SAÚDE DA MULHER

EPIDEMIOLOGIA, INTERVENÇÕES, CASOS CLÍNICOS E POLÍTICAS DE SAÚDE

Capítulo 14

INTELIGÊNCIA ARTIFICIAL NO RASTREAMENTO DO CÂNCER DE MAMA

CIBELE DE SOUZA E SILVA ARCANJO[1]
GABRIELY DE CAMARGO AGUIRRE[1]
GIOVANNA GAMBA CERQUEIRA[1]
GIOVANNA MARQUES SILVA[1]
GIOVANNI IGLESIA CASTILHO[1]
GUSTAVO BERETTA DE OLIVEIRA MATTOS[1]
IZABELLE DOS SANTOS AMARAL[1]
JULIA DO AMARAL PAES DE MELO[1]
MARIA CLARA HARUMI ALVES KAZIWARA[1]
MARIA EDUARDA CORRÊA MIYAKE POMPEO[1]
VITORIA MASOLA GALDEANO[1]

[1]*Discente - Medicina da Universidade Nove de Julho.*

Palavras-Chave: *Rastreamento; Câncer de Mama; Inteligência Artificial.*

10.59290/978-65-6029-246-8.14

EDITORA PASTEUR

Edição XXIV

INTRODUÇÃO

A digitalização da saúde tem sido intensamente impulsionada pelo uso estratégico de tecnologias baseadas em dados, que vêm transformando a forma como se oferece cuidado personalizado e se realiza assim a pesquisa médica. Parte fundamental dessa transformação está relacionada ao avanço das técnicas de inteligência artificial (IA), que têm assumido um papel central nesse cenário (COLLINS, 2019).

A IA refere-se a um campo abrangente da ciência que busca desenvolver sistemas computacionais capazes de interpretar dados, aprender com eles e gerar respostas inteligentes. Entre suas principais vertentes estão o machine learning (ML), as redes neurais artificiais (ANNs) e o deep learning (DL), amplamente empregados na construção de modelos preditivos (PMs) voltados e para a identificação de doenças e a projeção de desfechos clínicos (COLLINS, 2019; SFAKIANOUDIS *et al*, 2022). Tais modelos utilizam informações específicas, conhecidas como preditores, para estimar probabilidades relacionadas à saúde de um indivíduo — seja no diagnóstico de condições já existentes ou na previsão de eventos futuros. Essas ferramentas têm grande valor na prática clínica, pois auxiliam na tomada de decisões terapêuticas, na solicitação de exames e na definição de condutas preventivas (MOONS *et al*, 2019; CHEN *et al*, 2022).

Entre os principais diferenciais da IA no contexto médico está sua capacidade de processar grandes volumes de dados — os chamados big data — de maneira rápida e eficaz. Isso possibilita a identificação de padrões complexos que muitas vezes escapam à análise humana. Contudo, os algoritmos de IA e seus modelos preditivos ainda podem estar sujeitos a vieses de seleção, dependendo de como são desenvolvidos e aplicados, o que exige cautela e supervisão constante (BORMANN *et al*, 2021; KRAGH, 2021)

Na avaliação do desempenho desses modelos, alguns conceitos estatísticos são fundamentais. A sensibilidade representa a capacidade do modelo de identificar corretamente indivíduos com determinada condição (verdadeiros positivos), enquanto a especificidade mede sua capacidade de excluir corretamente aqueles sem a condição (verdadeiros negativos). Outro parâmetro essencial é a área sob a curva ROC (AUC –*Area Under the Curve*), que expressa a capacidade discriminativa geral do modelo, sendo que valores mais próximos de 1 indicam melhor desempenho (DA CUNHA *et al*, 2019; DEMBROWER *et al*, 2023; WANG *et al*. 2023).

