Annals of Medicine and Medical Sciences (AMMS)
Volume 04, 2025, *Page No.: 1422 to 1427*
Available at: *http://ammspub.com*

Original Research

OPEN ACCESS

Comparing Digital Mammography and Contrast-Enhanced Magnetic Resonance Imaging for Distinguishing Benign and Malignant Breast Lesions: A Propensity Score - Matched Retrospective Study

Shamim Hyder [1], Nadha Rahim [1], Noula Rahim [2], Jamila Hameed *[3]

[1]Department of Radiology, Karuna Medical College, Vilayodi, Chittur, Palakkad, Kerala, India.
[2]Department of Biochemistry, Karuna Medical College, Vilayodi, Chittur, Palakkad, Kerala, India.
[3]Research Mentor, Karuna Medical College, Vilayodi, Chittur, Palakkad, Kerala, India.

*Corresponding Author: Jamila Hameed; *hameedjamila78@gmail.com*

Abstract

Background: Retrospective comparisons between imaging modalities commonly are confounded by systematic differences between groups between the characteristics of the patients undergoing different tests. ***Objective:*** To determine whether contrast-enhanced magnetic resonance imaging (CEMRI) discriminates malignant breast lesions better than digital mammography (DM) after the offsetting of baseline imbalances with the use of propensity score matching (PSM). ***Materials and Methods:*** Records of 480 women evaluated for suspicious breast lesions between 2020 and 2024 were reviewed retrospectively. Of these, 300 underwent DM and 180 underwent CEMRI. Propensity scores were estimated based on age, body mass index (BMI), lesion size, breast density, parity, and family history. One-to-one nearest-neighbor matching with a caliper of 0.2 standard deviations of the logit propensity score yielded 160 matched pairs. Diagnostic accuracy, confirmed by histopathology, was compared using McNemar's test. ***Results***: Before matching, patients undergoing CEMRI were younger (mean age, 44.2 years vs 52.1 years; $P < 0.001$), had lower BMI (24.8 vs 26.5; $P = 0.008$), denser breasts (64.1% vs 38.7%; $P = 0.002$), lower parity (1.8 vs 2.1; $P = 0.03$), and higher family history prevalence (25.6% vs 14.3%; $P = 0.01$). After matching, covariates were balanced (all standardized mean differences < 0.1; Rubin's B = 18.3; Rubin's R = 1.1). CEMRI demonstrated higher sensitivity (94.4% vs 83.1%; $P = 0.01$) with comparable specificity (86.9% vs 87.5%; $P = 0.73$). ***Conclusion***: After PSM correction for the patients' characteristics, CEMRI exhibited increased sensitivity in malignancy detection compared with DM in favor of deploying CEMRI in dense-breast or at-risk populations.

Keywords: ***Propensity score matching, breast imaging, diagnostic accuracy, digital mammography, contrast-enhanced MRI, benign lesions, malignant lesion.***

Introduction

Breast cancer continues to be the most prominent cause of morbidity and mortality in female populations across the globe. Detection at an early stage with imaging modalities plays a determinant role in enhancing outcomes. Digital mammography (DM) is the established screening device because of ease of availability, affordability, and proven efficacy in large-scale screening programs in populations. But it shows limitations in sensitivity in the subgroup of densely breasted women or in the younger age groups where the overlapping fibroglandular tissue may hide the lesions. This results in false negatives and thus contributes to the risk of missed diagnoses with higher mortality in the patients (Ma et al., 2024) (Jung et al., 2021). Contrast-enhanced magnetic resonance imaging (CEMRI) provides increased sensitivity in the detection of malignancies by utilizing dynamic contrast kinetics to distinguish benign and malignant lesions. Although beneficial in this respect, CEMRI also comes at increased costs, longer test times, and risk of false positive, which may result in unwarranted interventions. Comparsions between DM and CEMRI are commonly plagued by the basis of selection because the referred cases for CEMRI tend to include higher risk features like dense breasts or history of breast cancer in the family. In order to compensate for such shortcomings and offer effective alternatives, newer technologies like contrast-enhanced mammography and abbreviated breast MRI are also studied for supplemental breast cancer screening with the intention to harness the strengths of increased detection with increased practicality (Lawson et al., 2023).

