

Validation of Combined Deep Learning Triaging and Computer-Aided Diagnosis in 2901 Breast MRI Examinations From the Second Screening Round of the Dense Tissue and Early Breast Neoplasm Screening Trial

Erik Verburg, MSc,* Carla H. van Gils, PhD,† Bas H.M. van der Velden, PhD,* Marije F. Bakker, PhD,† Ruud M. Pijnappel, MD, PhD,‡ Wouter B. Veldhuis, MD, PhD,‡ and Kenneth G.A. Gilhuijs, PhD*

Objectives: Computer-aided triaging (CAT) and computer-aided diagnosis (CAD) of screening breast magnetic resonance imaging have shown potential to reduce the workload of radiologists in the context of dismissing normal breast scans and dismissing benign disease in women with extremely dense breasts. The aim of this study was to validate the potential of integrating CAT and CAD to reduce workload and workup on benign lesions in the second screening round of the DENSE trial, without missing cancer.

Methods: We included 2901 breast magnetic resonance imaging scans, obtained from 8 hospitals in the Netherlands. Computer-aided triaging and CAD were previously developed on data from the first screening round. Computer-aided triaging dismissed examinations without lesions. Magnetic resonance imaging examinations triaged to radiological reading were counted and subsequently processed by CAD. The number of benign lesions correctly classified by CAD was recorded. The false-positive fraction of the CAD was compared with that of unassisted radiological reading in the second screening round. Receiver operating characteristics (ROC) analysis was performed and the generalizability of CAT and CAD was assessed by comparing results from first and second screening rounds.

Results: Computer-aided triaging dismissed 950 of 2901 (32.7%) examinations with 49 lesions in total; none were malignant. Subsequent CAD classified 132 of 285 (46.3%) lesions as benign without misclassifying any malignant lesion. Together, CAT and CAD yielded significantly fewer false-positive lesions, 53 of 109 (48.6%) and 89 of 109 (78.9%), respectively ($P = 0.001$), than radiological reading alone. Computer-aided triaging had a smaller area under the ROC curve in the second screening round compared with the first, 0.83 versus 0.76 ($P = 0.001$),

but this did not affect the negative predictive value at the 100% sensitivity operating threshold. Computer-aided diagnosis was not associated with significant differences in area under the ROC curve (0.857 vs 0.753, $P = 0.08$). At the operating thresholds, the specificities of CAT (39.7% vs 41.0%, $P = 0.70$) and CAD (41.0% vs 38.2%, $P = 0.62$) were successfully reproduced in the second round. **Conclusion:** The combined application of CAT and CAD showed potential to reduce workload of radiologists and to reduce number of biopsies on benign lesions. Computer-aided triaging (CAT) correctly dismissed 950 of 2901 (32.7%) examinations with 49 lesions in total; none were malignant. Subsequent computer-aided diagnosis (CAD) classified 132 of 285 (46.3%) lesions as benign without misclassifying any malignant lesion. Together, CAT and CAD yielded significantly fewer false-positive lesions, 53 of 109 (48.6%) and 89 of 109 (78.9%), respectively ($P = 0.001$), than radiological reading alone.

Key Words: computer-aided triaging, computer-aided diagnosis, deep learning, breast MRI, machine learning

(*Invest Radiol* 2023;58: 293–298)

Contrast-enhanced magnetic resonance imaging (MRI) may be used in combination with x-ray mammography to screen asymptomatic women for breast cancer. Supplemental MRI screening in women with extremely dense breasts improved the detection of cancer.¹ Similar observations were reported for women at increased lifetime risk. Nonetheless, breast MRI screening has lower specificity compared with mammography^{1–3} and it invokes additional workload.

To reduce the workload of breast magnetic resonance (MR) radiologists, researchers have focused on automated lesion detection.^{4,5} One focused on identifying normal scans using computer-aided triaging (CAT),⁶ Computer-aided diagnosis (CAD) of dynamic contrast-enhanced MRI^{7,8} and multiparametric MRI⁹ was found to further increase specificity.^{10–14}

A recently reported CAT—developed on data from 4783 MRI examinations from the first screening round of the DENSE trial—dismissed approximately 40% of normal breast examinations without dismissing malignant disease.⁶ In addition to CAT, CAD was developed on the same data to distinguish between 444 benign and 81 malignant lesions. It is yet unknown whether CAD is complementary to CAT to increase the positive predictive value (PPV) of MRI screening in women with extremely dense breasts while maintaining high negative predictive value (NPV) and minimizing the number of normal scans to be read by radiologists.

The aim of this study was to validate the potential of combining CAT with CAD in the second screening round of DENSE to minimize work load as well as minimizing the number of biopsies on benign lesions without dismissing malignant breast disease.

MATERIALS AND METHODS

We validated the potential impact of combined CAT and CAD in the second screening round of the DENSE trial and compared with to radiological reading without computer assistance. Impact is expressed

Received for publication July 12, 2022; and accepted for publication, after revision, September 10, 2022.

