Hong Kong J Radiol. 2026;29:Epub. Published 26 February 2026 | https://doi.org/10.12809/hkjr2417896

ORIGINAL ARTICLE

Improving Breast Cancer Detection in Screening Mammography with Artificial Intelligence Assistance: A Multi-reader Retrospective Study

PL Lam[1], D Fenn[1], EH Chan[2], EWS Fok[3], PH Lee[1], KM Kwok[2], LKM Wong[1], WS Mak[1], WP Cheung[1], WI Sit[1], WK Ng[1], GCY Chan[1], LW Lo[1], EPY Fung[1]

[1]*Department of Diagnostic and Interventional Radiology, Kwong Wah Hospital, Hong Kong SAR, China*
[2]*Department of Diagnostic and Interventional Radiology, Princess Margaret Hospital, Hong Kong SAR, China*
[3]*Department of Radiology and Organ Imaging, United Christian Hospital, Hong Kong SAR, China*

ABSTRACT

Introduction: *This study aimed to compare the performance of radiologists in screening mammography for breast cancer detection, with and without artificial intelligence (AI) assistance, including subgroup comparison between breast radiologists and general radiologists in Hong Kong.*

Methods: *This was a single-centre multi-reader retrospective study. A screening mammography test set was used (the Hong Kong Personal Performance in Mammographic Screening Scheme), comprising 80 mammograms with negative or benign findings and 36 mammograms with pathologically proven breast cancer acquired from December 2009 to December 2023. Radiologists' performance with and without AI assistance from a commercially available tool (Lunit INSIGHT MMG) was evaluated from December 2023 to April 2024. The two reading sessions were separated by a 4-week washout period. Study endpoints included sensitivity and specificity in the mammographic detection of breast cancer. The Obuchowski–Rockette model was used to estimate and compare diagnostic accuracy.*

Results: *A total of 16 radiologists completed the test set, including nine (56.3%) breast radiologists and seven (43.8%) general radiologists. Without AI assistance, the overall sensitivity and specificity in breast cancer detection were 73.3% and 89.9%, respectively. With AI assistance, both metrics improved significantly to 80.7% ($p = 0.007$) and 94.3% ($p < 0.001$), respectively. Subgroup analysis showed that breast radiologists demonstrated improved specificity from 87.6% to 92.6% ($p < 0.001$), while general radiologists acquired more sensitivity from 54.0% to 66.7% ($p < 0.001$) with the use of AI.*

Conclusion: *AI assistance significantly improved the diagnostic accuracy of breast radiologists and general radiologists in screening mammography for breast cancer detection.*

Key Words: *Artificial intelligence; Breast neoplasms; Mammography; Mass screening*

Correspondence: *Dr PL Lam, Department of Diagnostic and Interventional Radiology, Kwong Wah Hospital, Hong Kong SAR, China*
Email: lpl404@ha.org.hk

Submitted: 29 August 2024; Accepted: 9 December 2024. This version may differ from the final version when published in an issue.

Contributors: DF, EWSF and EPYF designed the study. DF, EWSF, PHL, KMK, LKMW, WSM, WPC, WIS, WKN, GCYC, LWL and EPYF acquired the data. PLL, DF, EHC, EWSF and EPYF analysed the data. PLL drafted the manuscript. All authors critically revised the manuscript for important intellectual content. All authors had full access to the data, contributed to the study, approved the final version for publication, and take responsibility for its accuracy and integrity.

Conflicts of Interest: All authors have disclosed no conflicts of interest.

Funding/Support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability: All data generated or analysed during the present study are available from the corresponding author on reasonable request.

Ethics Approval: This research was approved by the Central Institutional Review Board of Hospital Authority, Hong Kong (Ref No.: CIRB-2024-074-5). The requirement for informed consent from patients was waived by the Board due to the retrospective nature of the research.

Acknowledgement: The authors thank the Well Women Clinic of Tung Wah Group of Hospitals and radiologists from the Department of Diagnostic and Interventional Radiology of Kwong Wah Hospital for their support of this study.

Supplementary Material: The supplementary material was provided by the authors and some information may not have been peer reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by the Hong Kong College of Radiologists. The Hong Kong College of Radiologists disclaims all liability and responsibility arising from any reliance placed on the content. To view the file, please visit the journal online (https://doi.org/10.12809/hkjr2417896).

