**POLISH**
**JOURNAL** *of* **RADIOLOGY**

Review paper

# Artificial intelligence versus radiologists in detecting early-stage breast cancer from mammograms: a meta-analysis of paradigm shifts

Hashim Talib Hashim[1,A,C,E,G], Ahmed Qasim Mohammed Alhatemi[2,A,D,F], Motaz Daraghma[3,B,C,E], Hossam Tharwat Ali[4,A,C,D], Mudassir Ahmad Khan[5,A,E,F], Fatimah Abdullah Sulaiman[6,E,F], Zahraa Hussein Ali[7,E,F,G], Mohanad Ahmed Sahib[8,A,D,E,G], Ahmed Dheyaa Al-Obaidi[9,A E,F], Ammar Al-Obaidi[10,A,B,D,E]

[1]University of Warith AI-Anbiyaa, College of Medicine, Karbala, Iraq
[2]Al-Nassiryah Teaching Hospital, Thi Qar, Iraq
[3]Faculty of Medicine, Arab American University of Palestine, Jenin, Palestine
[4]Qena Faculty of Medicine, South Valley University, Qena, Egypt
[5]Department of Physiology, Shalamar Institute of Health Sciences, Lahore, Pakistan
[6]Al-Nassiryah Teaching Hospital, Nassiryah, Iraq
[7]University of Baghdad, College of Medicine, Baghdad, Iraq
[8]Department of Radiological Techniques, College of Health and Medical Techniques, Al-Mustaqbal University, Babylon, Iraq
[9]MBChB, University of Baghdad, College of Medicine, Baghdad, Iraq
[10]Department of Hematology/Oncology, University of Missouri-Kansas City, MO, USA

## Abstract

**Purpose**: Early detection of breast cancer is crucial for improving patient outcomes. With advancements in artificial intelligence (AI), there is growing interest in its potential to assist radiologists in interpreting mammograms for early cancer detection. AI algorithms offer the promise of increased accuracy and efficiency in identifying subtle signs of breast cancer, potentially complementing the expertise of radiologists and enhancing the screening process for early-stage breast cancer detection.

**Material and methods:** A systematic literature review was conducted to identify and select original research reports on breast cancer diagnosis by artificial intelligence versus conventional radiologists in using mammograms in accordance with the PRISMA guidelines. Data were analysed with Review Manager version 5.4. *P*-value and $I^2$ were used to test the significance of differences.

**Results:** This systematic review and meta-analysis included 8 studies with data from a total of 120,950 patients. Regarding the sensitivity of AI, the pooled analysis of 6 studies with sensitivities ranging from 0.70 to 0.89 yielded a sensitivity of 0.85. However, the sensitivity of the radiologists ranged from 0.63 to 0.85, with an overall sensitivity of 0.77. As for specificity, both radiologists and AI groups had closer results.

**Conclusions:** The comparison between AI systems and radiologists in detecting early-stage breast cancer from mammograms highlights the potential of AI as a valuable tool in breast cancer screening. While AI algorithms have shown promising results in terms of accuracy and efficiency, they should be viewed as complementary to radiologists rather than replacements.

**Key words:** artificial intelligence, radiologists, detecting early-stage breast cancer.

**Correspondence address:**
Dr. Hashim Talib Hashim, University of Warith Al-Anbiyaa, College of Medicine, Karbala, Iraq, e-mail: hashim.h.t.h@gmail.com

**Authors' contribution:**
**A** Study design · **B** Data collection · **C** Statistical analysis · **D** Data interpretation · **E** Manuscript preparation · **F** Literature search · **G** Funds collection

## Introduction

Breast cancer is the most common site-specific cancer in women and is the leading cause of death from cancer for women aged 20 to 59 years. It accounts for 26% of all newly diagnosed cancers in females and is responsible for 15% of the cancer-related deaths in women. Besides lung cancer, breast cancer is the most common cause of cancer related death in the world, the latest cancer statistics for the USA estimate that in 2022, 31% of cancer cases detected in women were breast cancer, with 43,250 cases resulting in death. This accounts for 15% of all cancer-related deaths [1,2]. Approximately 600 men and 40,000 women die of breast cancer annually according to the findings of the American Cancer Society [3].

About 15% of women who get breast cancer have a positive family history of it, but only 5-10% of breast cancers are inherited. The most significant risk factors for breast cancer are female gender and age [4].

