

Rendimiento de una herramienta de inteligencia artificial en biopsias percutáneas de mama

Performance of an artificial intelligence tool in percutaneous breast biopsies

Paulina Neira-Vallejos^{1*}, María C. Vial¹, Carolina Behnke¹, Marcelo Poblete-Becker¹, M. Rosario van Wersch² y Emilia Santelices²

¹Departamento de Radiología, Centro de la Mama de Clínica MEDS La Dehesa; ²Facultad de Medicina, Universidad de Los Andes. Santiago, Chile

Resumen

Introducción: El uso de herramientas de inteligencia artificial (IA) en ecografía mamaria ha mostrado un aumento en la exactitud diagnóstica del cáncer de mama. **Objetivo:** Evaluar el rendimiento diagnóstico de un software de inteligencia artificial (IA) en lesiones mamarias biopsiadas bajo guía ecográfica. **Material y métodos:** Estudio de pruebas diagnósticas prospectivo que incluyó los resultados de pacientes que concurren a realizarse una biopsia percutánea de mama bajo guía ecográfica y cuyas imágenes fueron analizadas y clasificadas por grado de sospecha, en primer lugar con una categoría ACR BI-RADS asignada por un radiólogo especialista y después por un software de IA. Posteriormente se correlacionaron con los resultados de histopatología obtenidos en la biopsia, que fue el estándar de referencia. **Resultados:** Se incluyeron 181 lesiones, de las cuales 80 fueron malignas y 101 benignas. La sensibilidad de la herramienta de IA fue del 95%, la especificidad del 58%, el valor predictivo positivo (VPP) del 64%, el valor predictivo negativo (VPN) del 93% y la exactitud de 0,74. Los radiólogos obtuvieron una sensibilidad del 98%, una especificidad del 19%, un VPP del 49%, un VPN del 95% y una exactitud de 0,54. **Conclusiones:** La herramienta de IA demostró una especificidad y un VPP superiores en comparación con los radiólogos en lesiones a las que se hizo biopsia, por lo que podría ser utilizada para disminuir el número de biopsias con resultado benigno.

Palabras clave: Ultrasonido. Mama. Biopsia core. Inteligencia artificial.

Abstract

Introduction: The use of artificial intelligence (AI) tools in breast ultrasound has shown an increase in the diagnostic accuracy of breast cancer. **Objective:** To evaluate the diagnostic performance of artificial intelligence (AI) software in the detection of breast cancer in breast lesions biopsied under ultrasound guidance. **Material and methods:** This prospective diagnostic testing study included the results of patients who underwent ultrasound-guided percutaneous breast biopsy. Images were analyzed and classified by degree of suspicion, first with an ACR BI-RADS category assigned by a specialist radiologist and then by AI software. These were subsequently correlated with the histopathology results obtained from the biopsy, which served as the reference standard. **Results:** 181 lesions were included in this study, of which 80 were malignant and 101 benign lesions. The sensitivity of the AI tool was 95%, specificity 58%, positive predictive value (PPV) 64%, negative predictive value (NPV) 93%, and accuracy 0.74. The values obtained by the radiologists were 98% sensitivity, 19% specificity, 49% PPV, 95% NPV, and 0.54 accuracy. **Conclusions:** The AI software demonstrated superior specificity and PPV compared to radiologists in lesions biopsied and could therefore be used to reduce the number of biopsies with benign results.

Keywords: Ultrasound. Breast. Core biopsy. Artificial intelligence.

*Correspondencia:

Paulina Neira-Vallejos

E-mail: paulina.neira@meds.cl

2810-6954 / © 2025 Sociedad Chilena de Radiología. Publicado por Permayer. Este es un artículo open access bajo la licencia CC BY-NC-ND (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Fecha de recepción: 02-05-2025

Fecha de aceptación: 01-09-2025

DOI: 10.24875/AJI.25000031

Disponible en internet: 28-04-2026

Austral J. Imaging. 2026;32(2):68-73

www.resochradi.com

Introducción

El cáncer de mama constituye la principal causa de mortalidad por cáncer en las mujeres en todo el mundo, lo que le confiere una gran relevancia¹.