中文摘要

利用人工智能輔助乳房X光檢查提高乳癌篩檢檢出率：一項多位閱片者回顧性研究

林栢麟、范德信、陳恩灝、霍泳珊、李璧希、郭勁明、黃嘉敏、麥詠詩、張偉彬、薛詠妍、吳詠淇、陳頌恩、羅麗雲、馮寶恩

引言：本研究旨在比較香港放射科醫生在乳房X光檢查篩檢乳癌時應用和不應用人工智能輔助兩種情況下的表現，並對乳腺放射科醫生和一般放射科醫生進行亞組比較。

方法：本研究為單中心多位閱片者回顧性研究。研究採用篩檢乳房X光攝影測試集（HKPERFORMS），此測試集包含於2009年12月至2023年12月期間採集的80例陰性或良性乳腺X光攝影影像及36例經病理證實為乳癌的乳腺X光攝影影像。研究於2023年12月至2024年4月期間評估了放射科醫生在應用和不應用商用人工智能輔助工具（Lunit INSIGHT MMG）兩種情況下的表現。兩次閱片之間相隔4週洗脫期。研究終點包括乳腺X光攝影檢測乳癌的敏感性和特異性。我們採用Obuchowski-Rockette模型評估及比較診斷準確性。

結果：共有16位放射科醫生完成了測試集，其中9名（56.3%）為乳腺放射科醫生，7名（43.8%）為一般放射科醫生。在未使用人工智能輔助的情況下，乳癌檢測的整體敏感性和特異性分別為73.3%和89.9%。使用人工智能輔助後，這兩項指標均顯著提高，分別達到80.7%（$p = 0.007$）和94.3%（$p < 0.001$）。亞組分析顯示，使用人工智能後，乳腺放射科醫生的特異性從87.6%提高到92.6%（$p < 0.001$），而一般放射科醫生的敏感性則從54.0%提高到66.7%（$p < 0.001$）。

結論：人工能輔助顯著提高了乳腺放射科醫生和一般放射科醫師在乳癌篩檢中應用乳房X光攝影的診斷準確率。

INTRODUCTION

In Hong Kong, breast cancer has been the most common malignancy among the female population since the early 1990s, with increasing incidence every year. It accounted for over a quarter (28.9%) of new cancer cases in 2023.[1] It was also the third leading cause of cancer deaths in women.[1] Fortunately, breast cancer can be curable in its early stages, with over 95% 5-year survival for patients with stage I disease.[2] Previous randomised controlled trials and meta-analyses have demonstrated the efficacy of screening mammography in detecting early-stage tumours and reducing breast cancer–related deaths.[3-6]

Breast screening programmes have been established in multiple developed economies worldwide. In Western countries, the American Cancer Society recommends that women consider annual mammography screening starting at the age of 40 years,[7] whereas in the United Kingdom, the National Health Service offers breast screening every 3 years for women aged between 50 and 71 years.[8] In Asian countries, such as Japan,[9] South Korea[10] and Singapore,[11] breast screening programmes have been in place for over a decade. In Hong Kong, the Centre for Health Protection recommends that women in the general population aged 44 to 69 years with an average risk of breast cancer consider mammography screening every 2 years.[12] Together with increased advocacy from non-profit organisations, which have heightened disease awareness among the public, screening mammography has become more popular.[13]

Like most tests, the diagnostic accuracy of screening mammography is not absolute. Sensitivity and specificity in breast cancer detection range between approximately 50% to 80% and about 80% to 90%, respectively, in the literature.[14-17] False-positive results lead to additional workup and the associated anxiety in patients, while false-negative results can delay treatment and worsen prognosis.[14]

Recent advancements in machine learning have led to the increased use of artificial intelligence (AI) in clinical radiology. Some studies, mainly conducted in Western countries, have shown promising results in employing AI-based tools to improve the diagnostic accuracy of screening mammography.[18-21]

AI-supported software has become more accessible and commercially available. To the best of our knowledge, there are no published studies evaluating the diagnostic performance of screening mammography with AI assistance in Hong Kong. The lack of established evidence in our local population could be a hurdle for radiologists to consider AI-assisted screening mammography. The external validity of previous research poses a major concern. Screening mammography tests employed in studies performed in Western countries were mainly selected from Caucasian patients.[22] Asian women, on the other hand, generally have different breast composition, with a higher prevalence of dense breasts. This can obscure abnormalities on mammograms, limiting the detection of breast cancer and reducing diagnostic accuracy.[23-25] Investigations on how AI-based tools could facilitate screening mammography using test sets derived from a local Asian population could bridge this data gap.

This study aimed to compare the performance of radiologists in screening mammography to detect breast cancer with and without AI assistance in the local population. Subgroup comparisons between breast radiologists and general radiologists were also performed.

