KRÓL, Agata, KWATERSKA, Katarzyna, KUTYŁOWSKI, Karol and ŁUCKIEWICZ, Paweł. The role of artificial intelligence in breast cancer screening as a supportive tool for radiologists. Quality in Sport. 2025;43:61336. eISSN 2450-3118.
https://doi.org/10.12775/QS.2025.43.61336
https://apcz.umk.pl/QS/article/view/61336

The journal has been awarded 20 points in the parametric evaluation by the Ministry of Higher Education and Science of Poland. This is according to the Annex to the announcement of the Minister of Higher Education and Science dated 05.01.2024, No. 32553. The journal has a Unique Identifier: 201398. Scientific disciplines assigned: Economics and Finance (Field of Social Sciences); Management and Quality Sciences (Field of Social Sciences).
Punkty Ministerialne z 2019 - aktualny rok 20 punktów. Załącznik do komunikatu Ministra Szkolnictwa Wyższego i Nauki z dnia 05.01.2024 Lp. 32553. Posiada Unikatowy Identyfikator Czasopisma: 201398. Przypisane dyscypliny naukowe: Ekonomia i finanse (Dziedzina nauk społecznych); Nauki o zarządzaniu i jakości (Dziedzina nauk społecznych).
The authors declare that there is no conflict of interest regarding the publication of this paper.
Received: 22.05.2025. Revised: 05.07.2025. Accepted: 05.07.2025. Published: 08.07.2025.

The role of artificial intelligence in breast cancer screening as a supportive tool for radiologists

Authors:

Agata Król – corresponding author
Dr Anna Gostynska Wolski Hospital, Independent Public Health Care Institution,
Marcina Kasprzaka 17, 01-211 Warsaw, Poland
https://orcid.org/0009-0006-1967-9684
e-mail: agatakrol517@gmail.com

Katarzyna Kwaterska
Mazovian Brodnowski Hospital,
Kondratowicza 8, 03-242 Warsaw, Poland
https://orcid.org/0009-0000-3157-5438
e-mail: katarzyna.kwaterska90@gmail.com

Karol Kutyłowski
Our Lady of Perpetual Help Hospital, Independent Health Care Institution,
Gdyńska 1/3 05-200 Wołomin, Poland
https://orcid.org/0009-0000-8592-1711
e-mail: kutylowski.karol@gmail.com

Paweł Łuckiewicz
Our Lady of Perpetual Help Hospital, Independent Health Care Institution,
Gdyńska 1/3 05-200 Wołomin, Poland
https://orcid.org/0009-0001-2550-1173
e-mail: pawel.luckiewicz@icloud.com

ABSTRACT

Introduction

Breast cancer is the most prevalent cancer in Polish women and worldwide. Still remains the second cause of cancer deaths in females. There are several risk factors, both modifiable and non-modifiable. Despite the vast majority is an adenocarcinoma, multiple molecular variants are distinguished with different management and prognosis. Early diagnosis is crucial for cancer-related burden and mortality reduction. For this reason several countries have implemented breast cancer screening programme. This imaging has some limitations associated with interpretative difficulties, huge number of examinations and shortage of staff. Artificial intelligence has been proposed for breast screening, both in assistance to radiologists and independently in the future.

Purpose and methods

The aim of this review is to provide an overview on clinical use of AI in breast cancer screening with focusing on impact on results quality and workload. The evaluation has been focused on the degree of AI participation in diagnostic process and its impact on screening or human performance.

Results

Analysed studies presented various ways of AI application in breast screening. Authors generally focused on cancer detection rate, false positive or negative results, recall rate and the workload involvement. Some improvements, especially in cancer detection rate have been noticed with significant workload reduction. Difficulties, possible errors and people's opinion were also highlighted.

Conclussion
AI algorithms find their potential application in breast cancer screening, mainly as a supportive tool for radiologists. However, in this moment, standalone AI interpretation does not seem to be accurate and safe enough, especially due to the lack of standarization and not fully convinced population. Future AI systems should take into account all relevant patients' data for better assessment of the examination. More prospective studies are needed to improve a knowledge about AI possibilities.