Breast cancer can be classified as benign, in situ carcinoma, or invasive carcinoma. Benign tumours are not classified as dangerous because they cause only a slight alteration in breast anatomy. In situ carcinomas are not dangerous if detected in the early stage and treated, because only the mammary duct lobules are affected and they do not spread to the other tissues. Invasive carcinoma is the most dangerous type because it can spread to any organ in the body [3].

Early detection is a crucial step to control and treat breast cancer. This aims at the detection of breast cancer very early in asymptomatic females to offer them a better chance of cure. Early breast cancer screening through mammography, ultrasound, or magnetic resonance imaging (MRI) has played an important role in helping detect breast cancer in the early stage, reduce the mortality rate, and improve prognosis; the mortality rate of breast cancer dropped by 40% from 1989 to 2017, which translates to 375,900 breast cancer deaths averted [1,2].

The prognosis of breast cancer is critically affected by early detection and treatment; for example, in 2020 more than 65% of breast cancer patients were diagnosed in the early stage of cancer and survived. There has been a significant decrease in breast cancer-related mortality in the United States between 1975 and 2000, and this is attributed to continued improvement in both screening mammography and treatment [4,5].

The clinical cure rate of breast cancer is highly optimistic and exceeds 90% if diagnosed in the early stage, and this rate decreases as the disease progresses, ranging from 50% to 70% in the middle stage, while treatment is typically not effective in the late stage [3].

Different breast cancer imaging modalities such as mammogram, US, MRI, and histopathological imaging can be used to detect and analyse the key features affecting the diagnosis and treatment.

Mammography represents the mainstay of breast cancer screening and a significant method for the detection and staging of breast cancer, evaluation of treatment efficacy, and follow-up examination. Mammography screening is one of the most widely used modalities for early breast cancer detection, and it has been shown to decrease mortality in multiple randomised clinical trials. Despite this, its performance is often unsatisfactory, with lower sensitivity (i.e. missing one in 8 cancers during interpretation) and very high false positive rates (i.e. < 30% of biopsies are malignant). Approximately 9-10% of the 40 million US women who undergo routine breast screening each year are recalled for additional diagnostic imaging; only 4% to 5% of women recalled are ultimately diagnosed as having breast cancer, and because of the downfall of mammography the need for other adjuvant imaging modalities is increased [2,6,7].

Radiologists view and interpret breast images produced by these modalities and use them for diagnosis. Analysing breast images remains difficult due to the high heterogeneity of breast tumours and the long hours worked by radiologists. Its benefit is dependent on subjective human interpretation to maximally extract all diagnostic information from the acquired images, and this can lead to misjudgement and misdiagnoses, which results in lower cancer detection sensitivity and specificity and large inter-reader variability. Thus, utilisation of new automatic methods to analyse all kinds of breast screening images to assist radiologists in interpreting images is required. To help overcome these clinical challenges, researchers have made great efforts to develop computer-aided detection and/or diagnosis (CAD) schemes of breast images to provide radiologists with decision-making support tools [4].

Artificial intelligence (AI) provides the capability of a computer system to interpret and analyse breast imaging, which can alleviate potential human errors.

AI applications have shown excellent performance in various image recognition tasks, and their use in breast cancer screening has been explored in numerous studies [3].

## Methodology

A systematic literature review was conducted to compare the diagnostic performance of AI with that of conventional radiologists in detecting breast cancer using mammograms, following the PRISMA guidelines. A meta-analysis was performed using Review Manager (RevMan) version 5.4. Pooled sensitivity and specificity with corresponding 95% confidence intervals (CIs) were calculated for both AI and radiologists. Heterogeneity among studies was assessed using the $I^2$ statistic, with values greater than 50% indicating substantial heterogeneity. The significance of heterogeneity was tested using the $\chi^2$ with a $p$-value of less than 0.10 (Figure 1).
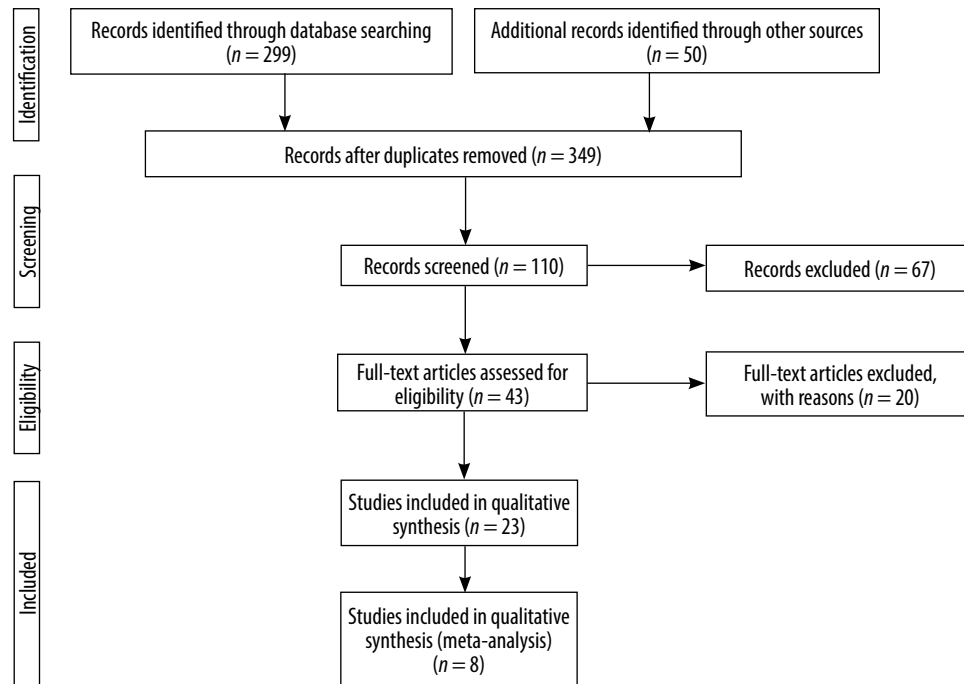
**Figure 1.** Flow chart of the studies selection process