El cribado con mamografía, con un rendimiento de 4,1 cánceres por 1000, ha permitido una pesquisa oportuna, con impacto en la disminución de la mortalidad por esta causa^{2,3}. Si bien la mamografía sigue siendo el estándar para la detección de esta patología, su rendimiento se ve limitado en cerca del 40% de la población femenina debido a que tienen tejido mamario denso, lo que afecta su sensibilidad⁴. La evaluación con ecografía, como método de cribado suplementario, ha permitido disminuir el enmascaramiento de lesiones en mamas densas, con tasas de detección de cáncer que oscilan entre 2,1 y 2,7 por 1000⁵. Sin embargo, el valor predictivo positivo (VPP) de las biopsias obtenidas en lesiones visibles únicamente con ecografía complementaria es bajo, entre el 2% y el 21%, lo que lleva a buscar métodos que mejoren los resultados en el proceso diagnóstico⁶. El uso de herramientas de inteligencia artificial (IA) en el cribado con mamografía puede mejorar el diagnóstico inicial de cáncer de mama⁷. También se ha investigado el uso de estas herramientas en ecografía mamaria, mostrando un aumento en la exactitud diagnóstica del cáncer de mama y disminuyendo la variabilidad de interpretación⁸. En noviembre de 2021, la Food and Drug Administration aprobó el uso del dispositivo KOIOS DS, que es un programa de IA para la evaluación de lesiones visibles en ecotomografía mamaria. Esta herramienta de IA se denomina «de soporte de decisiones» y utiliza métodos de aprendizaje profundo, con algoritmos avanzados de IA. En la práctica, el radiólogo elige dos imágenes ortogonales de la lesión y coloca una región de interés en el centro de la lesión, con lo que el *software* automáticamente dibuja un rectángulo alrededor de la lesión, la caracteriza en cuanto a forma y orientación, y entrega un nivel de sospecha.

El objetivo primario de este estudio es evaluar el rendimiento diagnóstico de una herramienta de IA, y los objetivos secundarios son compararlo con el rendimiento diagnóstico de los radiólogos y evaluar si su uso mejora la discriminación para la toma de biopsia en lesiones mamarias detectadas por ultrasonido.

Material y métodos

Estudio de pruebas diagnósticas prospectivo para evaluar el rendimiento diagnóstico de una herramienta

de IA. Se incluyeron todas las pacientes consecutivas a las que se hizo una biopsia guiada por ultrasonido por indicación de su médico tratante entre abril de 2023 y junio de 2024, en el centro de mama de una institución de salud privada.

En la mayoría de los casos, el médico tratante indicó la biopsia porque la lesión había sido clasificada como sospechosa (BIRADS [*Breast Imaging-Reporting and Data System*] 4 o 5) en la ecografía mamaria, pero también en algunos casos se solicitó para lesiones clasificadas como probablemente benignas, ya fuera porque estas eran palpables, por un riesgo aumentado para desarrollar cáncer de mama o por solicitud de la paciente.

Se excluyeron las lesiones con resultado de patología inflamatoria aguda y aquellas con resultado de cáncer conocido.

El radiólogo que efectuó el procedimiento asignó una categoría ACR (American College of Radiology) BIRADS clínica a la lesión en el momento de realizar la biopsia percutánea, y posteriormente, una vez ya registrada la categoría de sospecha clínica, se aplicó la herramienta de IA mediante el programa KOIOS DS; todo esto previo a conocer el resultado de la biopsia.

El *software* del programa de IA que usamos para este estudio fue la versión 3.6 KOIOS DS BREAST.

La clasificación del nivel de sospecha de KOIOS DS se divide en benigno, probablemente benigno, sospechoso y maligno, y se correlacionan con los niveles de sospecha de la clasificación ACR-BIRADS 5.^a edición⁹ en su sección de ultrasonido del siguiente modo:

- Categoría 2 ACR-BIRADS: 2 benigno del *software*.
- Categoría 3 ACR-BIRADS: 3 probablemente benigno del *software*.
- Categoría 4 ACR-BIRADS: 4A, 4B y 4C sospechoso del *software*.
- Categoría 5 ACR-BIRADS: 4C+ maligno del *software*.