Keywords: breast cancer, artificial intelligence, screening, mammography

INTRODUCTION

Breast cancer is the most prevalent cancer in Polish women and worldwide. Unfortunately, due to its malignancy, still remains the second cause of cancer deaths in females behind the lung cancer. According to the World Health Organization's data for 2022 - 2,3 million women were diagnosed with breast cancer and 670000 died globally. Women at any age after puberty may be affected with the incidence increasing in later life. Multiple factors are associated with increased risk of incidence, both modifiable and non-modifiable. The first group consists of obesity, alcohol consumption, smoking, physical inacticity, exogenous hormones like contraceptive pills or hormone replacement therapy. Unchangeable risk factors include an increasing age, which is the major factor and other such as genetic predispositions – BRCA1 and BRCA2 mutations or early menarche, having fewer children and no breast feeding. The characterization of cancers is based on morphology, immunohistochemistry, molecular subtype with staging due to tumour size, lymph node involvement and distant metastases. The main type of breast cancer is adenocarcinoma, but the disease course is very heterogenous due to multiple variants of molecular profiles. There are multiple divisions of types and subtypes, so we can distinguish three main subtypes depending on hormone receptor – estrogen (ER), progesterone (PR) and HER2 status, two according to cancer deriving cells – lobular or ductal and another in accordance with the severity of neoplastic process – in situ or invasive. Therefore, histopathological diagnosis is essential to determine correct management, prognosis and follow-up. Early diagnosis is crucial for cancer-related burden and mortality reduction and creates an opportunity for succesfull treatment and prolonged survival.

Taking into account all of the above facts and recognizing breast cancer as the global issue, several countries have implemented screening programmes addressed to asymptomatic women in particular age. The most common breast cancer screening test, also used in Poland, is a digital mammography in two projections – craniocaudal and mediolateral oblique. The investigation is aimed at Polish women aged 45-74 years at two-year screening intervals. Mammograms interpretation is challenging, even to the expert eye with the neccessity to balance between cancer misdiagnosis and overdiagnosis with too many false positives. Additionaly, dense breasts imaging has a limited sensitivity with relatively high number of overdetection. To minimise gaps and mistakes in diagnostic progress, European guidelines recommend a double reading that relies on the second assessment of primary diagnosed as normal mammogram by the other independent radiologist. If any abnormality is detected, during first or second look, further diagnosis is immediately performed, known as 'recall'. Getting called back does not mean that woman has breast cancer, but is neccessary to explain unclear imaging results and fewer than 1 in 10 patients are diagnosed with cancer. Additional imaging such as diagnostic mammogram or breast ultrasound may be performed. In case of suspicious finding, a biopsy is reccommended.

Screening programme is directed to the broad group and because of fortunately low prevalence of breast cancer, most mammograms are normal. This fact increases false negative results – radiologists are exposed to the risk of missing neoplastic signs. Otherwise they are qualifying normal images as suspect and deciding to recall healthy women for assessment. Implications of false positive results with consecutive further complicated diagnostic process creates a psychical burden for patients and needless costs for country. Additionally, workload associated with the amount of screens to asses prohibits radiologists focusing on demanding cases.

Nowadays, the concomitance of the increasing demand for imaging results and the radiologists' understaffing creates a challenge for scientists to discover new solutions. As with many areas of life, so too in healthcare, artificial intelligence (AI) is recognized to play a significant and hopeful role. Multiple applications in medicine have been found, especially in radiology, including interpretative and non-interpretative tasks. The purpose of AI is considered in breast imaging, chest radiograms to differentiate pneumothorax, pleural effusion, lung lesions etc. and with promising function in emergency radiology for rapid detection of sudden states such as pulmonary embolism, bowel obstruction, acute stroke or intracranial hemorrhage. Another areas may be supported by algorithms like workflow prioritization and optimisation. Currently, as technology advances, deep learning AI systems based on convolutional neural network are becaming increasingly accurate and most commonly used for image segmentation and classification. Global companies are conducting research of novel AI based software for imaging, including mammography. There are 16 available products for European market, directed to asymptomatic breast cancer screening population. Their effective implementation is possible since digital mammography have become a diagnostic standard in many countries.

PURPOSE AND METHODS

The aim of this review is to provide an overview on clinical use of AI in breast cancer screening with focusing on impact on results quality and workload. Some relevant factors, which determine patients safety are neccessary in risk/reward ratio assessment.

For the study, the most recent publications posted in the PubMed database and sources from Google Scholar browser were analysed, from 2020 to 2025. All articles were in English. The main assumption was to find prospective trials, which amount is limited and subsequently retrospective and meta-analyses. This review focuses on results depending on the degree of AI participation in diagnostic process. The clinical value assessment is based on impact on screening or human performance such as sensitivity, specificity, overdiagnosis, recall rate, cancer detection rate and the workload decrease. Some other aspects such as radiologists' and general population's opinion, and also possible diagnostic errors have been taken into account in the analysis.