From the *Image Sciences Institute, †Julius Center for Health Sciences and Primary Care, and ‡Department of Radiology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands.

Conflicts of interest and sources of funding: This study is financially supported by KWF, grant number UU-2014-7151, and used data acquired during the DENSE trial. The DENSE trial was supported by the regional screening organizations, Volpara Solutions, the Dutch Expert Centre for Screening, and the National Institute for Public Health and the Environment. The DENSE trial is financially supported by the University Medical Center Utrecht (project number: UMCU DENSE), the Netherlands Organization for Health Research and Development (ZonMw, project numbers: ZONMW-200320002-UMCU and ZonMW Preventie 50-53125-98-014), the Dutch Cancer Society (KWF Kankerbestrijding; project numbers: DCS-UU-2009-4348, UU-2014-6859, and UU2014-7151), the Dutch Pink Ribbon/A Sister's Hope (project number: Pink Ribbon-10074), Bayer AG Pharmaceuticals, Radiology (project number: BSP-DENSE), and Stichting Kankerpreventie Midden-West.

Correspondence to: Erik Verburg, MSc, Q.02.4.45, PO Box 85500, 3508 GA Utrecht, the Netherlands. E-mail: e.verburg-5@umcutrecht.nl; Kenneth G.A. Gilhuijs, PhD, Image Sciences Institute, Q.02.4.45, PO Box 85500, University Medical Center Utrecht, Utrecht University, Utrecht 3508 GA, the Netherlands. E-mail: k.g.a.gilhuijs@umcutrecht.nl

Copyright © 2022 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

ISSN: 0020-9996/23/5804-0293

DOI: 10.1097/RLI.0000000000000934

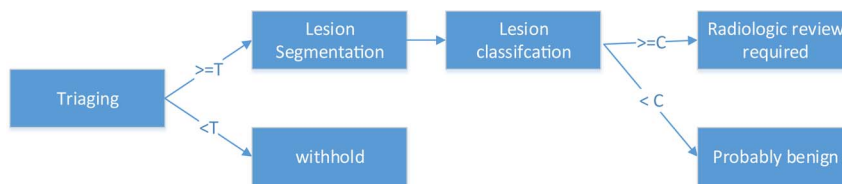


FIGURE 1. Combination of CAT and CAD applied to the second screening round of the DENSE trial. Breasts with probability of lesions lower than operating threshold T were dismissed by CAT for processing by CAD. If the probability of malignant disease was larger or equal to operating threshold C, the lesion was classified as malignant.

in terms of reduction in (1) MRI scans with normal anatomy read by radiologists and (2) false-positive referrals to further diagnostic work-up with additional MRI or biopsy. Both CAT and CAD were previously trained on MRI scans from the first screening round only.

First, we briefly describe the design of the DENSE trial, followed by description of the study participants, MRI acquisition parameters, unassisted radiological reading (ie, the reference standard), CAT and CAD, followed by the combination of the methods.

DENSE Trial

The DENSE trial (ClinicalTrials.gov: NCT01315015) investigates whether additional MRI screening of asymptomatic women with extremely dense breasts (ie, American College of Radiology Breast Imaging-Reporting and Data System category 4 measured with Volpara software) reduces the number of interval cancers.¹⁵ Participating women had extremely dense breasts without lesions suspected of malignancy on mammography. The first results of the DENSE trial confirmed the hypothesis of detection of additional breast cancers and the reduction of interval cancers. In the first round of screening, the cancer-detection yield with MRI after negative mammography was 79 in 4783 women, or 16.5 of 1000 screens.¹

The current validation study focused primarily on the screening data from the second round. No previous artificial intelligence (AI) studies have been performed on these data before. The screened data acquired in the first round were included in 2 previous AI studies, one on AI triaging,⁶ the other on CAD.¹⁴

Participants

Participants (between 50 and 75 years of age) were included from the national population-based mammography screening program. From the 4783 participants in the first MRI screening round of DENSE, 3436 women participated in the second MRI round between September 6, 2014, and April 17, 2019.¹⁶ To be eligible for the second MRI round, they had been participating in the national program, again with a normal mammography result (ie, no referral). Written informed consent was obtained from all women before screening. The trial was approved by the Dutch Minister of Health, Welfare, and Sport (2011/19 WBO, The Hague, the Netherlands). According to the Dutch law on population studies, the study was waived from ethical review by the local institutional review board.

MRI Acquisition

The MRI examinations were performed in 8 hospitals in the Netherlands using the same MRI protocol in each screening round. The protocol has been described in detail elsewhere.¹⁵ In short, T1-weighted images were acquired without fat suppression, followed by dynamic T1-weighted imaging, consisting of 1 precontrast series at high spatial resolution and 15 to 20 fast acquisitions after contrast administration. Four to 5 postcontrast series at high spatial resolution followed. Fat suppression was optional. In addition, diffusion-weighted series were acquired using 2 or 3 b-values. T2-weighted acquisition was optional.