METHODS

We developed a test set, the Hong Kong Personal Performance in Mammographic Screening Scheme (HKPERFORMS), to evaluate the diagnostic accuracy of radiologists in detecting breast cancer in the local Asian population with and without AI assistance. The test set comprised mammograms retrospectively selected from Asian adult female patients aged 40 years or above who underwent breast screening in a single well-woman clinic from December 2009 to December 2023. Exclusion criteria included symptomatic patients (e.g., those with a palpable breast mass), pregnant patients, and those with a history of breast implant augmentation surgery.

All studies in HKPERFORMS were two-dimensional (2D) screening full-field digital mammograms with standard craniocaudal and mediolateral oblique views. There were 80 mammograms showing negative or benign findings, confirmed as stable on subsequent mammographic follow-up for at least 3 years as assessed by breast radiologists recognised by the Hong Kong College of Radiologists (HKCR). There were 36 mammograms with pathologically proven breast cancer, including invasive ductal carcinoma, invasive lobular carcinoma, and ductal carcinoma in situ. Their mammographic appearances included mass (n = 21, 58.3%), calcifications (n = 6, 16.7%), architectural distortion (n = 5, 13.9%), and asymmetry (n = 4, 11.1%). The mammograms in the test set (n = 116) included breasts of varying densities: extremely dense (13.8%), heterogeneously dense (72.4%), scattered areas of fibroglandular density (12.1%), and almost entirely fatty (1.7%) [Figure 1]. Patient information and identifiers, such as name and age, were anonymised before compiled into the HKPERFORMS test set (Figure 2).

Reader Assessment

This was a single-centre study. Radiologists were recruited from an acute general hospital with subspecialty training in breast radiology accredited by the HKCR. They included breast radiologists and general radiologists. Breast radiologists were defined as radiologists with at least 3 months of subspecialty training recognised by the HKCR, or post-fellowship breast radiology training, and at least 500 screening mammograms read in the past year. General radiologists were defined as HKCR members or fellows actively practising in clinical radiology, but without dedicated subspecialty training in breast radiology.

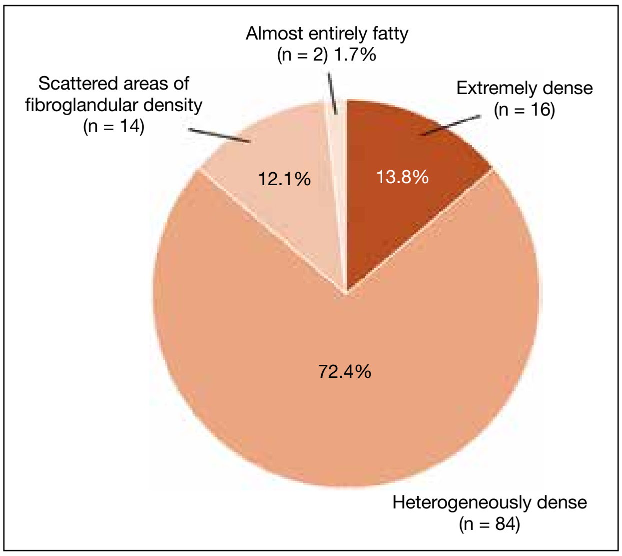

Figure 1. Proportion of breast densities in mammograms of the test set (n = 116).

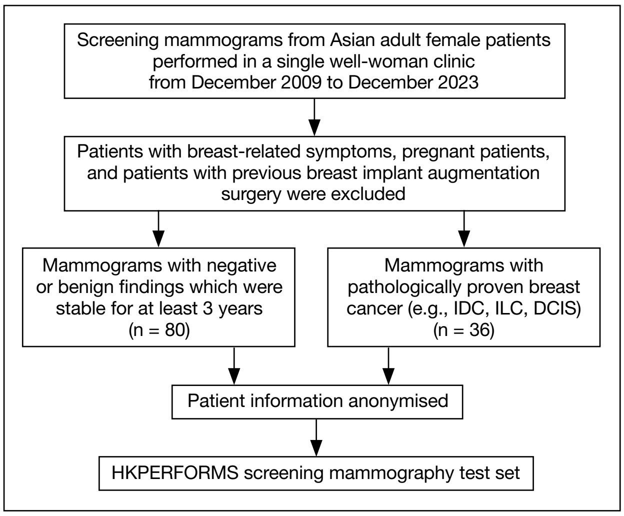

Figure 2. Development of the test set.
Abbreviations: DCIS = ductal carcinoma in situ; HKPERFORMS = Hong Kong Personal Performance in Mammographic Screening Scheme; IDC = invasive ductal carcinoma; ILC = invasive lobular carcinoma.

