

BREAST CANCER: MACHINE LEARNING AIMED AT
TRIAGING FOR SCREENING, DIAGNOSTIC WORKFLOW,
AND TREATMENT

Hui Wang

Breast cancer: machine learning aimed at triaging for screening, diagnostic workflow, and treatment

Hui Wang

Thesis, Utrecht University, The Netherlands

ISBN: 978-94-6483-533-5

DOI: 10.33540/2053

Printing: Ridderprint | www.ridderprint.nl

© Copyright 2023: Hui Wang

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, by photocopying, recording, or otherwise, without the prior written permission of the author. The copyright of the articles that have been published has been transferred to the respective journals.

Breast cancer: machine learning aimed at triaging for screening, diagnostic workflow, and treatment

Borstkanker: machine learning gericht op triage voor screening, diagnostische workflow en behandeling
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

dinsdag 12 december 2023 des ochtend te 10.15 uur

door

Hui Wang

geboren op 3 Juli 1990

te JiangSu, China

Promotor:

Prof. dr. ir. M.A. Viergever

Copromotoren:

Dr. K.G.A. Gilhuijs

Dr. ir. B.H.M. van der Velden

Beoordelingscommissie:

Prof. dr. M. Broeders

Prof. dr. P.J. van Diest (voorzitter)

Dr. M.R. Moman

Prof. dr. H.M. Verkooijen

Prof. dr. E. van der Wall

Table of Contents

<i>Chapter 1</i>	7
<i>Introduction</i>	7
<i>Chapter 2</i>	15
<i>Automated Rating of Background Parenchymal Enhancement in MRI of Extremely Dense Breasts without Compromising the Association with Breast Cancer in the DENSE Trial.....</i>	15
<i>Chapter 3</i>	31
<i>Assessing Quantitative Parenchymal Features on Baseline Dynamic Contrast-enhanced MRI and Cancer Occurrence in Women with Extremely Dense Breasts</i>	31
<i>Chapter 4</i>	51
<i>Towards Computer-Assisted Triaging of MRI-Guided Biopsy in Preoperative Breast-Cancer Patients.....</i>	51
<i>Chapter 5</i>	71
<i>Synchronous Breast Cancer: Phenotypic Similarities on MRI</i>	71
<i>Chapter 6</i>	87
<i>Summary and Discussion</i>	87
List of publications.....	100
Curriculum Vitae.....	103

Chapter 1

Introduction

Introduction

Breast cancer and breast cancer imaging

Breast cancer is the most frequent malignancy in female population (1) and primary cause of death among women worldwide (2). Proper management of breast diseases is therefore very important.

The most commonly used imaging modality for breast without doubt is mammography, while its advantage of high sensitivity is attenuated in women with a large relative fraction of fibroglandular tissue (FGT), i.e., dense breasts (3). Ultrasound is typically used to characterize the screen-detected breast lesions, or to identify and biopsy axillary lymph nodes in women who are suspected of having breast cancer (4). It is auxiliary to mammography, whereas it is highly dependent on the operator's experience and tuning of ultrasound parameters (5).

In contrast, MRI has higher sensitivity than mammography and ultrasound for detection of invasive breast cancer (6,7). Dynamic contrast enhanced MRI (DCE-MRI) has been regarded as the most powerful tool for breast cancer detection, but DCE-MRI is more expensive and more time-consuming, it is therefore only recommended for women who meet the indications.

Indications of breast MRI

1	Screening of women at high risk of breast cancer
2	Preoperative staging of newly diagnosed breast cancer (ipsilateral and contralateral)
3	Evaluation of the effect of neoadjuvant chemotherapy
4	Evaluation of women with breast implants
5	Occult primary breast carcinoma (search for breast cancer in patients with metastases and negative mammography and ultrasound)
6	Suspected local recurrence*
7	Problem solving (equivocal findings at mammography/ultrasound) *
*When needle biopsy cannot be performed Other new indications were recently proposed, such as nipple discharge and evaluation of lesions with uncertain malignant potential (so-called high-risk or B3 lesions) detected at mammography or ultrasound, and needle-biopsied under their guidance.	

According to the recommendations issued by the European Society of Breast Imaging (EUSOBI), the main indications of breast MRI are listed in **Table 1(8)**. The indications may vary in different countries due to differences in local guideline, insurance policy, health care system, etc.

Toward triaging of women for Breast MRI screening

Many models have been developed to evaluate the risk of developing breast cancer and to further determine whether the woman is eligible for receiving screening breast MRI. These models include IBIS also called Tyrer-Cuzick, BRCAPRO (based on internet, free for research and counseling use), and the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA), etc. (9–11). Evidence has shown that the Tyrer-Cuzick model is the most comprehensive one for predicting breast cancer risk (12). However, background parenchymal enhancement (BPE), a known risk factors of breast cancer, has not been taken into account and it is not well validated in any of these commonly used models. Therefore, incorporating novel and verified risk factors into the future models is indispensable, making efforts toward triaging of patients to breast MRI screening and realizing tailored screening plans such that a subgroup of women can really benefit from breast MRI screening.

Towards triaging for diagnostic workflow of breast cancer

After a breast cancer is detected, preoperative MRI is sometimes used to assess the tumor extent. Due to its superior sensitivity, MRI usually detects additional lesions undetected by mammography and ultrasound, around half of which are malignant (13). In this case, MRI-guided biopsy may be indicated to obtain pathological proof. Different from diagnostic MRI, MRI-guided biopsy requires dedicated equipment and trained personnel, and is relatively time and resource intensive. In addition, some cases are not eligible for MRI-guided biopsy due to difficulty in targeting lesions near the thorax. Management of these additional lesions would be greatly facilitated if an approach exists to identify malignancy with high specificity without MRI-guided biopsy.

Towards triaging for breast cancer treatment

When an additional breast cancer is detected simultaneously with an index breast cancer, but physically separated from the index breast cancer, we call them synchronous breast cancers (14) (**Figure 1**). It has been observed that discrepancies in prognostic markers between the index cancers and their synchronous counterparts may result in worse long-term survival and have

impact on systemic treatment of patients (15–18). Such women are usually evaluated by MRI, it is of great interest to know whether such discrepancy also exists in appearance of breast cancer on MRI (i.e., phenotype), and is also associated with prognosis. If so, the imaging phenotypes of breast cancer and the synchronous breast cancers on MRI can offer opportunity to triaging of women to breast cancer treatment.

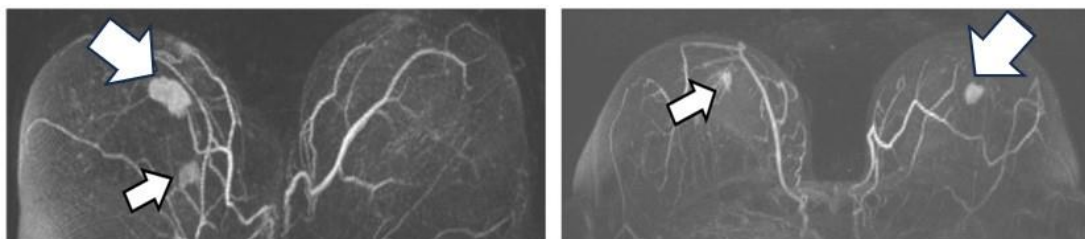


FIGURE 1. Images show the maximum intensity projections of the index breast cancer (thick arrow) and the corresponding synchronous cancer (thin arrow) in the ipsilateral and contralateral breast.

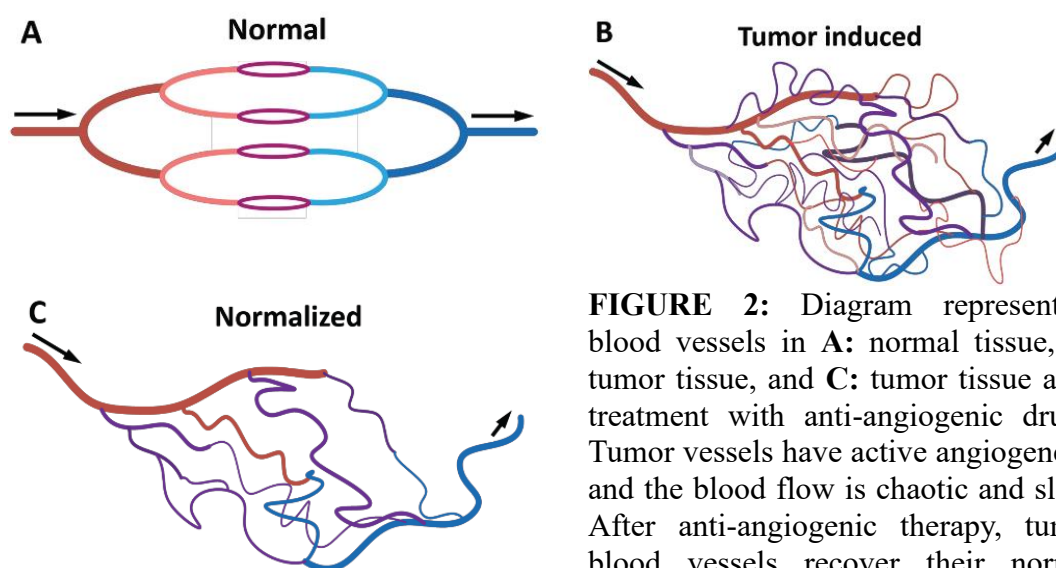


FIGURE 2: Diagram representing blood vessels in **A:** normal tissue, **B:** tumor tissue, and **C:** tumor tissue after treatment with anti-angiogenic drugs. Tumor vessels have active angiogenesis and the blood flow is chaotic and slow. After anti-angiogenic therapy, tumor blood vessels recover their normal functions. (Source: DOI: 10.18632/oncotarget.16482)

Mechanism of Breast MRI

As the most accurate imaging modality for breast cancer in all risk groups (19), the underlying mechanism of DCE-MRI to detect invasive breast cancer with high sensitivity is based on the fact that breast cancer cannot grow beyond 2 mm in size without creating new blood vessels to provide nutrients for the tumor (20). These newly created vessels resulting from angiogenesis are leaky. While the gadolinium-based contrast agents used in DCE-MRI are relatively large molecules, they can easily extravasate from such new leaky vessels, and thus rapidly accumulate in the breast cancer stroma (21). Therefore, DCE-MRI is

sensitive to changes in blood volume and vascular permeability, and is associated with angiogenesis (22) (**Figure 2**). Breast cancers are heterogenous in nature, the imaging phenotype of breast cancer on DCE-MRI reflecting the heterogeneous nature of angiogenesis can therefore serve as useful imaging biomarker.

Machine learning and computer-aided triaging

In daily clinical practice, the main undertaking of radiologists is medical image interpretation. The evaluation of breast lesions is largely qualitative and subjective in assessment of lesion shape, enhancement type, etc, resulting in inter- and intra-observer variations (23–25), this is a notable limitation. To overcome this shortcoming, a more reproducible evaluation method conducive for objective decision-making is necessary. Artificial intelligence (AI) provides the radiologists with opportunity to convert medical images to quantitative data (26,27). AI-aided systems such as Computer-Aided Detection (CADe) and Computer-Aided Diagnosis (CADx) have been under development for decades, and various machine learning models based on computer-extracted features have been developed for risk assessment, detection, diagnosis, prognosis, treatment response monitoring, etc. Although there is no one-size-fit-all when it comes to complicated clinical scenarios, machine learning provides an opportunity to improve the clinical decision-making in medical image interpretation (**Figure 3**).

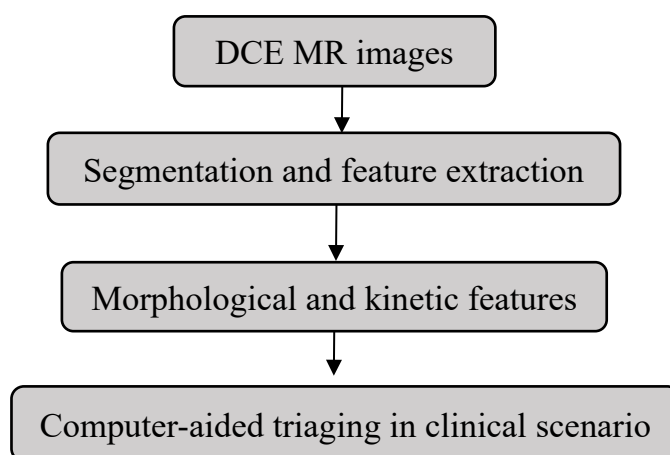


FIGURE 3. Schematic diagram of computer-aided triaging pipeline for breast cancers on DCE-MRI.

Thesis outline

The aim of this thesis is to investigate whether the imaging phenotype of healthy breast tissue and breast lesions on DCE-MRI have potential to improve the decision making in screening, diagnostic workflow and treatment.

In **Chapter 2**, we automatically assess the category of a known breast cancer risk factor - background parenchymal enhancement (BPE) in breast DCE-MRI without compromising the association between BPE and breast cancer occurrence.

In **Chapter 3**, we automatically identify quantitative properties of the breast parenchyma on baseline DCE-MRI scans and assess their association with breast cancer occurrence, so as to triage a subgroup of women with extremely dense breasts to ultimately receive a more tailored screening plan.

In **Chapter 4**, we identify malignant disease extension with near-perfect specificity in a consecutively series of preoperative patients originally indicated for a breast biopsy, as a means to realize computer-assisted triaging of women for MRI-guided breast biopsy of additional lesions in preoperative setting.

In **Chapter 5**, we investigate whether the MRI phenotype of breast cancer is different from that of additional breast cancer (i.e., synchronous cancer), and whether such a difference, if it exists, is associated with prognosis, so as to triage women to receive more tailored treatment.

In **Chapter 6**, we discuss the research findings, and prospects of our future research directions.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*. 2021;71(3):209–249.
2. Harbeck N, Gnant M. Breast cancer. *The Lancet*. Elsevier Ltd; 2017;389(10074):1134–1150.
3. Mann RM, Kuhl CK, Moy L. Contrast-enhanced MRI for breast cancer screening. *Journal of Magnetic Resonance Imaging*. 2019;50(2):377–390.
4. Houssami N, Ciatto S, Turner RM, Cody HS, MacAskill P. Preoperative ultrasound-guided needle biopsy of axillary nodes in invasive breast cancer: Meta-analysis of its accuracy and utility in staging the axilla. *Annals of Surgery*. 2011;254(2):243–251.
5. Kim HJ, Kim SM, Kim B, et al. Comparison of strain and shear wave elastography for qualitative and quantitative assessment of breast masses in the same population. *Scientific Reports*. 2018;8(1):1–11.
6. Morrow M, Harris JR. More Mastectomies: Is This What Patients Really Want? *Journal of Clinical Oncology*. 2009;27(25):4038–4040.
7. Houssami N, Ciatto S, Macaskill P, et al. Accuracy and surgical impact of magnetic resonance imaging in breast cancer staging: Systematic review and meta-analysis in detection of multifocal and multicentric cancer. *Journal of Clinical Oncology*. 2008;26(19):3248–3258.
8. Mann RM, Balleyguier C, Baltzer PA, et al. Breast MRI: EUSOBI recommendations for women’s information. *European Radiology*. 2015;25(12):3669–3678.
9. Saslow D, Boetes C, Burke W, et al. American cancer society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J Clin*. 2007;57(2):75–89.
10. Smith RA, Cokkinides V, Brawley OW. Cancer Screening in the United States, 2012 A Review of Current American Cancer Society Guidelines and Current Issues in Cancer Screening. 2012;62(2):129–142.
11. Antoniou AC, Cunningham AP, Peto J, et al. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: Updates and extensions. *British Journal of Cancer*. 2008;98(8):1457–1466.
12. Quante AS, Whittemore AS, Shriver T, Strauch K, Terry MB. Breast cancer risk assessment across the risk continuum: Genetic and nongenetic risk factors contributing to differential model performance. *Breast Cancer Research*. 2012;14(6):1–12.
13. Leung JWT. Second-look ultrasound: Only for biopsy or more? *European Journal of Radiology*. 2012;81(SUPPL1):87–89.

14. Jain S, Rezo A, Shadbolt B, Dahlstrom JE. Synchronous multiple ipsilateral breast cancers: Implications for patient management. *Pathology*. 2009;41(1):57–67.
15. Garimella V, Long ED, O’Kane SL, Drew PJ, Cawkwell L. Oestrogen and progesterone receptor status of individual foci in multifocal invasive ductal breast cancer. *Acta Oncologica*. 2007;46(2):204–207.
16. Buggi F, Folli S, Curcio A, et al. Multicentric/multifocal breast cancer with a single histotype: Is the biological characterization of all individual FOCI justified? *Annals of Oncology*. 2012;23(8):2042–2046.
17. Desmedt C, Fumagalli D, Pietri E, et al. Uncovering the genomic heterogeneity of multifocal breast cancer. *Journal of Pathology*. 2015;236(4):457–466.
18. Navale P, Bleiweiss IJ, Jaffer S, Nayak A. Evaluation of biomarkers in multiple ipsilateral synchronous invasive breast carcinomas. *Archives of Pathology and Laboratory Medicine*. 2019;143(2):190–196.
19. Kuhl CK, Strobel K, Bieling H, Leutner C, Schild HH, Schrading S. Supplemental breast MR imaging screening of women with average risk of breast cancer. *Radiology*. 2017;283(2):361–370.
20. Carmeliet P, Jain RK. Angiogenesis in cancer and other diseases. *Nature*. 2000;407(6801):249–257.
21. Knopp M V., Weiss E, Sinn HP, et al. Pathophysiologic basis of contrast enhancement in breast tumors. *Journal of Magnetic Resonance Imaging*. 1999;10(3):260–266.
22. Hylton N. Dynamic contrast-enhanced magnetic resonance imaging as an imaging biomarker. *Journal of Clinical Oncology*. 2006;24(20):3293–3298.
23. Kuhl CK, Schrading S, Leutner CC, et al. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. *Journal of Clinical Oncology*. 2005;23(33):8469–8476.
24. Arasu VA, Miglioretti DL, Sprague BL, et al. Population-based assessment of the association between magnetic resonance imaging background parenchymal enhancement and future primary breast cancer risk. *Journal of Clinical Oncology*. 2019;37(12):954–963.
25. Dontchos BN, Rahbar H, Partridge SC, et al. Are Qualitative Assessments of Background Parenchymal Enhancement, Amount of Fibroglandular Tissue on MR Images, and Mammographic Density Associated with Breast Cancer Risk? *Radiology*. 2015;276(2):371–380.
26. Giger ML, Karssemeijer N, Schnabel JA. Breast image analysis for risk assessment, detection, diagnosis, and treatment of cancer. *Annual Review of Biomedical Engineering*. 2013;15(April):327–357.
27. Giger ML. Machine Learning in Medical Imaging. *Journal of the American College of Radiology*. 2018;15(3):512–520.

Chapter 2

Automated Rating of Background Parenchymal Enhancement in MRI of Extremely Dense Breasts without Compromising the Association with Breast Cancer in the DENSE Trial

Based on:

Wang H, Van Der Velden BHM, Verburg E, Bakker MF, Pijnappel RM, Veldhuis WB, Gils CH van, Gilhuijs, KGA. Automated rating of background parenchymal enhancement in MRI of extremely dense breasts without compromising the association with breast cancer in the DENSE trial. Submitted

Abstract

Background and Objective: Background parenchymal enhancement (BPE) on dynamic contrast-enhanced MRI (DCE-MRI) as rated by radiologists is subject to inter- and intrareader variability. We aim to automate BPE category from DCE-MRI.

Materials and Methods: This study represents a secondary analysis of the Dense Tissue and Early Breast Neoplasm Screening trial. 4553 women with extremely dense breasts who received supplemental breast MRI screening in eight hospitals were included. Minimal, mild, moderate and marked BPE rated by radiologists were used as reference. Fifteen quantitative MRI features of the fibroglandular tissue were extracted to predict BPE using Random Forest, Naïve Bayes, and KNN classifiers. Majority voting was used to combine the predictions. Internal-external validation was used for training and validation. The inverse-variance weighted mean accuracy was used to express mean performance across the eight hospitals. Cox regression was used to verify non inferiority of the association between automated rating and breast cancer occurrence compared to the association for manual rating.

Results: The accuracy of majority voting ranged between 0.56 and 0.84 across the eight hospitals. The weighted mean prediction accuracy for the four BPE categories was 0.76. The hazard ratio (HR) of BPE for breast cancer occurrence was comparable between automated rating and manual rating (HR=2.12 versus HR=1.97, P=0.65 for mild/moderate/marked BPE relative to minimal BPE).

Conclusion: It is feasible to rate BPE automatically in DCE-MRI of women with extremely dense breasts without compromising the underlying association between BPE and breast cancer occurrence. The accuracy for minimal BPE is superior to that for other BPE categories.

Introduction

Women aged between 50 and 75 years in the Netherlands are screened for breast cancer every two years using mammography (1). Mammography is, however, less sensitive to detect breast cancer in women with extremely dense breasts (i.e., women in category d of the Breast Imaging Reporting and Data System (BIRADS) lexicon) (2). Women with extremely dense breast have three to six times higher risk of developing breast cancer compared to women with almost entirely fatty breasts (i.e., women in category a of the BIRADS lexicon) (3,4). A recent randomized controlled trial-- Dense Tissue and Early Breast Neoplasm Screening (DENSE) investigated whether MRI has complementary value for the detection of breast cancer in a mammography-screening population of women aged 50–75 years with extremely dense breasts. Based on the results of this national multi-institutional study, it would be helpful to identify other risk factors in addition to breast density that could be used to personalize breast cancer screening and reduce the workload of radiologists and costs for the community.

Evaluation of background parenchymal enhancement (BPE) has recently gained more attention due to its association with breast-cancer risk. The amount of BPE in dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) was shown to be an independent risk factor for breast cancer development (5–7). In a group of women at high risk of breast cancer but without a history of the disease, researchers found that those with mild, moderate or marked BPE had a nine times higher risk of developing breast cancer compared with those with minimal BPE (8). Moreover, BPE was found to be associated with treatment response (9–12) and prognosis (13–15). These findings suggest that BPE can potentially serve as an important biomarker of risk and prognosis. However, radiologists currently lack objective tools for assessing the category of BPE.

In clinical practice, BPE is evaluated qualitatively by radiologists using the BIRADS categories of minimal, mild, moderate, or marked. This assessment is, however, subject to intra- and inter- reader variability. Kappa values ranged from moderate to almost perfect for intra-reader agreement, and from fair to almost perfect for inter-reader agreement (6,16,17).

The aim of this study is to investigate whether automated rating of BPE category on DCE- MRI is feasible in women with extremely dense breasts using machine learning without compromising underlying associations between BPE and breast cancer occurrence.

Materials and Methods

Study population

This is a secondary analysis of data from the DENSE trial(1,18,19). The DENSE trial has been approved by the Dutch Minister of Health, Welfare and Sport (2011/19 WBO, The Hague, the Netherlands). The study including primary and secondary objectives was waived from ethical review by the local institutional review board based on the Dutch law on population studies. The DENSE trial was designed to compare the effectiveness of breast cancer screening with mammography alone versus mammography and MRI in women aged 50-75 years with extremely dense breasts (i.e., breast density category d at mammography) (20). The participants of this study were recruited from eight Dutch hospitals without abnormality on mammography at the time of inclusion. The MRI screenings were performed in three consecutive rounds starting in 2011 and concluding with the third-round screening in 2021. Written informed consent was obtained from all women who underwent MRI. In the current study, the first-round screening images only were analyzed. Women were excluded if they had bilateral cancer or had unilateral cancer and had previously undergone contralateral mastectomy.

TABLE 1. Imaging Parameters of the DCE-MRI Used in the DENSE Trial

Hospital ID	No. of women	MRI Device	Reconstructed Voxel size(mm ³)	Dimensions (Voxels)	FS
1	1579	Philips Achieva/ Ingenia	0.89*0.89*0.9	384*384*200	Yes
2	386	Siemens Magnetom Trio/ Skyra/ Prisma	0.80*0.80*1.0	448*448*176	No
3	237	Philips Achieva	0.89*0.89*0.9	384*384*200	Yes
4	490	Philips Ingenia	0.89*0.89*0.9	384*384*200	Yes
5	304	Philips Achieva/ Ingenia	0.89*0.89*0.9	384*384*200	Yes
6	517	Siemens Verio	0.85*0.85*1.0	448*448*176	No
7	402	Philips Ingenia	0.89*0.89*0.9	384*384*200	Yes
8	638	Siemens Skyra	0.80*0.80*1.0	448*448*160	No

DCE-MRI= Dynamic contrast-enhanced magnetic resonance imaging
DENSE=Dense Tissue and Early Breast Neoplasm Screening.
FS= fat suppression

MRI acquisition

The study participants underwent MRI screening using either a Philips 3.0T scanner (Achieva or Ingenia) or a Siemens 3.0T scanner (Trio, Verio, or Skyra). The MRI screening was conducted in the prone position and included a dynamic series with a precontrast image and 4 or 5 postcontrast images after using a gadolinium-based contrast medium (gadobutrol, Gadovist; Bayer Healthcare, Germany). Contrast agent was injected at a rate of 1 mL/s to a total dose of 0.1 mmol per kilogram of body weight. Fat suppression was optional during DCE-MRI acquisition. The imaging parameters have been described in detail elsewhere (21) and are summarized in **Table 1**.

Image analysis

The image analysis included deformable image registration (22), breast segmentation (23), FGT segmentation, and feature extraction. The FGT was automatically segmented from the precontrast images using an nnU-Net (24) convolutional neural network in an iterative learning approach (7). The image analysis was conducted using Python (version 2.7; Python Software Foundation, Beaverton, OR) and PyTorch (version 1.5).

Feature extraction

Fifteen quantitative features describing the spatiotemporal characteristics of FGT were extracted, including volumetric density, volumetric morphology (i.e., the three-dimensional shape of the FGT), and enhancement characteristics. These features have been described in great detail in a previous study (7) and are summarized in **Table 2**.

Manual rating of BPE category

In the DENSE trial, BPE category was evaluated by experienced breast radiologists with experience ranging between 5 and 23 years among 8 hospitals using the BIRADS MRI lexicon (25). The radiologists rated the BPE category in four categories: minimal, mild, moderate, and marked BPE. When BPE assessment between breasts of the same participant were inconsistent, the breast with the greater amount of BPE was used for the overall assessment. The experience of these radiologists in reading breast MRI ranged between 5 and 23 years at the start of the trial.