Atualmente, os avanços mais expressivos da IA têm ocorrido na área de diagnóstico, com destaque para a análise de imagens médicas, permitindo interpretações mais robustas e recomendações mais alinhadas à prática clínica (BORMANN *et al.*, 2021; MONTJEAN *et al.*, 2024; KIM *et al.*, 2024). Nesse contexto, portanto, a IA surge como uma ferramenta de grande potencial para o diagnóstico de precisão do câncer de mama. A doença é uma questão de saúde pública mundial e sua incidência vem aumentando em todo o mundo. No Brasil ocorrem em torno de 58 mil casos anualmente. No entanto, existem programas de triagem que permitem o diagnóstico e o tratamento precoces, que se divide em triagem pré-clínica e clínica. A triagem pré-clínica envolve o autoexame da mama, onde o diagnóstico é feito pela própria paciente, quando apresenta tumores com tamanhos possíveis de serem palpados, já a triagem clínica se baseia em mamografia e ultrassom combinados, além da ressonância magnética, tomossíntese digital da mama e mamografia com contraste (WANG *et al*, 2023).

Os fatores de risco para o câncer de mama podem ser classificados em não modificáveis e

modificáveis. Entre os fatores não modificáveis estão a idade, menarca precoce (antes dos 12 anos), menopausa tardia (após os 55 anos), primeira gravidez ocorrida após os 30 anos, nuliparidade, e o uso de contraceptivos hormonais combinados e realização de terapia hormonal na pós-menopausa. Já os fatores modificáveis, incluem sobrepeso ou obesidade, tabagismo, consumo de bebidas alcoólicas, sedentarismo e exposição à radiação ionizante. Apenas 5 a 10% dos casos de câncer de mama são genéticos. Os genes mais comumente mutados na doença são o BRCA 1 e o BRCA 2. Outros genes que podem estar envolvidos são: PALB2, CHEK2, BARD1, ATM, TP53 (FEBRASGO, 2021)

O tamanho do tumor e as condições dos linfonodos regionais são fatores prognósticos de maior importância. Quanto maior o tumor, maior a chance de metástase para linfonodos e menor a sobrevida.

As principais lesões suspeitas encontradas nos exames de imagem são: lesões nodulares com forma irregular, margens obscurecidas, indistintas ou espiculadas; presença de distorção do parênquima mamário; ecotextura heterogênea e presença de sombra acústica posterior; calcificações grosseiras heterogêneas, amorfas, pleomórficas finas e lineares finas (FEBRASGO, 2021)

Os métodos de diagnóstico variam a partir dessas lesões, sendo eles: (FEBRASGO, 2021)

Nódulos e distorções: biópsia core (tru-cut)

Microcalcificações e nódulos menores de 1 cm: biópsia a vácuo

Lesões sólido-císticas: biópsia a vácuo

Linfonodos: PAAF (agulha fina)

O BI-RADS é um sistema de classificação padronizado para laudos de exames de imagem da mama e é de extrema importância para definição de conduta, uniformidade na comunicação e estimativa de chance de neoplasia (INCA, 2022) (**Tabela 14.1**).

Tabela 14.1 Tabela BI-RADS

Categoria	Impressão diagnóstica	Recomendação	Risco de câncer (%)
0	Exame inconclusivo	Completar o estudo	Exame incompleto
1	Normal	Exame de rotina anual	0
2	Achado benigno	Exame de rotina anual	0
3	Achado provavelmente benigno	Realizar controle precoce (6, 12, 24, 36 meses)	<2
4	Achado suspeito	Prosseguir investigação: realizar biópsia	3-94
5	Achado altamente suspeito	Prosseguir investigação: realizar biópsia	>95
6	Achado investigado previamente e com resultado positivo (câncer)	Tratamento adequado	100

Fonte: Adaptado de INCA, 2022.