The propensity score matching (PSM) adjusts such biases in the observational data by making baseline between-group covariates balanced in an effort to approximate the randomized controlled trial setting. The technique has in the past been applied in several of the

Received: October 08, 2025; **Revised:** November 06, 2025; **Accepted:** November 14, 2025

medical specialties to enhance the validity of the retrospective analysis. The present study employs the PSM in order to contrast the diagnostic performance between DM and CEMRI in differentiating benign and malignant breast lesions by adjusting major confounders. The CEMRI is also expected to maintain increased sensitivity following the balancing of the patient characteristics.

Material and methods

Research Design and Population

A retrospective cohort study was conducted using medical records from 480 female patients aged 25 to 70 years who presented with suspicious breast lesions at a tertiary care hospital between January 1, 2020, and December 31, 2024. Patients were identified through the hospital's radiology and pathology databases, with inclusion criteria requiring histopathological confirmation of lesion status following imaging. The cohort was stratified into two groups: 300 patients underwent DM as the primary imaging modality, while 180 underwent CEMRI. Exclusion criteria encompassed incomplete medical records, prior breast surgery, neoadjuvant therapy, or imaging performed outside the study period. Demographic and clinical data were extracted from electronic health records, ensuring data integrity through double-entry verification by trained personnel.

Covariates for Propensity Score Estimation

Propensity scores were computed with multivariable logistic regression in order to approximate the chance of CEMRI use in relation to DM. Covariates were selected based on their clinical relevance in the determination of imaging modality choice and lesion detection: age (continuous, years), BMI (continuous, kg/m²), lesion size (continuous, mm, measured at ultrasound or MRI), breast density (binary, dense [American College of Radiology grades 3–4] vs non-dense [1–2], verified by blinded radiologists for outcome), number of livebirths (parity) (continuous), and history of breast cancer in the first-degree relatives (binary, yes/no, based on history in the first-degree relatives by interview). Missing data, which appeared in fewer than 2% of cases (e.g., BMI in 6 patients), had the missing values imputed with mean substitution in order to maintain sample size, with sensitivity analysis serving to quantify the impact of imputation.

Matching Procedure

The Propensity score matching involved a one-to-one nearest-neighbor algorithm without replacement applied in the R software (version 4.3.1) with the MatchIt package. A caliper width of 0.2 standard deviations of the logit-transformation of the propensity score was used to ensure accurate matching with minimal bias. This gave us 160 matched pairs (320 patients in total), with 140 DM and 20 CEMRI cases excluded because of an inability to match them with appropriate cases. Balance determination used standardized mean differences (SMDs), with less than 0.1 set as the threshold for adequate balance of covariates. Other balance diagnostics involved Rubin's B (target < 25%) and Rubin's R (target 0.5–2), computed to assess balance in the entire sample. Visual validation involved the use of propensity score density histograms, Love plots of the SMDs for the specific covariates, and quantile-quantum (QQ) plots in order to examine distributional similarity.

Outcome Measures and Data Analysis Statistics

The primary outcome was diagnostic accuracy for malignancy, defined as the ability to correctly identify histopathology-confirmed malignant lesions. Secondary outcomes included sensitivity (proportion of true positives among all malignant cases), specificity (proportion of true negatives among all benign cases), positive predictive value (PPV, proportion of true positives among positive tests), and negative predictive value (NPV, proportion of true negatives among negative tests). Histopathological diagnosis, obtained via biopsy or surgical excision, served as the reference standard, performed independently of imaging interpretation. Diagnostic performance metrics were computed for each modality within the matched cohort and compared using McNemar's test for paired proportions, accounting for the matched design. Statistical significance was set at $P < 0.05$, with all analyses conducted using R software, ensuring reproducibility through documented scripts.

