundos - 152 segundos) lo que supuso una diferencia no significativa. Por lo tanto, se asume que la IA puede tener una importante relevancia clínica ya que mejoró la evaluación de los casos erróneos. El aumento del AUROC fue mayor para los radiólogos con menor experiencia con mamografía, quienes son más propensos a adoptar nuevas técnicas para mejorar su rendimiento. El tiempo de lectura en la práctica clínica disminuiría un 4,5%, permitiendo dedicarles mayor atención a los casos más complejos.

Schaffter et al.(13)) elaboraron un estudio retrospectivo en el que se analizó a 153.588 mujeres de EE. UU. y Suecia. Se obtuvo un AUROC con el uso de IA de 0,858 (EE. UU.) y 0,903 (Suecia); y 66,2 % (EE. UU.) y 81,2 % (Suecia) en la sensibilidad de los radiólogos menor que la especificidad, del 90,5% (EE. UU.) y 98,5% (Suecia). La combinación de algoritmos de alto rendimiento y evaluaciones de radiólogos de EE. UU. dio como resultado un AUROC más alta de 0,942 y logró una especificidad significativamente mejorada (92,0%) con la misma sensibilidad. En este estudio la IA no superó la precisión diagnóstica de los radiólogos, pero la combinación IA más radiólogo sí que obtuvo un mejor rendimiento, lo que indica que esta herramienta será importante de cara al futuro para mejorar la precisión de la mamografía. Aun así, existe un

amplio margen de mejora en los algoritmos de IA mediante la valoración de pruebas de imágenes previas o de características adicionales de cada paciente, a la vez que se reducirá el gasto en el sistema de atención médica (Tabla 5).

Según la mayoría de los estudios analizados, la cooperación de los radiólogos y la IA produciría una reducción y mejora en la carga de trabajo. Sería posible descartar automáticamente los casos más sencillos de bajo riesgo permitiendo prestar mayor atención y cantidad de tiempo a los casos más tediosos, mejorando el flujo de trabajo del sistema médico (12,17,19,21,22). Se debe tener en cuenta que la detección mamográfica del CM no es la única función de los radiólogos, ya que es solo el primer paso de un proceso de diagnóstico, con una decisión final necesaria por su parte sobre la realización o no de biopsia (20). Además, los sistemas de IA encargados de predecir la respuesta al tratamiento podrán guiar estrategias terapéuticas personalizadas, así como enfoques de detección dirigidos en sistemas que identifiquen a las personas con mayor riesgo de desarrollar CM (22).

Los radiólogos se enfrentan a desafíos en el procesamiento de las pruebas de imagen, cuya interpretación se puede ver alterada por una elevada densidad mamaria, que impide la detección durante la evaluación inicial (12). Otro desafío son los FP, que desencadenan seguimientos y técnicas invasivas innecesarias, suponiendo un incremento del gasto público y de la carga de trabajo, además de las consecuencias psicológicas que puedan causar en los pacientes por este sobrediagnóstico y sobretratamiento, como pasa frecuentemente con el CDIS de bajo grado (5,23). Según *Shen et al.* (12), la colaboración de los radiólogos y la IA podría reducir la tasa de biopsias y de FP, mejorando la precisión diagnóstica actual. No obstante, esta herramienta no está exenta de una serie de riesgos y limitaciones.

En el caso de los estudios retrospectivos, la prevalencia del CM difiere a la de la práctica clínica, al verse enriquecida (13). Esto supone la presencia de un sesgo de “efecto de laboratorio” por parte de los radiólogos, al ser conscientes de la incrementada tasa de neoplasias a la que se enfrentaban (17). Además, los radiólogos cuentan con poca experiencia en el análisis de imágenes con IA

por lo que requerían la ayuda de técnicos experimentados. Así como tampoco es posible conocer el impacto en el rendimiento de los radiólogos si solo leyese los casos más complejos, al utilizar la IA únicamente para los casos sencillos (14). Cabe destacar que en los estudios analizados no se lleva a cabo un diagnóstico basado en el seguimiento y evolución de las pruebas de imagen, sino que se basa en una única imagen de cada paciente; como tampoco se evalúan aspectos como factores de riesgo, antecedentes familiares, mutaciones genéticas o mamografías previas (12,22).