O tratamento depende da imuno-histoquímica, sendo o subtipo molecular determinante para a definição da conduta, a estimativa da agressividade e tumoral e o estabelecimento do prognóstico. Os subtipos moleculares são: luminal A, luminal B, HER2 (associado ao pior

prognóstico) e triplo negativo. A escolha das modalidades terapêuticas dependem da idade da paciente, estadiamento do tumor, tipo histológico e subtipo molecular, comorbidades e a presença de mutações genéticas. Os tratamentos envolvem cirurgia, quimioterapia, radioterapia, terapia alvo e hormonioterapia (GOLDHIRSCH *et al*, 2013)

Assim, dada a importância de um rastreio mais eficaz, a IA tem emergido como uma ferramenta promissora para aprimorar a acurácia e a eficiência do rastreamento mamográfico, particularmente na detecção precoce do câncer de mama. Estudos retrospectivos de larga escala demonstraram que algoritmos de IA, especialmente aqueles baseados em redes neurais convolucionais (CNNs), são capazes de alcançar desempenho diagnóstico comparável ao de radiologistas experientes na leitura independente de mamografias de rastreamento, substituindo um dos radiologistas ou auxiliando na priorização de casos com maior probabilidade de malignidade. No entanto, a transição do uso investigacional para a adoção clínica rotineira depende de validação prospectiva rigorosa e contextualizada (DEMBROWER, 2023).

A IA já tem se consolidado como uma ferramenta valiosa no diagnóstico por imagem, demonstrando desempenho comparável ou até superior ao de especialistas humanos, especialmente em exames como a termografia infravermelha (WANG, 2023). No contexto da triagem do câncer de mama, a IA tem sido aplicada principalmente nas mamografias, o exame de rastreamento mais utilizado. Essa tecnologia tem mostrado grande potencial ao identificar tumores que poderiam ser ignorados por avaliadores humanos, o que contribui para o aumento da sensibilidade do exame e reduz a incidência de falsos positivos, evitando procedimentos invasivos desnecessários. Isso facilita o diagnóstico precoce, aumentando as chances de sucesso no tratamento (ZENG, 2024).

Contudo, a viabilidade da leitura autônoma por IA exige a integração de dados clínicos complementares, como antecedentes oncológicos, sintomas mamários atuais, achados de exames anteriores e características individuais de risco. Ignorar essas variáveis pode comprometer a acurácia diagnóstica e limitar a aplicabilidade clínica dos sistemas automatizados (DEMBROWER, 2023). Além disso, a IA não está isenta de falhas, com erros podendo ocorrer devido à transferência de algoritmos entre populações distintas ou atualizações tecnológicas que alteram as imagens, levando a interpretações equivocadas, como falsos negativos ou localização incorreta de lesões (ZENG, 2024).

Apesar dos avanços técnicos, a adoção clínica da IA como ferramenta autônoma ainda exige avaliação criteriosa. A eficácia de qualquer solução baseada em IA está intrinsecamente ligada ao seu desempenho independente, isto é, sua capacidade de interpretar exames sem supervisão direta e de forma confiável em cenários clínicos reais. Para que a IA possa efetivamente melhorar os desfechos da triagem mamográfica — seja aumentando a taxa de detecção de câncer, seja reduzindo falsos positivos ou otimizando o uso de recursos humanos — é necessário um desempenho robusto, e é validado prospectivamente e alinhado às necessidades operacionais dos programas de rastreamento populacional (YOON, 2023).

A IA também surge como apoio valioso ao trabalho dos radiologistas, que enfrentam uma constante pressão relacionada à carga de exames e às exigências institucionais. Com a integração da tecnologia, é possível realizar leituras únicas, duplas ou triplas dos exames, dependendo da infraestrutura local, aumentando a segurança no diagnóstico (HICKMAN, 2022). Apesar de os algoritmos tradicionais ainda não te-

rem se mostrado superiores ao desempenho humano em termos de custo-benefício e eficácia, os novos modelos baseados em aprendizado de máquina (ML) oferecem vantagens promissoras. Esses sistemas são capazes de delinear automaticamente áreas suspeitas de câncer, promovendo maior precisão. No entanto, o uso do ML em fluxos reais de trabalho ainda representa um desafio prático e é um dos principais obstáculos para sua adoção definitiva (HICKMAN, 2022).

Dessa forma, esforços futuros devem priorizar não apenas a avaliação técnica dos algoritmos, mas também a definição de estratégias seguras e escaláveis de implementação. Isso inclui a incorporação de protocolos de controle de qualidade contínuo, validação multicêntrica, vigilância pós-mercado e avaliação do impacto em indicadores clínicos e operacionais. O sucesso da IA como ferramenta autônoma na mamografia de rastreamento dependerá, portanto, da combinação entre rigor metodológico, governança clínica e monitoramento em tempo real de seu desempenho em larga escala (YOON, 2023).