Contrast agent was injected at rate of 1 mL/s to a total dose of 0.1 mmol of gadobutrol (Gadovist; Bayer AG, Leverkusen, Germany) per kilogram of body weight. Images were acquired using a 3-T MRI

unit; 5 hospitals used Philips MR devices (Eindhoven, the Netherlands), whereas the other 3 hospitals used Siemens devices (Erlangen, Germany).

Methods

Unassisted Radiological Reading

In the DENSE trial, breast MR examinations were read by trained breast MR radiologists (with experience from 5 to 23 years'). In short, MRI examinations were single read and scored according to the BI-RADS MRI lexicon.¹⁷ Only BI-RADS 3 lesions were double read (consensus reading); in these cases, MRI was repeated after 6 months. Women with BI-RADS 4 or BI-RADS 5 lesions were always recommended to undergo biopsy.

TABLE 1. Types of Lesions in the Second Screening Round

Benign lesions	314
Adenosis	2
Apocrine metaplasia	3
Atypical ductal hyperplasia	2
Cylindrical cell metaplasia	1
Cyst	3
Epithelia proliferation	1
Fibroadenoma	5
Fibrosis	8
Hemangioma	1
LCIS*	1
Lymph node	2
Mastopathy	4
Normal breast tissue	3
Papilloma	3
Periductitis	1
Sclerosis	5
Usual ductal hyperplasia	7
BI-RADS 2 (no biopsy)	225
BI-RADS 3 (no biopsy)	21
Unknown	16
Malignant lesions	20
Ductal carcinoma in situ	6
Invasive carcinoma (not otherwise specified)	8
Mixed invasive ductal and lobular carcinoma	2
Invasive lobular carcinoma	3
Invasive tubular carcinoma	1

Results were obtained from biopsy.

*In the DENSE trial, LCIS is considered a benign lesion,¹ conforming to Dutch and international guidelines.¹⁹

BI-RADS, breast imaging-reporting and data system; LCIS, lobular carcinoma in situ.

Downloaded from http://journals.lww.com/investigativeradiology by 51.15.134.125 on 04/14/2023

TABLE 2. Results of Triaging in First- and Second-Round Data

	First-Round Data	Second-Round Data	P
AUC	0.83 (0.80–0.85)	0.76 (0.72–0.81)	P = 0.001
Percentage of dismissed examinations without lesion	39.7% (30.0%–49.4%)	41.0% (30.4%–51.6%)	P = 0.70
Percentage of examinations with lesions triaged to radiological review	90.7% (86.7%–94.7%)	85.6% (79.2%–92.0%)	P = 0.07

AUC indicates area under the receiver operating characteristic curve.

Computer-Aided Triaging

The method previously developed⁶ to dismiss the largest number of normal breast MRI examinations without dismissing malignant disease was applied, without modifications, to the second screening round. In short, the probability of lesion presence was estimated using deep learning. This was done for each breast separately. The probability was established in 3 maximum intensity projection images of contrast-agent uptake in orthogonal directions (transversal, sagittal and coronal), and the 3 results were averaged. The probability per examination was equal to the highest probability in the left or right breast.

During model development on first screening round data, 8-fold internal-external validation was used; that is, in each fold, the data of 1 hospital were used as test data and the data of the remaining hospitals were used to train the convolutional neural network (CNN). Hence, 8 CNNs were developed (1 for each fold). When the probability of lesion presence was less than an operating threshold (established in the first screening round data), the breast examination was considered normal.⁶

Computer-Aided Diagnosis

Previously, a method was developed to distinguish between benign and malignant breast lesions on multiparametric MRI.¹⁴ In short, lesion segmentation was followed by feature extraction and classification

into benign or malignant groups. Lesion segmentation used constrained volume growing from a manually placed seed point¹⁸ at or near the lesion by a technical physician (E.V.) under supervision of a breast MR radiologist (W.B.V.). The features were extracted from the segmentation results and the MR images. In addition, clinical features were used (ie, age, body mass index, and BI-RADS).¹⁴ Training and testing were initially done on the first-round screening data only using Ridge-regression modeling with 10-fold cross-validation to estimate the probability of malignancy. In the current study, we retrained the Ridge-regression model on the first-round data and applied the model to the second-round screening data. An operating threshold in the probability was chosen in first screening round data at which all malignant lesions were correctly identified.

Combination of CAT and CAD

The current validation study applied CAT and CAD to the second round (Fig. 1), using the operating thresholds established in the first round.^{6,14} Scans considered to be normal by CAT were recorded and dismissed from further analysis. Scans considered to contain lesions were matched against the lesions detected by radiologists in the second screening round of the trial. These lesions were then offered to the CAD. Lesions considered to be benign and those considered to be malignant were recorded.