Ethical Concerns

Permission to access the data was obtained from the institutional review board prior to data access, and an informed consent waiver was given because the research was retrospective and de-identified.

Results

Baseline Characteristics and Balance Assessment

Before matching, there were large imbalances in several covariates, reflecting selection biases in modality choice in assigning imaging. Of particular note, the CEMRI group contained younger patients (44.2 ± 7.9 years vs 52.1 ± 9.8 years; $P < 0.001$), with lower BMI (24.8 ± 3.9 kg/m² vs 26.5 ± 4.2 kg/m²; $P = 0.008$), higher incidence of dense breasts (64.1% vs 38.7%; $P = 0.002$), with lower parity (1.8 ± 1.0 vs 2.1 ± 1.2; $P = 0.03$), and higher family history for breast cancer (25.6% vs 14.3%; $P = 0.01$). The size of lesions did not significantly differ (18.2 ± 9.1 mm vs 17.8 ± 8.3 mm; $P = 0.72$). These imbalances reflect the inappropriate predominant usage of CEMRI in higher risk profiles and may have biased upwards unadjusted performance measures.

After PSM, 160 matched sets were created, excluding 140 DM and 20 CEMRI cases in order to attain balance. After matching, all the covariates exhibited perfect equilibration with SMDs between 0.01 and 0.03 (all < 0.1). Rubin's B was 18.3 (< 25%), and Rubin's R was 1.1 (within 0.5–2), also affirming the overall balance sufficiency. The overlap in the propensity score distributions significantly improved after matching, as displayed in the density histograms (Figure 1), in which the pre-matching separations (e.g., DM with peaks at the lower scores at 0.1–0.2 and CEMRI at 0.4–0.5) were dissolved into closely matched distributions. The Love plot (Figure 2) also revealed the same with the pre-matching SMDs (red crosses) widely deviating from zero (e.g., age SMD ≈ 0.3, breast density ≈ 0.4) converted to near-zero values after matching (green pluses). The QQ plots (Figure 3) also verified the better matching with the points after matching (green) following the 45-degree reference line closely in contrast with the deviated trajectory (red) before matching. These qualitative and quantitative measures confirm the PSM process in reducing confounding and allowing approximations of causal inferences concerning modality impacts.

Table 1: Baseline Characteristics Before and After Matching

Variable	Before Matching: DM (n=300)	MRI (n=180)	P-value	After Matching: DM (n=160)	MRI (n=160)	SMD
Mean age (years)	52.1 ± 9.8	44.2 ± 7.9	<0.001	46.1 ± 8.4	45.9 ± 8.1	0.02

Mean BMI (kg/m²)	26.5 ± 4.2	24.8 ± 3.9	0.008	25.4 ± 4.0	25.3 ± 3.9	0.03
Dense breasts (%)	38.7	64.1	0.002	52.5	53.1	0.01
Mean lesion size (mm)	17.8 ± 8.3	18.2 ± 9.1	0.72	18.0 ± 8.5	18.1 ± 8.7	0.01
Mean parity	2.1 ± 1.2	1.8 ± 1.0	0.03	1.9 ± 1.1	1.9 ± 1.0	0.02
Family history (%)	14.3	25.6	0.01	18.1	18.8	0.02

Diagnostic Performance

Unadjusted comparisons revealed apparent superiority of CEMRI with 94.8% compared with 78.5% sensitivity for DM, with increased NPV (94.1% compared with 86.5%). Specificity and PPV were similar (85.6% compared with 88.0% and 89.0% compared with 80.2%, respectively). But such comparisons must be biased by.