Aunque los resultados pueden ser prometedores, se requiere una mayor validez externa para su implantación en la práctica clínica (20), por lo que será necesaria la realización de nuevos y amplios estudios prospectivos en un futuro (14,18,19,21,22), así como exámenes exhaustivos de los aspectos ético-legales y sociales sobre remplazar a un profesional médico por un software informático (18). Será necesario desarrollar y capacitar un nuevo equipo multidisciplinario formado por médicos y técnicos informáticos para incorporar el análisis por IA en las decisiones de la práctica clínica (22).

CONCLUSIONES

1. En este trabajo se ha conseguido realizar una aproximación a la literatura científica actual referida a la utilidad de la aplicación de la IA en el diagnóstico de CM.
2. En cuanto a las características generales de los textos utilizados en esta revisión, se puede observar el gran potencial que tiene en medicina la inteligencia artificial, sobre todo en el campo de la radiología, como herramienta diagnóstica complementaria a los médicos. Su importancia y utilidad continuará al alza a la vez que se siga avanzando en la digitalización de la práctica clínica.
3. La aplicación de técnicas de IA en la práctica clínica proporcionaría una serie de ventajas como una mejora en el flujo de trabajo reduciendo la carga laboral a los profesionales, con una probable disminución del gasto público y una mayor calidad en la precisión diagnóstica de la detección del cáncer de mama
4. La falta de experiencia de los radiólogos con esta herramienta, así como una prevalencia de cáncer de mama no aplicable a la práctica clínica y el no haber tenido en cuenta factores adicionales como los antecedentes o las mamografías previas y los factores de riesgo de los pacientes, suponen una serie de riesgos y limitaciones a tener en cuenta. Todavía existe un inmenso margen de mejora por delante, por lo que se hace necesaria la realización de amplios estudios prospectivos con el objetivo de perfeccionar y aumentar la evidencia de esta técnica.

PROYECTO DE INVESTIGACIÓN

El análisis de la revisión bibliográfica realizada para la elaboración de este trabajo de fin de grado me ha suscitado el interés para realizar un trabajo de investigación sobre esta área de conocimiento, que podría definirse en los puntos que desarrollo a continuación:

- **Hipótesis:** La detección del CM presenta una mayor rentabilidad diagnóstica utilizando la IA como herramienta complementaria a los radiólogos.
- **Objetivo:** Comparar la rentabilidad diagnóstica del CM utilizando la IA analizando paralelamente tanto sus ventajas como sus limitaciones.
- **Metodología y plan de trabajo:** Se realizará un estudio experimental en el que se estudian mujeres que se han realizado una prueba de cribado de CM.

En primer lugar, se realizará un análisis observacional de los datos de la prueba de cribado y posteriormente se compararán los resultados con diferentes modos de diagnóstico. Se estudiará un grupo de 300 mujeres entre los 50 y 69 años que asistirá a realizarse una mamografía de cribado para el CM. Dividiremos el total en tres grupos de 100 mujeres. Realizaremos un estudio paralelo de cada grupo con una estrategia diferente de detección. El primer grupo será estudiado por dos radiólogos, el método tradicional; el segundo grupo será estudiado por IA; y el tercer grupo por IA y un radiólogo. Los tres grupos cumplirán los mismos criterios de inclusión y exclusión.

En los grupos se incluirán mujeres multiétnicas de toda España. Se excluyen pacientes con implantes mamarios, antecedentes de CM y con mutaciones genéticas BRCA 1 y 2 conocidas. Se evaluará la sensibilidad, especificidad, área bajo la curva ROC, la tasa de falsos positivos y negativos, el valor predictivo positivo y negativo, así como el tiempo de lectura con cada estrategia.