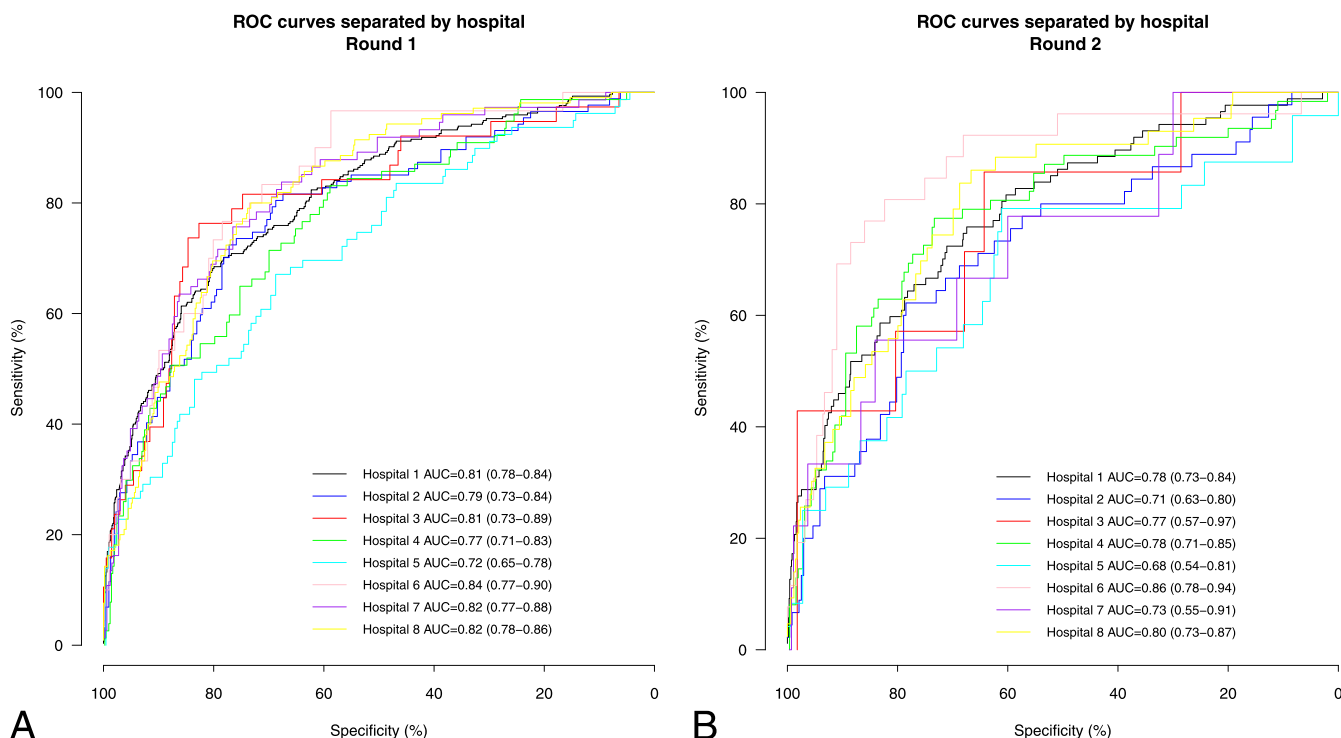


FIGURE 2. ROC curves of CAT for the task of distinguishing between examinations with lesions (benign and malignant) and examinations without lesions, applied to first (left) and second (right) screening-round data. The 95% confidence intervals are shown in the legend.

Downloaded from http://investigativeradiology.com/ on April 14, 2023

TABLE 3. Results of CAD and Radiological Reading in Data From the First and Second Screening Rounds of DENSE

	First-Round Data	Second-Round Data	P
AUC	0.86 (0.81–0.90)	0.75 (0.64–0.86)	P = 0.08
Benign lesions classified as benign by CAD	41.0% (36.3%–45.8%; 176/429)	38.2% (28.1%–49.1%; 34/89)	P = 0.62
PPV of CAD (BI-RADS 3–5)	23.6% (22.2%–25.1%; 77/326)	26.7% (23.6%–30.0; 20/75)	P < 0.001
PPV of radiological reading (BI-RADS 3–5)*	15.2% (14.4%–16.1%; 77/506)	18.4% (15.9%–21.1%; 20/109)	P < 0.001

In parentheses are the 95% confidence intervals.

*Percentages can differ from earlier publication¹⁶ because not all examinations of screening round 2 were included.

CAD indicates computer-aided diagnosis; AUC, area under the receiver operating characteristic curve; PPV, positive predictive value; BI-RADS, breast imaging-reporting and data system.

Avoiding bias

Because all participants in the second screening round were also screened in the first round, bias may occur when the round-1 model is validated on the round 2 data. Hence, the round 1 model was retrained on the round 1 data to avoid such bias, following the steps outlined below.