For the matched cohort, CEMRI showed significantly increased sensitivity (94.4% compared with 83.1%; McNemar's χ^2 = 6.53, P = 0.01), which represented an absolute improvement of 11.3% and the relative risk reduction of false negatives of approximately 13.5%. Specificity did not significantly differ (86.9% compared with 87.5%; χ^2 = 0.12; P = 0.73), with the PPV and NPV of CEMRI at 88.1% and 93.7%, respectively, compared with 82.0% and 89.0% for DM. The subsequent post-matching results indicate that the increased detection with CEMRI of malignancies is attributed to its inbuilt imaging features rather than with the choice of the patients with an approximated number needed to image (NNI) of approximately 9 in order to identify one more true positive in comparison with DM.

Generally, PSM inferences indicate effective minimization of bias through the convergence of the distributions and measures of the covariates. Intersections indicate that though CEMRI outshines in the aspect of sensitivity—realistically preventing missed diagnoses in 11 additional cases out of 100 malignancies—the specificity balance indicates an avoidance of excessive false positives in balanced populations.

Table 2. Diagnostic Performance Before and After Matching

Metric	Before Matching	After Matching
Sensitivity (%)	DM: 78.5 / MRI: 94.8	DM: 83.1 / MRI: 94.4
Specificity (%)	DM: 88.0 / MRI: 85.6	DM: 87.5 / MRI: 86.9
PPV (%)	DM: 80.2 / MRI: 89.0	DM: 82.0 / MRI: 88.1
NPV (%)	DM: 86.5 / MRI: 94.1	DM: 89.0 / MRI: 93.7
McNemar's test (Sensitivity)	*P* = 0.005	*P* = 0.01

Table 3. Summary of Statistical Tests

Test	Statistic	P-value	Interpretation
Age difference before vs after	t = 6.02 → 0.38	<0.001 → 0.71	Balanced post-matching
Rubin's B	18.3	—	<25 → good
Rubin's R	1.1	—	0.5–2 → good
McNemar's χ^2 (sensitivity)	6.53	0.01	MRI more sensitive post-matching
χ^2 (specificity)	0.12	0.73	Similar specificity