The point is to choose the threshold for the best sensitivity in comparison to the highest specificity. The area under the receiving operator characteristics (ROC) curve (AUC) value summarize the test's ability to distinguish disease. Taking this fact into account, many results is based on AUC to compare two or more studies. A value of 1.0 indicates excellent discrimination, while 0,5 is equivalent to random chance. The acceptable diagnostic accuracy aim is at AUC of 0,8 or more.

RESULTS

Analysed studies presented various ways of AI application in breast screening. Starting from its role as a reader aid, a kind of radiologist's assistant through concurrent decision support to an independent standalone reader. Several proposals for integrating AI in double-reader screening programmes have emerged. Some authors suggested its potential for exams triage, checking AI efficiency in negative or low risk mammograms detection without need for two doctors engaging. Otherwise, they have compared cases when AI replaced one radiologist in double-reading process to current diagnostic standard or accuracy of triple reading by two radiologists plus AI. Authors generally focused on cancer detection rate, false positive or false negative results, changes in recall rate and the workload involvement. Difficulties and possible errors have been also highlighted, taking into account population willingness.

One of the most common comercially available algorithm – Transpara – is based on categorization examinations into 10 groups depending on AI score from 1 to 10. An AI score of 1 indicated a low risk of abnormal fndings and 10 indicated high risk. [15] The majority of AI systems relies on similar principles.

Cancer detection rate

Supportive role of AI in diagnosis should come down to the increased detection of clinically relevant cancers while low percentage of false positive results and indolent cancers. Screen-detected cancer was defined as breast cancer diagnosed after a recall and within 6 months after the screening examination.

The Mammography Screening with Artificial Intelligence trial (MASAI) was a randomised controlled trial, in which the true effect of cooperation radiologist-AI and its influence on medical decision making could be studied [6]. Using AI-supported mammography resulted in 29% increase in cancer detection compared with standard double-reading without AI. Furthermore, higher detection of small, lymph-node negative, invasive cancers has been noted with AI, what gives hope for downstaging breast cancer by earlier detection and consequently treatment implementation with the most relevant for patient – better prognosis. Increased detection creates a risk of overdiagnosis and true, the MASAI trial confirmed more in situ cancers using AI. Roughly half of the extra detected ductal carcinoma in situ were of nuclear grade III and clinically relevant due to greater aggressiveness, but other half with intermediate risk was considered false positives. Dembrower et al. pointed out that double reading by one radiologist plus AI caused a 21% (868/4104) increase in the number of examinations with abnormal interpretation [10], which proved a human reader and AI synergism in better sensitivity for detecting breast cancers in mammograms. The AI score is adequately translated into further diagnosis, which was proved by the fact that 92.7% of the screen-detected and 40.0% of the interval cancers had an AI score of 10, representing the highest risk of breast cancer [15]. In addition, AI-based screening sensitivity was non-inferior to radiologist screening (p=0,02). The most approximate results concerned BI-RADS density 4 with slightly reduced AI sensitivity between 0.6 and 1.8 percentage points for BI-RADS densities 1–3 [14]. In terms of specificity, an increasing trends between 0.3 and 0.6 percentage points were achieved across all

BIRADS densities. Elías-Cabot et al. reported 2,4 more rates per 1000 detected while digital mammography reading with AI system [16]. Chang et al. in their prospective AI-STREAM study noticed a significantly higher cancer detection rate by 13.8% for breast radiologists using AI (n = 140 [5.70‰]) compared to those without AI (n = 123 [5.01‰]; $p < 0.001$) [27]. However, this study was based on comparison to single reading by radiologists, which is not a recommended procedure in Europe. Branco et al. determined in their study a greater specificity of diagnostics associated with AI 0.85 vs. without AI 0.67 ($p < 0.001$) and simultaneously decreased false-negative rates from 8,6% to 18% when AI algorithms were added to radiologists' evaluation of mammograms [19]. The overall assessment of specificity to sensitivity illustarted by the area under the receiving operator characteristics curve (AUC) has been done in multiple studies with improvement in case of radiologists alone vs. with AI assistance from 0.810 to 0.881 ($p < 0.001$). The combination of AI with radiologists resulted in a higher AUC of 0,942, concomitant with significantly improved specificity and overall accuracy [18]. Another study based on External Evaluation of 3 Commercial Artificial Intelligence Algorithms as Independent mammography readers confirmed high AUC values ranging from 0.920–0.956 with high sensitivities of 67.0–81.9% keeping the same specificity. From a different angle Friedewald et al. assessed AI as a triaging tool in prospective study and revealed no cancer detection increase, remaining at similar level to radiologists. This result differ from previous retrospective studies, which indicated more cancers diagnosed using AI tools [25].