Circularity	Ratio between the intersection of the FGT with a sphere of equivalent volume as the FGT and the volume of the FGT itself
Irregularity	Ratio between the surface of a sphere with the same volume as the FGT and the surface of the FGT
Density	Ratio between the volume of FGT and the volume of the whole breast
Top 10% /50%/90% intensity of early contrast uptake	Top 10%/ 50%/ 90% early enhancing intensity of the contralateral healthy breast (women with cancer) or mean value of top 10%/ 50%/ 90% of both breasts (women without cancer)
Top 10%/ 50%/ 90% volume of early enhancing voxels	Volume of the top 10%/ 50%/ 90% early enhancing voxels of the contralateral healthy breast (women with cancer) or mean value of top 10%/ 50%/ 90% of both breasts (women without cancer)
Top 10% /50%/ 90% intensity of late contrast uptake	Top 10% /50%/90% late enhancing intensity of the contralateral healthy breast (women with cancer) or mean value of top 10%/ 50%/ 90% of both breasts (women without cancer)
Top 10% /50%/ 90% volume of late enhancing voxels	Volume of the top 10% /50%/ 90% late enhancing voxels of the contralateral healthy breast (women with cancer) or mean value of top 10% /50%/ 90% of both breasts (women without cancer)
FGT= fibroglandular tissue	

Statistical analysis

Multiple imputation was used to impute missing values. Ten imputation sets were merged based on the mean imputed values (26–28). Outliers were defined based on Tukey’s method, i.e., 1.5 interquartile range below the 25th percentile or 1.5 interquartile range above the 75th percentile. Outliers were winsorized to the nearest whisker after which the feature values were normalized between 0 and 1.

Internal-external validation was used for training and validation. In short, seven of eight hospitals were used to train the machine-learning method using 5-fold cross validation, and the remaining hospital was then used to perform external validation. This process was repeated eight times, with each hospital serving as the external validation hospital once. The results were merged. In each internal validation round, bootstrapping was performed with 100 bootstrap cycles to assess the uncertainty. In each bootstrap cycle, three classifiers (i.e., Random Forest, Naïve Bayes, and KNN) were constructed separately to predict BPE category. Majority voting (29–31) was used to combine the prediction outcome

of these three classifiers.

The accuracy of the majority voting on the external validation set was used as metric to evaluate prediction performance in each bootstrap cycle. Inverse-variance weighted mean accuracy was used to merge the external validation performance of these eight hospitals (32).

Cox regression was used to verify if the association between automatically rated BPE and breast cancer occurrence – expressed as Hazard Ratio - was comparable to that observed between manual rating of BPE and breast cancer occurrence.

The statistical analyses were performed in R (version 4.1.2; R Foundation for Statistical Computing, Vienna, Austria), the classifier construction process was performed using the Caret package (version 6.0-92).

Results

Study population

In total, the data from 4783 women in the first round of the DENSE trial were available from the eight hospitals. Women with incomplete digital MRI data or who met the exclusion criteria were excluded (**Figure 1**). The mean age of the remaining 4553 study participants with extremely dense breasts was 56 years.

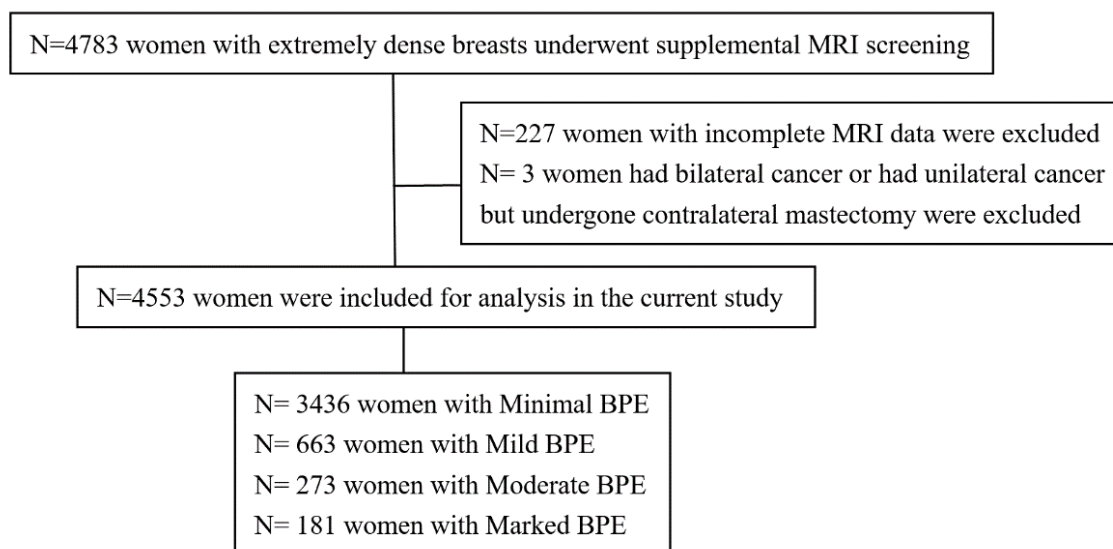


FIGURE 1: Flowchart of included women. BPE=Background parenchymal enhancement (BPE was assessed by radiologists)

Feature extraction

For 120 of the 4553 women, BPE category (assessed by radiologists) was not

available in the current study. For these 120 women, the scores were imputed. BPE categories minimal (75%) and mild (15%) were scored most often (**Table 3**). In hospital 6, women less often received BPE category ‘minimal’ (55%) compared to women from the other seven hospitals (67% to 80%). A higher fraction of women in hospital 6 had a BPE category of 'mild' (30%) compared to the fraction in other seven hospitals (10% to 17%) (**Figure 2**).

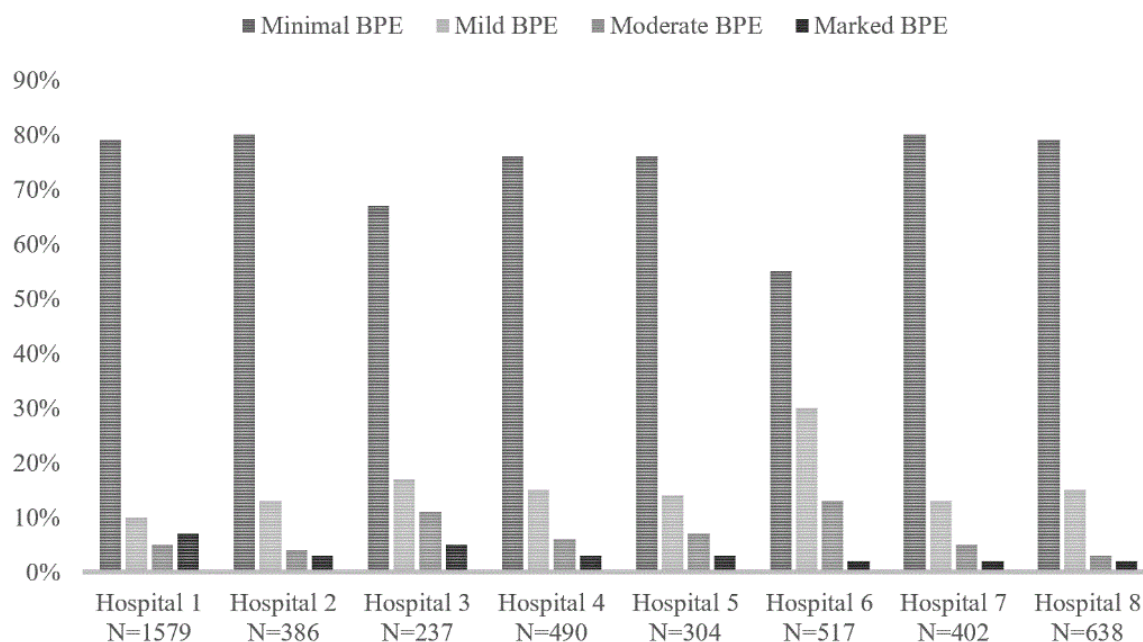


FIGURE 2. Percentage distribution of BPE categories in the eight hospitals participating in the DENSE trial

Age(mean ± SD), years	55.7 ± 6
BPE category	
Minimal	3436(75%)
Mild	663(15%)
Moderate	273(6%)
Marked	181(4%)
Numbers represent frequency unless stated otherwise; SD = standard deviation; BPE= background parenchymal enhancement	

Prediction of BPE category

The accuracy of the random forest, Naïve bayes, and k-NN classifiers to predict the BPE score in the eight hospitals ranged between 0.30 and 0.84 (**Table 4**). The accuracy of the majority voting ranged between 0.56 and 0.84 (**Table 4**). No obvious difference in accuracy was observed between centers, despite the use of scanners from different manufactures. The cross-validated prediction accuracy

of BPE based on majority voting was, however, somewhat lower in hospital 6 (0.56) compared to that in the other hospitals (0.68 to 0.84). The sensitivity of the majority voting method was highest for predicting the presence of 'minimal' BPE (0.95) compared to other BPE categories (0.13 to 0.36) (**Table 5**). This suggests that machine learning was most effective at accurately identifying minimal BPE among study participants. Conversely, the specificity of the majority voting was lowest for 'minimal' BPE (0.46) compared to the other BPE categories (0.95 to 0.97) (**Table 5**). The overall weighted mean accuracy established from the eight hospitals was 0.76.

TABLE 4. External validation accuracy of 100 times bootstrapping per hospital per classifier

Hospital ID	External validation size	Training size	Random forest	Naïve bayes	KNN	Majority Voting	Weighted accuracy
1	1579	2974	0.82 [0.81,0.82]	0.79 [0.79,0.79]	0.81 [0.81,0.81]	0.82 [0.81,0.82]	0.76
2	386	4167	0.82 [0.81,0.82]	0.57 [0.57,0.57]	0.80 [0.80,0.80]	0.80 [0.80,0.81]	
3	237	4316	0.71 [0.70,0.73]	0.72 [0.72,0.72]	0.72 [0.72,0.72]	0.71 [0.71,0.73]	
4	490	4063	0.80 [0.79,0.81]	0.76 [0.76,0.76]	0.82 [0.82,0.82]	0.8 [0.8,0.81]	
5	304	4249	0.80 [0.79,0.81]	0.75 [0.75,0.75]	0.79 [0.79,0.79]	0.79 [0.78,0.79]	
6	517	4036	0.56 [0.56,0.56]	0.57 [0.57,0.57]	0.57 [0.57,0.57]	0.56 [0.56,0.57]	
7	402	4151	0.83 [0.82,0.84]	0.83 [0.83,0.83]	0.84 [0.84,0.84]	0.84 [0.83,0.85]	
8	638	3915	0.71 [0.70,0.72]	0.30 [0.30,0.30]	0.69 [0.69,0.69]	0.68 [0.67,0.69]	
Numbers in column 4 to 7 represent mean accuracy [95% confidence interval]							

TABLE 5. Prediction performance in the four BPE categories

BPE category	Minimal	Mild	Moderate	Marked
Sensitivity	0.95 [0.95, 0.95]	0.16 [0.16, 0.17]	0.13 [0.12, 0.14]	0.36 [0.34,0.37]
Specificity	0.46 [0.45, .47]	0.95 [0.94, 0.95]	0.96 [0.96, 0.97]	0.97 [0.97, 0.97]
Numbers in column 2~5 represent median [95% confidence interval]				

Association with breast cancer occurrence

Among the included women, 122 of 4553 (3%) were diagnosed with breast cancer, and the mean time to cancer detection was 12.4 months. Regardless whether BPE was rated manually (HR=1.97, P<0.001) or automatically (HR=2.12, P<0.001), categories higher than minimal (i.e., mild/moderate/marked) were associated with higher breast cancer occurrence than category minimal (**Table 6**).

Additionally, the hazard ratio of BPE for breast cancer occurrence was comparable between automated rating and manual rating (HR=2.12 versus HR=1.97, P=0.65 for mild/moderate/marked BPE relative to minimal BPE).

TABLE 6. Association between BPE category and occurrence of breast cancer for manual scoring and scoring by machine learning

Characteristics	Hazard Ratio (HR)	95%CI	P value of HR	P value of comparing HRs
BPE rated by radiologist				
Minimal	Reference	-	-	0.65
Mild/Moderate/Marked	1.97	[1.37, 2.84]	<0.001	
BPE predicted by machine learning				
Minimal	Reference	-	-	0.65
Mild/Moderate/Marked	2.12	[1.43, 3.16]	<0.001	
BPE= Background parenchymal enhancement, CI= confidence interval				

Discussion

This study aimed to determine whether it is feasible to use machine learning to automatically evaluate BPE category on dynamic contrast-enhanced MRI (DCE-MRI) of women with extremely dense breasts without compromising associations with breast cancer occurrence. A combination of Random Forests,

Naïve Bayes, and KNN classifiers with majority voting was used to develop a machine-learning model that could predict BPE category with a cross-validated prediction accuracy of 0.76. The results indicate that it is feasible to automate the evaluation of BPE category in women with extremely dense breasts, although the accuracy for minimal BPE is superior to that for other BPE categories. The underlying association between BPE and breast cancer occurrence – expressed as hazard ratio – was not affected and was comparable with that of BPE derived from manual rating by radiologists in the DENSE trial.

The cross-validated prediction accuracy ranged between 0.56 and 0.84 across the eight hospitals. The lowest accuracy was observed in hospital 6 (0.56). This may be attributable to the fact that radiologists in hospital 6 typically assigned higher scores to BPE. Consequently, hospital 6 had the lowest percentage of women in the minimal BPE category (55%), while the sensitivity of the machine-learning model for this category was the highest. The percentage of women in the minimal BPE category ranged between 67% and 80% in the other hospitals.

In clinical practice, BPE is rated by radiologists according to the BIRADS lexicon, and is subject to inter- and intrareader variation (16), which may have hindered its role as predictive imaging biomarker of breast cancer. Quantitative evaluation of BPE can overcome this disadvantage. A previous study assessed quantitative parenchymal features at baseline DCE MRI and their association with breast cancer occurrence (7), using manual rating of BPE to confirm that such quantitative analysis is independently associated with breast cancer occurrence. The current study focuses explicitly on automated rating of BPE to help further reduce inter- and intra-observer variability.

Several other studies reported on automated prediction of BPE category from DCE MRI, but the study populations and methodology vary considerably. A recent study from Nam et al. investigated 794 patients with breast cancer who underwent preoperative breast MRI from 2014 to 2017. Deep learning was applied to automatically assess BPE category. The overall classification accuracy among the four BPE categories was 0.67 (33). In contrast, the current study focuses only on an unselected series of asymptomatic women with extremely dense breasts.

Borkowski et al. trained a deep convolutional neural network for the classification of BPE in 149 patients who underwent breast MRI from September

2013 to October 2015, yielding mean accuracy of 0.75 in external validation (34). Neither the study by Borkowski et al. nor the study by Nam et al. specifically focus on women with extremely dense breasts. Hence, it is difficult to assess the potential of these findings in a risk stratification tool for this specific screening population.

Sarah et al. recently reported on 3705 high-risk women ((i.e., women with >20% lifetime risk of breast cancer) using 5224 breast MRI examinations (35). Two deep learning models were constructed to distinguish between low (i.e., mild or minimal) and high (i.e., marked or moderate) BPE. Although the study uses a relatively large sample size, it also does not focus explicitly on women with extremely dense breasts.

The current study has some limitations. First, although the model was validated in different hospitals using internal-external validation, the results have not yet been validated in other external datasets, e.g., those from other countries. Wider validation of the model is recommended prior to clinical application. Secondly, the number of women in each BPE category was not balanced. Although this reflects the natural distribution of BPE in the population of Dutch women with extremely dense breasts, it may have impact on the machine learning classifier, especially in women with moderate and marked BPE. In the future, the model could be investigated for its efficacy in other populations of women at risk.

To conclude, it is feasible to rate BPE category automatically in contrast-enhanced MRI of women with extremely dense breasts without compromising the association between BPE and breast cancer occurrence, although the accuracy for minimal BPE is superior to that of other BPE categories.

References

1. Emaus MJ, Bakker MF, Peeters PHM, 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.
2. Wanders JOP, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Res Treat* 2017; 162:95–103.
3. Boyd NF, Guo H, Martin LJ. Mammographic Density and the Risk and Detection of

- Breast Cancer. *N Engl J Med* 2007; 356:227–236.
4. Boyd NF, Martin LJ, Yaffe MJ, et al. Mammographic density and breast cancer risk: current understanding and future prospects. *Breast Cancer Res* 2011; 13:1–12.
 5. Arasu VA, Miglioretti DL, Sprague BL, et al. Population-based assessment of the association between magnetic resonance imaging background parenchymal enhancement and future primary breast cancer risk. *J Clin Oncol* 2019; 37:954–963.
 6. King V, Brooks JD, Bernstein JL, et al. Background Parenchymal Enhancement at Breast MR Imaging and Breast Cancer Risk. *Radiology* 2011; 260:50–60.
 7. Wang H, Velden BHM Van Der, Verburg E, et al. Assessing Quantitative Parenchymal Features at Baseline Dynamic Contrast-enhanced MRI and Cancer Occurrence in Women with Extremely Dense Breasts. *Radiology* 2023; 308:e222841.
 8. Dontchos BN, Rahbar H, Partridge SC, et al. Are Qualitative Assessments of Background Parenchymal Enhancement, Amount of Fibroglandular Tissue on MR Images, and Mammographic Density Associated with Breast Cancer Risk? *Radiology* 2015; 276:371–380.
 9. Kim SY, Cho N, Choi Y, et al. Factors affecting pathologic complete response following neoadjuvant chemotherapy in breast cancer: Development and validation of a predictive nomogram. *Radiology* 2021; 299:290–300.
 10. Kim SY, Cho N, Shin SU, et al. Contrast-enhanced MRI after neoadjuvant chemotherapy of breast cancer: lesion-to-background parenchymal signal enhancement ratio for discriminating pathological complete response from minimal residual tumour. *Eur Radiol* 2018; 28:2986–2995.
 11. You C, Peng W, Zhi W, et al. Association Between Background Parenchymal Enhancement and Pathologic Complete Remission Throughout the Neoadjuvant Chemotherapy in Breast Cancer Patients. *Transl Oncol* 2017; 10:786–792.
 12. Hilal T, Covington M, Kosiorek HE, et al. Breast MRI phenotype and background parenchymal enhancement may predict tumor response to neoadjuvant endocrine therapy. *Breast J* 2018; 24:1010–1014.
 13. Zhang M, Sadinski M, Haddad D, et al. Background Parenchymal Enhancement on Breast MRI as a Prognostic Surrogate: Correlation With Breast Cancer Oncotype Dx Score. *Front Oncol* 2021; 10:1–7.
 14. Moliere S, Oddou I, Noblet V, et al. Quantitative background parenchymal enhancement to predict recurrence after neoadjuvant chemotherapy for breast cancer. *Sci Rep* 2019; 9:1–8.
 15. Bae MS, Chang JM, Cho N, et al. Association of preoperative breast MRI features with locoregional recurrence after breast conservation therapy. *Acta radiol* 2018; 59:409–417.
 16. Melsaether A, McDermott M, Gupta D, et al. Inter- And intrareader agreement for categorization of background parenchymal enhancement at baseline and after training.

Am J Roentgenol 2014; 203:209–215.

17. Tagliafico A, Bignotti B, Tagliafico G, et al. Quantitative evaluation of background parenchymal enhancement (BPE) on breast MRI. A feasibility study with a semi-automatic and automatic software compared to observer-based scores. *Br J Radiol* 2015; 88.

18. Veenhuizen SGA, Lange SV de, Bakker MF. Supplemental Breast MRI for Women with Extremely Dense Breasts: Results of the Second Screening Round of the DENSE Trial. *Radiology* 2021; 299:278–286.

19. Bakker MF, De Lange S V., Pijnappel RM, et al. Supplemental MRI screening for women with extremely dense breast tissue. *N Engl J Med* 2019; 381:2091–2102.

20. Spak DA, Plaxco JS, Santiago L, et al. BI-RADS fifth edition: A summary of changes. *Diagn Interv Imaging* 2017; 98:179–190.

21. Emaus MJ, Bakker MF, Peeters PHM, 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.

22. Klein S, Staring M, Murphy K, et al. elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Trans Med Imaging* 2010; 29:196–205.

23. van der Velden BHM, Dmitriev I, Loo CE, et al. Association between Parenchymal Enhancement of the Contralateral Breast in Dynamic Contrast-enhanced MR Imaging and Outcome of Patients with Unilateral Invasive Breast Cancer. *Radiology* 2015; 276:675–685.

24. Isensee F, Jaeger PF, Kohl SAA, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods* 2021; 18:203–211.

25. Morris EA. ACR Bi-Rads® Atlas — Breast MRI. *Am Coll Radiol* 2013:125–143.

26. Burns RA, Butterworth P, Kiely KM, et al. Multiple imputation was an efficient method for harmonizing the Mini-Mental State Examination with missing item-level data. *J Clin Epidemiol* 2011; 64:787–793.

27. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons, Inc.; 1987.

28. Marshall A, Altman DG, Holder RL, et al. Combining estimates of interest in prognostic modelling studies after multiple imputation: Current practice and guidelines. *BMC Med Res Methodol* 2009; 9:1–8.

29. Hastie R, Kameda T. The robust beauty of majority rules in group decisions. *Psychol Rev* 2005; 112:494–508.

30. Sorokin RD, Hays CJ, West R. Signal-detection analysis of group decision making. *Psychol Rev* 2001; 108:183–203.

31. Klein N, Epley N. Group discussion improves lie detection. *Proc Natl Acad Sci U S A* 2015; 112:7460–7465.

32. Lee CH, Cook S, Lee JS, et al. Comparison of Two Meta-Analysis Methods: Inverse-

Variance-Weighted Average and Weighted Sum of Z-Scores. *Genomics Inform* 2016; 14:173–180.

33. Nam Y, Park GE, Kang J, et al. Fully Automatic Assessment of Background Parenchymal Enhancement on Breast MRI Using Machine-Learning Models. *J Magn Reson Imaging* 2021; 53:818–826.

34. Borkowski K, Rossi C, Ciritsis A, et al. Fully automatic classification of breast MRI background parenchymal enhancement using a transfer learning approach. *Medicine (Baltimore)* 2020; 99:1–7.

35. Eskreis-Winkler S, Sutton EJ, D'Alessio D, et al. Breast MRI Background Parenchymal Enhancement Categorization Using Deep Learning: Outperforming the Radiologist. *J Magn Reson Imaging* 2022; 56:1068–1076.

Chapter 3

Assessing Quantitative Parenchymal Features on Baseline Dynamic Contrast-enhanced MRI and Cancer Occurrence in Women with Extremely Dense Breasts

Based on:

*Wang H, Van Der Velden BHM, Verburg E, Bakker MF, Pijnappel RM, Veldhuis WB, Gils CH van, Gilhuijs, KGA. Assessing Quantitative Parenchymal Features at Baseline Dynamic Contrast-enhanced MRI and Cancer Occurrence in Women with Extremely Dense Breasts. **Radiology.** 308(2):e222841 (2023)*

Abstract

Background and Objective: Automated identification of quantitative breast parenchymal enhancement features on dynamic-contrast enhanced MRI (DCE-MRI) could provide added value for assessing breast cancer risk in women with extremely dense breasts. Aim of this study is to automatically identify quantitative properties of the breast parenchyma on baseline DCE-MRI and assess their association with breast cancer occurrence in women with extremely dense breasts.

Materials and Methods: This study represents a secondary analysis of the Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial. MRIs were performed in eight hospitals between December 2011 and January 2016. Following segmentation of fibroglandular tissue, quantitative features (including volumetric density, volumetric morphology, and enhancement characteristics) of the parenchyma were extracted from baseline MRIs. Principal component analysis was applied to identify parenchymal measures with the greatest variance. Multivariable Cox proportional hazards regression was applied to assess the association between breast cancer occurrence and quantitative parenchymal features, followed by stratification of significant features into tertiles.

Results: A total of 4553 women (mean age, 55.7 years \pm 6 [SD]) with extremely dense breasts were included, of which 122 (2.7%) were diagnosed with breast cancer. Five principal components representing 96% of the variance were identified, and the component explaining the greatest independent variance (42%) consisted of MRI features relating to volume of enhancing parenchyma. Multivariable analysis showed that volume of enhancing parenchyma was associated with breast cancer occurrence (hazard ratio [HR], 1.09 [95% CI: 1.01, 1.18]; $P = .02$). Additionally, women in the high tertile of volume of enhancing

parenchyma showed a breast cancer occurrence two times that of women in the low tertile (HR, 2.09 [95% CI: 1.25, 3.61]; $P = .005$).

Conclusion: In women with extremely dense breasts, a high volume of enhancing parenchyma on baseline DCE-MRI was associated with increased occurrence of breast cancer as compared to a low volume.

Introduction

High breast density on mammography, expressed by the ratio of fibroglandular tissue (FGT) to fatty breast tissue, is a known risk factor for developing breast cancer (1,2). Women in the highest breast-density category (Breast Imaging Reporting and Data System category d, extremely dense breast) have three to six times higher risk than women with almost entirely fatty breasts (2,3).

Although mammography is effective to screen for early-stage breast cancer reducing overall mortality (4), the sensitivity is lower in women with extremely dense breasts than in women in the lower density category (1). As the most sensitive breast imaging modality, dynamic-contrast enhanced MRI (DCE-MRI) has been proposed to screen women with high risk of developing breast cancer (5,6).

The multi-institutional Netherlands-based Dense Tissue and Early Breast Neoplasm Screening (DENSE) randomized controlled trial demonstrated that supplemental MRI screening in women aged 50 – 75 years with extremely dense breasts resulted in fewer interval cancers than using mammography alone (7). If implemented in the Dutch national screening program—which offers biennial mammograms to women between 50 and 75 years of age (8)—nearly 82000 women could be eligible to receive supplemental MRI, which would considerably increase the workload of breast MRI radiologists.

Several studies aimed to reduce the workload of radiologists using computer-aided triaging (9) and computer-aided diagnosis (10). Another approach is to tailor the screening of individual women based on additional risk factors.

Breast-cancer risk models incorporate characteristics such as age and body mass index (BMI)(11). More recently, the mammography-derived percentage of breast-FGT (2,3,12) was used to triage women to MRI. Additionally, background parenchymal enhancement (BPE) was found an independent risk factor for breast cancer (13,14), but it is subject to intra- and interobserver variability (13). Therefore, it is of interest to know whether other properties of breast parenchymal enhancement can be automatically derived and used to help predict risk of breast cancer in women with extremely dense breasts. Image feature extraction via radiomics has shown promise for computer-aided diagnosis and prognosis of breast cancer (15,16), but whether features extracted from MRIs of

women with extremely dense breasts can be used to predict breast cancer in parallel with common clinical risk factors is not well established.

The aim of this study was to evaluate quantitative properties of the breast parenchyma on baseline DCE-MRI and assess the association between these features and breast cancer occurrence in women with extremely dense breasts who participated in the DENSE trial.