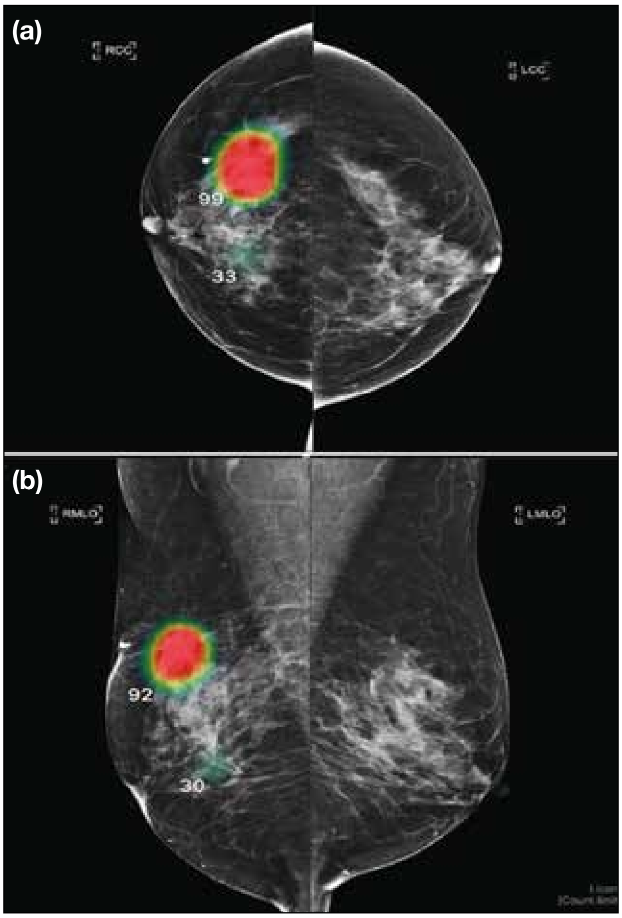

Figure 3. Screening mammogram of an adult female patient in (a) craniocaudal and (b) mediolateral oblique views with application of the artificial intelligence (AI)–based tool. A right breast upper outer quadrant mass has been colour-coded red, while central inner architectural distortion has been colour-coded green, indicating AI-perceived abnormal regions with different degrees of suspicion. Predicted probabilities of malignancy are also provided numerically. Subsequent biopsies of both lesions confirmed invasive ductal carcinoma of the right breast. The numbers on the images represent predicted probabilities of malignancy.
Abbreviations: LCC = left breast craniocaudal; LMLO = left breast mediolateral oblique; RCC = right breast craniocaudal; RMLO = right breast mediolateral oblique.

The recruited radiologists were blinded to all patient information and identifiers in the HKPERFORMS screening mammography test set. They assessed the mammograms under standardised conditions using dedicated software (Selenia Dimensions version 1.11; Hologic, Bedford [MA], US) with diagnostic-quality monitors (Coronis Uniti MDMC 12133; Barco, Kortrijk, Belgium) in accordance with department standards. Readers documented their screening results digitally (SurveyMonkey; SurveyMonkey, San Mateo [CA], US). Data to be entered included breast density, laterality, quadrant, depth, and presence or absence of architectural distortion if an abnormality was identified. Respondents were required to classify each study as benign or suspicious for malignancy.

All radiologists assessed the HKPERFORMS test set twice. In the first reading, they read the screening mammograms without AI assistance. In the second reading, additional data were provided by a commercially available AI-based tool (INSIGHT MMG version 1.1.7.3; Lunit, Seoul, South Korea),[26] which automatically highlighted regions perceived as abnormal with a colour-coded heatmap indicating the degree of suspicion. A predicted probability of malignancy was also presented numerically (Figure 3). Both pre– and post–AI-processed mammograms were available during the second reading. Respondents were instructed to record their screening results after reviewing all images. They were at liberty to follow or disregard the AI-based assessment entirely. A washout period of at least 4 weeks was observed between the two readings. The orders of the screening mammograms in the test set were different and randomised across the two sittings. Respondents who did not complete either reading were excluded from the study (Figure 4).