MÉTODO

Uma revisão bibliográfica foi conduzida com o objetivo de identificar textos relacionados ao uso da IA no contexto do rastreamento do câncer de mama, com foco na aplicação e eficácia dessa tecnologia na melhora dos desfechos clínicos de pacientes com a doença. Trata-se de uma revisão realizada no período de Setembro de 2024, por meio de pesquisa na base de dados PubMed. Foram utilizados os descritores: "*screnning AND breast AND artificial intelligence*". Desta busca foram encontrados 26 artigos, posteriormente submetidos aos critérios de seleção.

Os critérios de inclusão foram: artigos publicados no período de Janeiro de 2019 a Janeiro de 2024 e que abordavam as temáticas propostas para esta pesquisa em mulheres, estudos do tipo ensaio clínico, ensaio controlado randomizado e meta-análise, disponibilizados na íntegra. Os critérios de exclusão foram: artigos duplicados, títulos e resumos que não abordavam diretamente a proposta estudada e que não atendiam aos demais critérios de inclusão.

Após os critérios de seleção, 19 artigos foram excluídos pelo título e 2 pelo resumo, restando 5 artigos. Esses artigos foram submetidos à leitura minuciosa para a coleta de dados. Os resultados foram apresentados de forma a comparar métodos tradicionais de radiologistas e o uso da IA no rastreamento do câncer de ma-ma incluindo métricas como a "*Area Under the Curve*" (AUC) do índice de discriminação.

RESULTADOS E DISCUSSÃO

No primeiro estudo, foi observado que a taxa de detecção de câncer de mama pela inteligência artificial (IA) não foi inferior em até 4% em comparação com a leitura dupla realizada por dois radiologistas e, em alguns cenários, a IA apresentou desempenho superior aos radiologistas na detecção da patologia. Esses resultados sugerem que a IA possui grande potencial para ser integrada no processo de rastreamento mamográfico, especialmente quando sua implementação é cuidadosamente controlada, e seu desempenho monitorado de maneira contínua no mundo real. Além disso, a IA também pode ser aplicada de maneira eficaz ao gerenciamento de riscos, garantindo a eficiência do diagnóstico. Ademais, o estudo corroborou achados anteriores de que a utilização de IA associada à leitura radiológica contribuiu para a redução de biópsias desnecessárias, sem comprometer a precisão do diagnóstico. Esse benefício reforça a importância da IA na medicina diagnóstica, não apenas pelo aumento na acurácia da detecção, mas também pela sua capacidade de dimi-

nuir a sobrecarga de procedimentos invasivos, melhorando a experiência do paciente e a eficiência do sistema de saúde (DEMBROWER *et al.*, 2023)

No segundo estudo, foram analisados 16 estudos que englobaram um total de 1.108.328 exames realizados em 497.091 mulheres, incluindo seis estudos de leitores, sete estudos de coorte históricos sobre mamografia digital e quatro sobre tomossíntese digital da mama (DBT). Nos estudos sobre mamografia digital, as áreas sob a curva característica de operação do receptor (AUCs) demonstraram resultados significativamente mais altos para a inteligência artificial (IA) autônoma em comparação aos radiologistas, com uma diferença de AUC de 0,05 (0,87 vs 0,81; P = 0,002). No entanto, nos estudos de coorte históricos, essa diferença não se manteve significativa, com uma diferença de AUC de 0,06 (0,89 vs 0,96; P = 0,152). Em contrapartida, nos estudos sobre tomossíntese digital da mama, a IA demonstrou um desempenho consideravelmente superior aos radiologistas, com uma diferença de AUC de 0,11 (0,90 vs 0,79; P <0,001). Esses resultados indicam que, para a tomossíntese, a IA teve um impacto estatisticamente significativo na melhoria da precisão do diagnóstico. De maneira geral, independentemente do tipo de exame ou do delineamento do estudo, foi observada uma maior sensibilidade e uma menor especificidade para a IA autônoma em comparação aos radiologistas, reforçando a potencialidade da IA como ferramenta complementar ou até mesmo de aprimoramento no processo de diagnóstico radiológico (YOON *et al*, 2023).