Materials and Methods

Study Design and Patients

The current study is a secondary analysis of data from the prospective DENSE trial (ClinicalTrials.gov: NCT01315015) (7,8,17), which was approved by the Dutch Minister of Health, Welfare and Sport (2011/19 WBO [Wet Bevolkingsonderzoek], The Hague, the Netherlands). According to the Dutch law on population studies, the study including primary and secondary objectives was waived from ethical review by the local institutional review board.

The DENSE trial included consecutive patients who underwent MRI at eight hospitals in the Netherlands between December 2011 and January 2016. Women aged 50–75 years with extremely dense breasts (i.e., Breast Imaging Reporting and Data System category d at mammography without abnormality on mammography at the time of inclusion) (18) were eligible for inclusion. Women in the DENSE study were excluded in cases of standard contraindications to MR imaging (8).

Screening was conducted every 2 years and was performed in three consecutive rounds starting in December 2011 and concluding with the third-round screening in November 2021. Patients were excluded from the current study if they 1) had bilateral cancer, or 2) had unilateral cancer but had previously received contralateral mastectomy. Participant characteristics including age and BMI were acquired from electronic case report forms of the DENSE trial. Data generated or analyzed during the study are available from the corresponding author by request.

Previously (9), MRIs from 4581 participants of the DENSE trial were used to develop a deep learning model to differentiate between breasts with lesions and breasts without lesions. The current study included 4553 of these participants.

MRI acquisition

All women received MRI in prone position using a Philips 3.0T scanner (Achieva or Ingenia) or a Siemens 3.0T scanner (Trio, Verio, or Skyra) with a dedicated phased-array bilateral breast coil. The MRI included a dynamic series consisting of a precontrast image and four or five postcontrast images. Image acquisition was done according to fixed protocol which has been described in detail by Emaus et al. (1). The imaging parameters have been summarized in **Table 1**. In short, flip angle ranged between 10° and 20°, echo times ranged between 1.7 msec and 2.4 msec, and repetition times ranged between 3.3 msec and 5.5 msec. The macrocyclic gadolinium-based contrast agent gadobutrol (Gadovist; Bayer AG) was injected at a rate of 1 mL/s to a total dose of 0.1 mmol per kilogram of body weight. The use of fat suppression was optional.

TABLE 1. Imaging Parameters of the Dynamic Contrast-Enhanced MRI Used in the Dense Tissue and Early Breast Neoplasm Screening Trial

Medical Center and MRI Device	No. of women	Reconstructed Voxel Size(mm ³)	Dimensions (Voxels)	Fat Suppression
1	1579			
Philips Achieva		0.89* 0.89* 0.9	384* 384* 200	Yes
Philips Ingenia		0.89* 0.89* 0.9	384* 384* 200	Yes
2	386			
Siemens Magnetom Trio		0.80* 0.80*1.0	448* 448* 176	No
Siemens Skyra		0.80* 0.80*1.0	448* 448* 176	No
Siemens Prisma		0.80* 0.80*1.0	448* 448* 176	No
3 Philips Achieva	237	0.89* 0.89* 0.9	384* 384* 200	Yes
4 Philips Ingenia	490	0.89* 0.89* 0.9	384* 384* 200	Yes
5	304			
Philips Achieva		0.89* 0.89* 0.9	384* 384* 200	Yes
Philips Ingenia CX		0.89* 0.89* 0.9	384* 384* 200	Yes
6 Siemens Verio	517	0.85* 0.85* 1.0	448* 448* 176	No
7 Philips Ingenia	402	0.89* 0.89* 0.9	384* 384* 200	Yes
8 Siemens Skyra	638	0.80* 0.80* 1.0	448* 448* 160	No

Image analysis

Image analysis consisted of deformable image registration (19), breast segmentation (20), FGT segmentation, and feature extraction using baseline dynamic contrast-enhanced MRI. The FGT was segmented from the precontrast image using nnU-Net (21). After five cycles of iterative learning, all FGT segmentations were deemed acceptable. The workflow used for image analysis was adopted from a prior study (22). The nnU-Net was retrained on 427 randomly selected MRIs from the current dataset and the results were visually

inspected. The quality of the segmentations was assessed by H.W. (medical doctor with more than 6 years of experience in breast MRI) and B.V. (biomedical engineer with more than 11 years of experience) under supervision of W.V. (dedicated breast radiologist with more than 12 years of experience).

Spatio-temporal characteristics of the FGT were described using 15 quantitative features and extracted from the MRIs using Python (version 2.7; Python Software Foundation). These included volumetric density, volumetric morphology (the three-dimensional shape of the FGT), and enhancement characteristics (**Table 2**). These features were previously associated with patient characteristics and outcome in the symptomatic patient populations (20,22,23). We used features of the contralateral healthy breast in women who had breast cancer on one side, and the average features from left and right breasts in women who did not have breast cancer.

MRIs from the first screening round of the DENSE trial were evaluated by breast radiologists with experience ranging between 5 and 23 years among 8 hospitals (including W.V. with more than 12 years of experience in breast MRI). A maximum of two breast radiologists were involved in each participating hospital. The MRI images were assessed according to the Breast Imaging Reporting and Data System MRI lexicon. When confronted with uncertainty, independent double reading was performed by a second radiologist who was affiliated with another participating hospital. Potential discrepancies were resolved by consensus (8). BPE category was classified as minimal, mild, moderate and marked BPE (24). BPE assessment of the two breasts are typically consistent. In instances where discrepancies were observed, the greatest BPE category was used. Additional details about the DENSE trial can be found in the publication of Emaus et al. (8)

Code and trained neural network weights are at https://github.com/Lab-Translational-Cancer-Imaging/FGT_segmentation and https://github.com/Lab-Translational-Cancer-Imaging/Harmonize_DCE/blob/main/harmonize_cpe.py

Statistical analysis

The event of interest in this study was detection of breast cancer. Time-to-event was defined as the period between the first round of screening and the detection of breast cancer.

TABLE 2. Description of the 15 Quantitative Parenchyma Features Evaluated on Baseline MRI

Feature	Description
Circularity	Ratio between the intersection of the FGT with a sphere of equivalent volume as the FGT and the volume of the FGT itself
Irregularity	Ratio between the surface of a sphere with the same volume as the FGT and the surface of the FGT
Density	Ratio between the volume of FGT and the volume of the whole breast
Top 10% intensity of early contrast uptake	Top 10% early enhancing intensity of the contralateral healthy breast (women with cancer) or mean value of top 10% of both breasts (women without cancer)
Top 10% volume of early enhancing voxels	Volume of the top 10% early enhancing voxels of the contralateral healthy breast (women with cancer) or mean value of top 10% of both breasts (women without cancer)
Top 50% intensity of early contrast uptake	Top 50% early enhancing intensity of the contralateral healthy breast (women with cancer) or mean value of top 10% of both breasts (women without cancer)
Top 50% volume of early enhancing voxels	Volume of the top 50% early enhancing voxels of the contralateral healthy breast (women with cancer) or mean value of top 10% of both breasts (women without cancer)
Top 90% intensity of early contrast uptake	Top 90% early enhancing intensity of the contralateral healthy breast (women with cancer) or mean value of top 10% of both breasts (women without cancer)
Top 90% volume of early enhancing voxels	Volume of the top 90% early enhancing voxels of the contralateral healthy breast (women with cancer) or mean value of top 10% of both breasts (women without cancer)
Top 10% intensity of late contrast uptake	Top 10% late enhancing intensity of the contralateral healthy breast (women with cancer) or mean value of top 10% of both breasts (women without cancer)
Top 10% volume of late enhancing voxels	Volume of the top 10% late enhancing voxels of the contralateral healthy breast (women with cancer) or mean value of top 10% of both breasts (women without cancer)
Top 50% intensity of late contrast uptake	Top 50% late enhancing intensity of the contralateral healthy breast (women with cancer) or mean value of top 10% of both breasts (women without cancer)
Top 50% volume of late enhancing voxels	Volume of the top 50% late enhancing voxels of the contralateral healthy breast (women with cancer) or mean value of top 10% of both breasts (women without cancer)
Top 90% intensity of late contrast uptake	Top 90% late enhancing intensity of the contralateral healthy breast (women with cancer) or mean value of top 10% of both breasts (women without cancer)
Top 90% volume of late enhancing voxels	Volume of the top 90% late enhancing voxels of the contralateral healthy breast (women with cancer) or mean value of top 10% of both breasts (women without cancer)

FGT= fibroglandular tissue.

Multiple imputation was used to account for missing BMI data in 6 women diagnosed with cancer. Ten imputation data sets were generated and the mean value was used as the imputed value (2,3). Outliers were identified using Tukey’s method (i.e., the lower limit is 1.5 interquartile range [IQR] below the 25th percentile, and the upper limit is 1.5 IQR above the 75th percentile). All values greater than the upper limit were replaced with the upper limit and all values smaller than the lower limit with the lower limit, after which the feature values were normalized between 0 and 1.

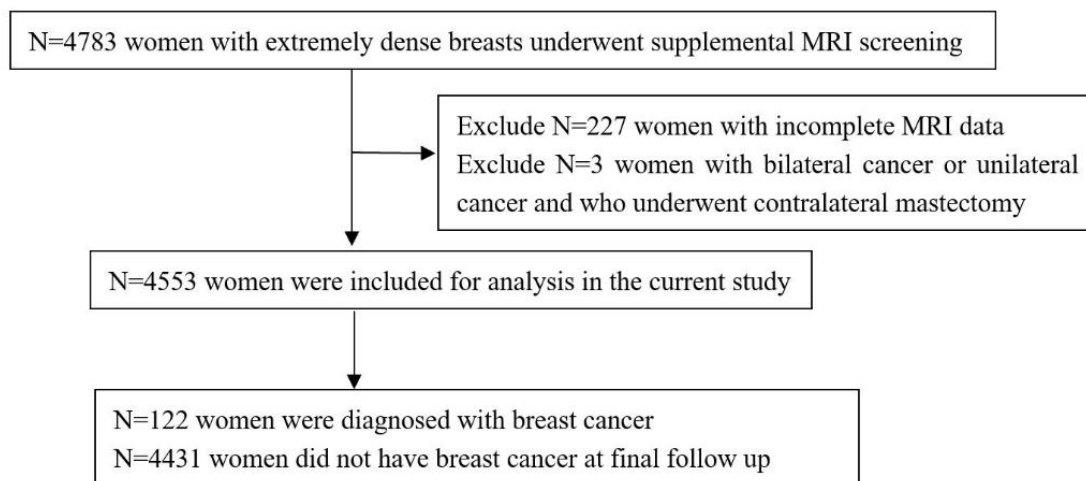


FIGURE 1. Flowchart of selection of women included.

Principal component analysis of the 15 quantitative parenchymal features was used to reduce the number of variables in the dataset while maximizing the variance of the components retained. Varimax rotation of initial estimates of principal components was then used to simplify the structure and increase the interpretability of factors (4,5). In this study, the components that together accounted for 96% of the cumulative variance were retained.

TABLE 3. Characteristics of Participants from the Dense Tissue and Early Breast Neoplasm Screening Trial included in the current study

	Total (N = 4553)	Women without cancer (N = 4431)	Women with cancer (N = 122)
Age (years)	55.7 ± 6	55.7 ± 6	55.9 ± 6.3
BMI (kg/m ²)	22.2 ± 2.5	22.2 ± 2.6	23.2 ± 2.9
BPE category			
Minimal	3436 (75%)	3362 (76%)	74 (61%)
Mild	663 (15%)	630 (14%)	33 (27%)
Moderate	273 (6%)	268 (6%)	5 (4%)
Marked	181 (4%)	171 (4%)	10 (8%)
Continuous variables are presented as mean ± standard deviation and categorical variables are presented as number participants with percentages in parentheses. BMI = body mass index, BPE = background parenchymal enhancement.			

Multivariable Cox proportional hazards regression was used to assess the association between quantitative parenchymal features in screening round one with breast-cancer occurrence. For patients who had already developed breast cancer at the time of the first screening, data were left-censored in the Cox regression (6).

TABLE 4. Principal Component Analysis Identification of Five Components Explaining 96% Variance

	PC1	PC2	PC4	PC3	PC5
MRI Features	Volume of enhancing parenchyma	Early contrast uptake	Late contrast uptake	Shape	FGT density on MRI
Variance	42%	20%	19%	11%	5%
Cumulative variance	42%	62%	80%	91%	96%
Variance represents separate variance explained by each principal component while cumulative variance represents the total variance explained by the current principal component and all prior components. FGT = fibroglandular tissue, PC = principal component.					

The principal component scores (which are a linear combination of the quantitative parenchymal features) that were found to be associated with breast-cancer occurrence after multivariable analysis were divided into tertiles (representing low, intermediate and high subgroups) and the cumulative hazard was calculated for each tertile, adjusting for clinical characteristics using inverse probability weighting (7). The differences in cumulative hazard between the inverse probability weighting-adjusted tertiles were evaluated via Kaplan-Meier survival curve analysis. Comparisons between the intermediate tertile survival curve and the low tertile survival curve as well as the high tertile survival curve and the low tertile survival curve were carried out using log-rank tests. The statistical analyses were performed by H.W. (6 years of experience conducting statistical analyses) under the supervision of B.V. (11 years of experience conducting statistical analyses) and K.G. (32 years of experience conducting statistical analyses) using R-Studio (version 4.1.2, R Foundation for Statistical Computing, Vienna, Austria). A *P* value of < .05 was considered to indicate statistical significance.

Results

Participant Characteristics

Of the 4783 MRI examinations included in the DENSE trial, three were excluded due to bilateral cancer or unilateral cancer with contralateral mastectomy history and 227 were excluded due to incomplete MRI data sets (**Figure 1**). A total of 4553 women (mean age, 55.7 years \pm 6 [SD]) with extremely dense breasts were

ultimately included in the study, 122 (2.7%) of whom were diagnosed with breast cancer detected on MRI, mammography, or as interval cancers (**Table 3**). Among the women diagnosed with breast cancer, 77/122 (63.1%) had breast cancer detected in the first round of screening and the remaining 45/122 (36.9%) women had a median time to cancer detection of 24 months ([IQR: 24- 47 months]; range: 17- 53 months).

TABLE 5. Univariable and Multivariable Cox Regression Analysis of Variables Associated with Occurrence of Breast Cancer

Features	Univariable analysis			Multivariable analysis		
	HR	95% CI	<i>P</i> value	HR	95% CI	<i>P</i> value
BPE category						
Minimal	Reference			Reference		
Mild	2.36	1.56, 3.55	<.001*	1.95	1.24, 3.07	.004*
Moderate	0.84	0.34, 2.08	.705	0.58	0.22, 1.52	.267
Marked	2.56	1.32, 4.95	.005*	1.39	0.62, 3.14	.425
Age (years)	1.01	0.98, 1.04	.377	1.05	1.02, 1.09	.003*
BMI (kg/m ²)	1.17	1.09, 1.25	< .001*	1.08	0.98, 1.18	.134
Principal Component [†]						
PC 1	1.06	1.03, 1.08	< .001*	1.09	1.01, 1.18	.023*
PC 2	1.07	1.02, 1.13	.006*	1.07	0.98, 1.18	.144
PC 3	1.04	0.95, 1.15	.396	1.02	0.89, 1.17	.781
PC 4	1.04	0.99, 1.10	.131	1.03	0.94, 1.13	.559
PC 5	1.18	1.05, 1.34	.007*	0.83	0.58, 1.21	.337

BPE = background parenchymal enhancement, BMI = body mass index, CI = confidence interval, HR = hazard ratio, PC = principal component. * *P* value < .05
[†] MRI features comprising PC1 were related to volume of enhancing parenchyma, PC2 were related to early contrast uptake of parenchyma, PC3 were related to the shape of parenchyma, PC4 were related to late contrast uptake of parenchyma, and PC5 were related to breast density.

TABLE 6. Cancer Detection in Participants Stratified by Tertiles of Volume of Enhancing Parenchyma

	Without cancer detected (<i>n</i> = 4431)	With cancer detected (<i>n</i> = 122)	Total
Low PC1	1494 (98.4%)	24 (1.6%)	1518
Intermediate	1472 (97.0%)	45 (3.0%)	1517
High PC1	1465 (96.5%)	53 (3.5%)	1518

PC = principal component. MRI features comprising PC1 were related to volume of enhancing parenchyma and PC1 was stratified into tertiles: low PC1 [IQR: -7.3, -4.9], intermediate PC1 [IQR: -2.7, -0.1], high PC1 [IQR: 3.7, 10.6].

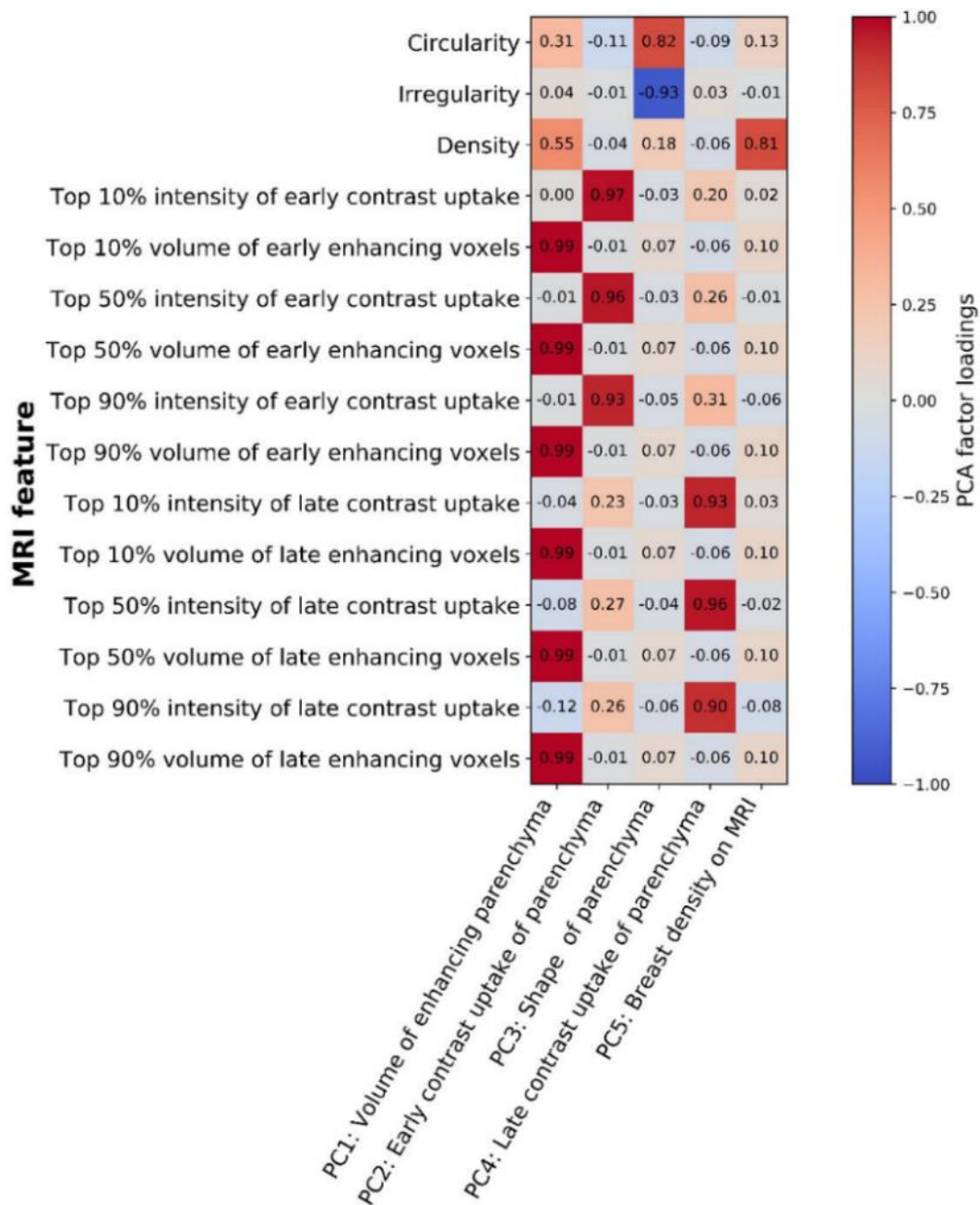


FIGURE 2. Heatmap showing the correlation between the 15 quantitative parenchyma features evaluated on baseline MRI and the principal components identified as accounting for 96% of the variance. The scale bar represents the principal component analysis (PCA) factor loading value where the darkest red color (value 1) represents a strong positive relationship and the darkest blue color (value -1) represents a strong negative relationship.

Quantitative FGT Features that Explain the Greatest Variance and Association with Breast Cancer Occurrence

Of the 15 quantitative FGT features, PCA identified five principal components which explained 96% of the cumulative variance. Individually, principal

component 1 explained 42% of the variance and was composed of MRI features relating to volume of enhancing parenchyma (**Figure 2, Table 4**). Principal component 2, 4, 3, and 5 explained 20%, 19%, 11% and 5% of the variance and were composed of MRI features relating to early contrast uptake of parenchyma, late contrast uptake of parenchyma, the shape of parenchyma, and breast density on MRI, respectively.

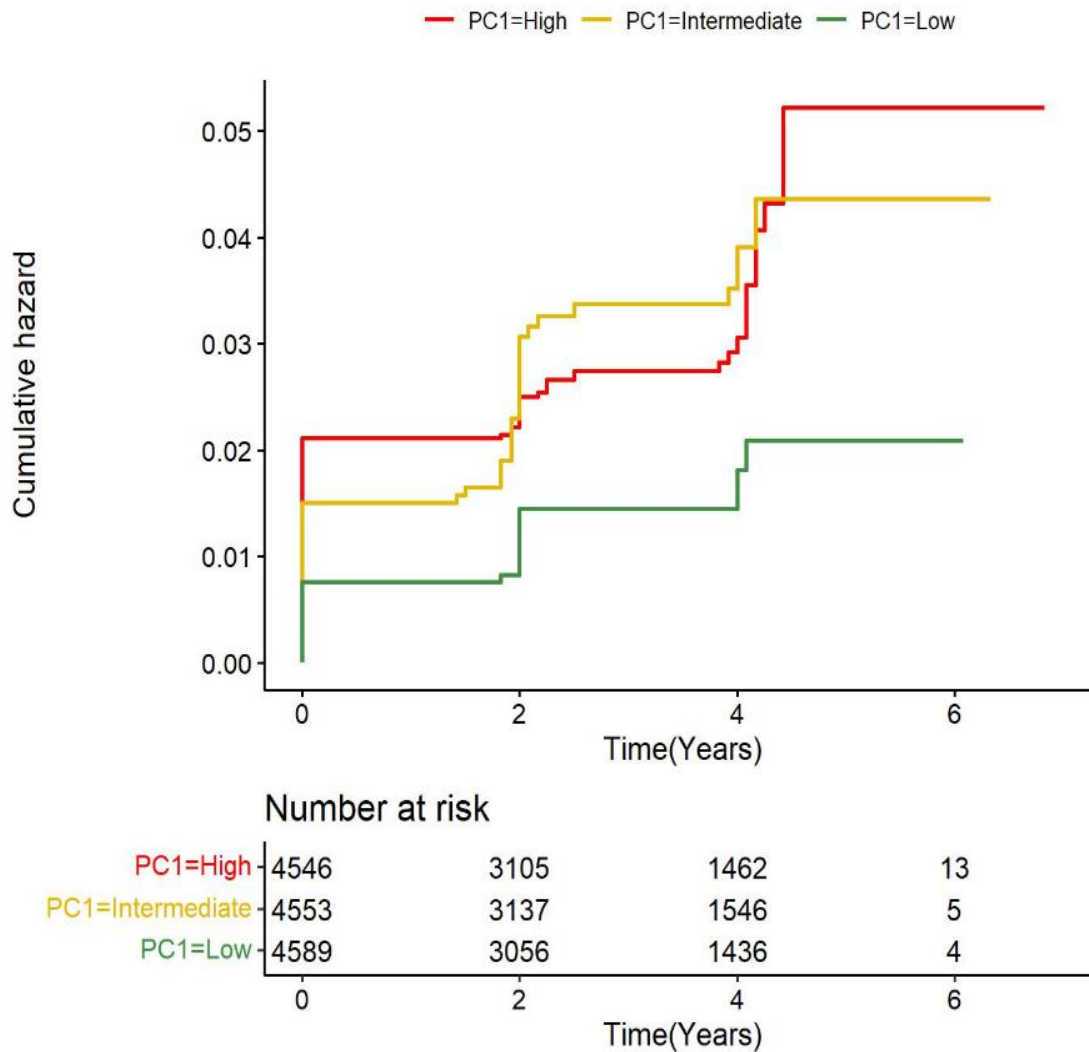


FIGURE 3. Kaplan-Meier curves for occurrence of breast cancer stratified by tertiles of volume of enhancing parenchyma. Women in the high tertile (red line) and the intermediate tertile (yellow line) of volume of enhancing parenchyma had a higher cumulative hazard for breast cancer occurrence than those in the low tertile (green line) (log-rank test $P = .001$ and $.014$, respectively). This analysis was adjusted for the clinical characteristics age, body mass index, and background parenchymal enhancement category using inverse probability weighting.

Multivariable analysis showed that principal component 1 (PC1, reflecting volume of enhancing parenchyma) was associated with occurrence of breast cancer (Hazard ratio [HR], 1.09 [95% CI: 1.01, 1.18]; $P = .02$) as was patient age (HR, 1.05 [95% CI: 1.02, 1.09]; $P = .003$) and the mild BPE category (HR, 1.95 [95% CI: 1.24, 3.07]; $P = .004$) (**Table 5**).

TABLE 7. Unadjusted and Adjusted Cox Regression Analysis of Tertiles of Volume of Enhancing Parenchyma for Association with Occurrence of Breast Cancer.