To train the CAT model on the data from the first screening round, internal-external validation was used, meaning that the model was trained on data from 7 hospitals and tested on data from the eighth hospital, alternating such that each hospital was used once as an external test set. Internal-external validation thus yielded 8 models, and each model was constructed without overlap of women in training and test set. The overall performance of CAT in the first screening round was then estimated by averaging the performance of the 8 models. To validate the CAT on the data from the second screening round, we used the same 8 internal-external validation folds from the first round. That is, each of the 8 CAT models from the first round was applied—without additional training—to the corresponding hospital in the second round. Hence, no woman in the first or second round was ever assessed by a model that included their training data. Again, the overall performance of the CAT in the second round was then assessed by averaging the performance of the 8 models.

To validate the CAD, we also took measures to avoid overlap in women in the training and test set: for each lesion detected in the second screening round, the first-round training data of that patient were removed from the CAD model before they were applied to the second-round screening data.

Although the BI-RADS scores of radiologists are used by the CAD,¹⁴ these scores were ignored in this combined CAT/CAD assessment to mimic prospective autonomous application where radiologists have not yet assigned BI-RADS scores. For this purpose, the CAD model was established for each hospital separately using the first-round screening data without the BI-RADS scores.

Statistics

The performance of combined CAT and CAD was compared with that of unassisted radiological reading during the DENSE trial. This was done as follows: CAT either detects no lesions in an examination or lesions in 1 or both breasts. These occurrences were counted separately, but the results are presented at examination level, that is, whether 1 or more lesions are present in both breasts in the same examination according to the CAT. In addition, the number of correctly classified benign lesions by the CAD was counted. The false-positive rate of the CAD was compared with the false-positive rate of unassisted radiological reading (ie, the rate of recalled suspicious lesions that turned out to be benign) using McNemar tests. A P value of less than 0.05 was considered statistically significant.

The reproducibility of CAT and CAD separately was established by comparing the results from the first and second screening rounds. For this purpose, differences in area under the receiver operating characteristic (ROC) curve (AUC) were tested using the paired Student t test (8

CAT models) or DeLong test (1 CAD model). The percentage of examinations dismissed and the percentage of examinations with lesions that would be offered to radiologists by CAT were recorded and compared using paired Student t test.

The CAD developed on the first round was applied to BI-RADS 3, 4, and 5 lesions in round 2. In addition to AUC, PPV and percentage of correctly classified benign lesions were compared between rounds using McNemar tests. It was verified that the NPV of CAT and CAD for malignant disease is 100%, as established in the first screening round.

RESULTS

DENSE Trial and Unassisted Radiological Reading

In total, 2901 (84.4%) MRI examinations of 3436 women in the second screening round were included. A total of 535 women were

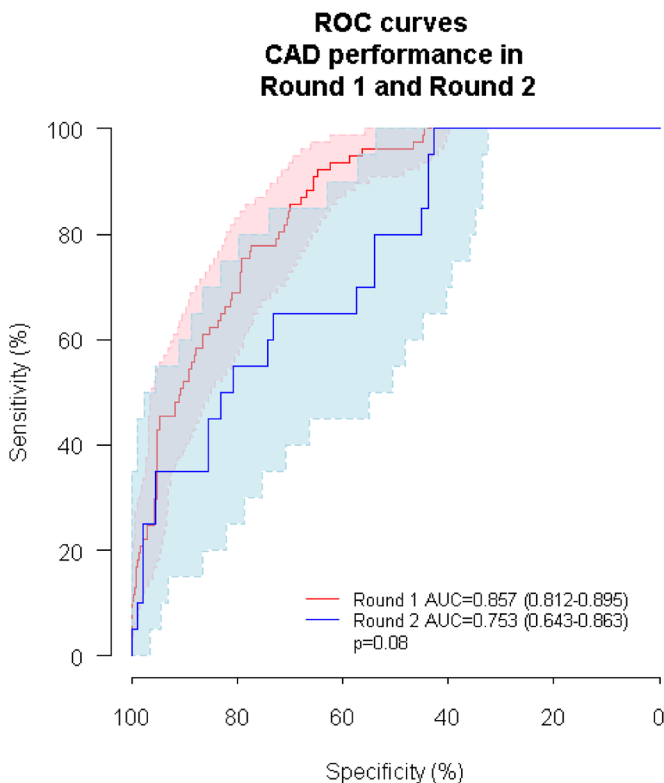


FIGURE 3. ROC curves of CAD for the task of distinguishing between benign and malignant lesions, applied to first and second screening round data. The shaded regions represent the 95% confidence intervals.