## Study selection, inclusion, and exclusion criteria

Comprehensive database searches including PubMed, Scopus, and Web of Science incorporated keywords such as "artificial intelligence", "radiologists", "breast cancer", "mammograms", and "diagnostic accuracy". All acquired studies from the search were integrated into Rayyan for subsequent analysis and deduplication, then they were screened and reviewed by title, abstract, and full text screening. Inclusion and exclusion criteria were predetermined. Studies were selected if they included the following:

- comparison of AI systems to radiologists in interpreting mammograms;
- sensitivity and specificity data;
- were published in peer-reviewed journals;
- involved human subjects.

Extracted data included study characteristics, non-English language-based articles, duplicated studies, book chapters, conference abstracts, studies employing animal subjects, or studies with no relevant data.

## Results

This systematic review and meta-analysis included 8 studies with data from a total of 120,950 patients (Tables 1 and 2). Regarding the sensitivity of AI, the pooled analysis of 6 studies with sensitivities ranging from 0.70 to 0.89 yielded a sensitivity of 0.85 (95% CI: [0.79–0.89]) (Figure 2). On the other hand, the sensitivity of radiologists ranged from 0.63 to 0.85 with an overall sensitivity of 0.77 (95% CI: [0.69–0.83]) (Figure 3). Significant heterogeneity was detected in both cases ($I^2 > 90\%$, $p < 0.01$).
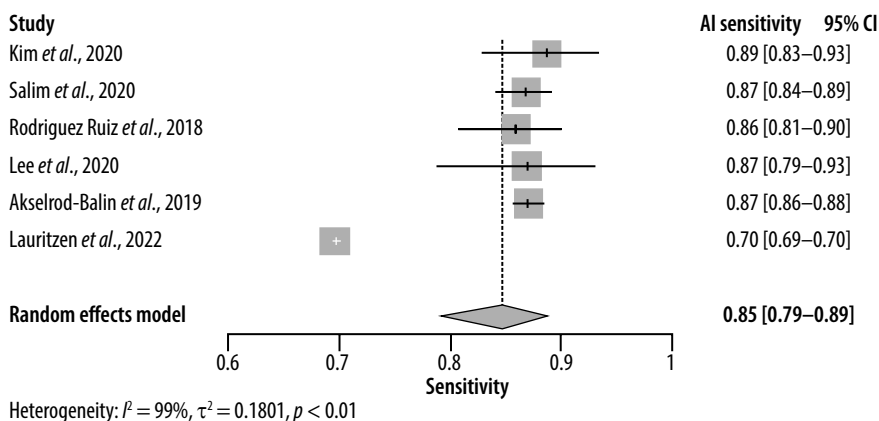
As for specificity, the radiologist and AI groups had closer results. The pooled specificity of AI was 0.89 (95% CI: [0.76–0.95]) ranging from 0.77 to 0.99 in 6 studies (Figure 4), whereas for the radiologists it was