O terceiro estudo avaliou a eficácia de um sistema de termografia infravermelha assistido por inteligência artificial (AI-IRT) para triagem precoce do câncer de mama nas mulhres. O Desempenho diagnóstico do modelo binário foi avaliado pelos seguintes métricas obtendo os seguintes resultados: Área sob a curva ROC (AUROC)- Interna igual a 0,9487 (IC 95%: 0,9231–0,9744) e externa igual a 0,9120 (IC 95%: 0,8460–0,9790); Acurácia- Interna igual a 86,27% (IC 95%: 78,04–92,29) e externa igual a 73,53% (IC 95%: 63,87–81,78); Sensibilidade- Interna igual a 93,25%, Externa igual a 84,51%; Especificidade- Interna igual a 80,26%, Externa igual a 83,87%; Valor Preditivo Positivo (PPV)- Interno igual a 83,52% e externo igual a 92,31%; Valor Preditivo Negativo (NPV)- Interno igual a 91,73% e externo igual a 70,27%. Também foi realizada a Comparação com os Leitores humanos através da Acurácia Diagnóstica, com resultados que favorecem o sistema AI-IRT: Acurácia média dos leitores humanos- 0,8088 (IC 95%: 0,7804–0,9229) e Acurácia do modelo AI-IRT- 0,8627 (IC 95%: 0,7804–0,9229). Em termos de Aplicações Clínicas, a sensibilidade foi de 100% para pacientes de alto risco, com especificidade ajustável conforme o limiar de risco. Ainda nesse tópico, foi analisado o custo comparativo por caso detectado, sendo, AI- IRT com custo de 146.100 RMB (US$ 21.266), ultrassonografia com 152.990 RMB (US$ 22.269) e, por fim, mamografia custando 419.600 RMB (US$ 61.077). Dessa forma, o sistema AI-IRT demonstrou alta precisão diagnóstica, superando leitores humanos em acurácia, e apresentou vantagens econômicas significativas, tornando-o uma ferramenta promissora para triagem precoce do câncer de mama (WANG *et al*, 2023).

O quarto estudo apresenta uma revisão sistemática sobre os erros cometidos por sistemas de IA na leitura de mamografias de rastreamento. O estudo utiliza a ferramenta de IA Transpara, desenvolvida para auxiliar na interpretação de mamografias utilizando algoritmos de aprendizado profundo (*Deep Learning*) treinados com grandes volumes de imagens para detectar potenciais áreas suspeitas de câncer de mama,

atribuir uma pontuação de risco para cada mamografia e ajudar radiologistas a priorizar exames com maior probabilidade de anormalidades. O Transpara é utilizado com diferentes limiares de pontuação de 1 a 10. Pontuação mais baixa (ex: 3): mais sensível, detecta mais lesões, mas com maior risco de falsos positivos; Pontuação mais alta (ex: 9): mais específica, menos falsos positivos, mas com maior risco de falsos negativos. Isso significa que ao mudar o ponto de corte da pontuação usada para definir se um exame é suspeito, o desempenho da IA varia em mais segurança ou mais sensibilidade, dependendo da escolha. Dessa forma, os resultados obtidos foram: Falsos negativos- Quando a IA não detecta anomalias presentes nas imagens. A taxa de falsos negativos aumentou com o aumento do limiar de positividade. No limiar Transpara 3, a taxa de falsos negativos foi de 0,02% (intervalo de confiança de 95%: 0,01% a 0,03%), enquanto no limiar Transpara 9, essa taxa subiu para 0,12% (intervalo de confiança de 95%: 0,06% a 0,26%); Falsos positivos - Quando a IA identifica anomalias que não estão presentes. A proporção de falsos positivos diminuiu conforme o aumento do limiar de positividade do algoritmo. Por exemplo, com o limiar Transpara 3, a taxa de falsos positivos foi de 71,83% (intervalo de confiança de 95%: 69,67 % a 73,90%), enquanto com o limiar Transpara 9, essa taxa caiu para 10,77% (intervalo de confiança de 95%: 8,34% a 13,79%); Erros de localização e classificação- quando a IA marca incorretamente a localização de uma anomalia ou erroneamente a gravidade ou tipo de anomalia. Esses erros foram reportados em menor frequência, com apenas um estudo incluindo cada tipo desses erros. Os resultados, portanto, indicam que os erros da IA na leitura de mamografias estão principalmente relacionados à precisão do teste, com variações esperadas não apenas pelo limiar de positividade, mas também pela versão do algoritmo e pela qualidade do estudo. A revisão destaca a importância de relatar outros tipos de erros da IA, como erros técnicos e de localização, para fornecer uma compreensão mais abrangente da utilidade da IA na prática de rastreamento do câncer de mama (ZENG *et al*, 2024).