Tertile	Unadjusted			Adjusted [†]		
	HR	95% CI	P value	HR	95% CI	P value
Low	Reference			Reference		
Intermediate	1.84	1.12, 3.03	.02*	2.15	1.30, 3.84	.003*
High	2.19	1.35, 3.55	.001*	2.09	1.25, 3.61	.005*

CI = confidence interval, HR = hazard ratio, PC = principal component. PC1 reflects volume of enhancing parenchyma and was stratified into tertiles: low PC1 [IQR: -7.3, -4.9], intermediate PC1 [IQR: -2.7, -0.1], high PC1 [IQR: 3.7, 10.6].
[†]Inverse probability weighting was used to adjust for age, body mass index, and background parenchymal enhancement. * P value < .05

Assessment of Breast Cancer Risk in Patients Stratified by PC1

When the principal component score reflecting volume of enhancing parenchyma was stratified into tertiles--representing low, intermediate, and high subgroups [IQR of the three tertiles of PC1 were provided in **Table 6**]--cancer was detected in 24/1518 (1.6%) women in the low PC1 tertile, 45/1517 (3.0%) women in the intermediate PC1 tertile, and 53/1518 (3.5%) of women in the high PC1 tertile (**Table 6**). In a multivariable model adjusted for age, BMI, and BPE category, women in the high PC1 tertile were found to have a higher occurrence of breast cancer than those in the lowest tertile (HR, 2.09 [95% CI: 1.25, 3.61], $P = .005$). Women with an intermediate PC1 also showed a greater occurrence of breast cancer compared with women in the lowest tertile (HR, 2.15 [95% CI: 1.30, 3.84], $P = .003$) (**Table 7**). Furthermore, women with a high and intermediate PC1 both showed an increasing risk for developing breast cancer over time compared to women with low PC1 (log-rank test, $P = .001$ and $P = .014$) (**Figure 3**). Representative MRIs in patients with a high and low PC1 are shown in **Figure 4**.

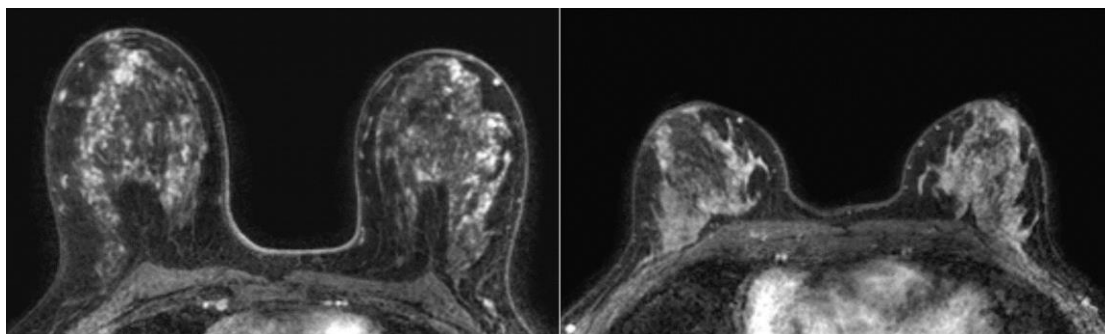


FIGURE 4. Representative dynamic contrast-enhanced MRIs (transverse plane) in patients in the low and high tertiles of volume of enhancing parenchyma. Left: Baseline MRI in a 50-year-old female with a body mass index (BMI) of 23 and ‘marked’ background parenchymal enhancement (BPE) who was stratified into the high tertile of volume of enhancing parenchyma. Cancer was detected on MRI in the second screening round. Right: A 51-year-old female with a BMI of 24 and ‘marked’ BPE who was stratified into the low tertile of volume enhancing parenchyma. No cancer was detected during 6 years of follow-up. Note, differences in high and low volumes of enhancing parenchyma are not visible to the naked eye on MRI.

Discussion

Automated identification of quantitative breast parenchymal enhancement features on dynamic-contrast enhanced MRI could provide added value for assessing breast cancer risk in women with extremely dense breasts. This secondary analysis of the Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial investigated whether quantitative parenchymal features on dynamic contrast-enhanced MRI were associated with breast cancer occurrence, independent of conventional clinical characteristics. Following adjustment for age, body mass index (BMI), and background parenchymal enhancement (BPE), multivariable analysis showed that breast cancer occurrence was greater in women with higher volumes of enhancing parenchyma compared to women with low volumes of enhancing parenchyma (Hazard ratio, 2.09 [95% CI: 1.25, 3.61]; $P = .005$). The range in age of included women (50-75 years) was relatively narrow. This may partly explain why statistical significance for age could not be demonstrated in univariable analysis, but significance was observed after correcting for BMI and BPE in multivariable analysis.

Several previous studies have investigated the association of BPE and fibroglandular tissue (FGT) on MRI with breast cancer risk. The meta-analysis conducted by Thompson et al. (31) showed that a higher level of BPE measured at breast MRI is associated with the presence of breast cancer in women with high risk, but not in women with average risk. The results of this meta-analysis

are difficult to compare to our study because women included in the DENSE trial all had negative screening mammography results. Consistent with our results, Acciavatti et al. (32) observed in a review study that higher BPE of the FGT on breast MRI is associated with a higher risk for breast cancer. In a retrospective study of women at high risk for breast cancer, Dontchos et al. (13) found that the 23 patients who developed cancer had greater qualitatively-measured BPE on MRI compared to case matched controls who did not develop cancer. However, they did not find an association between breast cancer development and pattern of BPE, amount of FGT on MRI, or mammographic density. In another retrospective study of 1533 women (41 of whom developed cancer) with an increased lifetime risk of breast cancer ($\geq 20\text{--}25\%$), Vreemann et al. (33) found no evidence of an association between quantitative measures of BPE and FGT on baseline MRI and overall breast cancer development, but indicated that high amounts of FGT and BPE on these MRIs lead to more false-positive findings. Niell et al. (34) and Saha et al. (35) also investigated the association between quantitative features and future breast cancer risk, and found that quantitative measures of BPE are predictive of breast cancer, but their research population is limited to women at high lifetime risk of breast cancer.

Unlike these other studies which focused on women at high-risk for breast cancer, the current study focused on women with extremely dense breasts in whom baseline MRI typically shows more false-positive findings. We found that volume of enhancing parenchyma on baseline MRI was independently associated with breast cancer occurrence. This finding may differ from those of other studies as in our study we adjusted for conventional characteristics known to be associated with breast cancer risk such as age, BMI, and BPE.

Our study has limitations. First, the follow up time was 6 years, consequently, only short-term associations between parenchymal features and occurrence of future breast cancer could be investigated. Second, information on race and ethnicity were not available for participants in the DENSE trial and thus the findings of this study may lack generalizability.

To conclude, our study demonstrated that quantitative parenchymal features on baseline dynamic contrast-enhanced MRI are independent predictors of breast cancer occurrence in women with extremely dense breasts. Additionally, women with the highest volumes of enhancing parenchyma had approximately two times the occurrence rate of breast cancer than those with the lowest volumes of

enhancing parenchyma. In order to develop personalized screening strategies for breast imaging, more focus is needed on construction and validation of generalizable risk prediction models on the basis of imaging markers.

References

1. Wanders JOP, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Research and Treatment*. 2017;162(1):95–103.
2. Boyd NF, Guo H, Martin LJ. Mammographic Density and the Risk and Detection of Breast Cancer. *New England Journal of Medicine*. 2007;356(3):227–236.
3. Boyd NF, Martin LJ, Yaffe MJ, Minkin S. Mammographic density and breast cancer risk: current understanding and future prospects. *Breast Cancer Research*. 2011;13(6):1–12.
4. Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE. Long-term effects of mammography screening: Updated overview of the Swedish randomised trials. *Lancet*. 2002;359(9310):909–919.
5. Saslow D, Boetes C, Burke W, et al. American cancer society guidelines for breast screening with MRI as an adjunct to mammography. *CA: A Cancer Journal for Clinicians*. 2007;57(7):75–89.
6. Lord SJ, Lei W, Craft P, et al. A systematic review of the effectiveness of magnetic resonance imaging (MRI) as an addition to mammography and ultrasound in screening young women at high risk of breast cancer. *European Journal of Cancer*. 2007;43(13):1905–1917.
7. Bakker MF, De Lange S V., Pijnappel RM, et al. Supplemental MRI screening for women with extremely dense breast tissue. *New England Journal of Medicine*. 2019;381(22):2091–2102.
8. Emaus MJ, Bakker MF, Peeters PHM, 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(2):527–537.
9. 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(1):29–36.
10. 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. *Investigative radiology*. 2020;55(7):438–444.

11. Schoemaker MJ, Nichols HB, Wright LB, et al. Association of Body Mass Index and Age with Subsequent Breast Cancer Risk in Premenopausal Women. *JAMA Oncology*. 2018;4(11):1–10.
12. Duffy SW, Morrish OWE, Allgood PC, et al. Mammographic density and breast cancer risk in breast screening assessment cases and women with a family history of breast cancer. *European Journal of Cancer*. 2018;88(2018):48–56.
13. Dontchos BN, Rahbar H, Partridge SC, et al. Are Qualitative Assessments of Background Parenchymal Enhancement, Amount of Fibroglandular Tissue on MR Images, and Mammographic Density Associated with Breast Cancer Risk? *Radiology*. 2015;276(2):371–380.
14. Arasu VA, Miglioretti DL, Sprague BL, et al. Population-based assessment of the association between magnetic resonance imaging background parenchymal enhancement and future primary breast cancer risk. *Journal of Clinical Oncology*. 2019;37(12):954–963.
15. Liu Z, Li Z, Qu J, et al. Radiomics of multiparametric MRI for pretreatment prediction of pathologic complete response to neoadjuvant chemotherapy in breast cancer: A multicenter study. *Clinical Cancer Research*. 2019;25(12):3538–3547.
16. Park H, Lim Y, Ko ES, et al. Radiomics signature on magnetic resonance imaging: Association with disease-free survival in patients with invasive breast cancer. *Clinical Cancer Research*. 2018;24(19):4705–4714.
17. Veenhuizen SGA, Lange SV de, Bakker MF. Supplemental Breast MRI for Women with Extremely Dense Breasts: Results of the Second Screening Round of the DENSE Trial. *Radiology*. 2021;299(2):278–286.
18. Spak DA, Plaxco JS, Santiago L, Dryden MJ, Dogan BE. BI-RADS fifth edition : A summary of changes. *Diagnostic and Interventional Imaging*. 2017;98(3):179–190.
19. Klein S, Staring M, Murphy K, Viergever M a., Pluim J. elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Transactions on Medical Imaging*. 2010;29(1):196–205.
20. van der Velden BHM, Dmitriev I, Loo CE, Pijnappel RM, Gilhuijs KGA. Association between Parenchymal Enhancement of the Contralateral Breast in Dynamic Contrast-enhanced MR Imaging and Outcome of Patients with Unilateral Invasive Breast Cancer. *Radiology*. 2015;276(3):675–685.
21. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*. Springer US; 2021;18(2):203–211.
22. Ragusi MAA, Velden BHM Van Der, Meeuwis C, et al. Long-term Survival in Breast Cancer Patients Is Associated with Contralateral Parenchymal

- Enhancement at MRI: Outcomes of the SELECT Study. *Radiology*. 2023;000:e22192.
23. Nie K, Chang D, Chen JH, Hsu CC, Nalcioglu O, Su MY. Quantitative analysis of breast parenchymal patterns using 3D fibroglandular tissues segmented based on MRI. *Medical Physics*. 2010;37(1):217–226.
 24. Morris EA. *ACR Bi-Rads® Atlas — Breast MRI*. American College of Radiology. 2013;
 25. Burns RA, Butterworth P, Kiely KM, et al. Multiple imputation was an efficient method for harmonizing the Mini-Mental State Examination with missing item-level data. *Journal of Clinical Epidemiology*. 2011;64(7):787–793.
 26. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: Current practice and guidelines. *BMC Medical Research Methodology*. 2009;9(1):1–8.
 27. Bookstein FL. A method of factor analysis for shape coordinates. *American Journal of Physical Anthropology*. 2017;164(2):221–245.
 28. Ringnér M. What is principal components analysis? *Nature Biotechnology*. 2006;26(3):303–304.
 29. Yefenof J, Goldberg Y, Wiler J, Mandelbaum A, Ritov Y. Self-reporting and screening: Data with right-censored, left-censored, and complete observations. *Statistics in Medicine*. 2022;41(18):3561–3578.
 30. Van Der Velden BHM, Elias SG, Bismeijer T, et al. Complementary value of contralateral parenchymal enhancement on DCE-MRI to prognostic models and molecular assays in high-risk ER+/HER2–breast cancer. *Clinical Cancer Research*. 2017;23(21):6505–6515.
 31. Thompson CM, Mallawaarachchi I, Dwivedi DK, et al. The association of background parenchymal enhancement at breast mri with breast cancer: A systematic review and meta-analysis. *Radiology*. 2019;292(3):552–561.
 32. Acciavatti RJ, Lee SH, Reig B, et al. Beyond Breast Density: Risk Measures for Breast Cancer in Multiple Imaging Modalities. *Radiology*. 2023;306(3):e222575.
 33. Vreemann S, Dalmis MU, Bult P, et al. Amount of fibroglandular tissue FGT and background parenchymal enhancement BPE in relation to breast cancer risk and false positives in a breast MRI screening program: A retrospective cohort study. *European Radiology*. 2019;29(9):4678–4690.
 34. Niell BL, Abdalah M, Stringfield O, et al. Quantitative measures of background parenchymal enhancement predict breast cancer risk. *American Journal of Roentgenology*. 2021;217(1):64–75.
 35. Saha A, Grimm LJ, Ghate S V, et al. Machine learning-based prediction of future breast cancer using algorithmically measured background parenchymal enhancement on high-risk screening MRI. *Journal of Magnetic Resonance*

Imaging. 2019;50(2):456–464.

Chapter 4

Towards Computer-Assisted Triaging of MRI-Guided Biopsy in Preoperative Breast-Cancer Patients

Based on:

*Wang H, Van Der Velden BHM, Ragusi MAA, Veldhuis WB, Viergever MA, Gilhuijs KGA. Toward Computer-Assisted Triaging of Magnetic Resonance Imaging-Guided Biopsy in Preoperative Breast Cancer Patients. **Investigative Radiology.** 256(7):442–449 (2021)*

Abstract

Background and Objective: Incidental MR-detected breast lesions (i.e., additional lesions to the index cancer) pose challenges in the preoperative

workup of patients with early breast cancer. We pursue computer-assisted triaging of MRI-guided breast biopsy of additional lesions at high specificity.

Materials and Methods: We investigated 316 consecutive female patients (aged 26 to 76 years, mean 54 years) with early breast cancer who received preoperative multiparametric breast MRI between 2013 and 2016. In total 82/316 patients (26%) had additional breast lesions on MRI. These 82 patients had 101 additional lesions in total, 51 were benign and 50 were malignant. We collected 4 clinical features, and 46 MRI radiomic features from T₁-weighted dynamic-enhanced (DCE) imaging, high-temporal-resolution DCE imaging (fast-DCE), T₂-weighted imaging, and diffusion-weighted imaging (DWI). A multiparametric computer-aided diagnosis (CAD) model using 10-fold cross validated ridge regression was constructed. The sensitivities were calculated at operating points corresponding to 98%, 95%, and 90% specificity. The model calibration performance was evaluated by calibration-plot analysis and goodness-of-fit tests. The model was tested in an independent testing cohort of 187 consecutive patients from 2017 and 2018 (aged 35 to 76 years, mean 59 years). In this testing cohort, 45/187 patients (24%) had 55 additional breast lesions in total, 23 were benign and 32 were malignant.

Results: The multiparametric CAD model correctly identified 48% of the malignant additional lesions with a specificity of 98%. At specificity 95% and 90%, the sensitivity was 62% and 72%, respectively. Calibration-plot analysis and goodness-of-fit tests indicated that the model was well fitted. In the independent testing cohort, the specificity was 96% and the sensitivity 44% at the 98%-specificity operating point of the training set. At operating points 95% and 90%, the specificity was 83% at 69% sensitivity, and the specificity was 78% at 81% sensitivity, respectively.

Conclusions: The multiparametric CAD model showed potential to identify malignant disease extension with near-perfect specificity in approximately half the population of preoperative patients originally indicated for a breast biopsy. In the other half, patients would still proceed to MRI-guided biopsy to confirm absence of malignant disease. These findings demonstrate the potential to triage MRI-guided breast biopsy.

Introduction

Magnetic resonance imaging (MRI) of the breast is used for preoperative staging of breast cancer and various other indications (1–8). Preoperative MRI is

typically associated with an increased detection rate of additional lesions, caused by the high sensitivity of MRI, while the specificity is limited (9). It detects additional lesions in 5.5% to 37% of the patients (10–13).

Targeted ultrasound examination and subsequent biopsy of the MRI-detected lesions is usually attempted first (14), but ultrasound does not always visualize these lesions to allow biopsy (15). It has been reported that between 11% and 77% of additional lesions detected on MRI are not visible on ultrasound (15–17), while the rate of malignancy can be as high as 54% (14).

If the lesion fails to visualize on ultrasound, MRI-guided biopsy is usually performed next (18). Although this procedure often leads to pathology proof, technical limitations exist, e.g., the difficulty to target lesions near the thorax or the nipple (19). In addition, the procedure is relatively time-consuming and resource-intensive (20). Management of additional lesions detected on preoperative MRI would be greatly facilitated if an approach exists to identify malignant disease with near-perfect accuracy without MRI guided biopsy, so that this subgroup of patients may be triaged directly to surgery.

Methods based on computer-aided diagnosis (CAD) can be pursued to achieve this goal. Several CAD models based on conventional MRI and multiparametric MRI have been reported to discriminate between benign and malignant lesions in the breast (21–29). However, the investigated populations vary. While some studies focused on multiparametric MRI of primary breast cancer, others focused on additional breast lesions using conventional dynamic contrast-enhanced (DCE) MRI only (30,31). To the best of our knowledge, it is currently unknown whether the additional information from multiparametric MRI can be used by CAD to triage preoperative breast biopsy at high specificity.

The aim of our study is to assess the sensitivity of computer-assisted triaging of MRI-guided breast biopsy at high specificity in patients with additional breast lesions on preoperative breast MRI.

Materials and Methods

Study population

The patients in this study were retrospectively collected from the University Medical Center Utrecht, the Netherlands. The Medical Research Ethics Committee of the UMC Utrecht waived ethical review. Analyses were performed

according to the Standards for Reporting of Diagnostic Accuracy Studies) (32).

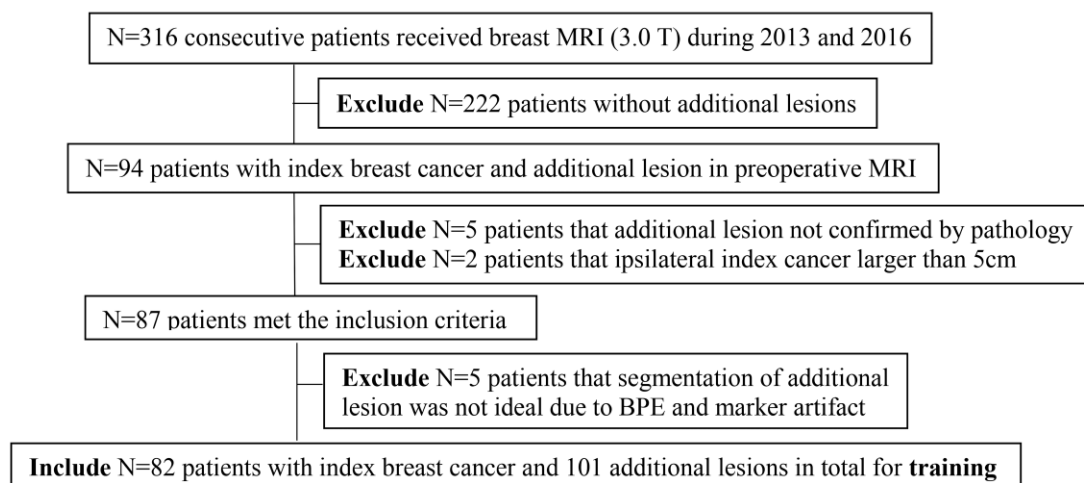


FIGURE 1.1 Flowchart of patient selection in the training cohort (BPE: Background Parenchymal Enhancement)

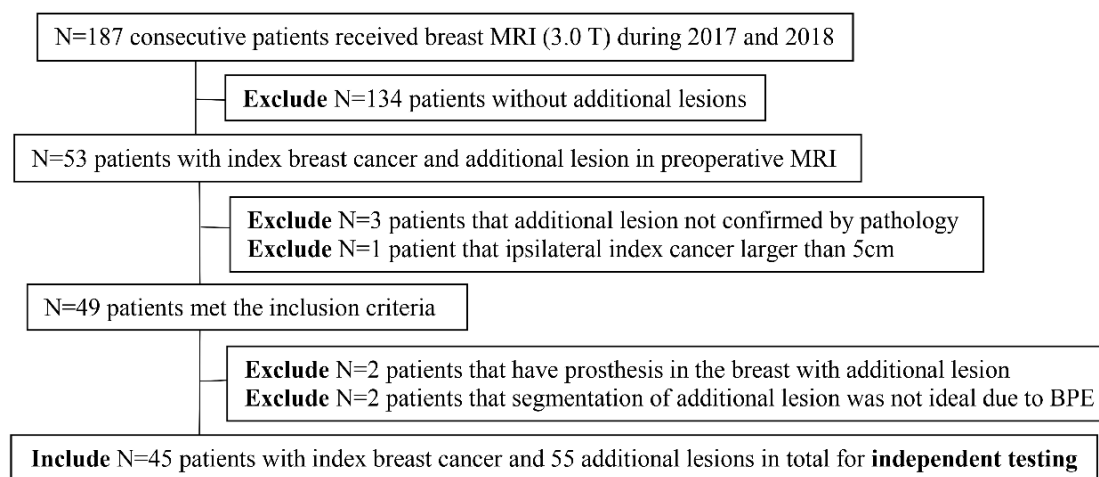


FIGURE 1.2 Flowchart of patient selection in an independent testing cohort (BPE: Background Parenchymal Enhancement)

We performed the study on consecutive patients who received a breast MRI (3.0T) examination between 2013 and 2016 (**Figure 1.1**). Patients were included if they satisfied the following conditions: 1) They were diagnosed with breast cancer and had at least one additional lesion in the ipsilateral or contralateral breast detected on preoperative MRI; 2) The additional lesion was assessed as malignant or benign by pathology; 3) The index cancer was smaller than 5cm if the additional lesion was in the ipsilateral breast (33,34).

We tested the CAD model, which will be described below, on an independent hold-out testing cohort of consecutive patients who received breast MRI (3T) in 2017 or 2018. The same inclusion criteria as those in the training set were used (**Figure 1.2**). The testing cohort was not available during model development,

and was hence not used for training.

TABLE 1. Description of the features		
DCE (39)		
1	Circularity	Ratio between the volume of a sphere with the same diameter as the lesion and the volume of the lesion
2	Irregularity	Ratio between the surface of a sphere with the same volume as the lesion and the surface of the lesion
3	Largest diameter	Largest diameter of the lesion
4	Volume	Volume of the lesion
5	Uptake	Average of top 50% enhancing voxel values $[(I1-I0)/I0]$ in the lesion; I0 and I1 are the pre-contrast and the first post-contrast signal intensities
6	Washout	Average of top 50% washing-out voxel values $[(I4-I1)/I1]$ in the lesion; I1 and I4 are the first and last post-contrast signal intensities
7	Signal enhancing ratio	Average of top 50% Signal-enhancing Ratio values $[(I1-I0)/(I4-I0)]$ in the lesion; I0, I1 and I4 are the, pre-contrast and first and last post-contrast signal intensities
8	Top uptake	Average uptake of the top 10% enhancing lesion voxels
9	Top washout	Average washout of the top 10% washing-out lesion voxels
10	Largest diameter high uptake	Largest diameter of the lesion with more than 100% washin
11	Volume high uptake	Volume of lesion with more than 100% washin
12	Largest diameter washout	Largest diameter of lesion in washout image
13	Volume washout	Volume of lesion in washout image
14	Smoothness1	Radial gradient index in 3D at the first subtraction series
15	Smoothness2	Variation of the radial gradient histogram values at the first subtraction series
16	Smoothness3	Mean of the radial gradient histogram values at the subtraction series where it takes its maximum value
17	Smoothness4	Variation of the radial gradient histogram values at the subtraction series where it takes its maximum value
18	Sharpness1	Mean of the sharpness of the lesion margin at the first subtraction series
19	Sharpness2	Variation in the sharpness of the lesion margin at the first subtraction series
20	Sharpness3	Mean of the sharpness of the lesion margin at the subtraction series where it takes its maximum value
21	Sharpness4	Variation of the sharpness of the lesion margin at the subtraction series where it takes its maximum value
22	Sharpness5	Variance of sharpness of the lesion margin at the subtraction series with maximum mean sharpness
T₂ (27)		

23	T ₂	Minimum intensity in lesion volume
24	T ₂ Q05/pec	5 th Percentile of the intensities present in the lesion volume
25	T ₂ Q10/pec	10 th Percentile of the intensities present in the lesion volume
26	T ₂ Q25/pec	25 th Percentile of the intensities present in the lesion volume
27	T ₂ Q50/pec	50 th Percentile of the intensities present in the lesion volume
28	T ₂ Q75/pec	75 th Percentile of the intensities present in the lesion volume
29	T ₂ Q90/pec	90 th Percentile of the intensities present in the lesion volume
30	T ₂ Q95/pec	95 th Percentile of the intensities present in the lesion volume
31	T ₂ maximum/pec	Maximum of the intensities present in the lesion volume
Fast-DCE (24)		
32	Maximum slope	Maximum slope of uptake contrast agent in lesion volume
33	Maximum enhancement	Maximal normalized intensity in lesion volume
34	Time maximum slope	Time between maximum slope of contrast uptake in descending aorta and lesion volume
35	Final slope (washout)	Intensity gradient at last time point of Fast-DCE
36	General slope (uptake)	Maximal slope of contrast uptake in lesion between time point aorta and any other time point during contrast uptake.
37	Time to enhancement	Time between maximum contrast uptake in descending aorta and start contrast uptake in lesion volume
ADC (27)		
38	Minimum ADC	Minimum intensity in lesion volume
39	Q05 ADC	5 th Percentile of the intensities present in the lesion volume
40	Q10 ADC	10 th Percentile of the intensities present in the lesion volume
41	Q25 ADC	25 th Percentile of the intensities present in the lesion volume
42	Q50 ADC	50 th Percentile of the intensities present in the lesion volume
43	Q75 ADC	75 th Percentile of the intensities present in the lesion volume
44	Q90 ADC	90 th Percentile of the intensities present in the lesion volume
45	Q95 ADC	95 th Percentile of the intensities present in the lesion volume
46	Maximum ADC	Maximum of the intensities present in the lesion volume
*pec = pectoral muscle, T ₂ intensities were normalized to the intensity of the pectoral muscle		

MRI acquisition

Patients underwent MRI in prone position using a Philips 3T MR scanner (Achieva or Ingenia) with a dedicated phased-array bilateral breast coil. The MRI protocols included T1-weighted high-spatial resolution DCE-MRI, high-temporal-resolution DCE imaging (fast-DCE), T2-weighted imaging, and diffusion-weighted imaging (DWI). DCE-MRI consisted of a high-spatial-resolution pre-contrast series, followed by a high-temporal-resolution series after contrast agent injection, followed by 4 or 5 high-spatial-resolution images.