**Table 1.** Baseline table for the included studies

| Authors, years [Ref.] | n.e | SnAI | SpAI | AUCAI | n.c | SnRad | SpRad | AUCRad |
|---|---|---|---|---|---|---|---|---|
| Kim *et al.*, 2020 [17] | 160 | 88.75 | 82 | 0.94 | | 75.27 | 71.61 | 0.81 |
| Watanabe *et al.*, 2019 [19] | 90 | | | 0.84 | | | | 0.814 |
| Salim *et al.*, 2020 [18] | 739 | 86.70 | 93 | | | 85.00 | 98.50 | |
| Rodriguez Ruiz *et al.*, 2019 [19] | 2652 | | | 0.84 | | | | 0.814 |
| Rodriguez Ruiz *et al.*, 2018 [24] | 240 | 0.86 | 79 | 0.89 | | 83.00 | 77.00 | 0.87 |
| Lee *et al.*, 2020 [22] | 100 | 87.00 | 79 | 0.92 | | 62.90 | 68.70 | 0.748 |
| Akselrod-Balin *et al.*, 2019 [23] | 2548 | 87.00 | 77 | 0.91 | | | | |
| Lauritzen *et al.*, 2022 [21] | 11,4421 | 69.70 | 99 | | | 70.80 | 98.10 | |

**Table 2.** Demographic data of the included studies

| Population | Women undergoing mammography | | | | | | |
|---|---|---|---|---|---|---|---|
| **Index tests** | AI and radiologists (ref: histopath diagnosis) | | | | | | |
| **Target** | Early breast CA | | | | | | |
| **Study: Author, year [Ref.]** | | **No of pts** | **Outcomes** | | | | |
| | | | **SnAI** | **Rad** | **SpAI** | **Rad** | **AUCAI** | **Rad** |

| Study: Author, year [Ref.] | | No of pts | SnAI | Rad | SpAI | Rad | AUCAI | Rad |
|---|---|---|---|---|---|---|---|---|
| Kim, 2020 [17] | South Korea data sets | Total 320 (n = 160) | | | | | 0.959 | 0.81 |
| | | | | | | | 0.97 | |
| | US data set | | | | | | 0.953 | |
| | UK data set | | | | | | 0.938 | |
| | Reader study | | 88.75 | 75.27 | 81.87 | 71.61 | 0.94 | |
| Watanabe, 2019 [19] | | Total 122 (n = 90) | | | | | 0.84 | 0.814 |
| | | | | | | | 0.6639 | 0.759 |
| Salim, 2020 [18] | Overall | Total 113,663 (n = 739) | 86.7 | 85 | 92.5 | 98.5 | | |
| | 1 | | 81.9 | 77.4 | 96.6 | 96.6 | 0.956 | |
| | 2 | | 67 | 80.1 | 96.6 | 97.2 | 0.922 | |
| | 3 | | 67.4 | | 96.7 | | 0.92 | |
| Dembrower, 2020 [13] | | Total 7364 (n = 547) | | | | | | |
| Rodriguez Ruiz, 2019 [16] | Overall | Total 1393 (n = 499) | | | | | 0.84 | 0.814 |
| | 1 | | 0.85 | 0.84 | | 0.49 | 0.783 | 0.769 (0.698, 0.840) |
| | 2 | | N/A | N/A | | N/A | 0.915 | 0.907 (0.854, 0.961) |
| | 3 | | 0.8 | 0.77 | | 0.79 | 0.879 | 0.858 (0.814, 0.901) |
| | 4 | | 0.85 | 0.77 | | 0.67 | 0.85 | 0.815 (0.767, 0.864) |
| | 5 | | 0.86 | 0.82 | | 0.54 | 0.825 | 0.787 (0.732, 0.841) |
| | 6 | | 0.81 | 0.83 | | 0.51 | 0.796 | 0.803 (0.763, 0.843) |
| | 7 | | 0.86 | 0.84 | | 0.68 | 0.852 | 0.860 (0.831, 0.889) |
| | 8 | | 0.75 | 0.76 | | 0.75 | 0.817 | 0.808 (0.752, 0.859) |
| | 9 | | 0.81 | 0.83 | | 0.73 | 0.861 | 0.841 (0.785, 0.897) |
| Sickles, 2002 [25] | | | | | | | | |
| Rodriguez Ruiz, 2018 [24] | | | 0.86 | 83 | 79 | 77 | 0.89 | 0.87 |
| Lee, 2022 [23] | Overall | Total 200 (n = 100) | 87 | 62.9 | 79 | 68.7 | 0.915 | 0.748 |
| Akselrod-Balin, 2019 [22] | | | 87 | | 77.3 | | 0.91 | |
| Lauritzen, 2022 [21] | | | 69.7 | 70.8 | 98.6 | 98.1 | | |



| Study | | AI sensitivity 95% CI |
|---|---|---|
| Kim et al., 2020 | | 0.89 [0.83–0.93] |
| Salim et al., 2020 | | 0.87 [0.84–0.89] |
| Rodriguez Ruiz et al., 2018 | | 0.86 [0.81–0.90] |
| Lee et al., 2020 | | 0.87 [0.79–0.93] |
| Akselrod-Balin et al., 2019 | | 0.87 [0.86–0.88] |
| Lauritzen et al., 2022 | | 0.70 [0.69–0.70] |
| **Random effects model** | | **0.85 [0.79–0.89]** |