Interval cancers detection

Interval cancers are defined as breast cancers diagnosed within 24 months after a negative screening or 6–24 months after a false-positive screening result [14,15] or by other authors - those that evade detection during routine screenings and are diagnosed symptomatically between screening rounds [26]. Therefore it is an important indicator on the efficiacy of screening programme. The biennial screening delivers the interval cancer rate between 0.8 and 3.0 per 1000 screened women [30]. Unfortunately, they tend to be more aggressive and associated with poorer outcomes. Many authors suggest AI possibilities in earlier detection and delayed diagnosis avoidance, with potential to identify up to 20-30% of interval cancers and next-round screen-detected cancers. [26] There is a speculation that these algorithms may detect subtle unidentified tumor features, resulting in higher accuracy for predicting future interval cancers and next-round screen-detected cancers than mammographic density with a substantial number of interval cancers detected, previously missed by radiologists. Lång et al. [30] in their retrospective study checked the AI abilities in the assessment of the interval cancers considered false negative by doctors. AI classified these examinations in the high risk scores, what suggest its role as a helpful tool to detect suspicious lesions prospectively. They found statistically significant correlation between classification groups of interval cancer and AI risk score ($p < 0,0001$) – 9.0 for minimal visible signs and 9.7 for false negative interval cancers. Retrospectively assessed mammograms of interval cancers through AI algorithm exposed 60,6% true negative images, while 26,3% minimal signs and 13,1% false negatives. Hence, 39,4% were considered visible at the time of mamography. Taking into account a percentage of correctly localised tumors and various settings of AI recall thresholds, the potential reduction of interval cancers in screening for score 9 was 19,3% [30]. Multiple recent studies about AI clinical application, such as prospective MASAI trial, remain in awaiting for outcome „interval cancer rate", which will be assessed after a 2-year follow-up. In addition, there is a need to combine mammograms findings with traditional risk factors obtained from medical records to predict breast cancer risk with promising results and in response to minimize the interval cancers rate. Hybrid models showed the highest diagnostic performance (AUC 0.70) compared to a clinical risk-factor based model (AUC 0.62–0.67) or image-only deep-learning model (AUC 0.68) [18], so modern teamwork may be crucial for diagnostic progress.

Recall rate

The cost-effective triage tool is characterized by high specificity, which means less false positive results, which is reflected in low percentage of unnecessary recalls. The MASAI trial revealed non-significant 8% increase in recall rate in the group with AI, but most of them were true positives [6]. However Koch et al. suggested that selecting all women with an examination with an AI score of 10 for direct recall without further adjustment yield an unacceptably high recall rate around 10%, 2–4 times higher than what is currently normal [15]. On the other hand, Elías-Cabot et al. reported not statistically relevant growth in recall rate after reading with AI system (+0,4; $p=0,373$) [16]. A new perspective has been offered by the ScreenTrustCAD prospective trial, which showed a 4% lower recall rate for double reading by one radiologist plus AI thanks to the juxtaposition of mammograms, medical history and AI information in subsequent consensus discussions [10]. Lauritzen et al. reported a 25,1% reduction of false-positive screenings using AI-based interpretetion compared with radiologist screening [14], leading to lower recall rate.