Imaging parameters of the high-spatial DCE-MRI were: repetition time ranging between 3.3 and 5.1 ms, flip angle 10° or 8°, and echo time 2 ms. Parameters of the high-temporal DCE-MRI were: repetition time 2.2 ms, flip angle 10° or 8°, and echo time 1.1 ms. Parameters of the T2-weighted imaging were: repetition time ranging between 5431 and 5959 ms, flip angle 90°, and echo time 60 or 100 ms. The DWI was performed using three b-values (0, 150, and 800 s/mm²). Contrast agent was injected at a rate of 1 ml/sec to a total dose of 0.1 mmol of macrocyclic GBCA gadobutrol (Gadovist®, Bayer AG, Leverkusen, Germany) per kilogram of body weight (35).

Image analysis

Image analysis mainly consisted of registration, semi-automatic lesion segmentation, and multiparametric MRI feature extraction, which has been described in detail before (27). In short, intensity-based image registration of DCE series was performed using Elastix (36). The additional lesions were semi-automatically segmented in 4-D by placing a seed point for constrained volume growing in the target lesion (37). Lesions were segmented by medical doctors HW and MAAR on the basis of lesion location, shape and size using radiological reports, followed by consensus reading by a breast MRI radiologist (WBV.) in case of ambiguity.

Image analysis was performed using MeVisLab (version 3.0, MeVis Medical Solution AG, Bremen, Germany), Python (version 2.7, Python Software Foundation, Beaverton, OR, USA), and MATLAB (version R2017a; Mathworks, Natick, MA, USA).

MRI features

Each lesion was described using 46 radiomic CAD features extracted from MRI: 22 DCE features representing morphology and enhancement kinetics (38,39), six fast-DCE features describing contrast uptake (24) nine T2 intensity features, and nine ADC features (**Table1**). Missing feature values in the training cohort were imputed by multiple imputation. Ten imputation sets were merged based on the mean imputed values (40–42). Features in the testing cohort were complete without missing values.

Number of additional lesions		Total N=101	Benign N=51	Malignant N=50
Patients	Age (years, mean±SD)	54.5±11.6	53.1±11.7	55.9±11.4

Additional lesions	BIRADS category			
	BIRADS 2/3	18(18)	17(33)	1(2)
	BIRADS 4	57(56)	29(57)	28(56)
	BIRADS 5/6	26(26)	5(10)	21(42)
	Laterality (ipsilateral/contralateral)			
	Ipsilateral	55(54)	21(41)	34(68)
	Contralateral	46(46)	30(59)	16(32)
	Visibility on Ultrasound			
	Visible	50(50)	18(35)	32(64)
Invisible	51(50)	33(65)	18(36)	
SD= standard deviation. Numbers represent frequency (percentage) unless stated otherwise.				

TABLE 2.2 Distribution of clinical features in the testing cohort

Number of additional lesions		Total N=55	Benign N=23	Malignant N=32
Patients	Age (years, mean \pm SD)	58.5 \pm 12	56.4 \pm 12.6	60 \pm 11.4
Additional lesions	BIRADS category			
	BIRADS 2/3	9(16)	9(39)	0(0)
	BIRADS 4	28(51)	13(57)	15(47)
	BIRADS 5/6	18(33)	1(4)	17(53)
	Laterality (ipsilateral/contralateral)			
	Ipsilateral	33(60)	9(39)	24(75)
	Contralateral	22(40)	14(61)	8(25)
	Visibility on Ultrasound			
	Visible	36(65)	12(52)	24(75)
Invisible	19(35)	11(48)	8(25)	
SD=standard deviation. Numbers represent frequency (percentage) unless stated otherwise.				

Clinical features

Clinical features were collected by reviewing the medical records, i.e., the ultrasound report, the MRI report, the biopsy report, and the surgical report. The clinical features were limited to those available preoperatively, including 1) age of the patient at diagnosis; 2) Breast Imaging Reporting and Data System (BIRADS) score of the additional lesion, grouped in three categories: 2-3, 4, 5-6; 3) the laterality of the additional lesion with respect to the index tumor (ipsilateral or contralateral); and 4) whether the additional lesion was visible on ultrasound.

CAD model training

The CAD model consisted of 50 features in total (46 MRI features and 4 clinical features). The influence of outliers in regression models is high when the sample

size is small (43). Considering the limited sample size in our study, outliers were corrected. They were defined using Tukey’s method, i.e., 1.5 IQR (interquartile range) below the lower quartile (Q1) or 1.5 IQR above the upper quartile (Q3). Outliers in feature values in the training cohort were winsorized to the nearest whisker (44,45), after which the features values were normalized to values between 0 and 1.

We used ridge regression with 10-fold cross validation (46) and partial area under curve (pAUC) as the loss function. We used pAUC as loss function to be able to optimize on high sensitivity at high specificity (i.e., 90% to 100% range) (47). We chose the regularization parameter lambda at the value where the pAUC is maximum between specificity 90% and 100% in the cross-validation. This lambda value was applied to the training set.

TABLE 3.1 Training performance of the CAD model at three operating points

Operating point	Training Performance
Specificity	Sensitivity Median[95%CI]
90%	72% [61%,86%]
95%	62% [45%,81%]
98%	48% [25%,73%]
CI = Confidence Interval, CAD = Computer-Aided Diagnosis	

TABLE 3.2 Testing performance of the CAD model at three operating points

Training	Testing performance	
Specificity	Sensitivity	Specificity
90%	81%	78%
95%	69%	83%
98%	44%	96%
CAD = Computer-Aided Diagnosis		

Three operating points on the receiver operating characteristic (ROC) curve were examined: at 90%, 95%, and 98% specificity, respectively. The 95% confidence intervals (95%CI) of the corresponding sensitivities on the fitted ROC curve were calculated by bootstrapping (48). Agreement between the observed probability of malignancy and predicted probability of malignancy was evaluated by calibration-plot analysis and Hosmer-Lemeshow tests (49). A p-value above 0.05 in the Hosmer-Lemeshow test indicated a good fit.

CAD model testing

The CAD model derived from model training was frozen and further tested in the independent testing cohort. Feature values in the testing cohort were

transformed to the range in the training cohort and normalized to values between 0 and 1 based on the upper and lower limits in the training cohort. The sensitivity and specificity in the testing cohort were calculated by applying the thresholds corresponding to 90%, 95%, and 98% specificity in the training cohort.

Differences in clinical variables between training and testing set were tested using Chi-squared tests for categorical features and Wilcoxon rank-sum tests for continuous features. All CAD model analyses and corresponding statistical analyses were performed in R (version 3.4.4, R Foundation for Statistical Computing, Vienna, Austria).

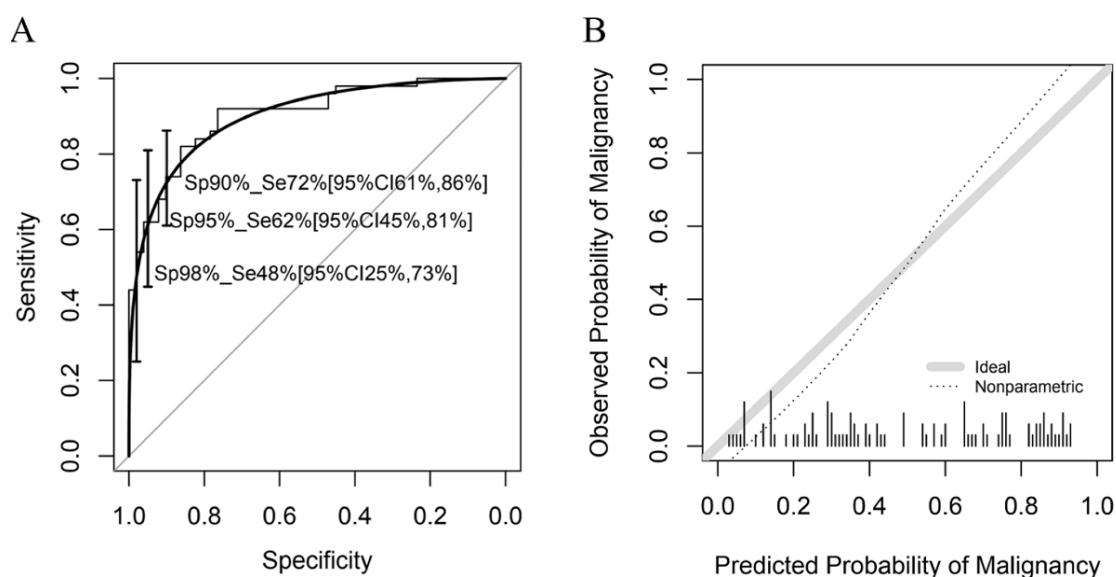


FIGURE 2A. Receiver operating characteristic (ROC) curve of the ridge regression model. Bars represent 95% CI of sensitivities at specificity 90%, 95% and 98% in the fitted ROC curve. Sp = Specificity, Se = Sensitivity. CI = Confidence Interval. **B.** Calibration plot of the training model. The diagonal gray reference line represents perfect calibration. The dashed curve represents the calibration of the model. The x-axis represents the predicted probability of malignancy of additional lesions and the y-axis represents the observed probability of malignancy.

Results

Patients and lesions

From a total of 316 consecutive patients who received breast MRI between 2013 and 2016, 87 patients (28%) met the initial inclusion criteria. Five patients were further excluded because the segmentation of the additional lesions did not result in satisfying demarcations owing to marked background parenchymal enhancement (BPE) or marker artifacts. Inaccurate segmentations caused by severe masking of the lesion margins by BPE could not be corrected by manual

segmentation; The margins were also not sufficiently clear to allow manual correction. Finally, 82/87 patients (94%) who met the inclusion criteria were included in the training cohort (**Figure 1.1**). These 82 patients had 101 additional lesions; 51 were benign and 50 were malignant (**Table 2.1**).

The independent hold-out testing cohort was collected separately after training: From a total of 187 consecutive patients who received breast MRI in 2017 or 2018, 49 patients (26%) met the initial inclusion criteria. Two patients were further excluded because of prostheses in the breast with the additional lesion, and two because the segmentation of the additional lesion was not accurate owing to marked BPE. Finally, 45/49 patients (92%) who met the inclusion criteria were included in the testing cohort (**Figure 1.2**). These 45 patients had 55 additional lesions; 23 were benign and 32 were malignant (**Table 2.2**).

No significant differences were found in clinical variables between training and testing set: age ($P=0.054$), BIRADS score ($P=0.651$), laterality ($P=0.618$), visibility on ultrasound ($P=0.081$)

CAD model training

The multiparametric CAD model identified 72% [61%, 86%] of the malignant lesions at 90% specificity, 62% [45%, 81%] of the malignant lesions at 95% specificity, and 48% [25%, 73%] of the malignant lesions at 98% specificity (**Table 3.1, Figure 2A**). The Hosmer-Lemeshow test indicated good fit of the model ($P=0.29$), which was demonstrated by the calibration plot (**Figure 2B**). The coefficients of the features are listed in **Figure 3**.

CAD model testing

We applied the training model to the independent testing cohort according to the thresholds from the training set at specificity of 90%, 95%, and 98%, resulting in sensitivity of 81%, 69%, 44% respectively, and specificity of 78%, 83%, 96% respectively (**Table 3.2, Figure 4A**). The sensitivities in the testing set fell within the 95%CI of the training performance. The Hosmer-Lemeshow test between predicted and actual probability indicated good fit on the testing set ($P=0.36$), which was demonstrated by the calibration plot (**Figure 4B**). A representative MRI case is shown in **Figure 5**.

Discussion

Incidental MR-detected breast lesions (i.e., additional lesions to the index cancer) pose challenges in the preoperative workup of patients with early breast cancer.

We constructed a multiparametric CAD model based on ridge regression to identify malignant additional breast lesions preoperatively with high specificity. The CAD model shows potential to spare patients an MRI-guided biopsy in approximately half the number of cases. In the other cases, patients would be referred to MRI-guided biopsy.

Quantitative CAD features may contain complementary information to that perceived by radiologists. For example, CAD features are representative of tissue phenotypes at genetic and molecular level (50). In recent radiogenomics study (50) in 295 breast-cancer patients, the radiomic features Uptake, Smoothness, and Sharpness were linked with expression of ribosomal proteins in the tumor. In addition, Smoothness and Shape were associated with gene expressions involved in the extracellular matrix and collagen production, indirectly relating to cancer progression. Especially Sharpness shows relatively high absolute Ridge regression coefficients in the current study, suggesting that the CAD discriminates between benign and malignant lesions using MRI markers related to tumor progression at biomolecular tissue level.

CAD is believed to have great potential to aid radiologists in better understanding the underlying tumor behavior (51,52). There have been several CAD models based on conventional MRI sequences or multiparametric MRI sequences for differential diagnosis of lesions detected in the breast. The diagnostic performance (AUC) ranged between 0.83 and 0.88 to distinguish between benign and malignant disease (21,31,53,54).

To the best of our knowledge, no CAD models have been reported to attempt triaging patients in a preoperative setting for MRI-guided biopsy. For this purpose, we trained the CAD model to optimize the pAUC to favor high specificity models. While AUC is a routine statistic to evaluate model performance, the pAUC statistic has gained increasing popularity because many clinical problems require high performance in the region of the ROC curve corresponding to high specificity (i.e., to minimize false-positive rate) (47). Hence, using the pAUC statistics over a limited range of high specificity (e.g., 90%-100%) may be more appropriate to consider than AUC over the whole range.

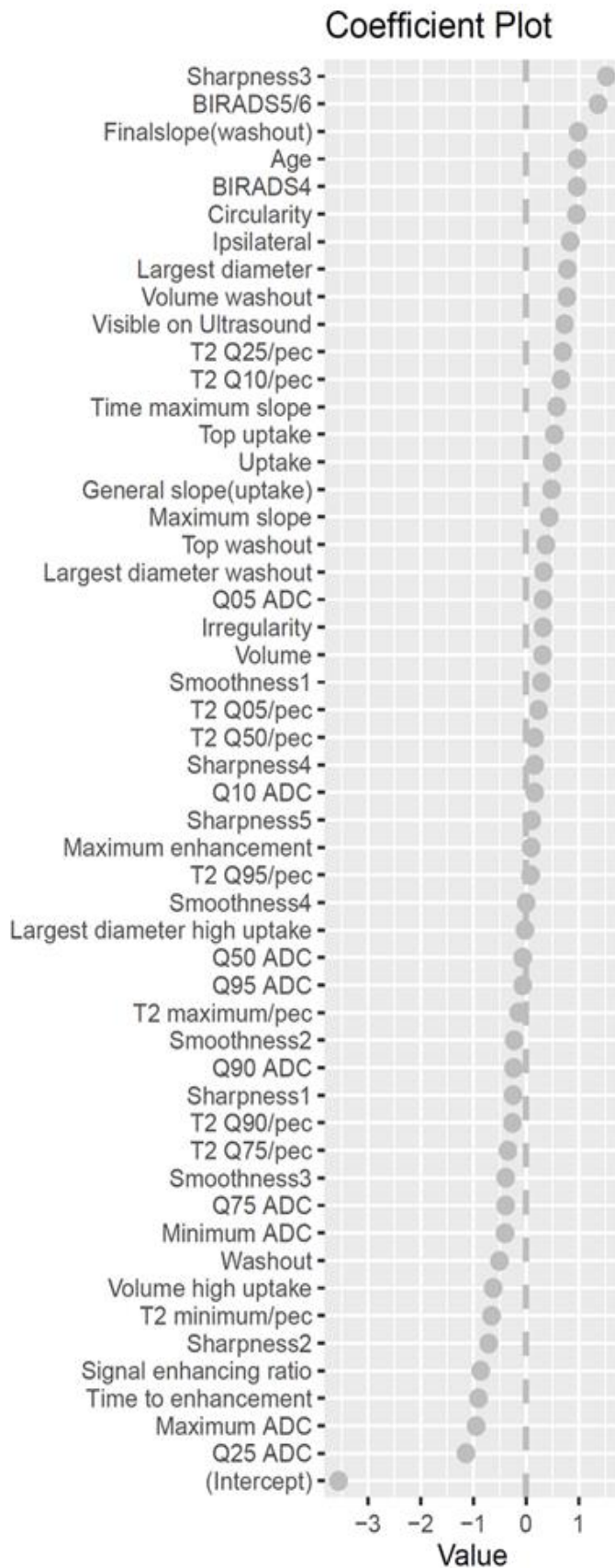


FIGURE3. Coefficients of features in ridge regression

The different levels of high specificity tested (90%, 95%, 98%) represent different probabilities to which the CAD can be set to correctly triage patients to surgery without biopsy (i.e., triage patients without causing extended surgery for false-positive disease extension). The rationale for focusing on high specificity (i.e., establishing that a lesion is malignant with high certainty, instead of establishing that it is not malignant with high certainty) is because we envision the CAD to be used to triage patients to MRI-guided biopsy or surgery: if the CAD shows that the lesion is malignant with certainty, the patient may triage directly to surgery (where receptor type can be determined). If, however, the CAD determines that the lesion is not malignant, the patient triages to MRI-guided biopsy to confirm absence of cancer, avoiding extended surgery due to false-positive findings. Either way, the result is that a number of MRI-guided biopsies would be avoided without missing cancer.

Ridge regression (46) is sometimes used for classification for reasons outlined below. In our study, it estimates the posterior probability of a case belonging to a certain class. By thresholding this posterior probability, the regression result is converted to a hard classification. There are many different classification methods possible. Because the aim of our study was not to perform comparative analysis of different classification models due to the relatively limited sample size, we used the prior knowledge that ridge regression is well suited for relatively small datasets; It minimizes the risk of overfitting, and typically shows optimal discrimination performance with the lowest prediction error among regression shrinkage strategies (55). Ridge regression indicates feature importance by the absolute magnitude of the feature coefficients. To trade off bias and variance of the model, we used 10-fold cross-validation (56) and tested our model in an independent dataset.

Features with high absolute coefficients included three features from the DCE series (Sharpness3, Circularity, Signal enhancing ratio), two features from fast-DCE (Time to enhancement, Final slope), two features from DWI (25th percentile ADC, Maximum ADC), and two clinical features (age at diagnosis, BIRADS-score). Conversely, T2 features yielded relatively small coefficient values, which was consistent with previously published results (27). A recent publication concludes that mean ADC is sufficient to improve diagnostic performance of breast MRI (57). The results from our multivariable analysis show that the 25th percentile of ADC and maximum ADC are important for differential diagnosis of additional breast lesions. An underlying reason for these differences in observation (apart from the differences in analysis) might be the difference in study population and population size. Our population (127 patients, 156 additional lesions) consists exclusively of preoperative breast-cancer patients with additional lesions, whereas the population of McDonald et al. (67 patients, 81 lesions) also contains patients who underwent breast MRI for other indications such as screening. Other studies also suggest value of other measures from the ADC histogram (27,58).

The false-positive lesion associated with the highest CAD probability for malignant disease in the testing cohort was Lobular Carcinoma in Situ (LCIS). LCIS is considered a non-malignant breast lesion, but it is associated with increased risk of developing invasive breast cancer (59,60). Although LCIS is classified as a benign lesion by histopathology, it is considered a direct precursor of invasive lobular carcinoma (61,62). The representation of LCIS on MRI is

subtle. LCIS typically shows focally distributed heterogeneous enhancement, or irregular mass with homogeneous enhancement (61). The reproducibility of the segmentation method was investigated in previous study (37) where two independent observers performed the semi-automated segmentation on 46 breast lesions, leading to negligible differences in segmentation results (adjusted R-square value 0.99). In addition, the results in the current study were verified in an independent testing database.

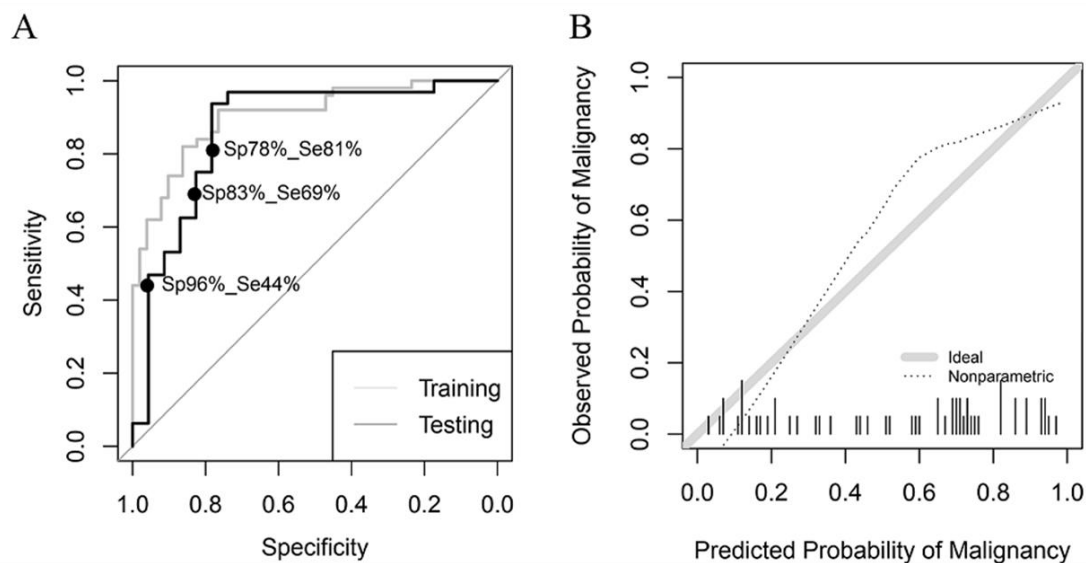


FIGURE 4A. Receiver operating characteristic (ROC) curve of the testing cohort. Black curve: ROC curve of the testing cohort. Gray curve: reference ROC curve of the training model, Sp = Specificity, Se = Sensitivity. **4B.** Calibration plot of the testing cohort. The diagonal gray reference line represents perfect calibration. The dashed curve represents the calibration of the model. The x-axis represents the predicted probability of malignancy of additional lesions and the y-axis represents the observed probability of malignancy.

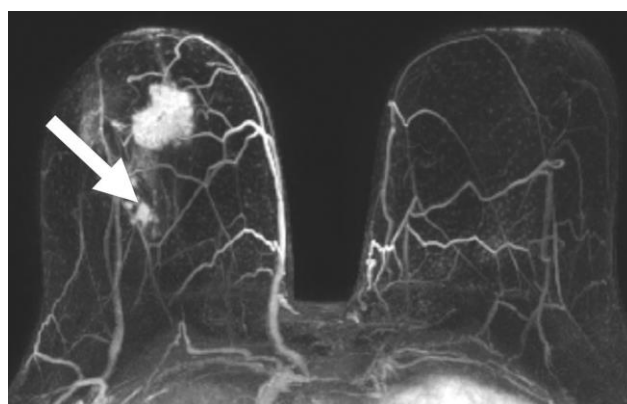


FIGURE 5. Example of a patient with an index breast cancer (infiltrating ductal carcinoma, IDC) and an additional lesion detected on the preoperative MRI (arrow). The image shows a maximum intensity projection of the first postcontrast T₁-weighted MRI minus the precontrast T₁-weighted MRI. Computer-aided diagnosis (CAD) classified the additional lesion (BIRADS 4) as malignant with predicted probability of malignancy of 88%. The lesion was confirmed on pathology as IDC.

There were some limitations in this study. First, the sample size was limited, which was mitigated in part by stringent regularization and by testing the performance in an independent cohort. In addition, although we tested our model in an independent cohort from our hospital, we have not yet tested it in other hospitals. Broader validation of the model is recommended to enable transition to clinical practice of CAD-based triaging of MRI-guided biopsy in preoperative breast-cancer patients.

To conclude, we presented a multiparametric CAD model that showed potential to identify malignant disease extension with near-perfect specificity in approximately half the population of preoperative patients originally indicated for a breast biopsy. In the other half, patients would still proceed to MRI-guided biopsy to confirm absence of malignant disease. These findings demonstrate the potential to triage MRI-guided breast biopsy.

References

1. Pediconi F, Miglio E, Telesca M, et al. Effect of preoperative breast magnetic resonance imaging on surgical decision making and cancer recurrence rates. *Investigative Radiology*. 2012;47(2):128–135.
2. Tillman GF, Orel SG, Schnall MD, Schultz DJ, Tan JE, Solin LJ. Effect of breast magnetic resonance imaging on the clinical management of women with early-stage breast carcinoma. *Journal of Clinical Oncology*. 2002;20(16):3413–3423.
3. Lee SG, Orel SG, Woo IJ, et al. MR imaging screening of the contralateral breast in patients with newly diagnosed breast cancer: Preliminary results. *Radiology*. 2003;226(3):773–778.
4. Esserman L, Hylton N, Yassa L, Barclay J, Frankel S, Sickles E. Utility of Magnetic Resonance Imaging in the Management of Breast Cancer: Evidence for Improved Preoperative Staging. *Journal of Clinical Oncology*. 1999;17(1):110–119.
5. Sung JS, Li J, Da Costa G, et al. Preoperative breast MRI for early-stage breast cancer: Effect on surgical and long-term outcomes. *American Journal of Roentgenology*. 2014;202(6):1376–1382.
6. Peters NHGM, Van Esser S, Van Den Bosch MAAJ, et al. Preoperative MRI and surgical management in patients with nonpalpable breast cancer: The MONET - Randomised controlled trial. *European Journal of Cancer*. Elsevier Ltd; 2011;47(6):879–886.
7. Turnbull L, Brown S, Harvey I, et al. Comparative effectiveness of MRI in breast cancer (COMICE) trial: a randomised controlled trial. *The Lancet*. Elsevier Ltd; 2010;375(9714):563–571.