Sensitivity

Heterogeneity: $I^2 = 99\%$, $\tau^2 = 0.1801$, $p < 0.01$

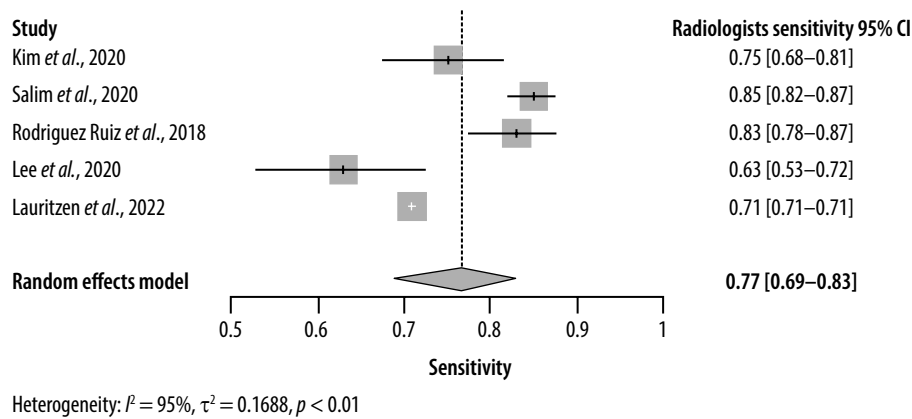**Figure 2.** The sensitivity of artificial intelligence (AI)

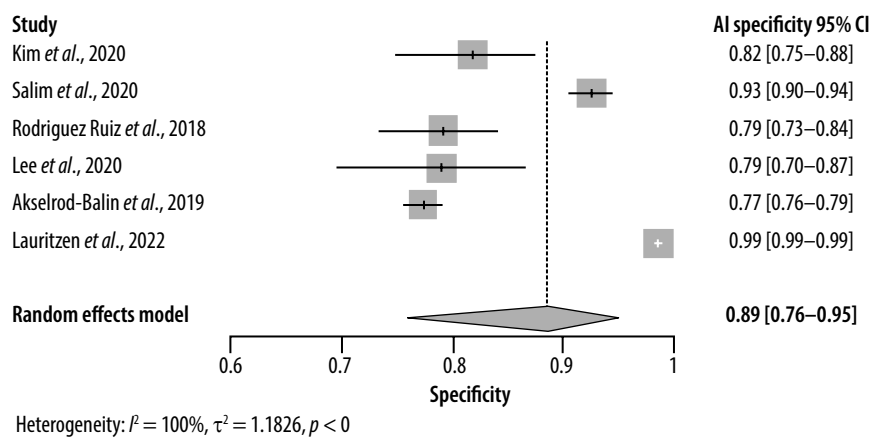**Figure 3.** The sensitivity of radiologists



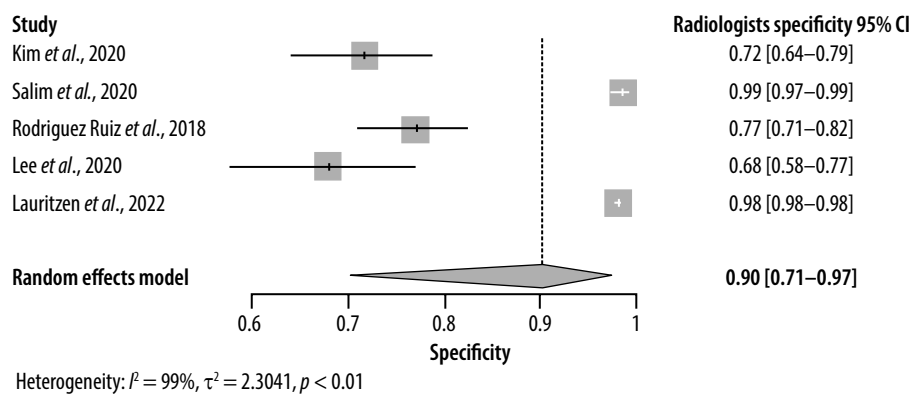**Figure 4.** The pooled specificity of artificial intelligence (AI)



**Figure 5.** The pooled specificity of radiologists

0.90 (95% CI: [0.71–0.97]) with a range of 0.68 to 0.99 in 5 studies (Figure 5). Significant heterogeneity was detected in both cases ($I^2 > 90\%$, $p < 0.01$).