Workload

As previously mentioned, AI can identify cases that require less time and predesignate cancer free-examinations as normal with no need for deeper radiologists interpretation. Several studies have assessed different thresholds efficiacy and safety. For example, setting the threshold at 5 resulted in approximately 50% workload reduction with 7% false-negatives, whereas the threshold at an AI score of 2 resulted in 17% workload reduction with 1% of missed cancers [18]. The establishment of threshold at 7 means the exclusion for reading 70% of studies and this setting did not result in a loss of sensitivity in the detection of cancers or an increase in false-positive recalls [17]. Friedewald et al. suggested the AI-modified workflow, which resulted in significantly shortened diagnostic delays - time to additional imaging was reduced by 25% to 19.1 days ($p<0.001$), while time to biopsy imaging decreased by 30% to 39.2 days compared to control group without AI impact [25]. Lauritzen et al. evaluated retrospectively that radiologists would have avoided reading images from 71585 screenings due to the exclusion of normal or suspicious mammograms, which corresponds to a 62.6% (71585 of 114421) workload reduction [14]. Also prospective MASAI trial confirmed 44,2% reduction in the screen-reading workload after AI implementation [6]. Workload reduction is reflected into more time for breast radiologists to spend on more difficult and demanding cases, minimazing the time spent on normal images analysis. Dembrower et al. estimated that replacing one radiologist with AI in a screening population of 100000 women, would save 100000 radiologist reads while increasing consensus discussions by 1562. Even if about five times prolonged consensus discussions versus independent read, the workload reduction would be substantial [10].

Difficult cases

Mammography imaging meets some interpreting difficulties. Lower sensitivity in dense breasts (with increased proportion of fibroglandular tissues) assessment is relevant problem due to the association with a four to six-fold increased risk for cancer in these breasts. Moreover, increased parenchymal density may decrease breast masses detection leading to increased number of interval cancers. Due to human eye limitations, deep-learning algorithms have been proposed as a more consistent assessment tool of breast denisity and as a more accurate breast cancer risk predictor. AI systems evaluation depends on pixel-based information embedded in mammograms, which are not perceptible to people. Authors have proved that AI convolutional neural network has a greater predictive potential than using breast density assessments by radiologists (odds ratio 4.42 vs. 1.67) with an overall accuracy of 72% [18]. Koch et al.

retrospectively found that AI identifed a large proportion of cancers in extremely dense breasts with significant correlation to AI score arisinig from the fact that all screen-detected cancers and almost half of the interval cancers among women with the highest density score had an AI score of 10 [15]. As mentioned above, women with dense breast are at higher risk to develop cancer, so depite the interpretative difficulties, their examinations need a scrupulous analysis. Gastounioti et al. proposed a risk model that incorporates age, automated breast density, mammographic features (i.e., suspicious microcalcifcations and masses) and bilateral parenchymal pattern diferences evaluated through comercially available AI systems. All of studied softwares demonstrated promising predictive performance in short-term breast cancer risk assessment (AUC=0.73–0.79) [13], giving a chance for better overall assessment in difficult cases.

AI disadvantages

However, it should be emphasized that despite the present delight in AI possibilities, algorithms are not unmistakable and have some limitations. Studies are conducted on different screening populations, using different imaging machines with various image acquisition settings, so the lack of generalization creates the risk of diagnostic errors. Mammography imaging has no standarization, so launching AI system may meet some technical problems, arising from the variation in mammographic images – different technicians, vendors and units, several proprietary post-processing software in images preparation. Current algorithms are not capable of comparing images across time and often have no possibilities to compare cranio-caudal and medio-lateral oblique views, while radiologists strictly rely on view-to-view correlation, contralateral comparison and prior images. Various ethnicity of the women is a challenging factor, which may have an impact on the algorithm sensitivity too. Previously proposed usage in dense breasts imaging is not supported by all authors. Lauritzen et al. noticed that the cancer detection rate has a downward trend with increasing BI-RADS density, possibly due to masking by fibroglandular tissue [14]. Additionaly, the amount of prospective studies including extremely dense breasts is still limited with poor external validation. The most common mistakes are false positive and false negative results. False positive proportion notably decreased with increasing detection threshold with statistically significant results – from 71,83% at 3 to 10,77% at 9. Simultaneously false negative proportion increased from 0,02% at threshold 3 to 0,12% at 9. These values transfer into a need to find a correct threshold for the best detection efficiacy. False negatives often arised from a trend that AI algorithms were more likely to miss smaller tumour than larger – median diameter of missed cancers varied from 7 to 25 mm.

Whereas false positives were misinterpreted calcifications [23]. Incorporating AI algorithms comes with unintended consequences like the detection of many in situ cancers rather than invasive cancers and particularly important – the regression of interpretative skills and decisiveness due to radiologists' overdependence on AI.