8. Li X, Abramson RG, Arlinghaus LR, et al. Multiparametric magnetic resonance imaging for predicting pathological response after the first cycle of neoadjuvant chemotherapy in breast cancer. *Investigative Radiology*. 2015;50(4):195–204.
9. Kuhl CK. The Changing World of Breast Cancer: A Radiologist's Perspective. *Investigative Radiology*. 2015;50(9):615–628.
10. Houssami N, Ciatto S, Macaskill P, et al. Accuracy and surgical impact of magnetic resonance imaging in breast cancer staging: Systematic review and meta-analysis in detection of multifocal and multicentric cancer. *Journal of Clinical Oncology*. 2008;26(19):3248–3258.
11. Braun M, Pölcher M, Schrading S, et al. Influence of preoperative MRI on the surgical management of patients with operable breast cancer. *Breast Cancer Research and Treatment*. 2008;111(1):179–187.
12. Barco I, Chabrera C, García-Fernández A, et al. Magnetic resonance imaging in the preoperative setting for breast cancer patients with undetected additional disease. *European Journal of Radiology*. 2016;85(10):1786–1793.
13. Liberman L, Morris EA, Kim CM, et al. MR Imaging Findings in the Contralateral Breast of Women with Recently Diagnosed Breast Cancer. *AJR American journal of roentgenology*. 2003;180(February):333–341.
14. Leung JWT. Second-look ultrasound: Only for biopsy or more? *European Journal of Radiology*. 2012;81(SUPPL1):87–89.
15. LaTrenta LR, Menell JH, Morris EA, Abramson AF, Dershaw DD, Liberman L. Breast lesions detected with MR imaging: Utility and histopathologic importance of identification with US. *Radiology*. 2003;227(3):856–861.
16. Beran LA, Liang W, Nims T, Paquelet J, Sickel-Santanello B. Correlation of targeted ultrasound with magnetic resonance imaging abnormalities of the breast. *American Journal of Surgery*. 2005;190(4):592–594.
17. Wiratkapun C, Duke D, Nordmann AS, et al. Indeterminate or Suspicious Breast Lesions Detected Initially with MR Imaging. Value of MRI-directed Breast Ultrasound. *Academic Radiology*. 2008;15(5):618–625.
18. Mann RM, Balleyguier C, Baltzer PA, et al. Breast MRI: EUSOBI recommendations for women's information. *European Radiology*. 2015;25(12):3669–3678.
19. Papalouka V, Kilburn-Toppin F, Gaskarth M, Gilbert F. MRI-guided breast biopsy: a review of technique, indications, and radiological–pathological correlations. *Clinical Radiology*. Elsevier; 2018;73(10):908.e17-908.e25.
20. Leung JWT. Utility of second-look ultrasound in the evaluation of MRI-detected breast lesions. *Seminars in Roentgenology*. Elsevier Inc.; 2011;46(4):260–274.
21. Truhn D, Schrading S, Haarbuerger C, Schneider H, Merhof D, Kuhl C. Radiomic versus Convolutional Neural Networks Analysis for Classification of Contrast-

- enhancing Lesions at Multiparametric Breast MRI. *Radiology*. 2019;290(3):290–297.
22. 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. *Investigative Radiology*. 2019;54(6):325–332.
 23. Van Zelst JCM, Vreemann S, Witt HJ, et al. Multireader Study on the Diagnostic Accuracy of Ultrafast Breast Magnetic Resonance Imaging for Breast Cancer Screening. *Investigative Radiology*. 2018;53(10):579–586.
 24. Dalmiş MU, Gubern-Mérida A, Vreemann S, Karssemeijer N, Mann R, Platel B. A computer-aided diagnosis system for breast DCE-MRI at high spatiotemporal resolution. *Medical Physics*. 2016;43(1):84–94.
 25. Ballesio L, Savelli S, Angeletti M, et al. Breast MRI: Are T2 IR sequences useful in the evaluation of breast lesions? *European Journal of Radiology*. 2009;71(1):96–101.
 26. Kuroki Y, Nasu K, Kuroki S, et al. Diffusion-weighted imaging of breast cancer with the sensitivity encoding technique: Analysis of the apparent diffusion coefficient value. *Magnetic Resonance in Medical Sciences*. 2004;3(2):79–85.
 27. 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. *Investigative radiology*. 2020;55(7):438–444.
 28. Dietzel M, Krug B, Clauser P, et al. A Multicentric Comparison of Apparent Diffusion Coefficient Mapping and the Kaiser Score in the Assessment of Breast Lesions. *Investigative Radiology*. 2020;Publish Ah(00):1–9.
 29. Jiang Y, Edwards A V., Newstead GM. Artificial Intelligence Applied to Breast MRI for Improved Diagnosis. *Radiology*. 2020;00(16):7–10.
 30. Deurloo EE, Peterse JL, Rutgers EJT. Additional breast lesions in patients eligible for breast-conserving therapy by MRI: Impact on preoperative management and potential benefit of computerised analysis. *European Journal of Cancer*. 2005;41:1393–1401.
 31. Deurloo EE, Muller SH, Peterse JL, Gilhuijs KGA. Clinically and mammographically occult breast lesions on MR images: Potential effect of computerized assessment on clinical reading. *Radiology*. 2005;234(3):693–701.
 32. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies¹. *Radiology*. 2015;277(3):826–832.
 33. Bold RJ. Surgical Management of Breast Cancer: Today and Tomorrow. *Cancer biotherapy & radiopharmaceuticals*. 2002;17(1):1–9.
 34. Kaviani A, Sodagari N, Sheikhabaei S, et al. From Radical Mastectomy to Breast-Conserving Therapy and Oncoplastic Breast Surgery: A Narrative Review Comparing Oncological Result, Cosmetic Outcome, Quality of Life, and Health Economy. *ISRN*

Oncology. 2013;2013:1–6.

35. Emaus MJ, Bakker MF, Peeters PHM, 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(2):527–537.

36. Klein S, Staring M, Murphy K, Viergever M a., Pluim J. elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Transactions on Medical Imaging*. 2010;29(1):196–205.

37. Alderliesten T, Schlieff A, Peterse J, et al. Validation of semiautomatic measurement of the extent of breast tumors using contrast-enhanced magnetic resonance imaging. *Investigative Radiology*. 2007;42(1):42–49.

38. Gilhuijs KGA, Deurloo EE, Muller SH, Peterse JL, Schultze Kool LJ. Breast MR imaging in women at increased lifetime risk of breast cancer: clinical system for computerized assessment of breast lesions initial results. *Radiology*. 2002;225(3):907–916.

39. Gilhuijs KGA, Giger ML, Bick U. Computerized analysis of breast lesions in three dimensions using dynamic magnetic-resonance imaging. *Medical Physics*. 1998;25(9):1647–1654.

40. Burns RA, Butterworth P, Kiely KM, et al. Multiple imputation was an efficient method for harmonizing the Mini-Mental State Examination with missing item-level data. *Journal of Clinical Epidemiology*. 2011;64(7):787–793.

41. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: Current practice and guidelines. *BMC Medical Research Methodology*. 2009;9(1):1–8.

42. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons, Inc.; 1987.

43. Cousineau D, Chartier S. Outliers detection and treatment : A review . *International Journal of Psychological Research*,. 2010;3(1):58–67.

44. Hellerstein JM. *Quantitative Data Cleaning for Large Databases*. 2008;1–42.

45. Tukey J. *Exploratory Data Analysis*. Addison-Wesley. Addison-Wesley; 1977.

46. Hoerl AE, Kennard RW. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*. 1970;42(1):80–86.

47. Ma H, Bandos AI, Rockette HE, Gur D. On use of partial area under the ROC curve for evaluation of diagnostic performance. *Statistics in Medicine*. 2013;32(20):3449–3458.

48. Robin X, Turck N, Hainard A. Package ‘ pROC ’. 2020;

49. Helmreich JE. *Regression Modeling Strategies with Applications to Linear Models, Logistic and Ordinal Regression and Survival Analysis (2nd Edition)*. *Journal of Statistical Software*. 2016;70(Book Review 2).

50. Bismeyer T, van der Velden BHM, Canisius S, et al. Radiogenomic Analysis of Breast Cancer by Linking MRI Phenotypes with Tumor Gene Expression. *Radiology*. 2020;296(2):277–287.
51. Li H, Zhu Y, Burnside ES, et al. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set. *npj Breast Cancer*. 2016;2(1):1–10.
52. Limkin EJ, Sun R, Dercle L, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Annals of Oncology*. 2017;28(6):1191–1206.
53. 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(11):1–16.
54. Bhooshan N, Giger M, Lan L. Combined Use of T2-weighted MRI and T1-weighted DCE-MRI in the Automated Analysis of Breast Lesions. *Magnetic Resonance in Medicine*. 2012;66(2):555–564.
55. van Smeden M, Moons KGM, de Groot JAH, et al. Sample size for binary logistic prediction models: Beyond events per variable criteria. *Statistical Methods in Medical Research*. 2019;28(8):2455–2474.
56. Kohavi R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *International Joint Conference on Artificial Intelligence*. 1995;
57. McDonald ES, Romanoff J, Rahbar H, et al. Mean Apparent Diffusion Coefficient Is a Sufficient Conventional Diffusion-weighted MRI Metric to Improve Breast MRI Diagnostic Performance: Results from the ECOG-ACRIN Cancer Research Group A6702 Diffusion Imaging Trial. *Radiology*. 2020;00:1–11.
58. Bickel H, Pinker K, Polanec S, et al. Diffusion-weighted imaging of breast lesions: Region-of-interest placement and different ADC parameters influence apparent diffusion coefficient values. *European Radiology*. *European Radiology*; 2017;27(5):1883–1892.
59. Thomas P. Diagnosis and management of high-risk breast lesions. *Journal of the National Comprehensive Cancer Network*. 2018;16(11):1391–1396.
60. King TA, Pilewskie M, Muhsen S, et al. Lobular carcinoma in situ: A 29-year longitudinal experience evaluating clinicopathologic features and breast cancer risk. *Journal of Clinical Oncology*. 2015;33(33):3945–3952.
61. Chu AJ, Cho N, Park IA, Cho SW. Features of pure lobular carcinoma in situ on magnetic resonance imaging associated with immediate re-excision after lumpectomy. *Journal of Breast Cancer*. 2016;19(2):199–205.
62. Cao D, Polyak K, Halushka MK, et al. Serial analysis of gene expression of lobular carcinoma in situ identifies down regulation of claudin 4 and overexpression of matrix metalloproteinase 9. *Breast Cancer Research*. 2008;10(5):1–10.

Chapter 5

Synchronous Breast Cancer: Phenotypic Similarities on MRI

Based on:
Wang H, Van Der Velden BHM, Chan HSM, Loo CE, Viergever MA, Gilhuijs KGA.
Synchronous Breast Cancer: Phenotypic Similarities on MRI.
***Journal of Magnetic Resonance Imaging.* 51(6):1858–1867(2020)**

Abstract

Background and Objective: Previous studies have shown discrepancies between index and synchronous breast cancer in histology and molecular phenotype. It is yet unknown whether this observation also applies to the MR imaging phenotype. We aim to investigate whether the appearance of breast cancer on MRI (i.e. phenotype) is different from that of additional breast cancer (i.e. synchronous cancer), and whether such difference, if it exists, is associated with prognosis.

Materials and Methods: 464 consecutive patients with early-stage ER+/HER2- breast cancer were included; 34/464 (7.3%) had 44 synchronous cancers in total (34 ipsilateral, 10 contralateral). We assessed imaging phenotype using 50 quantitative features from each cancer and applied principal component analysis (PCA) to identify independent properties. The degree of phenotype difference was assessed. Association between phenotype differences and prognosis in terms of Nottingham prognostic index (NPI) and PREDICT score were analyzed.

Results: PCA identified eight components in patients with ipsilateral synchronous cancer. Six out of eight were significantly different between index and synchronous cancer. These components represented features describing texture (Three Components, $P < 0.001$, $P < 0.001$, $P = 0.004$), size ($P < 0.001$), smoothness ($P < 0.001$), and kinetics ($P = 0.004$). Phenotype differences in terms of the six components were split in tertiles. Larger phenotype differences in size, kinetics and texture were associated with significantly worse prognosis in terms of NPI ($P = 0.019$, $P = 0.045$, $P = 0.014$), but not for PREDICT score ($P = 0.109$, $P = 0.479$, $P = 0.109$). PCA identified six components in patients with contralateral synchronous cancer. None were significantly different from the index cancer ($P = 0.178$, $P = 0.178$, $P = 0.178$, $P = 0.326$, $P = 0.739$, $P = 0.423$).

Conclusion: The MRI phenotype of ER+/HER2- breast cancer was different from that of ipsilateral synchronous cancer and large phenotype difference was associated with worse prognosis. No significant difference was found for synchronous contralateral cancer.

Introduction

Synchronous breast cancer refers to breast cancer detected simultaneously with an index breast cancer, but is physically separated (1). The incidence of synchronous breast cancer varies and is largely dependent on the criteria used in imaging and pathology (2,3). Synchronous breast cancer rate can reach as high as 38% (4–6).

Discrepancies in prognostic markers between the index cancers and their synchronous counterparts may have impact on systemic treatment of patients (7). It has been observed that patients with discrepant prognostic markers between index and the corresponding synchronous cancer have worse long-term survival than patients with congruent markers (8–12). It is yet unknown, however, whether the imaging phenotype of the index cancer also differs from that of the synchronous cancer, and if so, whether such difference is related to patient's prognosis.

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) has been regarded as the most sensitive method for detection of breast cancer, ranging between 89% and 100% (13,14). Quantitative analysis of the phenotype of breast cancer on MRI may extract subtle but reproducible information that is imperceptible to radiologists' eyes (15,16), thus providing more detail to compare phenotypes (17,18). The primary aim of this study was to determine whether the MRI phenotypes of index cancers and their synchronous counterparts differ in series of consecutive patients with early breast cancer. The second aim was to explore whether this difference, if it exists, is associated with patient prognosis.

Materials and Methods

Patients and lesions

This study was performed after approval of the institutional review board and with written informed consent of all patients. In total, 628 patients were collected. We retrospectively analyzed the prospectively collected data from the MARGINS study (Multi-modality Analysis and Radiological Guidance IN breast conServing therapy), which was conducted between 2000 and 2008, patients who were included after being diagnosed with early-stage breast cancer for which breast conserving therapy was indicated based on physical examination, mammography, and ultrasound, had an additional preoperative breast MRI. The index breast cancer was confirmed by fine-needle aspiration

cytology or core needle biopsy.

We evaluated patients with pathology-proven synchronous breast cancer. To eliminate the influence of intrinsic differences in terms of immunohistochemical (IHC) subtype of the index breast cancer, and due to limitation of sample size, we focused on patients with estrogen receptor positive and human epidermal growth factor receptor 2 negative (ER+/HER2-) primary cancer.

Clinicohistopathological variables included age at diagnosis, location of synchronous cancer (ipsilateral or contralateral), largest diameter of index and synchronous breast cancer, number of positive axillary lymph nodes, histologic grade, and IHC subtype of index cancer.

The number of positive lymph nodes was determined by sentinel node biopsy, and combined with axillary lymph node dissection where available. The cases were grouped into three categories: none, one to three, and four or more positive lymph nodes.

Histologic grade was assessed according to the Bloom and Richardson classification (19). Tumors were classified as estrogen receptor positive if more than 10% of the cells were stained positive. Tumors were classified as HER2 positive when scored at least 3 at IHC or when in situ hybridization demonstrated gene amplification, otherwise classified as HER2 negative.

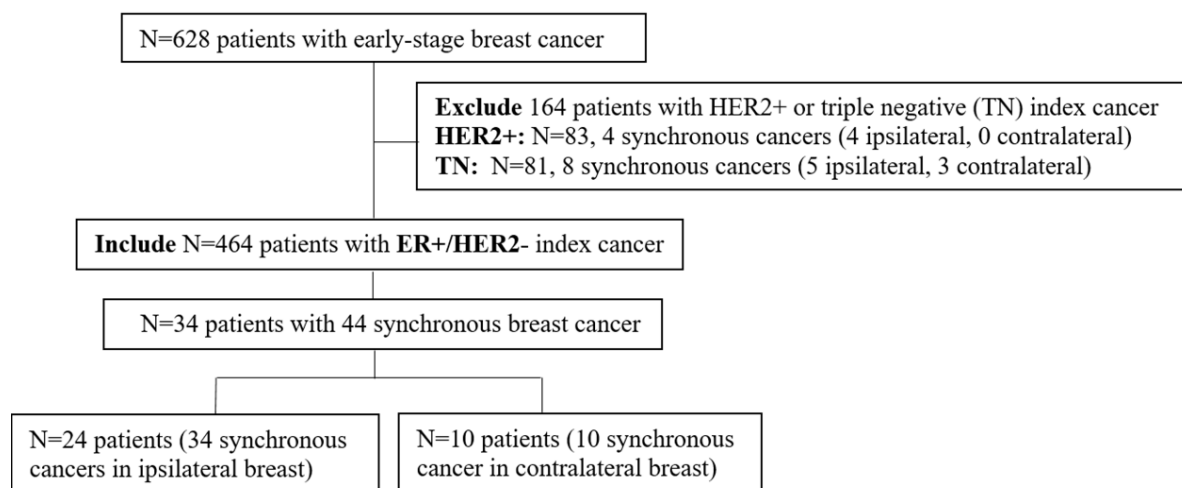


FIGURE 1. Flowchart of selection of patients included in this study

Imaging phenotype identification

Patients underwent MRI in prone position using a 1.5 T scanner (Magnetom;

TABLE 1. Feature list extracted from DCE-MRI			
ID	Texture Feature list	ID	Conventional Feature list
1	washin_Angular_Second_Moment	29	circularity
2	washin_Contrast	30	irregularity
3	washin_Correlation	31	volume
4	washin_Sum_of_Squares._Variance	32	largest_diameter
5	washin_Inverse_Difference_Moment	33	uptake_speed
6	washin_Sum_Average	34	washout
7	washin_Sum_Variance	35	SER
8	washin_Sum_Entropy	36	top_init_enhancement
9	washin_Entropy	37	top_late_enhancement
10	washin_Difference_Variance	38	vol_init_enhancement_GT100
11	washin_Difference_Entropy	39	ld_init_enhancement_GT100
12	washin_Measure_of_Correlation_1	40	volume_late_LT0
13	washin_Measure_of_Correlation_2	41	largest_diameter_late_LT0
14	washin_Maximal_Correlation_Coefficient	42	mean_sharpness
15	washout_Angular_Second_Moment	43	variation_sharpness
16	washout_Contrast	44	mean_sharpness_frame2
17	washout_Correlation	45	variation_sharpness_frame2
18	washout_Sum_of_Squares._Variance	46	variation_smoothness
19	washout_Inverse_Difference_Moment	47	mean_smoothness
20	washout_Sum_Average	48	std_rgh_val_frame2
21	washout_Sum_Variance	49	rad_grad_ind_frame2
22	washout_Sum_Entropy	50	lesion_to_nipple_relative_distance
23	washout_Entropy		
24	washout_Difference_Variance		
25	washout_Difference_Entropy		
26	washout_Measure_of_Correlation_1		
27	washout_Measure_of_Correlation_2		
28	washout_Maximal_Correlation_Coefficient		
DCE-MRI = Dynamic Contrast-Enhanced Magnetic Resonance Imaging			

Siemens Medical Systems, Erlangen, Germany) with a double-breast array coil. Five consecutive scans at intervals of 90 s were performed: one prior to and four after contrast administration. Contrast-enhanced scans were made after intravenous injection with the gadolinium-based contrast agent Gadoteridol (Prohance; Bracco- Byk Gulden, Konstanz, Germany) at 0.1 mmol/ kg body

weight. The following parameters were used: 3-D coronal T1-weighted sequence; repetition time 8.1 ms; echo time 4.0 ms; isotropic voxels 1.35 x 1.35 x 1.35 mm³, without fat suppression.

The index breast cancer and the corresponding synchronous cancer were segmented using a semi-automatic method that was previously reported (20-21). A dedicated breast radiologist (CL) with more than 15 years of experience manually checked the segmentations. The imaging phenotype of each segmentation was described using 50 features: 21 features representing the size, sharpness, smoothness, and enhancement kinetics (22-23). 28 texture features of wash-in and washout (24), and 1 feature describing the relative distance of the cancer to the nipple in relation to breast size in the inferior-posterior direction. The position of the cancer was defined by the center of mass of the segmentation. These features were extracted using in-house developed software in C++ and Python3.7.3, texture features were extracted using the Mahotas package (25). These 50 features were listed in **Table 1**.

Phenotype difference analysis

Association between phenotype difference and prognosis in terms of Nottingham prognostic index (NPI) and PREDICT score were analyzed. NPI was defined as

TABLE 2. Characteristics of patients and cancers				
	Features	Total (N =44)	Ipsilateral (N=34)	Contralateral
Patient	Age (years, mean ±SD)	53.5±7.5	52.5±7.8	57.1±5.4
Synchronous breast cancer	Largest Diameter (mm, mean ±SD)	11.9±3.7	11.3±3.4	13.9±3.8
Index cancer	Largest Diameter (mm, mean ±SD)	21.7±9.1	21.7±8.3	21.6±11.8
	Histological grade			
	Grade I	19(43)	14(41)	5(50)
	Grade II	23(52)	18(53)	5(50)
	Grade III	2(5)	2(6)	0(0)
	Lymph nodes positive			
	0	25 (57)	17(50)	8(80)
1 to 3	13(30)	12(35)	1(10)	
4 or more	6(13)	5(15)	1(10)	
SD = Standard deviation. Numbers represent frequency (percentage) unless stated otherwise.				

$(0.2 * S) + N + G$, where S represents the largest diameter of the index cancer in centimeter; N is 1 for no positive lymph node, 2 for 1 to 3 positive lymph nodes, and 3 for more than 3 positive lymph nodes; and G is the histologic grade (26). The PREDICT score was calculated through the PREDICT version 2.1 model (27), which is a breast cancer prognostication and treatment benefit prediction model, and estimates 10-year survival probability on the basis of patient age, tumor size, tumor grade, number of positive nodes, ER status, HER2 status, KI67 status, mode of detection, and adjuvant chemotherapy regimen.

TABLE 3. Ipsilateral group, PCA identified 8 components explaining 92% variance

	RC1	RC2	RC3	RC4	RC5	RC6	RC7	RC8
Eigenvalue	23.3	7.4	6.0	3.3	2.1	1.7	1.3	1.0
Variance%	27%	25%	11%	9%	7%	7%	4%	2%
Cumulative Variance%	27%	52%	63%	72%	79%	86%	90%	92%

PCA = Principal Component Analysis; RC= Rotated Component
 RC1=Texture1; RC2= Texture2+Size; RC3=Sharpness+Kinetics1; RC4= Smoothness;
 RC5= Texture3; RC6= Kinetics2; RC7= Texture4; RC8=Location

TABLE 4. Contralateral group, PCA identified 6 components explaining 92% variance

	RC1	RC2	RC3	RC4	RC5	RC6
Eigenvalue	24.7	7.2	5.8	4.2	2.4	1.6
Variance%	28%	27%	15%	9%	9%	4%
Cumulative Variance%	28%	55%	70%	79%	88%	92%

PCA = Principal Component Analysis; RC = Rotated Component; LD = Largest Diameter
 RC1=Texture1+Volume; RC2= Texture2+LD; RC3=Smoothness; RC4= Kinetics+ Location;
 RC5= Texture3+ Sharpness; RC6= Texture4

Statistical analysis

Outliers in feature values were winsorized to the nearest whisker (28). Principal component analysis (PCA) with varimax-rotation was performed. Components describing at least 90% cumulative variance were analyzed (29). The PCA yielded a score per component per lesion. These scores were compared between the index and synchronous cancers using the Wilcoxon rank sum test. Analysis was conducted independently for ipsilateral and contralateral synchronous cancers. Since multiple tests were performed, the Benjamini-Hochberg method was used to control the false discovery rate (FDR) (30). FDR-adjusted-P values lower than 0.05 were considered to be significant. Association between the

differences in these PCA scores and NPI and PREDICT scores were analyzed using the Wilcoxon rank sum test. All statistical analysis was performed using R version 3.5.2.

Results

Patients and lesions

Among a total of 628 patients, 464/628 (73.9%) patients had ER+/HER2- index cancer, 34/464 (7.3%) of whom had 44 synchronous breast cancers in total; 83/628 (13.2%) patients had HER2+ index cancer, 4/83 (4.8%) of whom had 4 synchronous breast cancers in total; 81/628 (12.9%) patients had triple negative index cancer, 6/81 (7.4%) of whom had 8 synchronous breast cancers in total (**Figure 1**). Finally, 34 patients with 44 synchronous breast cancers in total were included. Among the 44 synchronous cancers, 34 were in the ipsilateral breast and 10 were in the contralateral breast (**Figure1**). The average age of patients at diagnosis was 54 years. The average diameter of the index cancers and the synchronous cancers was 21.7mm and 11.9 mm, respectively (**Table 2**).

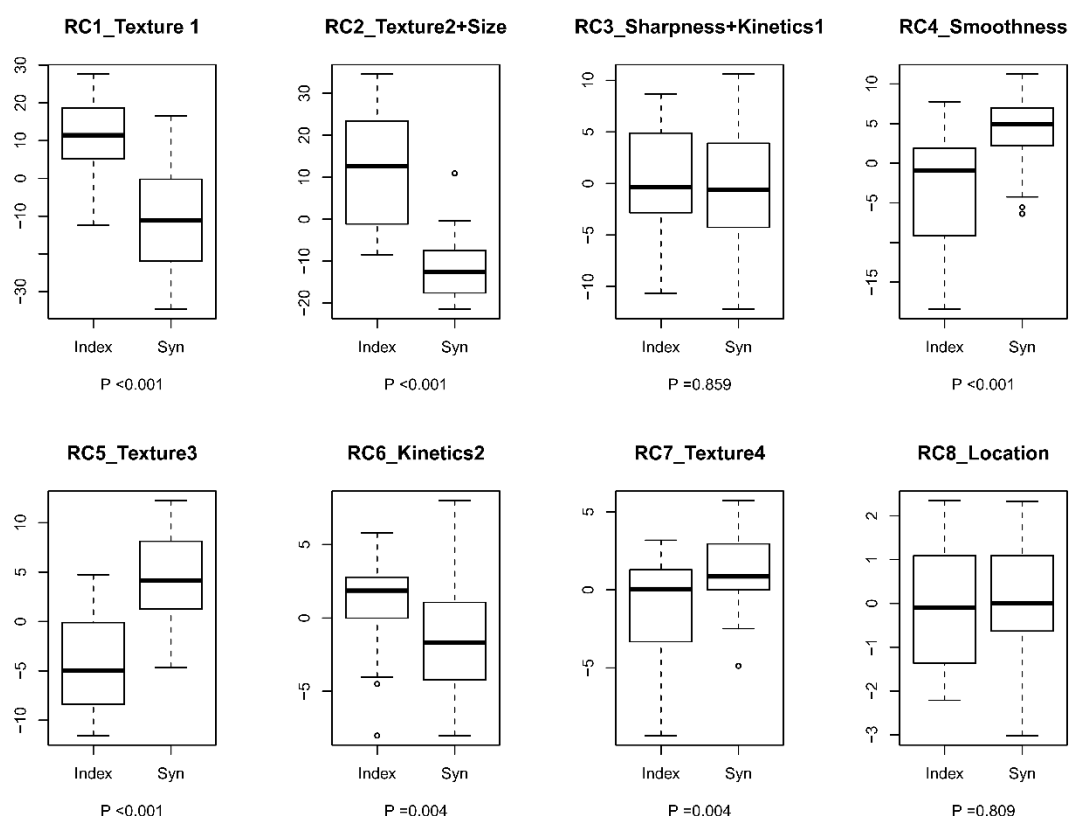


FIGURE 2. MRI phenotype of breast cancer (expressed in quantitative component score) (y-axis) for the index breast cancer and ipsilateral synchronous cancer (x-axis) (**RC**=Rotated Component, **Syn**=Synchronous breast cancer, **Index**=Index breast cancer).