Las categorías 1 y 6 ACR-BIRADS corresponden a sin hallazgos y a neoplasia maligna diagnosticada, respectivamente, por lo que el *software* KOIOS no las considera en su clasificación.

Todas las biopsias se hicieron con una aguja de 14 G, excepto dos que se hicieron con aguja de 10 G.

Participaron en la clasificación del nivel de sospecha de las lesiones cinco radiólogos con 4 a 27 años de experiencia en imágenes de mama (16 años promedio). Utilizaron ecografía manual en un equipo Logic E10s GE con transductor lineal de 4-20 MH.

Se registraron la edad de la paciente y el tamaño de la lesión biopsiada, y una vez informada la biopsia por

el médico patólogo se registró el resultado histológico, que fue considerado el estándar de referencia.

Se resguardó la confidencialidad mediante medidas de anonimato y el estudio fue aprobado por el comité de ética.

Análisis estadístico

Se evaluó la relación existente entre el resultado de la biopsia (estándar de referencia) y la clasificación de la lesión por el *software* y por el método tradicional ACR-BIRADS por el radiólogo, ambas hechas previamente a la biopsia. Se calcularon la sensibilidad, la especificidad, el VPP, el valor predictivo negativo (VPN) y la exactitud de la herramienta de IA y de los radiólogos, para evaluar el rendimiento diagnóstico mediante un método binario (benigno o sospechoso): para el *software*, «benigno» son las clasificaciones benigno y probablemente benigno, y «sospechoso» las clasificaciones sospechoso y maligno; para los radiólogos, «benigno» son las clasificaciones ACR-BIRADS 2 y 3, y «sospechoso» las clasificaciones ACR-BIRADS 4 y 5.

Se hizo un subanálisis para el grupo de lesiones no palpables. También se calculó qué porcentaje de las biopsias se habría evitado en caso de haber seguido la recomendación del *software* de IA.

Resultados

Se incluyeron 191 casos consecutivos, de los cuales nueve correspondieron a patología inflamatoria aguda y subaguda, incluyendo siete mastitis granulomatosas y dos procesos inflamatorios agudos, que se excluyeron porque las especificaciones del *software* indican que no sirve para evaluar patología inflamatoria aguda. También se excluyó una paciente que ya tenía diagnóstico de cáncer de mama mediante punción percutánea y se le solicitó una repetición para evaluación de marcadores. Así, finalmente fueron 181 los casos incluidos y 177 las pacientes, ya que cuatro tuvieron biopsias de más de una lesión.

El promedio de edad de las pacientes fue de 49 años (rango: 18-89 años). El tamaño de las lesiones biopsiadas fluctuó entre 4 y 60 mm, con un promedio de 18 mm. De las 181 lesiones, 98 fueron palpables. Hubo 80 lesiones malignas, 19 lesiones de alto riesgo y 82 lesiones benignas. Los resultados histológicos de las biopsias se detallan en la [tabla 1](#).

Las categorías de sospecha de las lesiones biopsiadas asignadas por los radiólogos y por el *software* se detallan en las [tablas 2 y 3](#).

Tabla 1. Resultados histológicos de las lesiones mamarias estudiadas mediante biopsia percutánea bajo guía de ultrasonido

Tipo de lesión	n = 181
Lesiones malignas	80
Carcinoma ductal infiltrante	60
Carcinoma lobulillar infiltrante	9
Carcinoma ductal <i>in situ</i>	3
Carcinoma papilar	5
Carcinoma mucinoso	2
Carcinoma tubular	1
Lesiones de alto riesgo	19
Papiloma intraductal	6
Cicatriz radiada	2
Tumor filodes	6
Lesión papilar atípica	1
Lesión fibroelástica atípica	2
Atipia plana	1
Hiperplasia ductal atípica	0
Hiperplasia lobulillar atípica	1
Lesiones benignas	82
Fibroadenoma	44
Fibrosis estromal	14
Necrosis grasa	4
Proceso inflamatorio crónico	5
Hiperplasia estromal pseudoangiomatosa	2
Hamartoma	1
Mastitis linfocitaria	2
Otras	10

El *software* clasificó 42/181 como BIRADS 2, y de estas ninguna fue maligna. Correspondieron a lesiones de alto riesgo 7/42 (dos lesiones papilares sin atipia y cinco tumores filodes benignos); los cinco tumores filodes fueron palpables y clasificados como BIRADS 4A por los radiólogos.