Opinion about AI

Last but not least, the opinion of people affected by this modern invention cannot be overlooked. Both women's and radiologists' perception matters for further research development and AI integration into mammography. Screening program is elective so depends on women's willingness to participate, which may be affected by potential concern about the impact of novel technologies on diagnostic process. Also population's opinion is crucial to determine boundaries within AI system is allowed to operate. Ongena et al. carried out a women opinion survey about the degree of AI involvement into mammograms reading. Nearly 80% agreed that human check is necessary in mammograms assessment and were against a fully independent use of AI-based diagnostics without the involvement of the radiologist. Majority approved that it is too premature to leave the interpretation of screening mammograms completely up to autonomously operating AI algorithms [21]. Almost half (41,7%) of asked women did not support AI application for triaging and selecting patients for second reading. The vast majority endorsed the combination of a radiologist as a first reader and an AI system as a second reader. A big dilemma remains with the responsibility for error. Some respondents with negative attitude toward AI agreed that the developer is responsible, however 45% were ambiguous as to whether the developer should be held responsible for errors. When it comes to the responsibility of the radiologist, 39% had no clear opinion and 38,7% strongly agreed or agreed. It is now held that AI may provide assistance and relief to the growing case load and increasing demands placed on radiologists. In a recent survey, a majority of radiologist respondents see AI as an opportunity to improve their practice all-round, including holding expectations for a lower error rate and interpretation time [2]. However, they have to be aware of the tendency to follow an erroneous AI suggestion, when they trust the AI system too much, especially those less experienced radiologists, who end up making changes to up to 48% of mammograms after findings provided by artificial intelligence [19]. Högberg et al. [20] studied Swedish breast radiologists view on integrating AI into their daily workflow. Most of them (80,8%) were positive about this idea, especially if they had a heavy screen-reading workload. They hoped for the solution to the scarcity of specialists and the improvement in the detection and consistency in screen-reading. Almost one-fifth (19,2%) were negative or uncertain about AI-

supported screening due to the fear of large numbers of false positives arising from misinterpreted calcifications, difficulty in interpreting AI-assessments and the risk of an increased workload related to the need to secondary look at benign findings. Radiologists were worried about the loss of competence due to a lack of continuous training on healthy mammograms, while AI would rule-out normal examinations. Moreover, AI application in triaging mammograms gave a qualitative feedback, which highlighted the
discomfort radiologists faced in interpreting AI-prioritized cases as normal, especially initially. However, after multiple rounds of exposure to AI prioritization, radiologists realized that relying on their own expertise was equally as important [25]. Doctors also found some possible difficulties for AI interpretation, which they take into account during analysis such as dense breasts, atypical masses, architectural distortion or postoperative changes. The most preferred option in their opinion remains using AI as replacement for one reader in double-reading [20]. Although AI will not ultimately replace radiologists, many in radiology believe that radiologists who work with AI will replace those that do not [1].

CONCLUSSION

AI algorithms find their potential application in breast cancer screening, mainly as a supportive tool for radiologists. It has been proven that AI systems deliver higher cancer detection rate, resulting in earlier diagnosis with more favorable prognosis. Also, it may detect unidentified tumor features with well correlation between AI score and interval cancers mammograms in retrospective studies. Recent prospective studies are awaiting for the assessment of interval cancer rate outcome and results will be presented in the close future. What is more, in case of dense breasts, AI using pixel-based analysis may point out masses, which remain imperceptible for human eye. The AI triaging maintains radiologists performance with workload reduction, particularly beneficial in case of shortage of staff. Those primarily concerned, so women and radiologists do not reject AI integration into screening process, but emphasize some boundaries that should be preserved. In this moment, standalone AI interpretation does not seem to be accurate and safe enough, especially due to the lack of standarization and generalization. Studies were conducted on limited populations with concrete ethnicity, various devices with different settings, which have a relevant impact on algorithms sensitivity. Future AI systems should take into account all relevant multi-source data such as prior images, contralateral images, data from the patient’s medical history and be able to integrate them into diagnostic process. All these factors highlight a demand for further especially prospective studies and create an opportunity for the improvement in breast cancer screening.