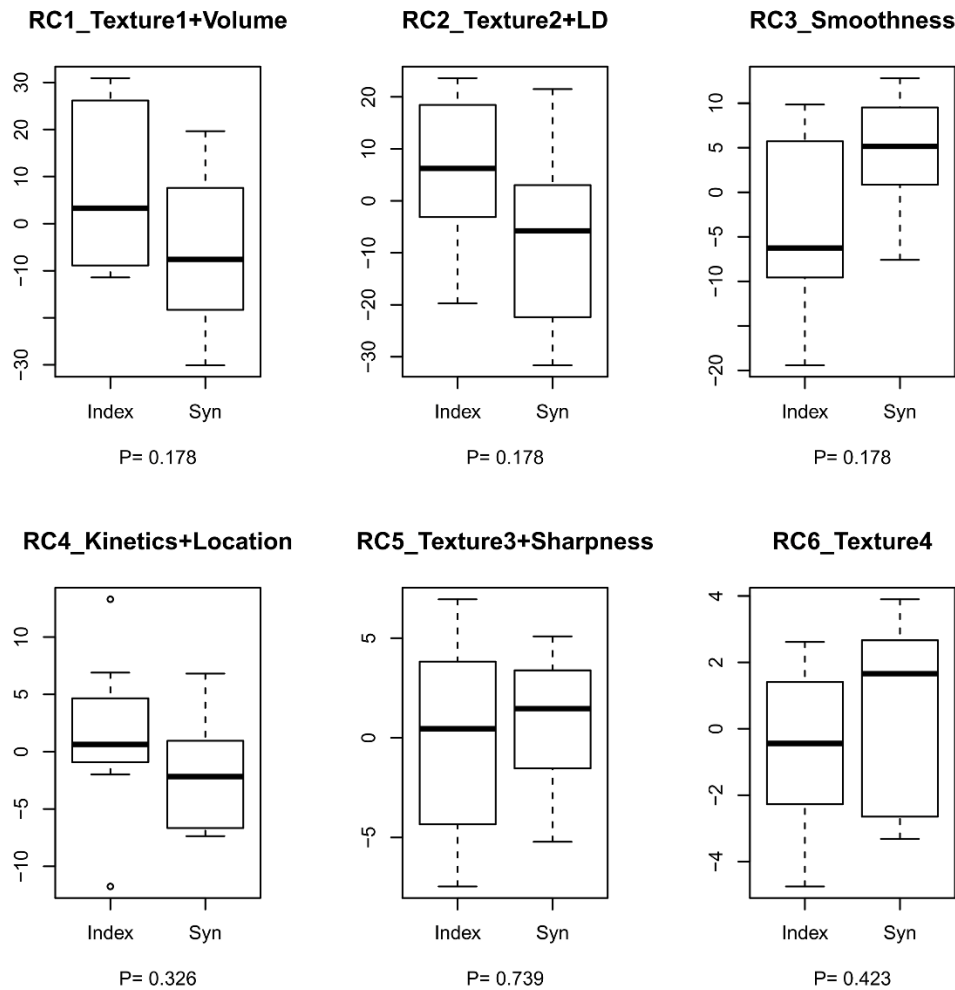


FIGURE 3. MRI phenotype of breast cancer (expressed in quantitative component score) (y-axis) for the index breast cancer and contralateral synchronous cancer (x-axis) (**RC** Rotated Component, **Syn** Synchronous breast cancer, **Index** Index breast cancer).

Imaging phenotype

Ipsilateral

For the patients with ipsilateral synchronous breast cancer, PCA identified eight components explaining 92% cumulative variance (**Table 3**). Components 1, 5 and 7 mainly represented texture, component 2 mainly represented texture and size of cancer, components 3 represented sharpness and uptake speed, component 4, 6, 8 represented smoothness, kinetics and relative distance of the cancer to the nipple, respectively.

Six out of eight components were significantly different between index and the synchronous cancers after FDR-adjustment. These components represented features describing texture (Component 1 (P<0.001), component 5 (P<0.001), and component 7 (P=0.004)), size (Component 2, P<0.001), smoothness

(Component 4, $P < 0.001$), and kinetics (Component 6, $P = 0.004$). Components 3 and 8 were not significantly different between index and the synchronous breast cancers ($P = 0.859$, $P = 0.809$) (**Figure 2**).

Contralateral

For the patients with contralateral synchronous cancer, PCA identified six components explaining 92% cumulative variance (**Table 4**). Components 1 represented texture and cancer volume, component 2 represented texture and largest diameter of cancer, component 3 represented smoothness, component 4 represented kinetics and relative distance to the nipple, component 5 represented sharpness and texture, component 6 represented texture. None of these six components were found to be significantly different between index and the synchronous breast cancer ($P = 0.178$, $P = 0.178$, $P = 0.178$, $P = 0.326$, $P = 0.739$, $P = 0.423$). (**Figure 3**).

TABLE 5. For patients with ipsilateral synchronous cancer, association between phenotype difference with prognosis.

	NPI			PREDICT		
	1 st Tertile	3 rd Tertile	P	1 st Tertile	3 rd Tertile	P
RC1	3.5(3.3,4.5)	4.5(3.3,4.7)	0.310	81% (67%, 83%)	66% (62%, 85%)	0.975
RC2	3.3(3.2,3.3)	4.5(3.5,4.6)	0.019	83% (73%, 84%)	66% (62%, 79%)	0.109
RC4	3.3(3.2,3.8)	4.5(4.1,4.6)	0.139	82% (71%, 84%)	63% (54%, 77%)	0.242
RC5	3.3(3.2,3.7)	3.5(3.3,4.5)	0.096	82% (71%, 84%)	80% (66%, 85%)	0.853
RC6	3.3(3.2,3.9)	4.5(3.4,4.7)	0.045	80% (69%, 85%)	73% (64%, 82%)	0.479
RC7	3.3(3.3,6)	4.5(3.5,4.6)	0.014	82% (71%, 90%)	66% (54%, 80%)	0.109

Numbers represent median (Q1, Q3) Q1, First quantile, Q3, Third quantile. 1stTertile and 3rdTertile means first and third tertile of phenotype difference. RC = Rotated Component; NPI = Nottingham Prognostic Index.

Phenotype difference and prognosis

For patients with ipsilateral synchronous cancer, the phenotype differences in the six components that were significantly different between index and the synchronous cancer were split in tertiles into small, medium and large differences. Compared with small phenotype difference, large phenotype difference in terms of lesion size and texture (component 2), kinetics (component 6) and texture (components 7) were associated with significantly higher NPI ($P = 0.019$, $P = 0.045$, $P = 0.014$ for component 2, 6, 7, respectively), while we did not find significantly different PREDICT score between small and large phenotype difference groups ($P = 0.109$, $P = 0.479$, $P = 0.109$ for component 2, 6, 7, respectively) (**Figure 4**, **Figure 5** and **Table 5**).

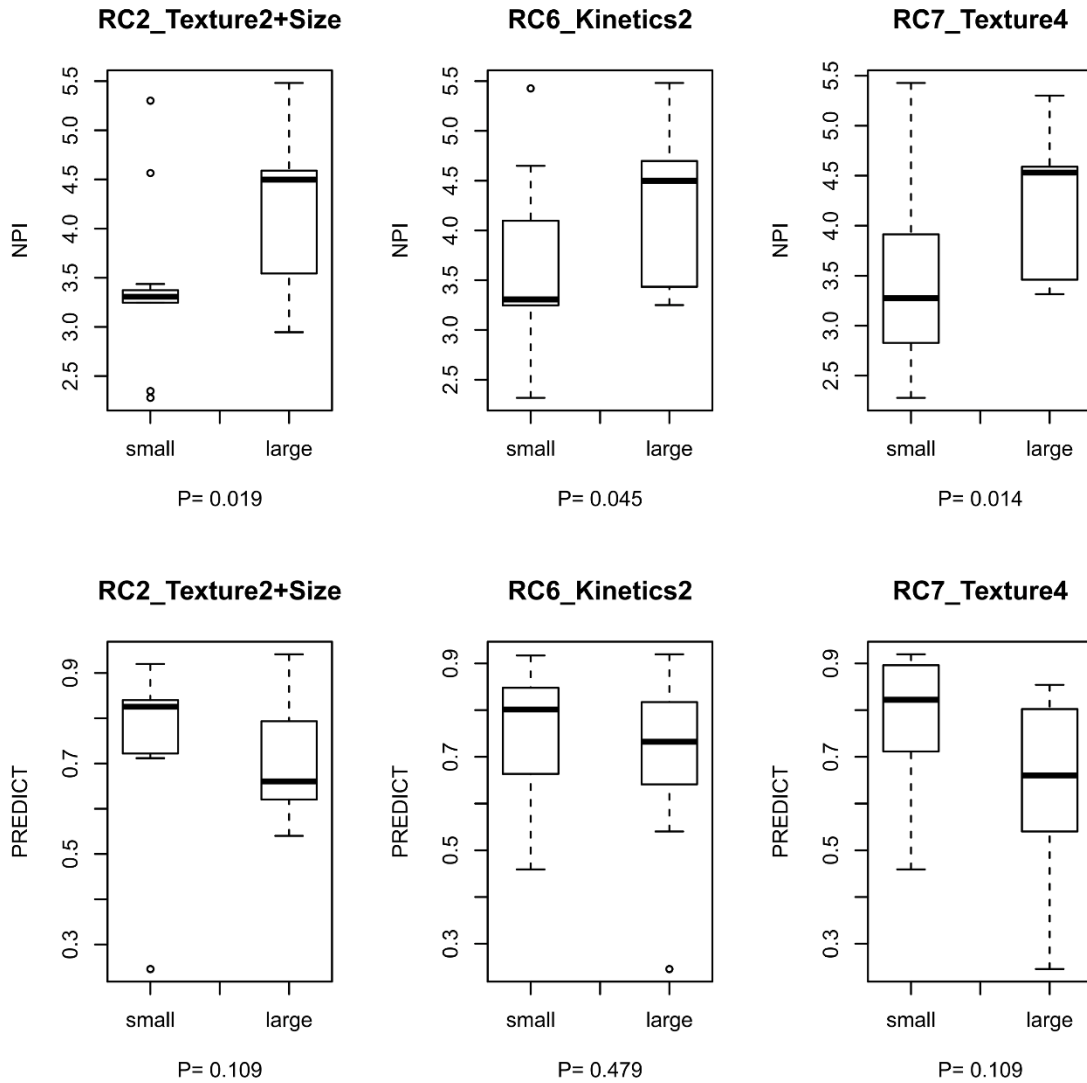


FIGURE 4. Association between Phenotype difference (Large and small group) with prognosis in terms of NPI and PREDICT score (NPI Nottingham Prognostic Index). P value indicates the significance of Wilcoxon rank sum test (RC Rotated Component, **small** 1st Tertile phenotype difference, **large** 3rd Tertile phenotype difference).

Discussion

In 34 patients with 44 synchronous breast cancers, we found that the imaging phenotype differed between index cancer and the corresponding synchronous cancers in the ipsilateral breast. Furthermore, patients with a large phenotype discrepancy between index and the ipsilateral synchronous cancer had relatively inferior prognosis in terms of NPI. In patients with contralateral synchronous breast cancer, no significant difference in imaging phenotype was observed.

The proportion of synchronous tumor foci detected on MRI varies considerably. In our study, we found 34/464 (7.3%) patients with synchronous breast cancer, which is

consistent with prior studies showing a frequency of 6% to 34% (31).

Our results indicated that the size of ER+/HER2- index breast cancers were larger than that of ipsilateral synchronous breast cancers. PCA identified eight components, component 2 was related to size and were indeed significantly different between index and synchronous cancer. In addition to size, texture, smoothness, and kinetics were also significantly different between index and ipsilateral synchronous cancer.

Synchronous breast cancer could result from intramammary spreading of index breast cancer with a similar phenotype. It could also develop independently, originating from separate progenitor cells and having a different phenotype (32). The discrepancy between index breast cancer and synchronous cancer observed on the ipsilateral side in our study is in agreement with the reported discrepancies in histological tumor grade (10), tumor type (33), and molecular phenotype (12).

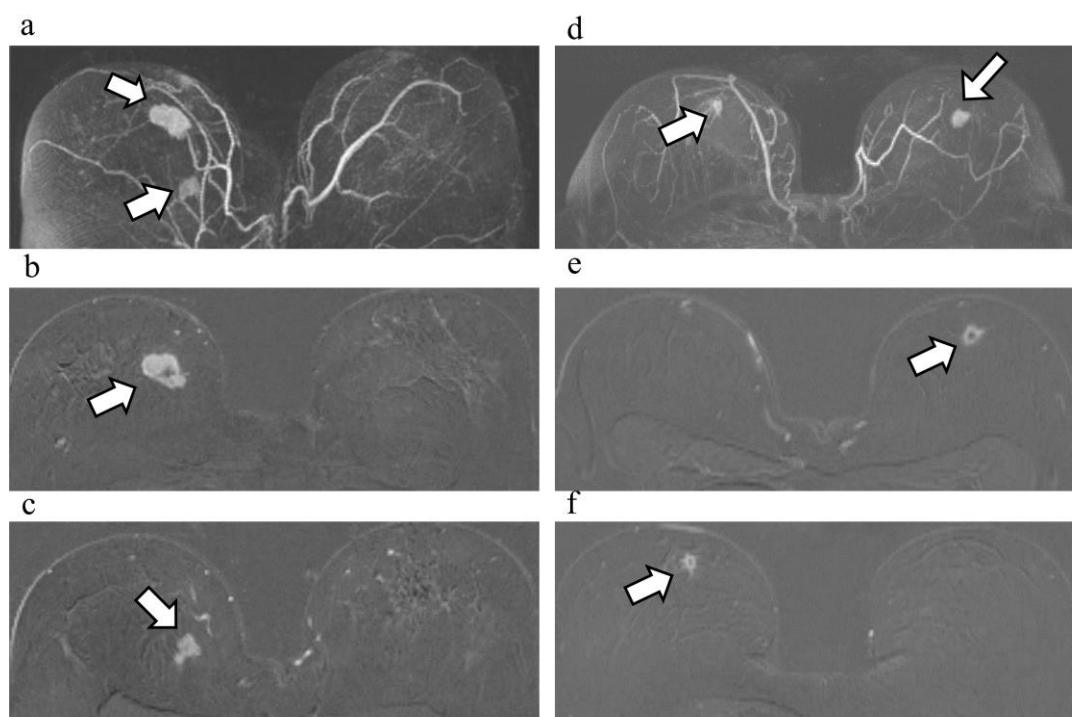


FIGURE 5. The top row shows the maximum intensity projection, the middle row shows the index breast cancer and the third row shows the synchronous cancer. **a-c** Index breast cancer and ipsilateral synchronous breast cancer in DCE-MRI. **Patient A**, 52 years old, largest diameter of index cancer and ipsilateral synchronous cancer were 28 mm and 20 mm respectively. Nottingham Prognostic Index is 4.56, and PREDICT score is 73.2% (10 year survival probability), which indicated that patientA has relatively bad prognosis. **d-f** Index breast cancer and contralateral synchronous breast cancer in DCE-MRI. **Patient B**, 61 years old, largest diameter of index cancer and contralateral synchronous cancer were 14 mm and 12 mm respectively. Nottingham Prognostic Index is 3.28, and PREDICT score is 87.6% (10 year survival probability), which indicated that patient B has relatively good prognosis.

In this study we used computer-extracted descriptions of the phenotype of the breast cancers. It has been increasingly accepted that quantitative features extracted from radiological images contain more detailed information than those perceived by radiologists in qualitative studies (34). These features represent phenotypes of the tissues that might reflect underlying information such as genetics. Since we have found significant phenotype differences between index and the synchronous breast cancers, the question arose whether such phenotype differences have the potential to serve as a non-invasive indicator of long-term prognosis before treatment, so as to provide insight into individualized treatment. For the six components that significantly differ between index and the synchronous breast cancers, the results indicated that larger differences of the phenotype in terms of size, kinetics and texture were indicative of worse prognosis in terms of NPI for ER+/HER2- breast cancer. Although the corresponding PREDICT score was not significantly different between small and large phenotype difference, to some extent, this could be attributed to the relatively small sample size.

We did not find a statistically significant discrepancy between index breast cancer and the synchronous cancer on the contralateral side. On one hand, this may be ascribed to the small sample size in our study. On the other hand, this result is in line with the literature, describing an agreement between index cancer and synchronous cancer on the contralateral side in terms of tumor associated antigens: bilateral breast cancers have been subjected to similar hormonal, environmental, and genetic influences during tumorigenesis (6,35). Therefore, it is reasonable that tumor phenotype in synchronous bilateral breast cancer may display similar biological characteristics. It should be noted, however, that these results are based on a limited amount of data.

The main limitation of this study is small sample size. Because of this, we only investigated patients with ER+/HER2- index breast cancer that form the majority (~70%) of the breast cancer population. For patients with HER2+ and triple negative index cancer, we lack statistical power. Nonetheless, it would be interesting to expand these analyses to HER2+ and triple negative breast cancer patients.

In conclusion, the MRI phenotype of ER+/HER2- breast cancer was significantly different from that of ipsilateral synchronous breast cancer, and a large phenotype difference was associated with relatively worse prognosis in terms of

NPI. Significant phenotype differences were not found for contralateral synchronous cancers.

References

1. Jain S, Rezo A, Shadbolt B, Dahlstrom JE. Synchronous multiple ipsilateral breast cancers: Implications for patient management. *Pathology*. 2009;41(1):57–67.
2. Newman LA, Sahin AA, Bondy ML, et al. A case-control study of unilateral and bilateral breast carcinoma patients. *Cancer*. 2001;91(10):1845–1853.
3. Pekar G, Gere M, Tarjan M, Hellberg D, Tot T. Molecular phenotype of the foci in multifocal invasive breast carcinomas: Intertumoral heterogeneity is related to shorter survival and may influence the choice of therapy. *Cancer*. 2014;120(1):26–34.
4. Tot T, Pekár G. Multifocality in “basal-like” breast carcinomas and its influence on lymph node status. *Annals of Surgical Oncology*. 2011;18(6):1671–1677.
5. Tot T, Gere M, Pekár G, et al. Breast cancer multifocality, disease extent, and survival. *Human Pathology*. Elsevier Inc.; 2011;42(11):1761–1769.
6. Kollias J, Pinder SE, Denley HE, et al. Phenotypic similarities in bilateral breast cancer. *Breast Cancer Research and Treatment*. 2004;85(3):255–261.
7. Srigley JR, Amin MB, Delahunt B, et al. Protocol for the examination of specimens from patients with invasive carcinoma of renal tubular origin. *Archives of Pathology and Laboratory Medicine*. 2010;134(4):1–32.
8. Garimella V, Long ED, O’Kane SL, Drew PJ, Cawkwell L. Oestrogen and progesterone receptor status of individual foci in multifocal invasive ductal breast cancer. *Acta Oncologica*. 2007;46(2):204–207.
9. Boros M, Marian C, Moldovan C, Stolnicu S. Morphological heterogeneity of the simultaneous ipsilateral invasive tumor foci in breast carcinoma: A retrospective study of 418 cases of carcinomas. *Pathology Research and Practice*. Elsevier GmbH.; 2012;208(10):604–609.
10. Buggi F, Folli S, Curcio A, et al. Multicentric/multifocal breast cancer with a single histotype: Is the biological characterization of all individual FOCI justified? *Annals of Oncology*. 2012;23(8):2042–2046.
11. Desmedt C, Fumagalli D, Pietri E, et al. Uncovering the genomic heterogeneity of multifocal breast cancer. *Journal of Pathology*. 2015;236(4):457–466.
12. Navale P, Bleiweiss IJ, Jaffer S, Nayak A. Evaluation of biomarkers in multiple ipsilateral synchronous invasive breast carcinomas. *Archives of Pathology and Laboratory Medicine*. 2019;143(2):190–196.
13. Turnbull L, Brown S, Harvey I, et al. Comparative effectiveness of MRI in breast cancer (COMICE) trial: a randomised controlled trial. *The Lancet*. Elsevier Ltd; 2010;375(9714):563–571.
14. Bedrosian I, Mick R, Orel SG, et al. Changes in the surgical management of patients with breast carcinoma based on preoperative magnetic resonance imaging. *Cancer*. 2003;98(3):468–473.
15. Li H, Zhu Y, Burnside ES, et al. MR imaging radiomics signatures for Predicting

the risk of Breast cancer recurrence as given by research Versions of MammaPrint, Oncotype DX, and PaM50 gene assays. *Radiology*. 2016;281(2):382–391.

16. Tobias H, Merkle EM, Reiner CS, et al. Reproducibility of dynamic contrast-enhanced MR imaging: Part II. comparison of intra- and interobserver variability with manual region of interest placement versus semiautomatic lesion segmentation and histogram analysis. *Radiology*. 2013;266(3):812–821.

17. Grimm LJ. Breast MRI radiogenomics: Current status and research implications. *Journal of Magnetic Resonance Imaging*. 2016;43(6):1269–1278.

18. Mazurowski MA. Radiogenomics: What It Is and Why It Is Important. *Journal of the American College of Radiology*. American College of Radiology; 2015;12(8):862–866.

19. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer a study of 1409 cases of which 359 have been followed for 15 years. *British Journal of Cancer*. 1957;11(3):359–377.

20. Dmitriev ID, Loo CE, Vogel W V, Pengel KE, A Gilhuijs KG. Fully automated deformable registration of breast DCE-MRI and PET/CT. *Phys Med Biol*. 2013;58(4):1221–1233.

21. 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(1):42–49.

22. Gilhuijs KGA, Giger ML, Bick U. Computerized analysis of breast lesions in three dimensions using dynamic magnetic-resonance imaging. *Medical Physics*. 1998;25(9):1647–1654.

23. Gilhuijs KGA, Deurloo EE, Muller SH, Peterse JL, Schultze Kool LJ. Breast MR imaging in women at increased lifetime risk of breast cancer: clinical system for computerized assessment of breast lesions initial results. *Radiology*. 2002;225(3):907–916.

24. Manjunath BS, Ma W-Y. Texture Features for Image Retrieval. *Image Databases*. 2003. bl 313–344.

25. Coelho LP. Mahotas : Open source software for scriptable computer vision. *Journal of Open Research Software*. 2013;1–13.

26. Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham prognostic index in primary breast cancer. *Breast Cancer Research and Treatment*. 1992;22(3):207–219.

27. Candido dos Reis FJ, Wishart GC, Dicks EM, et al. An updated PREDICT breast cancer prognostication and treatment benefit prediction model with independent validation. *Breast Cancer Research*. *Breast Cancer Research*; 2017;19(1):1–13.

28. Hellerstein JM. Quantitative Data Cleaning for Large Databases. United Nations Economic Commission for Europe. 2008;42.

29. Rose W. Principal component analysis: Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2010;2(4):433–459.

30. Chatelain F. A tutorial on multiple testing: False discovery control. *EAS Publications Series*. 2016;78–79:163–178.

31. Houssami N, Ciatto S, Macaskill P, et al. Accuracy and surgical impact of magnetic resonance imaging in breast cancer staging: Systematic review and meta-analysis in

detection of multifocal and multicentric cancer. *Journal of Clinical Oncology*. 2008;26(19):3248–3258.

32. Middleton LP, Vlastos G, Mirza NQ, Evasingleary S, Sahin AA. Multicentric mammary carcinoma: Evidence of monoclonal proliferation. *Cancer*. 2002;94(7):1910–1916.

33. Choi Y, Kim EJ, Seol H, et al. The hormone receptor, human epidermal growth factor receptor 2, and molecular subtype status of individual tumor foci in multifocal/multicentric invasive ductal carcinoma of breast. *Human Pathology*. Elsevier Inc.; 2012;43(1):48–55.

34. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. *Radiology*. 2016;278(2):563–577.

35. Lundy J, Mishriki Y, Varma AO, Kufe D, Schuss A. Tumor-associated antigens in bilateral breast cancer. *Journal of Surgical Oncology*. 1987;35(1):24–29

Chapter 6

Summary and Discussion

English Summary

We aimed to investigate whether the imaging phenotype of healthy breast tissue and breast lesions on dynamic contrast-enhanced MRI (DCE-MRI) have potential to support the decision making in screening, diagnostic workflow and treatment.

Traditional biomarkers, such as histological tumor grade, tumor type, and molecular subtype (receptor status) are dependent on the invasive sampling of the areas suspicious of breast cancer. In contrast, MRI not only investigates the whole tumor in vivo, therefore reducing the potential risk of sampling errors, but also evaluates the surrounding tissues (e.g., background parenchymal enhancement [BPE]). MRI has therefore been routinely used in many state-of-the-art breast imaging units.

In the past decade, there have been tremendous advances in medical imaging and revolutionary development in artificial intelligence, which have contributed to the progress in clinical decision-making. In the current thesis, we explored artificial intelligence applications that can be seamlessly integrated into the current clinical workflow, and explored the potential of imaging biomarkers derived from dynamic contrast-enhanced DCE-MRI to support the decision-making in screening, diagnostic workflow and treatment.

In **Chapter 2**, we automated BPE rating from DCE-MRI using machine learning without compromising the underlying association between manual rating of BPE and breast cancer occurrence. A combination of Random Forests, Naïve Bayes, and KNN classifiers with majority voting was used to predict BPE category with a cross-validated prediction accuracy of 0.76. The hazard ratio of BPE for breast cancer occurrence was comparable between automated rating and manual rating (Hazard Ratio =2.12 versus Hazard Ratio =1.97, P=0.65) for categories “mild, moderate, marked” relative to category “minimal”.

In **Chapter 3**, we automatically identified quantitative properties of the breast parenchyma on baseline DCE MRI scans and assessed their association with breast cancer occurrence. After adjustment for age, body mass index, and BPE, multivariable analysis showed that breast cancer occurrence was greater in women with higher volumes of enhancing parenchyma compared to that in women with low volumes of enhancing parenchyma (Hazard Ratio=2.09, P =0.005).

In **Chapter 4**, we constructed a multiparametric computer-aided diagnosis (CAD) model based on ridge regression to preoperatively identify malignant additional breast lesions with high specificity. The CAD model showed potential to identify malignant disease extension with near-perfect specificity in approximately half the population of preoperative patients originally indicated for a breast biopsy. If implemented in clinical practice, these women could be triaged directly to surgery while patients in the other half would still need to proceed to MRI-guided biopsy to confirm absence of malignant disease. These findings demonstrate the potential to triage women to MRI-guided breast biopsy.