Las lesiones que el *software* clasificó como BIRADS 3 fueron 21/181, de las cuales 4/21 resultaron malignas (un cáncer papilar *in situ* y tres carcinomas ductales infiltrantes, dos de estos con grado de diferenciación 3 y uno grado 2), con tamaños de 36, 19, 30 y 32 mm cada uno. Estas cuatro lesiones fueron clasificadas como sospechosas (BIRADS 4 o 5) por los radiólogos.

Dentro de las 21 lesiones clasificadas como BIRADS 3 por el *software* hubo tres lesiones de alto riesgo (un papiloma, una atipia epitelial plana y un tumor filodes), que fueron clasificadas como sospechosas por los radiólogos ([Fig. 1](#)).

Hubo ocho lesiones clasificadas como BI-RADS 3 por los radiólogos y como BIRADS 4 por el *software* de IA, de las cuales siete resultaron benignas, sin lesiones de alto riesgo y una lesión (1/8) que resultó maligna (carcinoma ductal infiltrante con grado de diferenciación 3), con un tamaño de 20 mm.

Tabla 2. Categorías de sospecha de las lesiones biopsiadas asignadas por los radiólogos y por el *software* de inteligencia artificial

BIRADS radiólogos	Benignas	Malignas	N.º por categoría	BIRADS software*	Benignas	Malignas	N.º por categoría
2	0	0	0	2	42	0	42
3	20	1	21	3	17	4	21
4	79	47	126	4	38	50	88
5	2	32	34	4c+	4	26	30
Total	-	-	181	-	-	-	181

* *Software* de inteligencia artificial KOIOS DS.

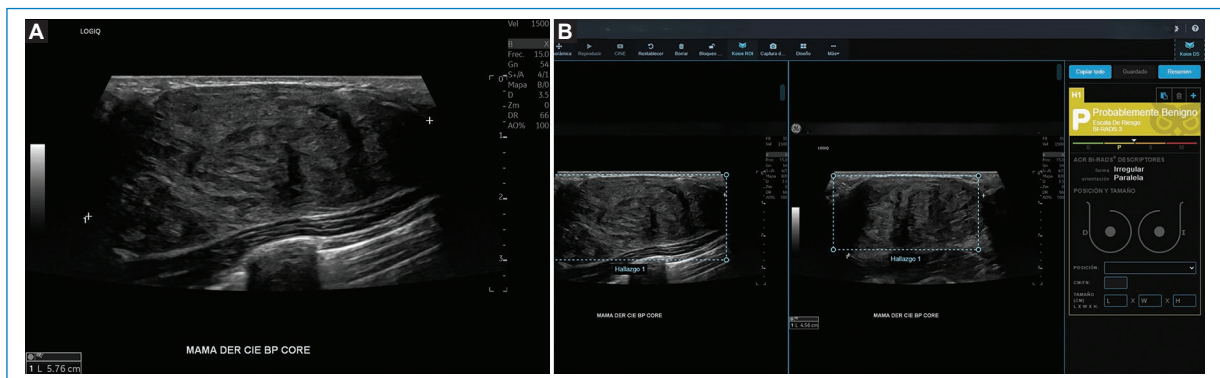
BIRADS 2: de aspecto benigno; BIRADS 3: probablemente benigno; BIRADS 4: sospechoso de neoplasia maligna; BIRADS 5 o 4c+: altamente sospechoso de neoplasia.

Tabla 3. Clasificación binaria (benigna o sospechosa) de las lesiones biopsiadas asignada por los radiólogos y por el *software* de inteligencia artificial

BIRADS radiólogos	Benignas	Malignas	N.º por categoría	BIRADS software*	Benignas	Malignas	N.º por categoría
2 y 3	20	1	21	2 y 3	59	4	63
4 y 5	81	79	160	4 y 4c+	42	76	118

* *Software* de inteligencia artificial KOIOS DS.

BIRADS 2 y 3: benigno; BIRADS 4 y 5 o 4c+: sospechoso.