Estimating the area under the curve (AUC) for both AI and radiologists revealed a beneficial effect of AI with a pooled AUC of 0.89 (95% CI: [0.86–0.92]) with significant heterogeneity ($I^2 = 92\%$, $p < 0.01$) (Figure 6) compared to 0.82 (95% CI: [0.80–0.83]) for the radiologists, without significant heterogeneity ($I^2 = 48\%$, $p = 0.10$) (Figure 7).
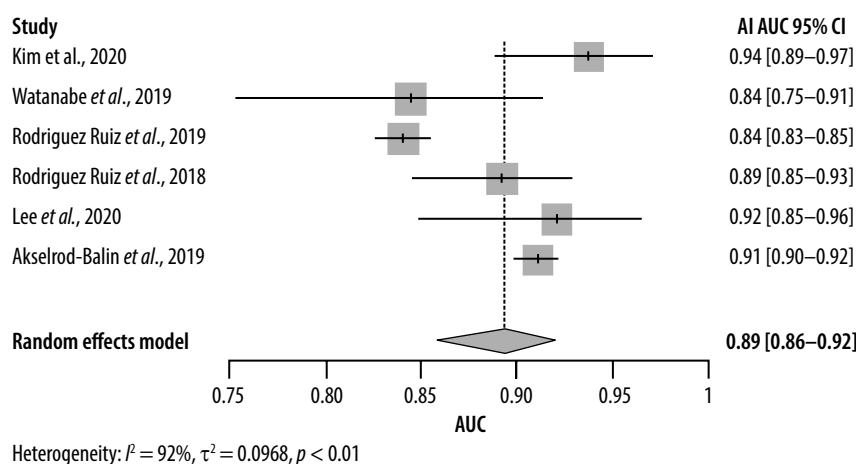
| Study | | AI AUC 95% CI |
|---|---|---|
| Kim et al., 2020 | | 0.94 [0.89–0.97] |
| Watanabe et al., 2019 | | 0.84 [0.75–0.91] |
| Rodriguez Ruiz et al., 2019 | | 0.84 [0.83–0.85] |
| Rodriguez Ruiz et al., 2018 | | 0.89 [0.85–0.93] |
| Lee et al., 2020 | | 0.92 [0.85–0.96] |
| Akselrod-Balin et al., 2019 | | 0.91 [0.90–0.92] |
| **Random effects model** | | **0.89 [0.86–0.92]** |

Heterogeneity: $I^2 = 92\%$, $\tau^2 = 0.0968$, $p < 0.01$

**Figure 6.** The area under the curve (AUC) for artificial intelligence

| Study | | Rad AUC 95% CI |
|---|---|---|
| Kim et al., 2020 | | 0.81 [0.74–0.87] |
| Watanabe et al., 2019 | | 0.81 [0.71–0.89] |
| Rodriguez Ruiz et al., 2019 | | 0.81 [0.80–0.83] |
| Rodriguez Ruiz et al., 2018 | | 0.87 [0.82–0.91] |
| Lee et al., 2020 | | 0.75 [0.65–0.83] |
| **Random effects model** | | **0.82 [0.80–0.83]** |

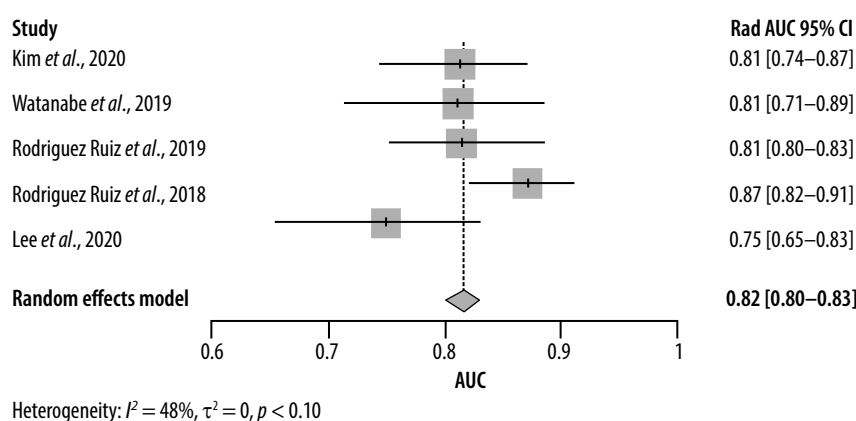Heterogeneity: $I^2 = 48\%$, $\tau^2 = 0$, $p < 0.10$