In **Chapter 5**, we investigated the imaging phenotype of the index cancer and its synchronous cancers derived from DCE-MRI, and found that the MRI phenotype of ER+/HER2– breast cancer was different from that of ipsilateral synchronous cancer and a large phenotype difference was associated with worse prognosis. Such phenotype differences may have the potential to serve as a noninvasive indicator of long-term prognosis before treatment, so as to provide insight into individualized treatment.

To conclude, MRI has been an indispensable imaging modality and the indication for breast MRI has been expanding. Appropriate use of breast MRI may facilitate the personalization of breast disease management. In the current thesis, we constructed practical machine learning models which may play a potential role in future clinical decision-making in order to save resources and to reduce the number of invasive interventions. However, these models need to be further validated prospectively on independent data from other hospitals.

Nederlandse samenvatting (Dutch Summary)

We onderzochten of het beeldfenotype van gezond borstweefsel en borstlaesies op dynamische MRI met contrast (DCE-MRI) potentie heeft om de besluitvorming over screening, diagnostische workflow en behandeling te verbeteren.

Traditionele biomarkers, zoals histologische tumorgraad, tumortype en moleculair subtype (receptorstatus) zijn afhankelijk van de invasieve bemonstering van de gebieden die verdacht zijn voor borstkanker. MRI daarentegen onderzoekt niet alleen de gehele tumor in vivo, waardoor het potentiële risico van steekproeffouten wordt verkleind, maar evalueert ook de omliggende weefsels (bijv. achtergrond parenchymaankleuring [BPE]). MRI wordt daarom routinematig gebruikt in veel ziekenhuizen.

In het afgelopen decennium is er enorme vooruitgang geboekt op het gebied van medische beeldvorming en zijn er revolutionaire ontwikkelingen geweest op het gebied van kunstmatige intelligentie, wat heeft bijgedragen aan de vooruitgang in klinische besluitvorming. In dit proefschrift streven we naar een effectieve integratie van kunstmatige intelligentie in de klinische praktijk en onderzoeken we het potentieel van biomarkers voor beeldvorming afgeleid van DCE-MRI om de besluitvorming op het gebied van screening, diagnostische workflow en behandeling te ondersteunen.

In **Hoofdstuk 2** hebben we de BPE-classificatie van DCE-MRI geautomatiseerd met behulp van machine learning zonder afbreuk te doen aan de onderliggende associatie tussen handmatige classificatie van BPE en het optreden van borstkanker. Een combinatie van Random Forests, Naïve Bayes en KNN-classifiers met consensusstemming werd gebruikt om de BPE-categorie te voorspellen met een gecrossvalideerde voorspellingsnauwkeurigheid van 0,76. De hazard ratio van BPE voor het optreden van borstkanker was vergelijkbaar tussen geautomatiseerde classificatie en handmatige classificatie (Hazard Ratio =2.12 versus Hazard Ratio =1.97, P=0.65) voor categorieën "mild, matig, gemarkeerd" ten opzichte van categorie "minimaal".

In **Hoofdstuk 3** identificeerden we automatisch kwantitatieve eigenschappen van het borstparenchym op baseline DCE MRI-scans en beoordeelden we hun

verband met het optreden van borstkanker. Na correctie voor leeftijd, body mass index en BPE toonde multivariabele analyse aan dat borstkanker vaker voorkwam bij vrouwen met grotere volumes aankleurend parenchym in vergelijking met vrouwen met kleine volumes aankleurend parenchym (Hazard Ratio=2,09, P =0.005).

In **Hoofdstuk 4** hebben we een multiparametrisch computerondersteund diagnosemodel (CAD-model) geconstrueerd op basis van ridgeregressie om preoperatief kwaadaardige additionele borstlaesies te identificeren met een hoge specificiteit. Het CAD-model toonde potentieel om kwaadaardige uitbreiding van de ziekte te identificeren met een bijna perfecte specificiteit in ongeveer de helft van de populatie van preoperatieve patiënten die oorspronkelijk waren geïndiceerd voor een borstbiopsie. Deze patiënten zouden in de toekomst na validatie rechtstreeks naar chirurgie kunnen worden verwezen terwijl patiënten in de andere helft van de populatie nog steeds zouden worden doorverwezen naar MRI-geleide biopsie om de afwezigheid van kwaadaardige ziekte te bevestigen. Deze bevindingen tonen het potentieel aan om patiënten te triageren naar MRI-geleide borstbiopsie door middel van AI.

In **Hoofdstuk 5** onderzochten we het beeldfenotype van de indexkanker en van de synchrone kankers afgeleid uit DCE-MRI, en vonden we dat het MRI-fenotype van ER+/HER2- borstkanker verschilde van dat van ipsilaterale synchrone kanker en dat een groot verschil in fenotype geassocieerd was met een slechtere prognose. dergelijke fenotypeverschillen kunnen mogelijk dienen als een niet-invasieve indicator van de lange-termijn prognose vóór behandeling, om zo keuzes te ondersteunen voor geïndividualiseerde behandeling.

Concluderend kan worden gesteld dat MRI een onmisbare beeldvormingsmodaliteit is en dat de indicatie voor borst-MRI steeds verder is uitgebreid. Adequaat gebruik van MRI voor de borst kan de personalisatie van borstziekten vergemakkelijken. In dit proefschrift hebben we enkele praktische modellen geconstrueerd gebaseerd op machine-learning die kunnen bijdragen aan klinische besluitvorming om zorgcapaciteit effectiever te benutten en het aantal invasieve interventies te reduceren. Deze modellen moeten echter prospectief verder extern gevalideerd worden in andere ziekenhuizen.

General discussion

The aim of this thesis was to use machine learning to support triaging decisions in screening, diagnostic workflow and treatment by extracting imaging biomarkers from healthy breast tissue and breast lesions on dynamic contrast-enhanced MRI (DCE-MRI).

Biomarkers from enhancement of healthy breast parenchyma

The enhancement of normal breast tissue on breast MRI is called background parenchymal enhancement (BPE). BPE varies between individuals and changes with time even in the same individual as BPE is sensitive to endogenous hormonal changes as well as exogenous factors (1–3). In clinical practice, the radiologists qualitatively classify BPE as minimal, mild, moderate, or marked, according to the Breast Imaging Reporting and Data System (BI-RADS) (4), which is subject to inter- and intrareader variation, and is not efficient in terms of improving the working efficiency of the radiologists. These limitations have highlighted the importance of objective evaluation and interpretation of BPE. Hence, in **Chapter 2**, we automated the rating of BPE in women with extremely dense breasts implementing objective evaluation of BPE. Some other studies have used deep learning to automate prediction of BPE category from DCE-MRI (5–7), but their research populations and methodology are quite different from ours. Hence, it is difficult to assess the potential of these findings in a risk stratification tool for this specific screening population. As far as we know, our study is the first one focusing only on an unselected series of asymptomatic women with extremely dense breasts, and to predict BPE category with sufficient accuracy to not compromise the underlying association between BPE and breast cancer occurrence. Therefore, our model has the potential to autopopulate the BPE categories in breast MRI reports, which can greatly alleviate radiologists' workload and improve the interpretation accuracy by avoiding inter and intrareader variability.

From biological perspective, BPE may be associated with vascular permeability regulated by endogenous hormones (8–11) and may represent tissue at risk for neoplasia (12–15). Several studies have reported associations between BPE and breast cancer occurrence with conflicting conclusions (12,16–19). A systematic review and meta-analysis of 18 studies conducted by Thompson et al. indicated that a higher level of BPE measured at breast MRI is associated with the presence of breast cancer in women with high risk (20). A population-wide cohort of women undergoing breast MRI were collected from 46 radiology facilities in the

United States by Arasu et al. (21). They found that BPE is associated with future invasive breast cancer independent of breast density, and suggested that BPE should be incorporated into risk prediction models for women receiving breast MRI. Then in **Chapter 3**, we further identified quantitative properties of BPE on baseline DCE-MRI, assessed their association with breast cancer occurrence, and found that a high volume of enhancing parenchyma on baseline DCE-MRI was associated with increased occurrence of breast cancer as compared with a low volume of enhancing parenchyma, which could help more easily interpret the quantitative properties of BPE. The results of chapter 2 and 3 in this thesis have demonstrated the feasibility of objective evaluation and interpretation of BPE.

Biomarkers from enhancement of breast lesion

Breast cancer is a heterogenous disease, which can present inter-tumor and intra-tumor heterogeneity (22). Intra-tumor heterogeneity can further manifest spatially and temporally (23,24). Typically, key clinical decisions are made based on established histopathological markers obtained from biopsy, e.g., tumor grade, nodal status, expression of estrogen receptor (ER), progesterone receptor (PR), and HER2 status, etc., but the limited tissue samples may undersample the breast lesions (spatial heterogeneity), and the breast lesion itself may progress over time and be affected by systematic treatment (temporal heterogeneity), so there is an urgent need to explore useful markers of intra-tumor heterogeneity that can be used to augment established markers for personalized management of breast disease.

Evolution of medical imaging and artificial intelligence (AI) has provided a noninvasive method for whole-tumor assessment by quantifying the appearance (phenotype) of breast lesions on DCE-MRI (25,26). Identifying the imaging phenotype of breast cancer on DCE-MRI and investigating its association with patient outcome can provide added value for established diagnostic and prognostic biomarkers.

Triaging for diagnosis of breast cancer

Incidental MR-detected breast lesions (i.e., additional lesions to the index cancer) pose challenges in preoperative patients with early breast cancer. We constructed a multiparametric computer-aided diagnosis (CAD) model based on ridge regression to identify malignant additional breast lesions preoperatively with high specificity in **Chapter 4**. Features with high absolute coefficients included

those from the DCE series (Sharpness, Circularity, Signal enhancing ratio), fast-DCE (Time to enhancement, Final slope), DWI (25th percentile ADC, Maximum ADC), and clinical features (age at diagnosis, BIRADS-score). Conversely, T₂ features yielded small coefficient values, which was consistent with previously published results (27). The CAD model shows potential to spare patients an MRI-guided biopsy in approximately half the number of cases. In the other cases, patients would be referred to MRI-guided biopsy to confirm absence of malignant disease. These findings demonstrate the potential to triage MRI-guided breast biopsy. To the best of our knowledge, no CAD models have been reported to attempt triaging patients in a preoperative setting for MRI-guided biopsy.

Triaging for prognosis of breast cancer

For patients with synchronous breast cancer, we found significant phenotype differences between index breast cancer and the ipsilateral synchronous counterpart in **Chapter 5**. The potential reason is that synchronous breast cancers in the same breast may originate from separate progenitor cells and thus have different phenotypes (28). The discrepancy observed in our study is in agreement with the reported discrepancies in histological tumor grade (29), tumor type (30), and molecular phenotype (31). The question arose whether such phenotype differences have the potential to serve as a noninvasive biomarker of long-term prognosis prior to treatment, so as to provide insight into individualized treatment. We further found that larger differences of the phenotype in terms of size, kinetics, and texture were indicative of worse prognosis in terms of the Nottingham Prognostic Index (NPI). However, that these results are based on a limited amount of data.

Limitations and future prospectives

Among the modalities for evaluating breast cancer, DCE-MRI achieves the highest sensitivity, but its specificity is not superior. Standard MRI protocols mainly provide morphological and part of functional information of breast tissue. In contrast, multiparametric MRI quantifies and visualizes the functional properties at cellular and molecular levels and further clarifies the progression of breast cancer, providing information which cannot be obtained from DCE-MRI alone, and therefore improving the diagnostic accuracy (32,33). In addition, utilization of ultrahigh field strength (7.0 T) provides an increased signal-to-noise-ratio, thus allows for a higher spatial and temporal resolution, and has

showed improvement in the sensitivity and specificity to 100% and 90% respectively (33).

Although the specificity can be improved through adoption of multiparametric MRI and higher field strength, MRI has been only indicated as a supplemental screening modality to mammography for women at high risk (greater than 20% lifetime risk) due to its disadvantages of being more time-consuming and higher cost. Advances in MRI acquisition technique has made it possible to shorten the breast MRI protocols, decrease the image acquisition time and shorten image interpretation time (34) (Figure 1), resulting in application of abbreviated and ultrafast MRI protocols, which bring lower cost and higher efficiency, and make MRI more available and accessible to women with intermediate risk of breast cancer, including women with dense breasts.

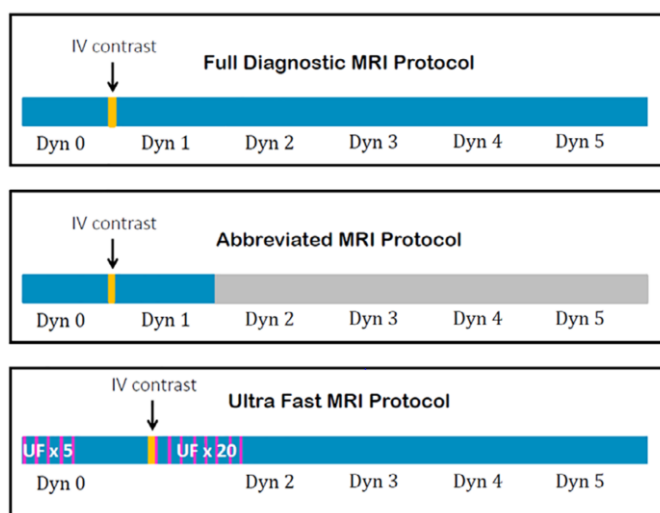


FIGURE 1: Diagram describing three MRI protocols including a full DCE-MRI, an abbreviated MRI, and an ultrafast MRI protocol. The full DCE-MRI is the default protocol with one precontrast and multiple postcontrast sequences. The abbreviated MRI protocol includes one precontrast and one postcontrast sequence. The ultrafast MRI protocol includes several high temporal resolution sequences during the pre- and postcontrast phase. (Source: DOI: 10.1002/jmri.26878)

Manual assessment of breast lesions on MRI by radiologists is limited by many factors, such as clinical experience, image quality, large amounts of image data and the consequent fatigue, complexity of the disease, etc. Therefore, several AI models have been constructed to assist radiologists in interpretation of breast MRI. AI-aided systems not only can be implemented for MRI, but also for mammography. The first commercial CAD system in mammograph was FDA approved in 1998 with the aim to provide the radiologists with a second opinion to enhance radiologists' diagnostic accuracy. In addition, mammography is the only modality with a proven mortality benefit, leading to the adoption of mammography-based worldwide screening programs. The biggest disadvantage of mammography is the occurrence of false positives due to the overlap of normal fibroglandular breast tissue. Contrast enhanced spectral mammography

(CESM) or contrast enhanced dual-energy mammography (CEDM) is a new technology, in which the iodine-based contrast agent is intravenously injected two minutes before the acquisition of the mammographic images, resulting in the acquisition of a low-energy and high-energy image. Previous studies have proved that CESM shows a better diagnosis of suspected lesions compared with conventional mammography, ultrasound, and breast MRI. However, the detection of hypervascular breast lesions using CESM is poorer than breast MRI (35).

Machine learning is a subfield of AI, it typically uses human-engineered features as the basis of learning. Deep learning is a class of machine learning methods that learns features directly from data and characterized by the use of neural networks with many layers (36). The development in AI has contributed tremendous performance augmentation in medical imaging analysis. AI can be used to help solve various tasks such as cancer detection, risk assessment, diagnostic workflow and triage, response to neoadjuvant chemotherapy, etc. However, there are many obstacles to overcome before the AI models can be smoothly implemented to clinical scenarios. For example, generalizability of the AI algorithms in breast imaging is a big issue to be confronted with, which requires careful assessment of the application scenario and keeping a close eye on the performance. Another concern is the understanding of radiologists and the trust from patients. The radiologists need to fully understand the underlying mechanisms of the AI and know how to integrate it into clinical scenario but not totally rely on the predictions alone, and the patients need to trust and support the applications. In high-risk decision-makings, there has been a call for better understanding of AI, such approaches are called explainable artificial intelligence (XAI) (37). Researchers and physicians have been increasingly using XAI to explain the results of the algorithms and the AI models. Large amounts of work remain to be done to explain to the patients on how AI works and make sure that they are aware of the pros and con, so that the clinical options can be communicated effectively and efficiently.

References

1. Taron J, Fleischer S, Preibsch H, Nikolaou K, Gruber I, Bahrs S. Background parenchymal enhancement in pregnancy-associated breast cancer: a hindrance to diagnosis? *European Radiology*. *European Radiology*; 2019;29(3):1187–1193.
2. Brooks JD, Sung JS, Pike MC, et al. MRI background parenchymal enhancement, breast density and serum hormones in postmenopausal women. *International Journal of Cancer*. 2018;143(4):823–830.
3. Su MYL. Effects of tamoxifen and aromatase inhibitors on breast tissue enhancement in dynamic contrast-enhanced breast MR imaging: A longitudinal intraindividual cohort study. *Breast Diseases*. 2014;25(4):317–319.

4. Morris EA. ACR Bi-Rads® Atlas — Breast MRI. American College of Radiology. 2013;
5. Nam Y, Park GE, Kang J, Kim SH. Fully Automatic Assessment of Background Parenchymal Enhancement on Breast MRI Using Machine-Learning Models. *Journal of Magnetic Resonance Imaging*. 2021;53:818–826.
6. Borkowski K, Rossi C, Ciritsis A, Marcon M. Fully automatic classification of breast MRI background parenchymal enhancement using a transfer learning approach. *Medicine*. 2020;99(29):1–7.
7. Eskreis-Winkler S, Sutton EJ, D’Alessio D, et al. Breast MRI Background Parenchymal Enhancement Categorization Using Deep Learning: Outperforming the Radiologist. *Journal of Magnetic Resonance Imaging*. 2022;56:1068–1076.
8. Giess CS, Yeh ED, Raza S, Birdwell RL. Background parenchymal enhancement at breast MR imaging: Normal patterns, diagnostic challenges, and potential for false-positive and false-negative interpretation. *Radiographics*. 2014;34(1):234–247.
9. Tice JA, Miglioretti DL, Li CS, Vachon CM, Gard CC, Kerlikowske K. Breast density and benign breast disease: Risk assessment to identify women at high risk of breast cancer. *Journal of Clinical Oncology*. 2015;33(28):3137–3143.
10. Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: Development and validation of a new predictive model. *Annals of Internal Medicine*. 2008;148(5):337–347.
11. Kuhl C. The current status of breast MR imaging. Part I. Choice of technique, image interpretation, diagnostic accuracy, and transfer to clinical practice. *Radiology*. 2007;244(2):356–378.
12. Dontchos BN, Rahbar H, Partridge SC, et al. Are Qualitative Assessments of Background Parenchymal Enhancement, Amount of Fibroglandular Tissue on MR Images, and Mammographic Density Associated with Breast Cancer Risk? *Radiology*. 2015;276(2):371–380.
13. van der Velden BHM, Dmitriev I, Loo CE, Pijnappel RM, Gilhuijs KGA. Association between Parenchymal Enhancement of the Contralateral Breast in Dynamic Contrast-enhanced MR Imaging and Outcome of Patients with Unilateral Invasive Breast Cancer. *Radiology*. 2015;276(3):675–685.
14. Kumar AS, Chen DF, Au A, et al. Biologic significance of false-positive magnetic resonance imaging enhancement in the setting of ductal carcinoma in situ. *American Journal of Surgery*. 2006;192(4):520–524.
15. King V, Brooks JD, Bernstein JL, Reiner AS, Pike MC, Morris EA. Background Parenchymal Enhancement at Breast MR Imaging and Breast Cancer Risk. *Radiology*. 2011;260(1):50–60.

16. Telegrafo M, Rella L, Stabile Ianora AA, Angelelli G, Moschetta M. Breast MRI background parenchymal enhancement (BPE) correlates with the risk of breast cancer. *Magnetic Resonance Imaging*. Elsevier Inc.; 2016;34(2):173–176.
17. Melsaether A, Pujara AC, Elias K, et al. Background parenchymal enhancement over exam time in patients with and without breast cancer. *Journal of Magnetic Resonance Imaging*. 2017;45(1):74–83.
18. DeLeo MJ, Domchek SM, Kontos D, Conant E, Chen J, Weinstein S. Breast MRI fibroglandular volume and parenchymal enhancement in BRCA1 and BRCA2 mutation carriers before and immediately after risk-reducing salpingo-oophorectomy. *American Journal of Roentgenology*. 2015;204(3):669–673.
19. Watt GP, Sung J, Morris EA, et al. Association of breast cancer with MRI background parenchymal enhancement: the IMAGINE case-control study. *Breast Cancer Research*. *Breast Cancer Research*; 2020;22(1):1–12.
20. Thompson CM, Mallawaarachchi I, Dwivedi DK, et al. The association of background parenchymal enhancement at breast mri with breast cancer: A systematic review and meta-analysis. *Radiology*. 2019;292(3):552–561.
21. Arasu VA, Miglioretti DL, Sprague BL, et al. Population-based assessment of the association between magnetic resonance imaging background parenchymal enhancement and future primary breast cancer risk. *Journal of Clinical Oncology*. 2019;37(12):954–963.
22. Chitalia RD, Rowland J, McDonald ES, et al. Imaging Phenotypes of Breast Cancer Heterogeneity in Preoperative Breast Dynamic Contrast Enhanced Magnetic Resonance Imaging (DCE-MRI) Scans Predict 10-Year Recurrence. *Clinical Cancer Research*. 2019;862–870.
23. McGranahan N, Swanton C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell*. Elsevier Inc.; 2015;27(1):15–26.
24. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. *Nature Reviews Clinical Oncology*. 2018;15(2):81–94.
25. Parekh VS, Jacobs MA. Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI. *npj Breast Cancer*. Springer US; 2017;3(1):1–8.
26. Li H, Zhu Y, Burnside ES, et al. MR imaging radiomics signatures for Predicting the risk of Breast cancer recurrence as given by research Versions of MammaPrint, Oncotype DX, and PaM50 gene assays. *Radiology*. 2016;281(2):382–391.
27. 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. *Investigative radiology*. 2020;55(7):438–444.

28. Middleton LP, Vlastos G, Mirza NQ, Evasingletery S, Sahin AA. Multicentric mammary carcinoma: Evidence of monoclonal proliferation. *Cancer*. 2002;94(7):1910–1916.
29. Buggi F, Folli S, Curcio A, et al. Multicentric/multifocal breast cancer with a single histotype: Is the biological characterization of all individual FOCI justified? *Annals of Oncology*. 2012;23(8):2042–2046.
30. Choi Y, Kim EJ, Seol H, et al. The hormone receptor, human epidermal growth factor receptor 2, and molecular subtype status of individual tumor foci in multifocal/multicentric invasive ductal carcinoma of breast. *Human Pathology*. Elsevier Inc.; 2012;43(1):48–55.
31. Navale P, Bleiweiss IJ, Jaffer S, Nayak A. Evaluation of biomarkers in multiple ipsilateral synchronous invasive breast carcinomas. *Archives of Pathology and Laboratory Medicine*. 2019;143(2):190–196.
32. Pinker K, Bickel H, Helbich TH, et al. Combined contrast-enhanced magnetic resonance and diffusion-weighted imaging reading adapted to the “breast Imaging Reporting and Data System” for multiparametric 3-T imaging of breast lesions. *European Radiology*. 2013;23(7):1791–1802.
33. Pinker K, Bogner W, Baltzer P, et al. Clinical application of bilateral high temporal and spatial resolution dynamic contrast-enhanced magnetic resonance imaging of the breast at 7 T. *European Radiology*. 2014;24(4):913–920.
34. Sheth D, Giger ML. Artificial intelligence in the interpretation of breast cancer on MRI. *Journal of Magnetic Resonance Imaging*. 2019;1–15.
35. Blum KS, Antoch G, Mohrmann S, Obenauer S. Use of low-energy contrast-enhanced spectral mammography (CESM) as diagnostic mammography-proof of concept. *Radiography*. Elsevier Ltd; 2015;21(4):352–358.
36. Chartrand G, Cheng PM, Vorontsov E, et al. Deep learning: A primer for radiologists. *Radiographics*. 2017;37(7):2113–2131.
37. van der Velden BHM, Kuijf HJ, Gilhuijs KGA, Viergever MA. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical Image Analysis*. Elsevier B.V.; 2022;79:102470.

List of publications

Wang H, Van Der Velden BHM, Chan HSM, Loo CE, Viergever MA, Gilhuijs KGA. Synchronous Breast Cancer: Phenotypic Similarities on MRI. *Journal of Magnetic Resonance Imaging.* 51(6):1858–1867(2020)

Wang H, Van Der Velden BHM, Ragusi MAA, Veldhuis WB, Viergever MA, Gilhuijs KGA. Toward Computer-Assisted Triaging of Magnetic Resonance Imaging-Guided Biopsy in Preoperative Breast Cancer Patients. *Investigative Radiology.* 256(7):442–449 (2021)

Wang H, Van Der Velden BHM, Verburg E, Bakker MF, Pijnappel RM, Veldhuis WB, Gils CH van, Gilhuijs, KGA. Assessing Quantitative Parenchymal Features at Baseline Dynamic Contrast-enhanced MRI and Cancer Occurrence in Women with Extremely Dense Breasts. *Radiology.* 308(2):e222841 (2023)

Wang H, Van Der Velden BHM, Verburg E, Bakker MF, Pijnappel RM, Veldhuis WB, Gils CH van, Gilhuijs, KGA. Automated Rating of Background Parenchymal Enhancement in Breast MRI from 4553 Women in the DENSE Trial Using Machine Learning. Submitted

Wang H, Van Der Velden BHM, Chan HSM, Loo CE, Viergever MA, Gilhuijs KGA. Interpretation of phenotype of synchronous breast cancer in relation to the index cancer. *European Congress of Radiology 2020* (Poster presentation)

Wang H, Van Der Velden BHM, Ragusi MAA, Veldhuis WB, Viergever MA, Gilhuijs KGA. Feasibility of computer-assisted triaging of MRI-guided biopsy in preoperative breast-cancer patients. *European Congress of Radiology 2021*(Oral presentation)

Curriculum Vitae

Hui Wang was born on 3rd July 1990 in China. She majored in clinical medicine and obtained her M.D. from Soochow University. She then moved to the Netherlands and started her PhD journey at the Image Sciences Institute under the supervision of dr. Kenneth Gilhuijs, dr. ir. Bas van der Velden and prof. dr. ir. Max Viergever. Her research aims to support triaging decisions in screening, diagnostic workflow, and treatment by extracting imaging biomarkers from fibroglandular tissue and breast lesions on breast MRI. The relevant results are presented in the current thesis.