BIBLIOGRAFIA

1. Ibrahim A, Gamble P, Jaroensri R, Abdelsamea MM, Mermel CH, Chen PHC, et al. Artificial intelligence in digital breast pathology: Techniques and applications. *Breast*. 2020 Feb 1;49:267–73.
2. Gucalp A, Traina TA, Eisner JR, Parker JS, Selitsky SR, Park BH, et al. Male breast cancer: a disease distinct from female breast cancer. Vol. 173, *Breast Cancer Research and Treatment*. Springer New York LLC; 2019. p. 37–48.
3. International BMR. Retracted: Global Increase in Breast Cancer Incidence: Risk Factors and Preventive Measures. Vol. 2023, *BioMed research international*. 2023. p. 9872034.
4. Alsharif WM. The utilization of artificial intelligence applications to improve breast cancer detection and prognosis. Vol. 44, *Saudi Medical Journal*. Saudi Arabian Armed Forces Hospital; 2023. p. 119–27.
5. Freeman K, Geppert J, Stinton C, Todkill D, Johnson S, Clarke A, et al. Use of artificial intelligence for image analysis in breast cancer screening programmes: Systematic review of test accuracy. *The BMJ*. 2021 Sep 1;374.
6. Hu Q, Giger ML. Clinical Artificial Intelligence Applications: Breast Imaging. Vol. 59, *Radiologic Clinics of North America*. W.B. Saunders; 2021. p. 1027–43.
7. Tang A, Tam R, Cadrin-Chênevert A, Guest W, Chong J, Barfett J, et al. Canadian Association of Radiologists White Paper on Artificial Intelligence in Radiology. Vol. 69, *Canadian Association of Radiologists Journal*. Canadian Medical Association; 2018. p. 120–35.
8. Mendelson EB. Artificial intelligence in breast imaging: Potentials and limitations. Vol. 212, *American Journal of Roentgenology*. American Roentgen Ray Society; 2019. p. 293–9.

9. Bitencourt A, Daimiel Naranjo I, Lo Gullo R, Rossi Saccarelli C, Pinker K. AI-enhanced breast imaging: Where are we and where are we heading? Vol. 142, *European Journal of Radiology*. Elsevier Ireland Ltd; 2021.
10. Tagliafico AS, Piana M, Schenone D, Lai R, Massone AM, Houssami N. Overview of radiomics in breast cancer diagnosis and prognostication. *Breast*. 2020 Feb 1;49:74–80.
11. Dembrower K, Crippa A, Colón E, Eklund M, Strand F. Artificial intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study. *Lancet Digit Health*. 2023 Oct 1;5(10):e703–11.
12. Shen Y, Shamout FE, Oliver JR, Witowski J, Kannan K, Park J, et al. Artificial intelligence system reduces false-positive findings in the interpretation of breast ultrasound exams. *Nat Commun*. 2021 Dec 1;12(1).
13. Schaffter T, Buist DSM, Lee CI, Nikulin Y, Ribli D, Guan Y, et al. Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms. *JAMA Netw Open*. 2020 Mar 2;3(3):E200265.
14. Raya-Povedano JL, Romero-Martín S, Elías-Cabot E, Gubern-Mérida A, Rodríguez-Ruiz A, Álvarez-Benito M. AI-based Strategies to Reduce Workload in Breast Cancer Screening with Mammography and Tomosynthesis: A Retrospective Evaluation. *Radiology*. 2021 Jul 1;300(1):57–65.
15. Dembrower K, Wåhlin E, Liu Y, Salim M, Smith K, Lindholm P, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. *Lancet Digit Health*. 2020 Sep 1;2(9):e468–74.
16. Wang X, Chou K, Zhang G, Zuo Z, Zhang T, Zhou Y, et al. Breast cancer pre-clinical screening using infrared thermography and artificial intelligence: a prospective, multicentre, diagnostic accuracy cohort study. *Int J Surg*. 2023 Oct 1;109(10):3021–31.