Downloaded from http://investigativeradiology.com/ on 04/14/2023

excluded because their data could not be retrieved in full from participating hospitals. Unassisted by CAT or CAD, radiologists reported 334 lesions in 303 (of 2901) women. Three women had 3 lesions and 25 had 2 lesions. Twenty lesions were malignant, and 314 were benign (Table 1). The lesions were scored BI-RADS 2 ($n = 225$), BI-RADS 3 ($n = 21$), BI-RADS 4 ($n = 82$), and BI-RADS 5 ($n = 6$).

Computer-Aided Triaging

The performance of CAT in the second screening round is shown in Table 2. Computer-aided triaging showed a smaller AUC in the second screening round than in the first screening round (0.76 vs 0.83, $P = 0.001$) (Fig. 2). We found no evidence of differences in performance at the operating threshold ($P = 0.70$). In the second round, 41.0% (95% confidence interval [CI], 30.4–51.6) of the examinations without any lesions would be dismissed compared with 39.7% (95% CI, 30.0–49.4) in the first screening round.⁶ The percentage of examinations with lesions that would continue to radiological review was also not different ($P = 0.07$) between the second and first screening rounds (85.6% [95% CI, 79.2–92.0] vs 90.7% [95% CI, 86.7–94.7], respectively). No examinations with malignant disease were dismissed, that is, NPV = 100%.

Computer-Aided Diagnosis

Computer-aided diagnosis, applied on all lesions in the dataset, classified 34 lesions (12 BI-RADS 3 and 22 BI-RADS 4) correctly as benign and 75 as malignant (7 BI-RADS 3, 62 BI-RADS 4, and 6 BI-RADS 5), of which 20 (17 BI-RADS 4 and 3 BI-RADS 5) were malignant at histology. An increase in PPV was observed compared with

unassisted radiological reading in both screening rounds ($P < 0.001$) (Table 3).

At the established operating threshold, the CAD shows no difference in results in classifying benign lesions in the second screening round compared with the first round ($P = 0.08$) (Table 3, Fig. 3).

Combination of CAT and CAD

Combined CAT and CAD confirm potential to dismiss a subset of examinations and lesions for further assessment without missing any malignant disease (Fig. 4). Computer-aided triaging would dismiss 950 of 2901 (32.7%) examinations. Of 950 dismissed examinations, 38 (4.0%) contained 1 or more benign lesions. None were malignant.

In the remaining 1951 examinations, 265 examinations contained 285 lesions. Computer-aided diagnosis classified 132 of 285 (46.3%) of these lesions as benign. No malignant lesions were called benign. At best, the combination of CAT and CAD would yield 53 of 109 (48.6%) false-positive referrals to additional MRI screening or biopsy, compared with 89 of 109 (78.9%) for radiologists without computer assistance ($P = 0.001$).

DISCUSSION

Adding MRI to breast cancer screening programs for women with extremely dense breasts will result in increased workload for radiologists. To reduce this workload and also the false-positive rate, methods were developed for CAT and CAD.^{6,14} Here, we show that CAT and CAD developed on data from the first screening round of the DENSE trial reproduce robustly in the second screening round: combined CAT and CAD have the potential to reduce the workload of radiologists by 32.7% (950/2901) by