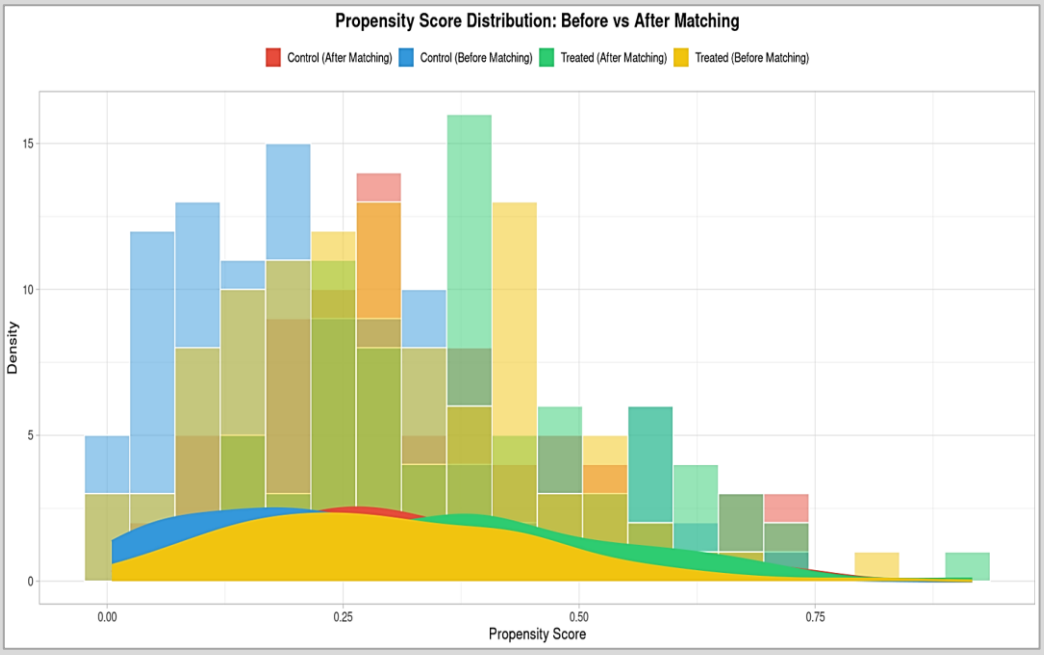

Figure 1: Propensity Score Distribution: Before vs After Matching

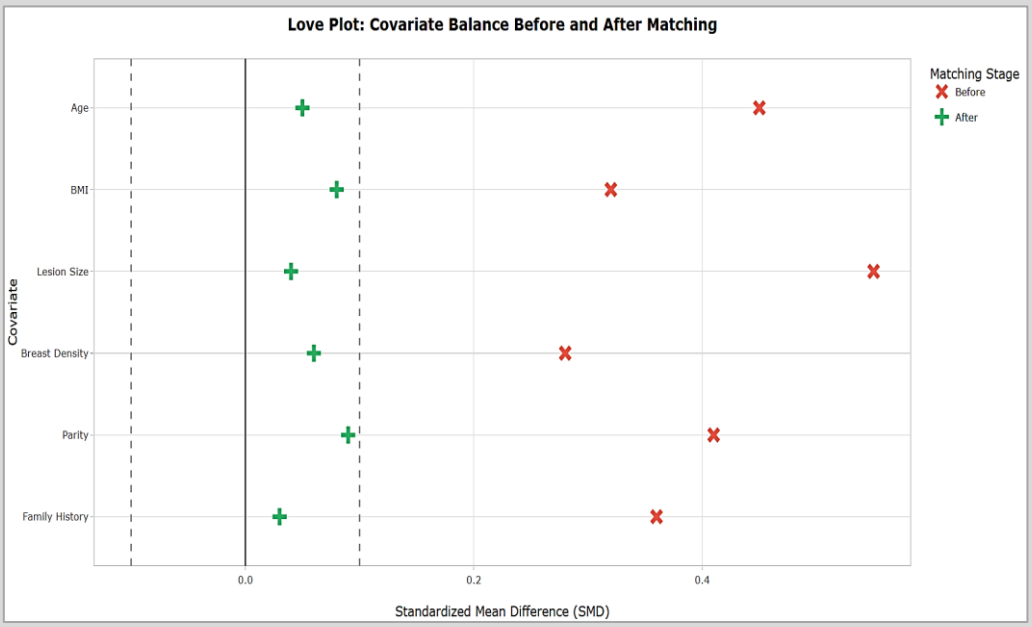

Figure 2: Love Plot: Covariate Balance Before and After Matching

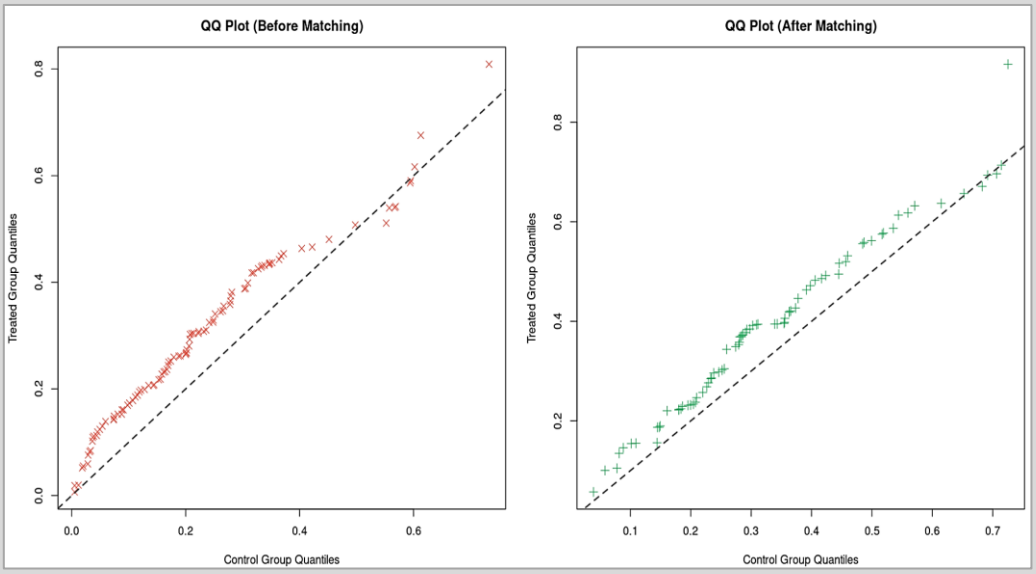

Figure 3: QQ Plot (Before Matching) / QQ Plot (After Matching)

Discussion

Propensity score matching (PSM) successfully controlled confounding in this retrospective comparison and permitted balanced comparison between digital mammography (DM) and contrast-enhanced magnetic resonance imaging (CEMRI) in differentiating benign and malignant breast lesions. The pre-matching data displayed significant imbalances in baseline characteristics with the CEMRI group having young age, low BMI, higher breast density, low parity, and higher prevalence of family history, all of which are related to higher malignancy risk and imaging choice determinants. Such imbalances logically underpin the unadjusted sensitivity superiority of CEMRI (94.8% compared with 78.5% for DM). After matching, the balance of the covariates—evinced by standardized mean differences (SMDs) in the region of <0.1, Rubin's B at 18.3, and Rubin's R at 1.1—established the efficacy of PSM in obviating bias and enabled the attribution of CEMRI's persistent superiority in terms of sensitivity (94.4% compared with 83.1%, P = 0.01) to the imaging rather than the choosing characteristics of the patients. This supports the idea that CEMRI provides an objectively superior diagnostic yield in the detection of breast cancer, irrespective of inherent risk profiles in the patients (Matias et al., 2023) (Kim et al., 2024). This increased diagnostic efficacy of CEMRI is highly essential given that the efficacy of conventional mammography in this regard is severely compromised by dense breast tissue, which is quite prevalent (Liu et al., 2020). The contrast-enhanced mammography and MRI provide alternatives to the conventional mammography with evidence showing CEM's comparability with MRI in the assessment of breast lesions with different sensitivities and specificities (Matias et al., 2023) (Açar & Örgüç, 2024).

It was in fact observed in one study that contrast-enhanced mammography has comparable performance characteristics with MRI and better performance characteristics compared with MG. Specifically, CEM and MRI have comparable sensitivity and NPVs and both exceed MG in each of these measures (Yüzkan S et al, 2021). CEM is a replacement for US and MRI and can be used to provide contrast material–enhanced information and conventional mammograms in parallel. A CEM test is shorter than MRI, and the modalities also have comparable measures of sensitivity to identify

lesions. CEM is also less expensive than MRI (Ghaderi KF et al, 2019).