**Figure 7.** The area under the curve (AUC) for radiologists (Rad)

## Discussion

This meta-analysis of 8 studies involving 120,950 patients compared the diagnostic performance of AI and radiologists, demonstrating that AI exhibits a higher pooled sensitivity (0.85 vs. 0.77 for radiologists) and similar specificity (0.89 vs. 0.90 for radiologists), with significant variability observed in both sensitivity and specificity across studies. AI also showed a superior pooled AUC of 0.89 compared to 0.82 for radiologists, indicating better overall diagnostic accuracy despite variability, suggesting AI's potential to surpass human radiologists in certain contexts.

Similarly, Al-Karawi et al. [3] suggest that AI techniques, particularly deep learning (DL), are increasingly applied in various medical fields, including breast imaging, exhibiting robust performance in tasks such as image recognition. This enhances prospects for applications in in vitro diagnosis, rehabilitation, medical imaging, and prognosis. They also found that despite challenges such as multitasking limitations, ongoing advancements in DL-based systems for breast imaging, such as digital breast tomosynthesis and ultrasound, show rapid deve-

lopment. These systems aid in detecting, classifying, and predicting breast diseases, thereby improving diagnostic efficiency and treatment efficacy. Furthermore, AI's capabilities offer fast computation, repeatability, and objective data analysis, reducing medical professionals' workload and enhancing diagnostic accuracy and treatment outcomes [2,8].

Aligning with these results, a study conducted by Zhang et al. [9] supports the use of AI for breast cancer diagnosis through dual-modal deep polynomial networks, which proved superior to other methods in enhancing breast tumour classification performance. This suggests that AI-based techniques can approach the accuracy of breast tumour biopsies, proposing a dual-modal AI framework for diagnosis. However, experimental findings highlight dual-modal deep polynomial networks as the most effective framework, indicating AI's potential for streamlined breast tumour classification.

Similar results were found by Cè et al. [10], who highlighted significant advancements in AI-based tools for personalised patient management. The integration of AI into clinical workflows is expected to benefit women, radiologists, and healthcare systems by improving diag-

nostic accuracy and streamlining workload. AI facilitates the development of personalised screening protocols and aids in early lesion detection, thereby reducing overdiagnosis risks.

Gilbert *et al.* [11] recommend that AI systems be used to mitigate challenges related to the development, expansion, or continuation of screening programs in countries with insufficient experienced breast radiologists. In such situations, AI systems could be utilised as either a primary or secondary reader, thereby mitigating the impact of radiologist shortages.

Conversely, recent research assessing differences in diagnostic performance indicates that radiologists' interpretations of mammography are superior to AI analyses in clinical practice, particularly when considering clinical symptoms and prior mammograms. These findings have been supported by similar studies that explored the role of AI in clinical practices. AI should support rather than replace radiologists, especially in complex cases. Additionally, appropriate validation of AI scoring thresholds is essential for effective clinical use. AI-based image analysis differs from the comprehensive evaluation performed by radiologists, which includes consideration of clinical symptoms and comparisons with previous mammograms.

Radiologists detected subtle or ambiguous findings that AI often missed. Instances were observed where AI failed to identify clearly visible lesions, indicating limitations in AI algorithms and suggesting that AI's threshold scores may not reliably indicate malignancy, highlighting the need for refined score classifications [12-15].

However, this meta-analysis highlights the promising potential of collaborative use of AI algorithms alongside radiologists in enhancing diagnostic accuracy and efficiency in breast cancer detection, necessitating careful consideration of its limitations and validation in clinical practice. Further research efforts should prioritise the refinement of AI algorithms, validation of their clinical utility, and exploration of optimal strategies for integrating AI into routine clinical workflows to maximise its potential impact on patient care.

## Disclosures

1. Institutional review board statement: Not applicable.
2. Assistance with the article: None.
3. Financial support and sponsorship: None.
4. Conflicts of interest: None.

## References

1. Warrier S, Tapia G, Goltsman D, Beith J.. An update in breast cancer screening and management. Womens Health (Lond) 2016; 12: 229-239.

2. Jones MA, Islam W, Faiz R, Chen X, Zheng B. Applying artificial intelligence technology to assist with breast cancer diagnosis and prognosis prediction. Front Oncol 2022; 12: 980793. DOI: 10.3389/fonc.2022.980793.