Background information of the recruited radiologists, including prior subspecialty training in breast radiology and experience in reporting breast imaging, was collected. All responses submitted electronically were

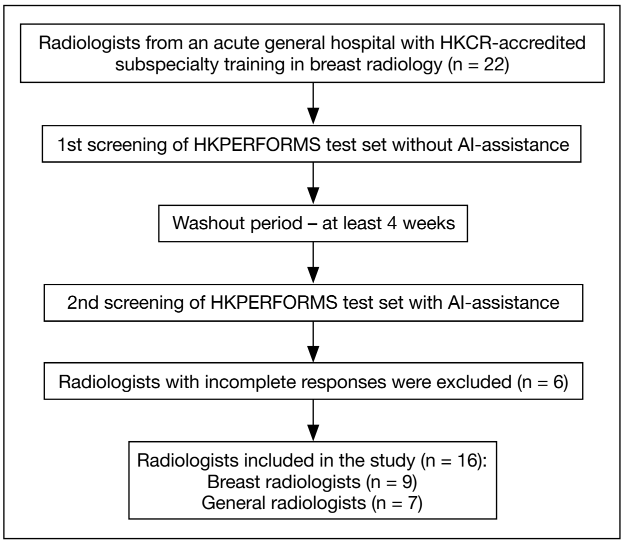

Figure 4. Assessment of screening mammograms in the test set (n = 22).
Abbreviations: AI = artificial intelligence; HKCR = Hong Kong College of Radiologists; HKPERFORMS = Hong Kong Personal Performance in Mammographic Screening Scheme.

anonymised and a random computer-generated number was assigned to each radiologist. Researchers were blinded to the identity of the respondents.

Statistical Analysis

Statistical analysis was performed using R (macOS version 4.4.1; R Core Team, Vienna, Austria).[27] Study endpoints of diagnostic accuracy included sensitivity and specificity in the mammographic detection of breast cancer. The Obuchowski–Rockette model was used to estimate and compare diagnostic accuracy.[28] A p value of < 0.05 was considered statistically significant.

This manuscript was prepared in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines.

RESULTS

Overall Performance

A total of 22 radiologists were invited to participate in this study; six respondents who did not complete the HKPERFORMS screening mammography test set were excluded, resulting in 16 radiologists completing the test set (Figure 4). Without AI assistance, the mean sensitivity and specificity for detecting breast cancer were 73.3% and 89.9%, respectively. With AI assistance, there was significant improvement in diagnostic accuracy, with the mean sensitivity and specificity increasing to 80.7% (p = 0.007) and

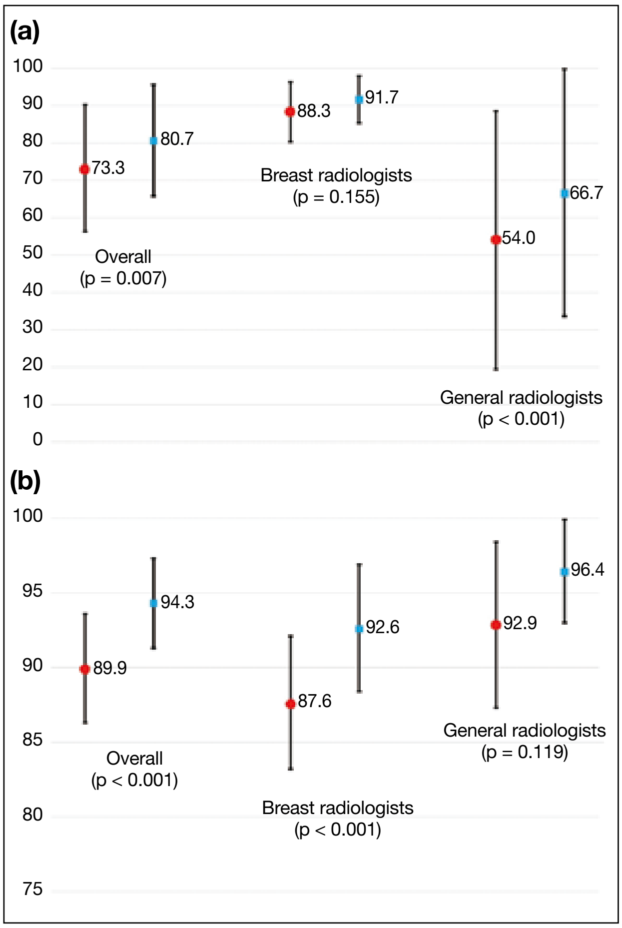

Figure 5. Dot plots and 95% confidence intervals showing (a) sensitivity and (b) specificity of all radiologists, breast radiologists, and general radiologists in screening mammography for breast cancer detection without (red circles) and with (blue squares) artificial intelligence assistance.

94.3% (p < 0.001), respectively (Figure 5 and online supplementary Table).