**Figura 1. A:** ecografía hecha por un radiólogo que clasificó el nódulo como BIRADS 4. **B:** clasificación del *software* de inteligencia artificial del mismo nódulo como probablemente benigno. El resultado histológico de la biopsia core fue: «hallazgos histológicos concordantes con lesión fibroepitelial con zonas de estroma hiper celular». Posteriormente se operó y la biopsia quirúrgica resultó «tumor filodes benigno».

La sensibilidad de la herramienta de IA fue del 95%, la especificidad del 58%, el VPP del 64%, el VPN del 93% y la exactitud de 0,74. La sensibilidad de los radiólogos fue del 98% (con valores por radiólogo que variaron entre el 96% y el 100%), la especificidad del 19% (entre el 13% y el 25%), el VPP del 49% (entre el 35% y el 56%), el VPN del 95% (entre el 75% y el 100%) y la exactitud de 0,54 (entre 0,43 y 0,58).

En el subanálisis incluyendo solo las lesiones no palpables, la IA tuvo una sensibilidad del 100%, una especificidad del 47%, un VPP del 46%, un VPN del 100% y una exactitud de 0,63, mientras que para los radiólogos la sensibilidad fue del 100%, la especificidad del 24%, el VPP del 41%, el VPN del 100% y la exactitud de 0,50.

Si no se hubiese hecho biopsia a las lesiones categorizadas como benignas y probablemente benignas

por el *software* de IA, pero clasificadas como sospechosas por los radiólogos (50/181), se habría evitado el 27,6% de las biopsias con resultado benigno y se habría perdido un 2,2% de los cánceres. Si solo no se hubiese hecho biopsia a las lesiones clasificadas como benignas (BIRADS 2) por el *software* y como sospechosas por el radiólogo (32/181), se habría evitado el 17,7% de las biopsias y no se hubiese perdido ninguna lesión maligna.

Discusión

Nuestros resultados muestran un buen desempeño de la herramienta de IA para diferenciar entre benigna y sospechosa en las lesiones mamarias visibles por ultrasonido, con una sensibilidad ligeramente inferior a la de los radiólogos, pero con una exactitud mayor que la de estos.

Hay varias publicaciones que han mostrado en general buenos resultados con el uso de estas herramientas de IA. El primer estudio, de Mango et al.⁸, de carácter retrospectivo, incluyó 900 lesiones mamarias detectadas con 15 evaluadores, y evaluó el uso del *software* KOIOS DS como herramienta de apoyo a la decisión en el diagnóstico de imágenes de ecografía mamaria convencional de nódulos mamarios, con una sensibilidad del 98% para el *software* y del 94% para los clínicos, y una especificidad mayor para el *software*, del 50% frente al 40%. En otro estudio publicado por Berg et al.¹⁰ en 2021, también de carácter retrospectivo, que incluyó 319 pacientes y nueve evaluadores, los resultados mostraron una curva ROC (*Receiver Operating Characteristic*) del *software* de 0,77, sin diferencia significativa con la de los radiólogos, que fue de 0,82.

Hay estudios que muestran que las herramientas de IA tendrían más utilidad frente a lectores poco experimentados^{11,12}. En el estudio de Park et al.¹¹, los radiólogos menos experimentados mejoraron significativamente su VPN con la asistencia de IA, mientras que los radiólogos más experimentados mejoraron su VPP. En el trabajo de Heller et al.¹² no hubo diferencia significativa en cuanto a exactitud, VPN, VPP, sensibilidad y especificidad entre radiólogos y el *software* de IA, pero sí mejoró la exactitud, particularmente en las lesiones de bajo nivel de certeza, mejorando el VPP (24,7% vs. 19,3%) y la especificidad (57,8% vs. 44,6%).

Un estudio con una población similar a la nuestra es el de Browne et al.¹³, que incluyó un total de 403 lesiones o nódulos mamarios a los que se realizó biopsia percutánea. Se evaluó el rendimiento diagnóstico de una herramienta de IA y de los radiólogos, utilizando como punto de referencia el resultado histológico de la biopsia,

y se evidenció que se habría reducido de manera significativa el número de biopsias benignas realizadas en lesiones clasificadas por el *software* como BIRADS 3, sin dejar de diagnosticar ninguna lesión maligna.