O quinto artigo analisado apresenta uma revisão sistemática e meta-análise sobre a aplicação de *Machine Learning* (ML) na triagem por mamografia para diagnóstico de câncer de mama. Foram incluídos 32 estudos, totalizando 23.804 imagens, com foco em três métodos de ML: redes neurais convolucionais (CNN), redes neurais artificiais (ANN) e máquinas de vetor de suporte (SVM). De modo geral, o desempenho dos modelos de ML foi o seguinte: Sensibilidade média: 91,4% (intervalo de confiança [IC]: 87,8–94,0%); Especificidade média: 91,6% (IC: 87,7–94,3%); Área sob a curva ROC (AUC): 0,945 (IC: 0,910–0,967). Avaliando separadamente cada modelo de ML obteve-se: Redes Neurais Convolucionais (CNN)- Sensibilidade de 96,1% (IC: 94,4–97,4%), Especificidade de 95,0% (IC: 91,3–97,3%) e AUC igual a 0,974 (IC: 0,957–0,984); Redes Neurais Artificiais (ANN)- Sensibilidade de 83,7% (IC: 77,7–88,4%), Especificidade de 89,4% (IC: 81,8–94,1%) e AUC igual a 0,881 (IC: 0,827–0,922); Máquinas de Vetor de Suporte (SVM) - Sensibilidade de 88,9% (IC: 80,5–94,0%), Especificidade de 84,3% (IC: 73,2–91,4%) e AUC igual a 0,913 (IC: 0,843–0,956). Esses números mostram que modelos de ML, especialmente as CNNs, apresentam alta precisão na triagem por mamografia. Entretanto, os autores destacam que, apesar dos resultados promissores, são necessários mais estudos prospectivos para avaliar o desempenho longitudinal dos sistemas de ML na triagem por mamografia (HICKMAN *et al.*, 2021).

CONCLUSÃO

A inteligência artificial aplicada ao rastreamento mamográfico tem demonstrado desempenho promissor, com algoritmos capazes de igualar ou superar radiologistas experientes na detecção precoce do câncer de mama, aumentando a acurácia diagnóstica e reduzindo falsos positivos. No entanto, sua incorporação plena à prática clínica exige validação multicêntrica, integração com dados clínicos e vigilância contínua de seu desempenho em contextos reais. A IA deve ser compreendida como uma aliada ao julgamento médico, contribuindo para a eficiência dos programas de triagem e para a ampliação do acesso ao diagnóstico precoce. Seu impacto sustentável dependerá da convergência entre inovação tecnológica, evidência científica e responsabilidade ética.

REFERÊNCIAS BIBLIOGRÁFICAS

BORMANN, C. L. *et al*. Deep learning early warning system for embryo culture conditions and embryologist performance in the ART laboratory. *Journal of Assisted Reproduction and Genetics*, v. 38, n. 7, p. 1641-1646, jul. 2021. DOI: 10.1007/s10815-021-02198-x.

CHEN, Z. *et al*. Artificial intelligence in the assessment of female reproductive function using ultrasound: a review. *Journal of Ultrasound in Medicine*, v. 41, n. 6, p. 1343-1353, jun. 2022. DOI: 10.1002/jum.15827.