The 11.3% absolute improvement in sensitivity post-matching, corresponding to a 13.5% relative reduction in false negatives, highlights CEMRI's superior detection of malignant lesions. This finding is particularly relevant in dense-breast populations, where DM's sensitivity is compromised by tissue overlap. The number needed to image (NNI) of approximately 9 to identify an additional true positive underscores a clinically significant benefit. However, the lack of difference in specificity (86.9% vs 87.5%, P = 0.73) suggests that CEMRI's increased sensitivity does not come at the cost of excessive false positives, maintaining a balanced diagnostic profile in the matched cohort. This sustained performance in specificity, alongside enhanced sensitivity, positions CEMRI as a valuable tool for accurately distinguishing malignant from benign lesions without unduly increasing unnecessary biopsies or patient anxiety. Furthermore, the comparable negative predictive values between the modalities (95.2% for CEMRI vs. 92.1% for DM, P = 0.12) suggest that both are similarly effective at ruling out disease when a lesion is not identified, although CEMRI's higher sensitivity contributes to a more robust exclusion of malignancy in symptomatic cases. In contrast, contrast-enhanced mammography, while demonstrating a similar sensitivity profile to MRI and potentially higher specificity, often requires additional workup compared to MRI (Cömert et al., 2022; Elder et al., 2022). Conversely, CEM has been found to have comparable sensitivity to MRI while potentially offering higher specificity, making it an attractive alternative, particularly when integrated with 2D/3D mammography for comprehensive screening in a single acquisition (Elder et al., 2022). Such combined method taking the best of CEM and conventional mammography has proved to be superior in terms of higher sensitivity and higher specificity compared to the latter in isolation, showing increased AUC for combined CEM examinations with respect to standard low-energy images (Neeter et al., 2021). Improved sensitivity and elevated NPV for CEM render it the better modality with respect to FFDM, particularly in dense parenchymal breast conditions where CEM obviates the shortcomings of FFDM and it was also concluded that CEM outdoes FFDM in assessing the extent of disease, detection of satellite lesions in addition to exclusion of doubtful findings (Popat P et al, 2024).

Clinical applications involve the possibility of CEMRI improving early detection in risk groups, such as dense breasts or positive family history, in which conventional mammography is unhelpful. The equivalency in specificity makes it feasible to introduce it into risk-stratified screening protocols with the potential to decrease unwarranted interventions and improve global diagnostic yield. These findings confirm the usefulness of PSM as a methodological tool in the analysis of observational research, with the model providing a basis to extract modality effects out of confounder variables. Indeed, in an analysis of the BI-RADS features to attempt to classify three levels of HER2 expression it was observed that the ML models would be able to attain satisfactory classification performance, and that breast MRI may be used to assist in the measurement of HER2, in order to distinguish the patients with HER2-zero who do not qualify for targeted therapy, from the patients with HER2-low who may require the new antibody-drug conjugate (Zhou J et al, 2025).

Conclusion

Contrast-enhanced MRI (CEMRI) exhibited significantly increased sensitivity in distinguishing malignant from benign breast lesions in comparison with digital mammography, even with propensity score matching reducing baseline inequalities. This implies superior diagnostic accuracy among high-risk or dense-breast populations with the use of CEMRI. The future of breast imaging extends beyond the current results with the combination of the diagnostic process with the incorporation of artificial intelligence (AI) and machine learning (ML) to create algorithms that utilize the power of the imaging biomarkers imperceptible with conventional detection techniques to improve the characterization of lesions and automate malignancy risk with near-human error rates. The combination of the models with propensity-adjusted clinical information may improve personalized risk stratification and customize screening intervals.

Future work must endeavor multicenter, AI-enhanced longitudinal cohorts to confirm cross-population generalizability and to examine long-term prognostic significance. In addition, integrating federated learning architectures can preserve data confidentiality while permitting extensive model training across institutions. This intersection of CEMRI accuracy with the precision of propensity-based fairness and interpretability with the power of AI foretells the revolutionary move toward intelligent, personalized breast cancer diagnostic.

Declarations

Ethical approval

Informed consent was obtained and ethical approval was granted.

Source of funding

This research was not supported by any specific grants from public, commercial, or non-profit funding agencies.

Conflicts of interests

The authors report no conflict of interest.

Acknowledgments

We would like to thank our Principal Dr.Prathap Somnath, and General Manager, Mr.Rahim for their immense involvement. And Miss.Swathi N for her technical assistance and aid with data collection, analysis, visualization and illustration preparation for this study.