DISCLOSURES

Author's contribution

Conceptualization: Agata Król, Katarzyna Kwaterska

Methodology: Agata Król

Software: Karol Kutyłowski, Paweł Łuckiewicz

Check: Katarzyna Kwaterska, Paweł Łuckiewicz

Formal analysis: Agata Król, Karol Kutyłowski

Investigation: Katarzyna Kwaterska, Paweł Łuckiewicz

Resources: Agata Król, Paweł Łuckiewicz

Data curation: Karol Kutyłowski, Paweł Łuckiewicz

Writing - rough preparation: Agata Król, Katarzyna Kwaterska, Paweł Łuckiewicz

Writing - review and editing: Agata Król, Karol Kutyłowski

Visualization: Agata Król, Katarzyna Kwaterska

Supervision: Agata Król, Karol Kutyłowski, Paweł Łuckiewicz

Project administration: Agata Król

All authors have read and agreed with the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed consent statement

Not applicable.

Data availability statement

Not applicable.

Conflicts of Interest Statement

The authors declare that there are no commercial and financial conflicts of interest associated with this review work.

REFERENCES

1. Katsura C, Ogunmwonyi I, Kankam HK, Saha S. Breast cancer: presentation, investigation and management. Br J Hosp Med (Lond). 2022;83(2):1-7. doi:10.12968/hmed.2021.0459
2. Wilkinson L, Gathani T. Understanding breast cancer as a global health concern. Br J Radiol. 2022;95(1130):20211033. doi:10.1259/bjr.20211033
3. Mello-Thoms C, Mello CAB. Clinical applications of artificial intelligence in radiology. Br J Radiol. 2023;96(1150):20221031. doi:10.1259/bjr.20221031
4. Gallée L, Kniesel H, Ropinski T, Götz M. Artificial intelligence in radiology - beyond the black box. Künstliche Intelligenz in der Radiologie – jenseits der Black-Box. Rofo. 2023;195(9):797-803. doi:10.1055/a-2076-6736
5. Katzman BD, van der Pol CB, Soyer P, Patlas MN. Artificial intelligence in emergency radiology: A review of applications and possibilities. Diagn Interv Imaging. 2023;104(1):6-10. doi:10.1016/j.diii.2022.07.005
6. Hernström V, Josefsson V, Sartor H, 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(3):e175-e183. doi:10.1016/S2589-7500(24)00267-X
7. Dang LA, Chazard E, Poncelet E, et al. Impact of artificial intelligence in breast cancer screening with mammography. Breast Cancer. 2022;29(6):967-977. doi:10.1007/s12282-022-01375-9
8. Gjesvik J, Moshina N, Lee CI, Miglioretti DL, Hofvind S. Artificial Intelligence Algorithm for Subclinical Breast Cancer Detection. JAMA Netw Open. 2024;7(10):e2437402. Published 2024 Oct 1. doi:10.1001/jamanetworkopen.2024.37402

9. Sechopoulos I, Teuwen J, Mann R. Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art. Semin Cancer Biol. 2021;72:214-225. doi:10.1016/j.semcancer.2020.06.002
10. 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 [published correction appears in Lancet Digit Health. 2023 Oct;5(10):e646. doi: 10.1016/S2589-7500(23)00181-4.]. Lancet Digit Health. 2023;5(10):e703-e711. doi:10.1016/S2589-7500(23)00153-X
11. Díaz O, Rodríguez-Ruíz A, Sechopoulos I. Artificial Intelligence for breast cancer detection: Technology, challenges, and prospects. Eur J Radiol. 2024;175:111457. doi:10.1016/j.ejrad.2024.111457
12. Soliman A, Li Z, Parwani AV. Artificial intelligence's impact on breast cancer pathology: a literature review. Diagn Pathol. 2024;19(1):38. Published 2024 Feb 22. doi:10.1186/s13000-024-01453-w
13. Gastounioti A, Desai S, Ahluwalia VS, Conant EF, Kontos D. Artificial intelligence in mammographic phenotyping of breast cancer risk: a narrative review. Breast Cancer Res. 2022;24(1):14. Published 2022 Feb 20. doi:10.1186/s13058-022-01509-z
14. Lauritzen AD, Rodríguez-Ruiz A, von Euler-Chelpin MC, et al. An Artificial Intelligence-based Mammography Screening Protocol for Breast Cancer: Outcome and Radiologist Workload. Radiology. 2022;304(1):41-49. doi:10.1148/radiol.210948
15. Koch HW, Larsen M, Bartsch H, Kurz KD, Hofvind S. Artificial intelligence in BreastScreen Norway: a retrospective analysis of a cancer-enriched sample including 1254 breast cancer cases. Eur Radiol. 2023;33(5):3735-3743. doi:10.1007/s00330-023-09461-y
16. Elías-Cabot E, Romero-Martín S, Raya-Povedano JL, Brehl AK, Álvarez-Benito M. Impact of real-life use of artificial intelligence as support for human reading in a population-based breast cancer screening program with mammography and tomosynthesis. Eur Radiol. 2024;34(6):3958-3966. doi:10.1007/s00330-023-10426-4
17. Raya-Povedano JL. AI in breast cancer screening: a critical overview of what we know. Eur Radiol. 2024;34(7):4774-4775. doi:10.1007/s00330-023-10530-5
18. Yoon JH, Kim EK. Deep Learning-Based Artificial Intelligence for Mammography. Korean J Radiol. 2021;22(8):1225-1239. doi:10.3348/kjr.2020.1210