17. Rodríguez-Ruiz A, Krupinski E, Mordang JJ, Schilling K, Heywang-Köbrunner SH, Sechopoulos I, et al. Detection of breast cancer with mammography: Effect of an artificial intelligence support system. *Radiology*. 2019 Mar 1;290(3):305–14.
18. Salim M, Wåhlin E, Dembrower K, Azavedo E, Foukakis T, Liu Y, et al. External Evaluation of 3 Commercial Artificial Intelligence Algorithms for Independent Assessment of Screening Mammograms. *JAMA Oncol*. 2020 Oct 1;6(10):1581–8.
19. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature*. 2020 Jan 2;577(7788):89–94.
20. Wu N, Phang J, Park J, Shen Y, Huang Z, Zorin M, et al. Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening. *IEEE Trans Med Imaging*. 2020 Apr 1;39(4):1184–94.
21. Kim HE, Kim HH, Han BK, Kim KH, Han K, Nam H, et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. *Lancet Digit Health*. 2020 Mar 1;2(3):e138–48.
22. Hickman SE, Baxter GC, Gilbert FJ. Adoption of artificial intelligence in breast imaging: evaluation, ethical constraints and limitations. Vol. 125, *British Journal of Cancer*. Springer Nature; 2021. p. 15–22.
23. Alsharif WM. The utilization of artificial intelligence applications to improve breast cancer detection and prognosis. Vol. 44, *Saudi Medical Journal*. Saudi Arabian Armed Forces Hospital; 2023. p. 119–27.
24. Fundación Española de Senología y Patología Mamaria. *Manual de práctica clínica en senología*, 2019.

ANEXOS

FIGURA 1. Diagrama de flujo PRISMA para ilustrar la recopilación bibliográfica.