COLLINS, G. S *et al.* Reporting of artificial intelligence prediction models. *The Lancet*, v. 393, n. 10181, p. 1577-1579, 20 abr. 2019. DOI: 10.1016/S0140-6736(19)30037-6.

DA CUNHA, G. N. *et al*. Rastreamento do câncer de mama: modelo de melhoria do acesso pelo uso de mamógrafos móveis. *Revista Panamericana de Salud Pública*, v. 43, e19, 6 fev. 2019. DOI: 10.26633/RPSP.2019.19.

DEMBROWER, K. *et al*. Artificial intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study. *Lancet Digital Health*, v. 5, n. 10, p. e703-e711, out. 2023. DOI: 10.1016/S2589-7500(23)00153-X.

FEDERAÇÃO BRASILEIRA DAS ASSOCIAÇÕES DE GINECOLOGIA E OBSTETRÍCIA (FEBRASGO). Rastreamento e propedêutica do câncer de mama. São Paulo: FEBRASGO, 2021. (Protocolo FEBRASGO Ginecologia, n. 89. Comissão Nacional Especializada em Imaginologia Mamária).

GOLDHIRSCH, A. *et al*. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Annals of Oncology*, v. 24, n. 9, p. 2206-2223, set. 2013.

HICKMAN, S. E. *et al*. Machine learning for workflow applications in screening mammography: systematic review and meta-analysis. *Radiology*, v. 302, n. 1, p. 88-104, jan. 2022. DOI: 10.1148/radiol.2021210391.

INSTITUTO NACIONAL DE CÂNCER (INCA). Dados e números sobre câncer de mama: relatório anual 2022. Rio de Janeiro: INCA, set. 2022. Disponível em: http://www.inca.gov.br/mama. Acesso em: 28 abr. 2025.

KIM, H. M. *et al*. Improved prediction of clinical pregnancy using artificial intelligence with enhanced inner cell mass and trophectoderm images. *Scientific Reports*, v. 14, n. 1, p. 3240, 8 fev. 2024. DOI: 10.1038/s41598-024-52241-x.

KRAGH, M. F *et al*. Embryo selection with artificial intelligence: how to evaluate and compare methods? *Journal of Assisted Reproduction and Genetics*, v. 38, n. 7, p. 1675-1689, jul. 2021. DOI: 10.1007/s10815-021-02254-6.

MONTJEAN, D. *et al*. Automated single-sperm selection software (SiD) during ICSI: a prospective sibling oocyte evaluation. *Medical Sciences*, v. 12, n. 2, p. 19, 27 mar. 2024. DOI: 10.3390/medsci12020019.

MOONS, K. G. M. *et al*. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Annals of Internal Medicine*, v. 170, n. 1, p. W1-W33, 1 jan. 2019. DOI: 10.7326/M18-1377.

SFAKIANOUDIS, K. *et al*. Reporting on the value of artificial intelligence in predicting the optimal embryo for transfer: a systematic review including data synthesis. *Biomedicines*, v. 10, n. 3, p. 697, 17 mar. 2022. DOI: 10.3390/biomedicines10030697.

WANG, X**.** *et al*. Breast cancer pre-clinical screening using infrared thermography and artificial intelligence: a prospective, multicentre, diagnostic accuracy cohort study. *International Journal of Surgery*, v. 109, n. 10, p. 3021-3031, 1 out. 2023. DOI: 10.1097/JS9.0000000000000594.

YOON, J. H. *et al*. Standalone AI for breast cancer detection at screening digital mammography and digital breast tomosynthesis: a systematic review and meta-analysis. *Radiology*, v. 307, n. 5, p. e222639, jun. 2023. DOI: 10.1148/radiol.222639.

ZENG, A. *et al*. Frequency and characteristics of errors by artificial intelligence (AI) in reading screening mammography: a systematic review. *Breast Cancer Research and Treatment*, v. 207, n. 1, p. 1-13, ago. 2024. DOI: 10.1007/s10549-024-07353-3.