19. Branco PESC, Franco AHS, de Oliveira AP, et al. Artificial intelligence in mammography: a systematic review of the external validation. Rev Bras Ginecol Obstet. 2024;46:e-rbgo71. Published 2024 Sep 4. doi:10.61622/rbgo/2024rbgo71
20. Högberg C, Larsson S, Lång K. Anticipating artificial intelligence in mammography screening: views of Swedish breast radiologists. BMJ Health Care Inform. 2023;30(1):e100712. doi:10.1136/bmjhci-2022-100712
21. Ongena YP, Yakar D, Haan M, Kwee TC. Artificial Intelligence in Screening Mammography: A Population Survey of Women's Preferences. J Am Coll Radiol. 2021;18(1 Pt A):79-86. doi:10.1016/j.jacr.2020.09.042
22. Zhang H, Lin F, Zheng T, et al. Artificial intelligence-based classification of breast lesion from contrast enhanced mammography: a multicenter study. Int J Surg. 2024;110(5):2593-2603. Published 2024 May 1. doi:10.1097/JS9.0000000000001076
23. Zeng A, Houssami N, Noguchi N, Nickel B, Marinovich ML. Frequency and characteristics of errors by artificial intelligence (AI) in reading screening mammography: a systematic review. Breast Cancer Res Treat. 2024;207(1):1-13. doi:10.1007/s10549-024-07353-3
24. Larsen M, Aglen CF, Lee CI, et al. Artificial Intelligence Evaluation of 122 969 Mammography Examinations from a Population-based Screening Program. Radiology. 2022;303(3):502-511. doi:10.1148/radiol.212381
25. Friedewald SM, Sieniek M, Jansen S, et al. Triaging mammography with artificial intelligence: an implementation study. Breast Cancer Res Treat. 2025;211(1):1-10. doi:10.1007/s10549-025-07616-7
26. Fisches ZV, Ball M, Mukama T, et al. Strategies for integrating artificial intelligence into mammography screening programmes: a retrospective simulation analysis. Lancet Digit Health. 2024;6(11):e803-e814. doi:10.1016/S2589-7500(24)00173-0
27. Chang YW, Ryu JK, An JK, et al. Artificial intelligence for breast cancer screening in mammography (AI-STREAM): preliminary analysis of a prospective multicenter cohort study. Nat Commun. 2025;16(1):2248. Published 2025 Mar 6. doi:10.1038/s41467-025-57469-3
28. Sabani A, Landsmann A, Hejduk P, et al. BI-RADS-Based Classification of Mammographic Soft Tissue Opacities Using a Deep Convolutional Neural Network. Diagnostics (Basel). 2022;12(7):1564. Published 2022 Jun 28. doi:10.3390/diagnostics12071564

29. Zonderland H, Smithuis R. BI-RADS for mammography and ultrasound 2013. The Radiology Assistant Updated version. 2013
30. Lång K, Hofvind S, Rodríguez-Ruiz A, Andersson I. Can artificial intelligence reduce the interval cancer rate in mammography screening?. Eur Radiol. 2021;31(8):5940-5947. doi:10.1007/s00330-021-07686-3
31. Salim M, Wåhlin E, Dembrower K, et al. External Evaluation of 3 Commercial Artificial Intelligence Algorithms for Independent Assessment of Screening Mammograms. JAMA Oncol. 2020;6(10):1581-1588. doi:10.1001/jamaoncol.2020.3321
32. Çorbacıoğlu ŞK, Aksel G. Receiver operating characteristic curve analysis in diagnostic accuracy studies: A guide to interpreting the area under the curve value. Turk J Emerg Med. 2023;23(4):195-198. Published 2023 Oct 3. doi:10.4103/tjem.tjem_182_23