En nuestro estudio hay 50 casos que el *software* categorizó como BIRADS 2 o 3 y los radiólogos clasificaron como BIRADS 4 o 5. De estos 50 casos, 46 tuvieron resultado benigno y los otros cuatro resultaron ser cáncer y fueron clasificados como sospechosos por los radiólogos. Si no se hubiese hecho biopsia a las lesiones categorizadas como benignas y probablemente benignas por el *software* de IA, pero clasificadas como sospechosas por los radiólogos, se habría evitado un 27,6% de biopsias con resultado benigno, pero con pérdida de un 2,2% de diagnósticos de cáncer. En el caso de que solo no se hubiesen hecho las biopsias clasificadas como BIRADS 2 por el *software*, y que fueron clasificadas como sospechosas por el radiólogo, se habría evitado el 17,7% de biopsias benignas y no se habría perdido ninguna lesión maligna. Este último resultado va en la misma línea de lo publicado por Guldogan et al.¹⁴, quienes describe que con la aplicación de una herramienta de IA hubiesen evitado un 11% de biopsias benignas de lesiones clasificadas como BIRADS 2.

En nuestros resultados es notoria la mayor especificidad y el mejor VPP del *software* en comparación con los radiólogos, lo cual puede ser atribuido en parte a que estos últimos clasificaron un número más alto de lesiones de mayor tamaño, principalmente palpables, como sospechosas y el *software* las clasificó como benignas. Como mostró el subanálisis que hicimos de lesiones no palpables, la especificidad de los radiólogos mejora del 19% al 24% al no considerar las lesiones palpables, pero considerando esto ninguno de los tumores filodes habría sido diagnosticado si se hubiese seguido la clasificación de la herramienta de IA, ya que fueron categorizados como BIRADS 2 y 3 por el *software*. Todos ellos eran palpables y los radiólogos los categorizaron como sospechosos. Esta conducta se explica porque el ACR define la categoría BIRADS 3 solo para lesiones no palpables. Si bien el hecho de que una lesión sea palpable es un criterio clínico más que radiológico, en general los radiólogos eligen biopsiar las lesiones palpables aunque tengan una apariencia benigna en las imágenes¹⁵. Hay autores que cuestionan este criterio, ya que puede llevar a un importante número de resultados falsos positivos. En la serie de Raza et al.¹⁶, el VPN de la ecografía en la evaluación de nódulos palpables probablemente benignos fue alto, del 99,4%, y por tanto concluyen que el

seguimiento es una alternativa aceptable a la biopsia, similar al caso de los nódulos no palpables.

Entre las limitaciones de nuestro estudio se encuentra que la herramienta de IA se utilizó en una muestra seleccionada de pacientes con indicación de biopsia, por lo que los resultados obtenidos del rendimiento de esta solo son aplicables a poblaciones similares y no al conjunto de las mujeres que asisten a ecografía mamaria por diversas indicaciones. Otra limitación es el escaso número de casos, que no permitió hacer otros subanálisis del rendimiento de la herramienta, por ejemplo con respecto al tipo molecular de los cánceres.

Conclusiones

La aplicación de una herramienta de IA demostró un buen rendimiento diagnóstico en lesiones que fueron biopsiadas, con mayores especificidad y VPP que los radiólogos, por lo que podría tener un impacto positivo en la toma de decisión clínica de realización de biopsia, complementando la evaluación radiológica tradicional que se ha utilizado hasta ahora, y con el potencial de disminuir las biopsias falsas positivas.

Financiamiento

Los autores declaran no haber recibido financiamiento para este estudio.

Conflicto de intereses

Los autores declaran no tener conflicto de intereses.

Consideraciones éticas

Protección de personas y animales. Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

Confidencialidad, consentimiento informado y aprobación ética. Los autores han seguido los protocolos de su centro sanitario/institución para acceder a los datos de las historias clínicas. Se ha obtenido el consentimiento informado de las pacientes y se cuenta

con la aprobación del Comité de Ética. Se han seguido las recomendaciones de las guías SAGER.

Declaración sobre el uso de inteligencia artificial.

Los autores declaran que no se utilizó ningún tipo de inteligencia artificial generativa para la redacción ni la creación de contenido de este manuscrito.