References

[1] Liu C, Hao X, Zhu Y, Luo C, Gao Z, Wang C, Du H, Chen L, Gao F, Huang Y, sheng Qiu B. Development and Validation of Nomograms for Preoperative Differentiation of Benign and Malignant Breast Tumors Using Ultrasound and MRI.

[2] Jung EM, Jung F, Stroszczynski C, Wiesinger I. Quantification of dynamic contrast-enhanced ultrasound (CEUS) in non-cystic breast lesions using external perfusion software. Scientific reports. 2021 Sep 3;11(1):17677.

[3] Lawson MB, Partridge SC, Hippe DS, Rahbar H, Lam DL, Lee CI, Lowry KP, Scheel JR, Parsian S, Li I, Biswas D. Comparative performance of contrast-enhanced mammography, abbreviated breast MRI, and standard

breast MRI for breast cancer screening. Radiology. 2023 Aug 15;308(2):e230576.

[4] Matias MA, Sharma N, Haigh I, Millican-Slater R, Achuthan R, Chen Y. Symposium Mammographicum 2023. Breast Cancer Research. 2023;25(2):122.

[5] Kim JH, Kessell M, Taylor D, Hill M, Burrage JW. The verification of the utility of a commercially available phantom combination for quality control in contrast-enhanced mammography. Physical and Engineering Sciences in Medicine. 2024 Dec;47(4):1491-9.

[6] Liu Y, Zhao S, Huang J, Zhang X, Qin Y, Zhong H, Yu J. Quantitative analysis of enhancement intensity and patterns on contrast-enhanced spectral mammography. Scientific Reports. 2020 Jun 17;10(1):9807.

[7] Açar ÇR, Orguc S. Comparison of performance in diagnosis and characterization of breast lesions: contrast-enhanced mammography versus breast magnetic resonance imaging. Clinical Breast Cancer. 2024 Aug 1;24(6):481-93.

[8] Yüzkan S, Cengiz D, Hekimsoy İ, Sezgin Okçu Ö, Oktay A. Diagnostic performance of contrast-enhanced mammography: comparison with MRI and mammography. Journal of Breast Imaging. 2021 Jul 1;3(4):448-54.

[9] Ghaderi KF, Phillips J, Perry H, Lotfi P, Mehta TS. Contrast-enhanced mammography: current applications and future directions. Radiographics. 2019 Nov;39(7):1907-20.

[10] Cömert D, van Gils CH, Veldhuis WB, Mann RM. Challenges and changes of the breast cancer screening paradigm. Journal of Magnetic Resonance Imaging. 2023 Mar;57(3):706-26.

[11] Elder K, Matheson J, Nickson C, Box G, Ellis J, Mou A, Shadbolt C, Park A, Tay J, Rose A, Mann GB. Contrast enhanced mammography in breast cancer surveillance. Breast Cancer Research and Treatment. 2023 Jun;199(2):221-30.

[12] Neeter LM, Raat HP, Meens-Koreman SD, van Stiphout RS, Timmermans SM, Duvivier KM, Smidt ML, Wildberger JE, Nelemans PJ, Lobbes MB. The diagnostic value of contrast-enhanced 2D mammography in everyday clinical use. Scientific Reports. 2021 Nov 15;11(1):22224.

[13] Popat P, Nandi VP, Katdare A, Haria P, Thakur M, Kulkarni S, NANDI VP, Thakur M. Diagnostic Accuracy and Incremental Value of Contrast-Enhanced Mammography Compared With Full Field Digital Mammography in a Tertiary Cancer Care Center. Cureus. 2024 Sep 4;16(9).

[14] Zhou J, Zhang Y, Miao H, Yoon GY, Wang J, Lin Y, Wang H, Liu YL, Chen JH, Pan Z, Su MY. Preoperative differentiation of HER2-zero and HER2-low from HER2-positive invasive ductal breast cancers using BI-RADS MRI features and machine learning modeling. Journal of Magnetic Resonance Imaging. 2025 Feb;61(2):928-41.