Subgroup Analysis

Among the respondents, nine (56.3%) were breast radiologists and seven (43.8%) were general radiologists. The experience of the breast radiologists is shown in Figure 6. Without AI assistance, the mean sensitivity of the breast radiologists (88.3%) was significantly higher than that of the general radiologists (54.0%) in identifying breast cancer (p = 0.017). There was no significant difference in the mean specificity between the two groups (breast radiologists: 87.6% vs. general radiologists: 92.9%; p = 0.051). Using the AI-based tool, there was significant improvement in the specificity of the breast radiologists (from 87.6% to 92.6%; p < 0.001) and the sensitivity of the general radiologists (from 54.0% to 66.7%; p < 0.001). No significant changes in

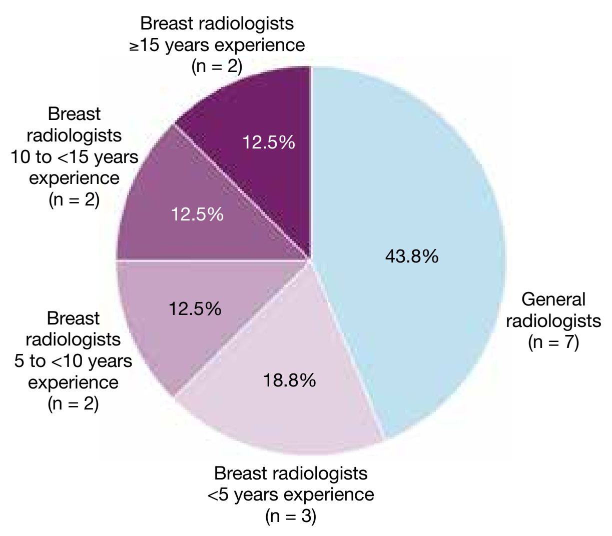

Figure 6. Proportion of breast radiologists and general radiologists included in the study (n = 16).

the sensitivity of breast radiologists and the specificity of general radiologists were observed after using the AI-based tool (Figure 5 and online supplementary Table).

DISCUSSION

Diagnostic Accuracy Without Artificial Intelligence Assistance

Without assistance from the AI-based tool, the diagnostic accuracy of the breast radiologists included in this study was comparable to figures reported in the literature, with both sensitivity and specificity exceeding 85%.[15-17] In contrast, general radiologists were less likely to detect breast malignancy, with a sensitivity of about 54%. Screening tests with low sensitivity lead to a higher proportion of false-negative results, potentially leading to false reassurance and missed opportunities for early diagnosis and treatment.[14] These findings highlight the importance of dedicated training in breast radiology.[29,30] The HKCR Mammography Statement outlines the standards for radiologists involved in screening. These include a minimum of 3 months of subspecialty training in breast radiology, interpretation of at least 500 screening mammograms annually, and ongoing participation in continuing medical education and multidisciplinary meetings.[31]

Improved Performance with Artificial Intelligence Assistance

There were significant improvements in overall sensitivity and specificity in breast cancer detection when radiologists in this study performed AI-assisted screening mammography. This echoed previous studies which demonstrated improved diagnostic accuracy in AI-assisted mammography readings.[18-21] Subgroup analysis further showed that the benefits of AI assistance differed between general radiologists and breast radiologists.

For general radiologists, there was significant improvement in sensitivity, from approximately 50% when screening unaided to over 65% with the use of AI-based tool. A previous study also demonstrated reduced variability in screening results and increased inter-reader reliability with AI assistance.[32] This indicates that utilising AI could yield more expertise-independent results. AI could act as an extra pair of eyes. Radiologists could refer to colour-coded heatmaps generated by AI-based software after initial mammography assessment to reduce the probability of missing breast cancer.[26]

Among the breast radiologists, there was improvement in specificity, while sensitivity in detecting breast cancer remained similar with and without AI assistance. The crux of screening lies in striking a balance between sensitivity and specificity. Tests with high sensitivity but low specificity may lead to over-investigation, resulting in unnecessary stress and interventions for patients.[14] While the specificity of the breast radiologists in breast cancer detection was satisfactory without AI assistance, it improved from over 85% to over 90% with the use of the AI-based tool without compromising sensitivity. Increased specificity in screening mammography would reduce call-back rates, avoid unwarranted workups for patients, and decrease the workload for radiologists.[20,33] A study by Raya-Povedano et al[34] revealed a reduction of over 70% in radiologists' workload following the implementation of AI-based strategies. Additionally, AI tools could be helpful to prioritise screening mammograms with suspected malignancy. Such abnormal studies could be flagged for earlier reporting by radiologists, expediting subsequent workup and treatment. Furthermore, placing flagged studies at the beginning of a screening session could minimise the risk of missed breast cancers due to reader fatigue. With the burgeoning demand for screening mammography in Hong Kong, AI-based tools could potentially alleviate the stress faced by radiologists.