Referencias

1. 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 Cancer J Clin.* 2021;71:209-49. doi: 10.3322/caac.21660.
2. Nelson HD, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L. Effectiveness of breast cancer screening: systematic review and meta-analysis to update the 2009 US preventive services Task Force recommendation. *Ann Intern Med.* 2016;164:244-55.
3. Sprague BL, Miglioretti DL, Lee CI, Perry H, Tosteson AAN, Kerlikowske K. New mammography screening performance metrics based on the entire screening episode. *Cancer.* 2020;126:3289-96. doi: 10.1002/cncr.32939.
4. Mann RM, Athanasiou A, Baltzer PAT, Camps-Herrero J, Clauser P, Fallenberg EM, et al. European Society of Breast Imaging (EUSOBI). Breast cancer screening in women with extremely dense breasts recommendations of the European Society of Breast Imaging (EUSOBI). *Eur Radiol.* 2022;32:4036-45. doi: 10.1007/s00330-022-08617-6.
5. Berg WA, Vourtsis A. Screening breast ultrasound using handheld or automated technique in women with dense breasts. *J Breast Imaging.* 2019;1:283-96.
6. Hussein H, Abbas E, Keshavarzi S, Fazelzad R, Bukhanov K, Kulkarni S, et al. Supplemental breast cancer screening in women with dense breasts and negative mammography: a systematic review and meta-analysis. *Radiology.* 2023;306:e221785.
7. Hernström V, Josefsson V, Sartor H, Schmidt D, Larsson AM, Hofvind S, et al. Screening performance and characteristics of breast cancer detected in the Mammography Screening with Artificial Intelligence trial (MASAI): a randomised, controlled, parallel-group, non-inferiority, single-blinded, screening accuracy study. *Lancet Digit Health.* 2025;7:e175-83. doi: 10.1016/S2589-7500(24)00267-X.
8. Mango VL, Sun M, Wynn RT, Ha R. Should we ignore, follow, or biopsy? Impact of artificial intelligence decision support on breast ultrasound lesion assessment. *Am J Roentgenol.* 2020;214:1445-52.
9. Mendelson EB, Böhm-Vélez M, Berg WA, Whitman GJ, Feldman MI, Madjar H. ACR-BI-RADS ultrasound. En: *ACR BI-RADS Atlas: Breast Imaging Reporting and Data System.* Reston (VA): American College of Radiology; 2013.
10. Berg WA, Gur D, Bandos AI, Nair B, Gizienski TA, Tyma CS, et al. Impact of original and artificially improved artificial intelligence-based computer-aided diagnosis on breast US interpretation. *J Breast Imaging.* 2021;3:301-11.
11. Park YM, Kim EK, Lee JH, Ryu JH, Han SS, Choi SJ, et al. Palpable breast masses with probably benign morphology at sonography: can biopsy be deferred? *Acta Radiol.* 2008;49:1104-11.
12. Heller SL, Wegener M, Babb JS, Gao Y. Can an artificial intelligence decision aid decrease false-positive breast biopsies? *Ultrasound Q.* 2021;37:10-5.
13. Browne JL, Pascual MA, Pérez J, Salazar S, Valero B, Rodríguez I, et al. AI: can it make a difference to the predictive value of ultrasound breast biopsy? *Diagnostics (Basel).* 2023;13:811.
14. Guldogan N, Taskin F, Icten GE, Yilmaz E, Turk EB, Erdemli S, et al. Artificial intelligence in BI-RADS categorization of breast lesions on ultrasound: can we omit excessive follow-ups and biopsies? *Acad Radiol.* 2024;31:2194-202. doi: 10.1016/j.acra.2023.11.031.
15. Mahboubi-Fooladi Z, Sabahi M, Amin-Astani S, Khazaei M, Ghomi Z. Attitudes of practicing radiologists toward the management of palpable circumscribed breast masses. *J Breast Imaging.* 2023;5:297-305.
16. Raza S, Chikarmane SA, Neilsen SS, Zorn LM, Birdwell RL. BI-RADS 3, 4, and 5 lesions: value of US in management — follow-up and outcome. *Radiology.* 2008;248:773-81. doi: 10.1148/radiol.2483071786.