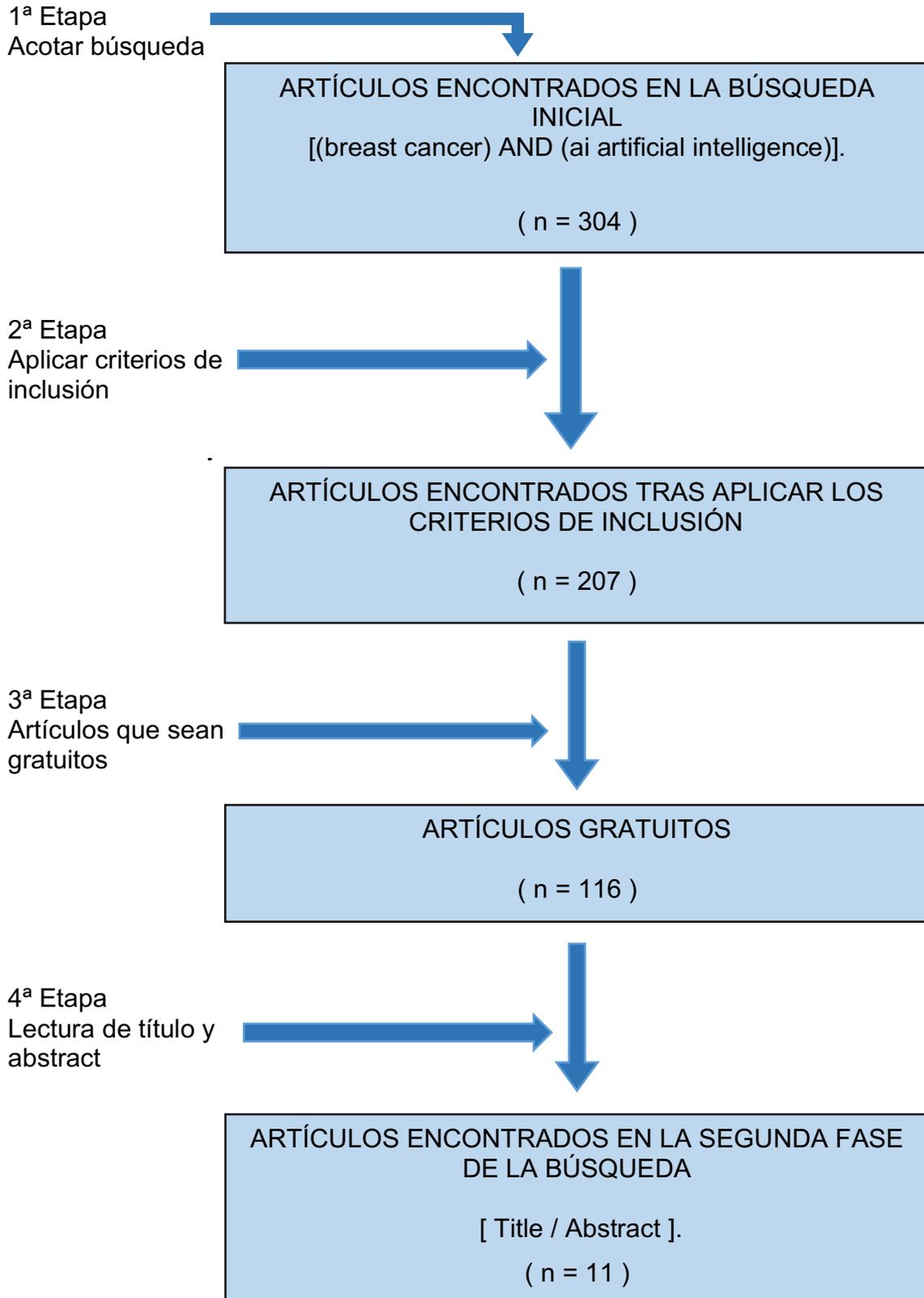


TABLA 1. Factores de riesgo para desarrollar cáncer de mama (CM) (24).

FACTORES DE RIESGO PARA DESARROLLAR CM
- Sexo femenino
- Raza blanca
- Edad
- Peso y grasa corporal
- Talla
- Historia personal de patología mamaria benigna
- Densidad mamaria
- Densidad mineral ósea
- Factores hormonales: <ul style="list-style-type: none"> ✓ Niveles de estrógenos y andrógenos elevados ✓ Terapia Hormonal Sustitutiva ✓ Anticonceptivos orales ✓ Menarquia precoz y menopausia tardía
- Factores reproductivos <ul style="list-style-type: none"> ✓ Nuliparidad y multiparidad ✓ Mayor edad en el primer embarazo
- Historia de cáncer de mama <ul style="list-style-type: none"> ✓ Antecedentes personales de cáncer de mama o CDIS ✓ Antecedentes familiares de CM ✓ Mutaciones genéticas hereditarias
- Exposición a radiación ionizante terapéutica
- Estilos de vida: <ul style="list-style-type: none"> ✓ Alcohol ✓ Tabaco ✓ Factores dietéticos ✓ Trabajo nocturno
- Factores medioambientales

TABLA 2. Clasificación inmunohistoquímica del cáncer de mama (CM) (24).

SUBTIPO	RE / RP	HER2	Ki - 67	PRONÓSTICO
Luminal A	RE + / RP +	-	Bajo (menor del 14%)	Bueno
Luminal B	RE + / RP +	-	Alto (mayor del 14 %)	Bueno
Her 2	RE - / RP -	+	Alto	Malo
Basal-like	RE - / RP -	-	Alto	El de peor pronóstico

RE – Receptores de estrógenos; **RP** – Receptores de progesterona;

HER2 – Receptor 2 del factor de crecimiento epidérmico humano; **Ki-67** – Antígeno Ki-67

TABLA 3. Identificación de los estudios, tabla de elaboración propia.