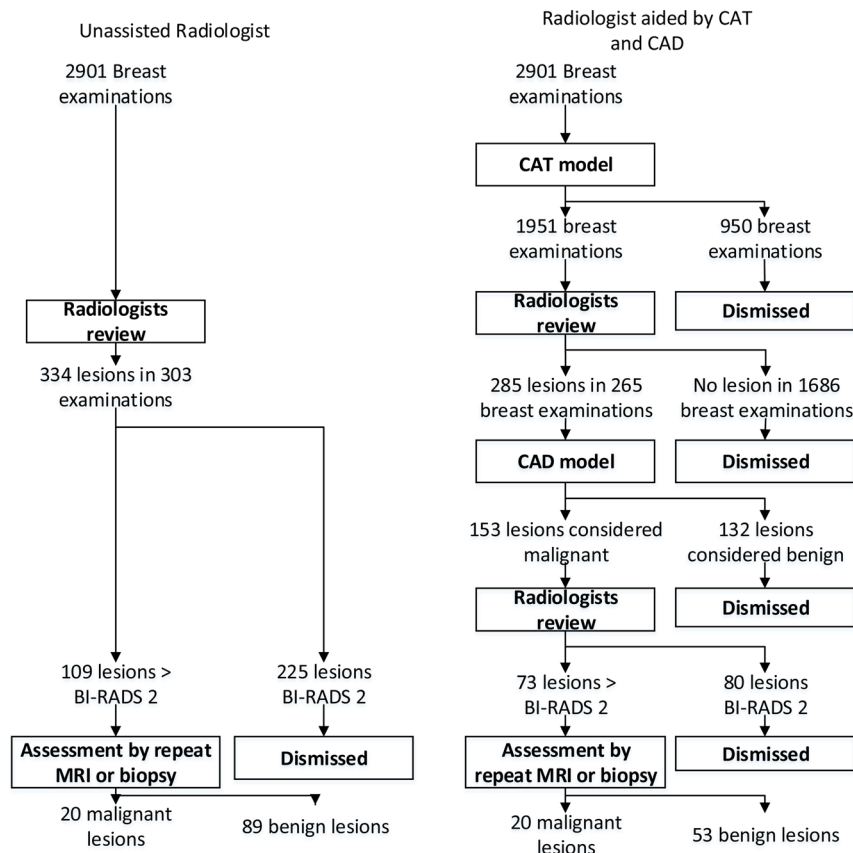


FIGURE 4. A comparison of the workload of the radiologist in terms of reading and workup with and without computerized analysis. Note: Two benign lesions that CAD would classify as probably malignant were dismissed by CAT, which explains why CAD classifies only 73 of the 75 as malignant, as stated in the CAD Results section.

dismissing normal examinations without dismissing cancers. In the remaining scans considered to require reading, 132 of 285 (46.3%) lesions were correctly identified as benign without missing cancers.

The combination of methods shows potential to reduce false-positive referrals by 40.4% (36/89).

Of the 89 benign lesions referred to biopsy or additional MRI after unassisted radiological reading, 3 benign lesions could be dismissed by CAT (3.4%), followed by 33 by CAD (36.0%), totaling 39.4%. To the best of our knowledge, no other groups than Verburg et al¹⁴ and den Dekker et al²⁰ have reported on reduction of false-positive referrals in the MR screening of women with extremely dense breasts. For other breast MRI indications, on the basis of smaller and more heterogeneous populations, reductions of 12 of 24 (50.0%) with CAD²¹ and 17 of 24 (70.8%) have been reported using proton MR spectroscopy.²² Although several studies reported on false-positive rates for CAD, none described the number of correctly identified normal breasts.^{5,23–27}

The strategy to detect presence of cancer indirectly by training a CNN on BI-RADS score has also been used by other investigators, for example, in mammography.²⁸

Although the percentage of dismissed examinations without lesion did not differ between screening rounds ($P = 0.70$), the AUC showed a difference ($P = 0.001$). This indicates that differences are present in the appearance of lesions in round 1 and those in round 2. This may be caused by differences between a prevalent screening round and an incident screening round: lesions in the incident round became visible in a time span of 2 years, where lesions in the prevalent round may have existed for a longer period of time. The observation that the AUC differed, but not the NPV, indicates that the difference in lesion appearance between first and second screening round primarily affected the shape of the lower part of the ROC curve, but not the upper, high-sensitivity, part.

Whereas the concept of automatically dismissing normal breast MRI examinations is attractive to reduce radiologist workload, and may be feasible from a technical perspective, challenges remain to clinical implementation. In current practice, every screening image has to be interpreted by a trained physician. Before the required paradigm shift for implementation would be accepted by patients, clinicians and policymakers must address multiple issues like safety, accountability, and quality.²⁹

This study also has limitations. Participants of the second round of the DENSE trial also participated in the first round; data were acquired in the same hospitals using the same MRI devices with identical sequences. The number of malignant lesions in the second screening round was limited. Although steps were taken to minimize the impact of potential bias, to further investigate robustness, methods should be tested on data acquired under more various conditions.

Future studies could focus on application of presented methodology to other screening populations, such as women at high lifetime risk of developing breast cancer. Also, the level of automation can be further increased; current operator involvement was 2-fold: providing manual location of lesions for the CAD and manual identification of BI-RADS 2 lesions after CAD. Future research will focus on automating these steps as well.

In conclusion, combining CAT and CAD has the potential to both reduce workload and reduce the number of biopsies without dismissing malignant breast disease.

ACKNOWLEDGMENTS

The authors thank the registration team of the Netherlands Comprehensive Cancer Organisation (IKNL) for the collection of data for the Netherlands Cancer Registry.

REFERENCES

- Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI screening for women with extremely dense breast tissue. *N Engl J Med*. 2019;381:2091–2102.
- Saadatmand S, Geuzinge HA, Rutgers EJ, et al. MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc): a multicentre, randomised, controlled trial. *Lancet Oncol*. 2019;20:1136–1147.