Limitations

The HKPERFORMS test set was enriched with abnormal mammograms, and the proportion of cases with biopsy-proven breast cancer was not representative of routine screening practice or the general population.[1,2]

Although respondents were instructed to interpret each individual mammogram as an independent screening case, their diagnostic accuracy might have been negatively influenced by the study design. Second, test sets used in the sittings with and without AI assistance were identical. Despite a washout period of at least 4 weeks with randomisation of the image order, radiologists might have recalled the proportion of normal to abnormal cases, potentially introducing bias in the second sitting. Third, all mammograms in the test set were 2D full-field digital mammograms. In recent years, three-dimensional mammography or digital breast tomosynthesis (DBT) has become more popular, with evidence showing improved diagnostic accuracy compared with traditional 2D mammography. Studies on AI-assisted DBT have shown non-inferior or improved sensitivity and specificity in detecting breast cancer.[35,36] Our study did not investigate DBT performance, which remains a potential direction for further research. Finally, this was a single-centre study with limited sample size. The performance and influence of AI may vary among radiologists with differing levels of experience across diverse clinical settings. Further large-scale multi-centre investigations would provide a more comprehensive assessment.

CONCLUSION

This multi-reader study evaluated the potential of AI to aid breast cancer detection using HKPERFORMS, an original screening mammography test set developed from a local Asian female population with a high incidence of dense breasts. The results demonstrated that diagnostic accuracy in screening mammography was improved across radiologists with varying levels of experience in breast radiology when supported by AI-based tools.

REFERENCES

1. Centre for Health Protection, Department of Health, Hong Kong SAR Government. Breast Cancer. 23 Jan 2026. Available from: https://www.chp.gov.hk/en/healthtopics/content/25/53.html. Accessed 2 Feb 2026.
2. Kwong A, Mang OW, Wong CH, Chau WW, Law SC; Hong Kong Breast Cancer Research Group. Breast cancer in Hong Kong, Southern China: the first population-based analysis of epidemiological characteristics, stage-specific, cancer-specific, and disease-free survival in breast cancer patients: 1997-2001. Ann Surg Oncol. 2011;18:3072-8.
3. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L, et al. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet. 2006;368:2053-60.
4. Duffy SW, Tabár L, Chen HH, Holmqvist M, Yen MF, Abdsalah S, et al. The impact of organized mammography service screening on breast carcinoma mortality in seven Swedish counties. Cancer. 2002;95:458-69.
5. Tabár L, Vitak B, Chen HH, Yen MF, Duffy SW, Smith RA. Beyond randomized controlled trials: organized mammographic screening substantially reduces breast carcinoma mortality. Cancer. 2001;91:1724-31.
6. Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL. Efficacy of screening mammography. A meta-analysis. JAMA. 1995;273:149-54.
7. American Cancer Society. American Cancer Society recommendations for the early detection of breast cancer. Available from: https://www.cancer.org/cancer/types/breast-cancer/screening-tests-and-early-detection/american-cancer-society-recommendations-for-the-early-detection-of-breast-cancer.html. Accessed 20 Aug 2024.
8. National Health Service, Department of Health and Social Care, United Kingdom Government. Breast screening (mammogram). Available from: https://www.nhs.uk/conditions/breast-screening-mammogram/. Accessed 20 Aug 2024.
9. Hamashima CC, Hattori M, Honjo S, Kasahara Y, Katayama T, Nakai M, et al. The Japanese guidelines for breast cancer screening. Jpn J Clin Oncol. 2016;46:482-92.
10. Shin DW, Yu J, Cho J, Lee SK, Jung JH, Han K, et al. Breast cancer screening disparities between women with and without disabilities: a national database study in South Korea. Cancer. 2020;126:1522-9.
11. Loy EY, Molinar D, Chow KY, Fock C. National Breast Cancer Screening Programme, Singapore: evaluation of participation and performance indicators. J Med Screen. 2015;22:194-200.
12. Cancer Expert Working Group on Cancer Prevention and Screening, Centre for Health Protection, Department of Health, Hong Kong SAR Government. Recommendations on Prevention and Screening for Breast Cancer for Health Professionals. June 2020. Available from: https://www.chp.gov.hk/files/pdf/breast_cancer_professional_hp.pdf. Accessed 20 Aug 2024.
13. Hong Kong Breast Cancer Foundation. What is breast cancer. Available from: https://www.hkbcf.org/en/breast_cancer/main/422/. Accessed 20 Aug 2024.
14. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. Br J Cancer. 2013;108:2205-40.
15. Hollingsworth AB. Redefining the sensitivity of screening mammography: a review. Am J Surg. 2019;218:411-8.
16. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Likelihood ratios for modern screening mammography. Risk of breast cancer based on age and mammographic interpretation. JAMA. 1996;276:39-43.
17. Lehman CD, Wellman RD, Buist DS, Kerlikowske K, Tosteson AN, Miglioretti DL, et al. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. JAMA Intern Med. 2015;175:1828-37.
18. Dembrower K, Crippa A, Colón E, Eklund M, Strand F; ScreenTrustCAD Trial Consortium. Artificial intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study. Lancet Digit Health. 2023;5:e703-11.
19. Lång K, Josefsson V, Larsson AM, Larsson S, Högberg C, Sartor H, et al. Artificial intelligence–supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. Lancet Oncol. 2023;24:936-44.
20. Lauritzen AD, Lillholm M, Lynge E, Nielsen M, Karssemeijer N, Vejborg I. Early indicators of the impact of using AI in mammography screening for breast cancer. Radiology. 2024;311:e232479.