CM – Cáncer de mama ; FP – Falsos positivos ; IA – Inteligencia artificial

AUTOR	AÑO	PAÍS	REVISTA	DISEÑO	OBJETIVO
Karin Dembrower et al.	2023	Suecia	Lancet Digit Health	Ensayo clínico prospectivo	Evaluar la utilidad de la IA en la detección del CM y en los hallazgos FP
Xuefei Wang et al.	2023	China	International Journal of Surgery	Estudio de cohortes prospectivo, multicéntrico	Evaluar el rendimiento del diagnóstico con IA frente al método convencional
Alejandro Rodríguez-Ruiz et al.	2018	Países Bajos	Radiology	Retrospectivo	Comparar rendimiento diagnóstico de CM de radiólogos con ayuda de IA con radiólogos sin ayuda.
Yiqiu Shen et al.	2021	EEUU	Nature communications	Retrospectivo	Plantear el uso de IA en la identificación de lesiones malignas de imágenes de mama para reducir los falsos positivos
Tomas Schaffter et al.	2020	EEUU	JAMA Network Open	Retrospectivo	Evaluar si la IA puede superar las limitaciones de interpretación de la mamografía
Hoy-Eun Kim et al.	2020	EEUU	Lancet Digit Health	Retrospectivo	Explorar si un algoritmo de IA podría beneficiar a los radiólogos a mejorar la precisión en el diagnóstico del CM
Jose Luis Raya-Povedano	2021	España	Radiology	Retrospectivo	Evaluar si mediante el uso de IA se podría reducir la carga de trabajo sin reducir la precisión en la detección del CM
Mattie Salim et al.	2020	Suecia	JAMA Oncology	Retrospectivo	Evaluación de estrategias de detección de CM comparando rendimiento diagnóstico de tres algoritmos de IA (IA1, IA2 y IA3), de los radiólogos y de su combinación.
Scott Mayer McKinney	2020	EEUU	Nature	Retrospectivo	Evaluar el rendimiento diagnóstico de una IA frente al de los radiólogos

Nan Wu et al.	2020	EEUU	IEEE Transactions on Medical Imaging	Retrospectivo	Contribuir en el desarrollo de redes neuronales que apoyen a los radiólogos en la interpretación de los exámenes de detección del CM.
Sarah E. Hickman et al.	2021	UK	British Journal of Cancer	Revisión bibliográfica	Explorar sobre el uso de la IA en patología mamaria y sus principales desafíos éticos, técnicos y legales que se han identificado hasta ahora.
Asma Ibrahim et al.	2019	Reino Unido	Breast	Revisión bibliográfica	Explorar la aplicación de la IA en la patología mamaria
Wala M. Alsharif	2023	Arabia Saudí	Saudí Med Journal	Revisión bibliográfica	Revisar las aplicaciones de la IA para la identificación, segmentación y categorización de lesiones, densidad mamaria y evaluación del riesgo de cáncer.
Karoline Freeman et al.	2021	Reino Unido	British Medical Journal	Revisión bibliográfica	Examinar la precisión de la IA para la detección del CM en la práctica de screening mamográfico
Qiyuan Hu et al.	2021	EEUU	Radiológica Clinics of North America	Revisión bibliográfica	Valorar el uso de la IA en la detección, diagnóstico y pronóstico del CM
An Tang et al.	2018	Canada	Canadian Association por Radiologists Journal	Revisión bibliográfica	Informar sobre el impacto de la IA en la radiología canadiense
Ellen B. Mendelson	2019	EEUU	American Journal of Roentgenology	Revisión bibliográfica	Valorar las potenciales aplicaciones de la IA en las imágenes de patología mamaria y sus limitaciones.
Almir Bitencourt et al.	2021	EEUU	European Journal of Radiology	Revisión bibliográfica	Revisar el conocimiento actual y las aplicaciones futuras de las imágenes mamarias mejoradas por IA en la práctica clínica
Alberto Stefano Tagliafico et al.	2019	Italia	Breast	Revisión bibliográfica	Valorar la importancia y el potencial de la radiómica en el diagnóstico y pronóstico del cáncer de mama

TABLA 4. Material y métodos de los estudios seleccionados, tabla de elaboración propia.