3. Al-Karawi D, Al-Zaidi S, Helael KA, Obeidat N, Mouhsen AM, Ajam T, et al. A review of artificial intelligence in breast imaging. Tomography 2024; 10: 705-726.

4. Shah R, Rosso K, Nathanson SD. Pathogenesis, prevention, diagnosis and treatment of breast cancer. World J Clin Oncol 2014; 5: 283-298.

5. Madani M, Behzadi MM, Nabavi S. The role of artificial intelligence in advancing breast cancer detection using different imaging modalities: a systematic review. Cancers (Basel) 2022; 14: 5334. DOI: 10.3390/cancers14215334.

6. 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; 3: e200265. DOI: 10.1001/jamanetworkopen.2020.0265.

7. Morgan MB, Mates JL. Applications of artificial intelligence in breast imaging. Radiol Clin N Am 2021; 59: 139-148.

8. Zhang Q, Song S, Xiao Y, Chen S, Shi J, Zheng H. Dual-mode artificially-intelligent diagnosis of breast tumours in shear-wave elastography and B-mode ultrasound using deep polynomial networks. Med Eng Phys 2019; 64: 1-6.

9. Cè M, Caloro E, Pellegrino ME, Basile M, Sorce A, Fazzini D, et al. Artificial intelligence in breast cancer imaging: risk stratification, lesion detection and classification, treatment planning and prognosis – a narrative review. Explor Target Antitumor Ther 2022; 3: 795-816.

10. Gilbert FJ, Astley SM, Gillan MG, Agbaje OF, Wallis MG, James J, et al. Single reading with computer-aided detection for screening mammography. N Engl J Med 2008; 359: 1675-1684.

11. Choi WJ, An JK, Woo JJ, Kwak HY. Comparison of diagnostic performance in mammography assessment: radiologist with reference to clinical information versus standalone artificial intelligence detection. Diagnostics (Basel) 2022; 13: 117. DOI: 10.3390/diagnostics13010117.

12. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, Broeders M, Gennaro G, Clauser P, et al. Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists. J Natl Cancer Inst 2019; 111: 916-922.

13. 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; 2: e468-e474. DOI: 10.1016/S2589-7500(20)30185-0.

14. Al-Obaidi MN, Al-Obaidi AD, Hashim HT, Al Sakini AS, Abd AM, Rashed RH, et al. Assessing breast self-examination knowledge and practices among women in Iraq: a cross-sectional study. Health Sci Rep 2024; 7: e2137. DOI: 10.1002/hsr2.2137.

15. Hashim HT, et al. Assessment of breast cancer risk among Iraqi women in 2019. BMC Womens Health 2021; 21: 412. DOI: 10.1186/s12905-021-01557-1.

16. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, Broeders M, Gennaro G, Clauser P, et al. Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists. J Natl Cancer Inst 2019; 111: 916-922.

17. 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; 2: e138-e148. DOI: 10.1016/S2589-7500(20)30003-0.

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; 6: 1581-1588.

19. Watanabe AT, Lim V, Vu HX, Chim R, Weise E, Liu J, et al. Improved cancer detection using artificial intelligence: a retrospective evaluation of missed cancers on mammography. J Digit Imaging 2019; 32: 625-637.

20. Lauritzen AD, Rodríguez-Ruiz A, von Euler-Chelpin MC, Lynge E, Vejborg I, Nielsen M, et al. An artificial intelligence-based mammography screening protocol for breast cancer: outcome and radiologist workload. Radiology 2022; 304: 41-49.

21. Akselrod-Ballin A, Chorev M, Shoshan Y, Spiro A, Hazan A, Melamed R, Barkan E, et al. Predicting breast cancer by applying deep learning to linked health records and mammograms. Radiology 2019; 292: 331-342.

22. Lee JH, Kim KH, Lee EH, Ahn JS, Ryu JK, Park YM, et al. Improving the performance of radiologists using artificial intelligence-based detection support software for mammography: a multi-reader study. Korean J Radiol 2022; 23: 505-216.

23. Rodríguez-Ruiz A, Krupinski E, Mordang JJ, Schilling K, Hey-wang-Köbrunner SH, Sechopoulos I, Mann RM. Detection of breast cancer with mammography: effect of an artificial intelligence support system. Radiology 2016; 290: 305-314.

24. Sickles EA, Wolverton DE, Dee KE. Performance parameters for screening and diagnostic mammography: specialist and general radiologists. Radiology 2002; 224: 861-869.