21. Ng AY, Oberije CJ, Ambrózay É, Szabó E, Serfőző O, Karpati E, et al. Prospective implementation of AI-assisted screen reading to improve early detection of breast cancer. Nat Med. 2023;29:3044-9.
22. Chen Y, Gale A. Performance assessment using standardized data sets: the PERFORMS scheme in breast screening and other domains. In: Samei E, Krupinski EA, editors. The Handbook of Medical Image Perception and Techniques. 2nd ed. Cambridge, England: Cambridge University Press; 2018: 328-42.
23. Bao C, Shen J, Zhang Y, Zhang Y, Wei W, Wang Z, et al. Evaluation of an artificial intelligence support system for breast cancer screening in Chinese people based on mammogram. Cancer Med. 2023;12:3718-26.
24. Yan H, Ren W, Jia M, Xue P, Li Z, Zhang S, et al. Breast cancer risk factors and mammographic density among 12518 average-risk women in rural China. BMC Cancer. 2023;23:952.
25. Jackson VP, Hendrick RE, Feig SA, Kopans DB. Imaging of the radiographically dense breast. Radiology. 1993;188:297-301.
26. 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:e138-48.
27. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2020.
28. Hillis SL, Obuchowski NA, Berbaum KS. Power estimation for multireader ROC methods: an updated and unified approach. Acad Radiol. 2011;18:129-42.
29. Trieu PD, Lewis SJ, Li T, Ho K, Wong DJ, Tran OT, et al. Improving radiologist's ability in identifying particular abnormal lesions on mammograms through training test set with immediate feedback. Sci Rep. 2021;11:9899.
30. Miglioretti DL, Gard CC, Carney PA, Onega TL, Buist DS, Sickles EA, et al. When radiologists perform best: the learning curve in screening mammogram interpretation. Radiology. 2009;253:632-40.
31. Hong Kong College of Radiologists. Hong Kong College of Radiologists Mammography Statement. Revised 25 August 2015. Available from: https://www.hkcr.org/templates/OS03C00336/case/lop/HKCR%20Mammography%20Statement_rev20150825.pdf. Accessed 20 Aug 2024.
32. Pacilè S, Lopez J, Chone P, Bertinotti T, Grouin JM, Fillard P. Improving breast cancer detection accuracy of mammography with the concurrent use of an artificial intelligence tool. Radiol Artif Intell. 2020;2:e190208.
33. Kim YS, Jang MJ, Lee SH, Kim SY, Ha SM, Kwon BR, et al. Use of artificial intelligence for reducing unnecessary recalls at screening mammography: a simulation study. Korean J Radiol. 2022;23:1241-50.
34. 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;300:57-65.
35. Goldberg JE, Reig B, Lewin AA, Gao Y, Heacock L, Heller SL, et al. New horizons: artificial intelligence for digital breast tomosynthesis. Radiographics. 2022;43:e220060.
36. Park EK, Kwak S, Lee W, Choi JS, Kooi T, Kim EK. Impact of AI for digital breast tomosynthesis on breast cancer detection and interpretation time. Radiol Artif Intell. 2024;6:e230318.