AUTOR	Material y métodos				
	Criterios de inclusión	Criterios de exclusión	Muestra	Método de valoración	Criterios evaluados
Karin Dembrower et al.	Mujeres que asisten a pruebas de cribado	Implantes mamarios, mutaciones genéticas, y antecedentes de CM	Mujeres entre 40 y 74 años (n = 58.344)	Doble ciego. Lectura doble por 2R (estrategia convencional) vs 1°. Lectura doble por 1R +IA 2°. Lectura única por IA. 3°. Lectura triple por 2R + IA	Precisión diagnóstica de cada estrategia.
Xuefei Wang et al.	Pacientes de 18 a 80 años y mamografía o ecografía en los últimos 3 meses	Embarazadas, lactantes, galactorrea, alergia al látex, dermatopatías, y antecedentes de patología mamaria	Pacientes aleatorios de centros de salud (n=2202)	Doble ciego. Se clasifica mediante BI-RADS y puntuación según el número de lesiones malignas.	S, E, VPP, VPN, TFN, TFP y precisión diagnóstica.
Alejandro Rodríguez-Ruiz et al.	Mujeres asintomáticas que asistían al cribado	Mujeres con implantes y/o antecedentes de CM	Mujeres de 39 a 89 años. (n=240)	Doble ciego. Se asignaba una clasificación de BI-RADS (del 1 al 5) y una probabilidad de malignidad (del 1 al 100)	AUROC, sensibilidad, especificidad y tiempo de lectura con y sin ayuda de IA.
Yiqiu Shen et al.	Pacientes que se realizaron una prueba de imagen entre 2012 y 2019.	Pruebas indicadoras de malignidad y puntuación BI-RADS de 1-2	(n=143.203)	Lectura por 10 R expertos en CM comparado con un método híbrido de diagnóstico (R + IA)	Rendimiento diagnóstico con IA, de los R y del modelo híbrido: AUROC, S, E, tasa de biopsia, VPP y VPN
Tomas Schaffter et al.	—	—	Mujeres de EE. UU. (n=85.580) y de Suecia (n=68.008) N total = 153.588	126 equipos de R de 44 países frente a algoritmos de IA	AUC, S y E

Hoy-Eun Kim et al.	—	—	Pacientes de 50,3 años de media	Doble ciego. Lectura por 14 R frente al algoritmo de IA	—
Jose Luis Raya-Povedano	—	—	Mujeres de 58 años de media de Córdoba (DS ± 6) (n=15.987)	—	Se evalúa la carga de trabajo, la S y la tasa de retirada usando la IA y mediante la estrategia convencional
Mattie Salim et al.	Todas las mujeres de 40 a 74 años diagnosticadas de CM entre 2008 y 2015	Antecedentes de CM, implantes mamarios o cáncer diagnosticado doce meses después de la detección	CASOS: Mujeres de 60 años de media con CM (n=739). CONTROL : Mujeres de 54 años de media sanas (n=8066).	—	AUROC, S, E
Scott Mayer McKinney	Mujeres de 50-70 años que se habían sometido a mamografía	—	2 cohortes, una de UK (n=25.856) y otra de EE. UU. (n=3.097)	En UK cada prueba es valorada por 2 R y en EE. UU. solo por 1.	S y E
Nan Wu et al.	Pacientes que se han sometido a mamografía	—	(n=141.473)	—	Valorar el AUROC de los R, de la IA y de la combinación de ambos

AUROC – Área bajo la curva ROC; **CM** – Cáncer de mama; **DS** – Desviación estándar; **E** – Especificidad; **R** – Radiólogos ; **S** – Sensibilidad; **TFN** – Tasa falsos negativos ; **TFP** – Tasa falsos positivos ; **VPN** – Valor predictivo negativo; **VPP** – Valor predictivo positivo.

TABLA 5. Resultados de los estudios, tabla de elaboración propia.