- Menezes GL, Knuttel FM, Stehouwer BL, et al. Magnetic resonance imaging in breast cancer: a literature review and future perspectives. *World J Clin Oncol*. 2014;5:61–70.
- Maicas C, Carneiro G, Bradley AP, et al. *Deep Reinforcement Learning for Active Breast Lesion Detection from DCE-MRI*. Cham, Switzerland: Springer International Publishing; 2017:665–673.
- Vignati A, Giannini V, De Luca M, et al. Performance of a fully automatic lesion detection system for breast DCE-MRI. *J Magn Reson Imaging*. 2011;34:1341–1351.
- Verburg E, van Gils CH, van der Velden BHM, et al. Deep learning for automated triaging of 4581 breast MRI examinations from the DENSE trial. *Radiology*. 2022;302:29–36.
- Gilhuijs KGA, Giger ML, Bick U. Computerized analysis of breast lesions in three dimensions using dynamic magnetic-resonance imaging. *Med Phys*. 1998;25:1647–1654.
- Honda E, Nakayama R, Koyama H, et al. Computer-aided diagnosis scheme for distinguishing between benign and malignant masses in breast DCE-MRI. *J Digit Imaging*. 2016;29:388–393.
- Rahbar H, Partridge SC. Multiparametric MRI imaging of breast Cancer. *Magn Reson Imaging Clin N Am*. 2016;24:223–238.
- Dalmiş MU, Gubern-Mérida A, Vreemann S, et al. Artificial intelligence-based classification of breast lesions imaged with a multiparametric breast MRI protocol with ultrafast DCE-MRI, T2, and DWI. *Invest Radiol*. 2019;54:325–332.
- Truhn D, Schrading S, Haarbuerger C, et al. Radiomic versus convolutional neural networks analysis for classification of contrast-enhancing lesions at multiparametric breast MRI. *Radiology*. 2019;290:290–297.
- Gallego-Ortiz C, Martel AL. Using quantitative features extracted from T2-weighted MRI to improve breast MRI computer-aided diagnosis (CAD). *Plos One*. 2017;12:e0187501.
- Bhooshan N, Giger M, Lan L, et al. Combined use of T2-weighted MRI and T1-weighted dynamic contrast-enhanced MRI in the automated analysis of breast lesions. *Magn Reson Med*. 2011;66:555–564.
- Verburg E, van Gils CH, Bakker MF, et al. Computer-aided diagnosis in multiparametric magnetic resonance imaging screening of women with extremely dense breasts to reduce false-positive diagnoses. *Invest Radiol*. 2020;55:438–444.
- Emaus MJ, Bakker MF, Peeters PH, et al. MR imaging as an additional screening modality for the detection of breast Cancer in women aged 50–75 years with extremely dense breasts: the DENSE trial study design. *Radiology*. 2015;277:527–537.
- Veenhuizen SGA, de Lange SV, Bakker MF, et al. Supplemental breast MRI for women with extremely dense breasts: results of the second screening round of the DENSE trial. *Radiology*. 2021;299:278–286.
- Morris EA, Comstock CE, Lee CH, et al. ACR BI-RADS® magnetic resonance imaging. *ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System*. 5. 2013.
- Alderliesten T, Schlieff A, Peterse J, et al. Validation of semiautomatic measurement of the extent of breast tumors using contrast-enhanced magnetic resonance imaging. *Invest Radiol*. 2007;42:42–49.
- Intergraal Kankercentrum Nederland. Oncoline. 2017. https://www.oncoline.nl/index.php?pagina=/richtlijn/item/pagina.php&id=41189&richtlijn_id=1069. Accessed September 1, 2019.
- den Dekker BM, Bakker MF, de Lange SV, et al. Reducing false-positive screening MRI rate in women with extremely dense breasts using prediction models based on data from the DENSE trial. *Radiology*. 2021;301:283–292.
- Lehman CD, Peacock S, DeMartini WB, et al. A new automated software system to evaluate breast MR examinations: improved specificity without decreased sensitivity. *Am J Roentgenol*. 2006;187:51–56.
- Clauser P, Marcon M, Dietzel M, et al. A new method to reduce false positive results in breast MRI by evaluation of multiple spectral regions in proton MR-spectroscopy. *Eur J Radiol*. 2017;92:51–57.
- Gubern-Mérida A, Martí R, Melendez J, et al. Automated localization of breast cancer in DCE-MRI. *Med Image Anal*. 2015;20:265–274.
- Dalmiş MU, Vreemann S, Kooi T, et al. Fully automated detection of breast cancer in screening MRI using convolutional neural networks. *J Med Imaging*. 2018;5:014502.
- Renz DM, Böttcher J, Diekmann F, et al. Detection and classification of contrast-enhancing masses by a fully automatic computer-assisted diagnosis system for breast MRI. *J Magn Reson Imaging*. 2012;35:1077–1088.
- Chang Y-C, Huang Y-H, Huang C-S, et al. Computerized breast lesions detection using kinetic and morphologic analysis for dynamic contrast-enhanced MRI. *Magn Reson Imaging*. 2014;32:514–522.
- Ayatollahi F, Shokouhi SB, Mann RM, et al. Automatic breast lesion detection in ultrafast DCE-MRI using deep learning. *Med Phys*. 2021;48:5897–5907.
- Schönenberger C, Hejduk P, Cirtsis A, et al. Classification of mammographic breast microcalcifications using a deep convolutional neural network: a BI-RADS-based approach. *Invest Radiol*. 2021;56:224–231.
- Joe BN. AI to dismiss normal breast MRI scans and reduce workload. *Radiology*. 2022;302:37–38.