AUTOR	RESULTADOS
Karin	1. 250 CM detectados con lectura doble por 2R
Dembrower et al.	2. Lectura doble IA + R se consideró superior con 261 CM detectados (PR = 1.04) 3. Lectura única por IA se consideró no inferior pero no superior con 246 CM detectados (PR=0.98) 4. Triple lectura por IA + 2 R se consideró superior con 269 CM detectados (PR=1.08)
Xuefei Wang et al.	IA vs Radiólogos = Sensibilidad 84,5% - 89,2%; Especificidad 83,8% - 63%, VPP 92,3% - 84,7%, VPN 70% - 73%, TFN 15,5% - 10,8%, TFP 16,1% - 37%. Mayor precisión y rendimiento diagnóstico de la IA que de los radiólogos. Mamografía basada en IA consigue un AUROC de 0,9.
Alejandro Rodríguez-Ruiz et al.	Los R mejoraron su rendimiento de detección con IA, obteniendo un AUROC de 0.89 frente a 0.87 sin ayuda de IA. La S fue 3 puntos porcentuales y la E 2 puntos porcentuales más alta con el uso de la IA. El tiempo de lectura obtuvo una diferencia no significativa sin IA (146 segundos) y con IA (149 segundos)
Yiqiu Shen et al.	AUROC con la IA de 0,976 frente a 0,924 de los R. Sensibilidad de los R de 90,1%, y con la IA del 94,5%. Especificidad de los R del 80,7% frente a un 88% de la IA. VPP del radiólogo de 27,1% frente a una mejora de la IA del 5,4%
Tomas Schaffter et al.	La IA logró un AUROC de especificidad de 0,858 (EE. UU.) y 0,903 (Suecia) y 66,2 % (EEUU) y 81,2 % (Suecia) en la sensibilidad de los R. La combinación de IA + R dio como resultado un AUROC más alta de 0,942 y logró una E significativamente mejorada (92,0%) con la misma S.
Hoy-Eun Kim et al.	El rendimiento independiente de la IA obtuvo un AUROC de 0,959. El rendimiento de los R sin asistencia de IA fue de 0.810, inferior al rendimiento con asistencia de IA, de 0,881.
Jose Luis Raya-Povedano	1º La detección basada en IA podría reducir la carga de trabajo hasta un 70% sin reducir la sensibilidad un 5% o más. 2º El uso de la IA reduciría un 30%% la carga de trabajo, mejoraría un 25% la sensibilidad y reduciría un 27% la tasa de retirada

Mattie Salim et al.	AUROC: IA1 0.956, IA2 0.922 y IA3 0.920. Sensibilidades: IA1 81,9%, IA2 67,4%, IA3 77,4%, R1 77,4% y R2 80,1%. La combinación de la IA + R logro una sensibilidad máxima del 88,6% con una especificidad del 93%, por IA1. Menor rendimiento en mujeres jóvenes y en densidades mamarias elevadas
Scott Mayer McKinney	En UK, la IA obtuvo una mejora del 1,2% de la sensibilidad y una mejora del 2,7% de especificidad sobre el primer radiólogo; frente al segundo radiólogo obtuvo un resultado de no inferioridad tanto en sensibilidad como en especificidad. En EE. UU., LA IA obtuvo una mejora del 5,7% en la especificidad y del 9,4% de sensibilidad.
Nan Wu et al.	La IA logró un AUROC de 0,876. Los AUROC alcanzados por los lectores individuales variaron de 0,705 a 0,860. Los híbridos de IA + R lograron un AUROC promedio de 0,891. Según estos resultados la IA se puede utilizar como complemento a los R mejorando la detección de CM.

AUROC – Área bajo la curva ROC; **CM** – Cáncer de mama; **E** – Especificidad; **IA** – Inteligencia Artificial; **PR** – Proporción relativa; **R** – Radiólogo; **TFN** – Tasa falsos negativos; **TFP** – Tasa falsos positivos ; **VPN** – Valor predictivo negativo; **VPP** – Valor predictivo positivo