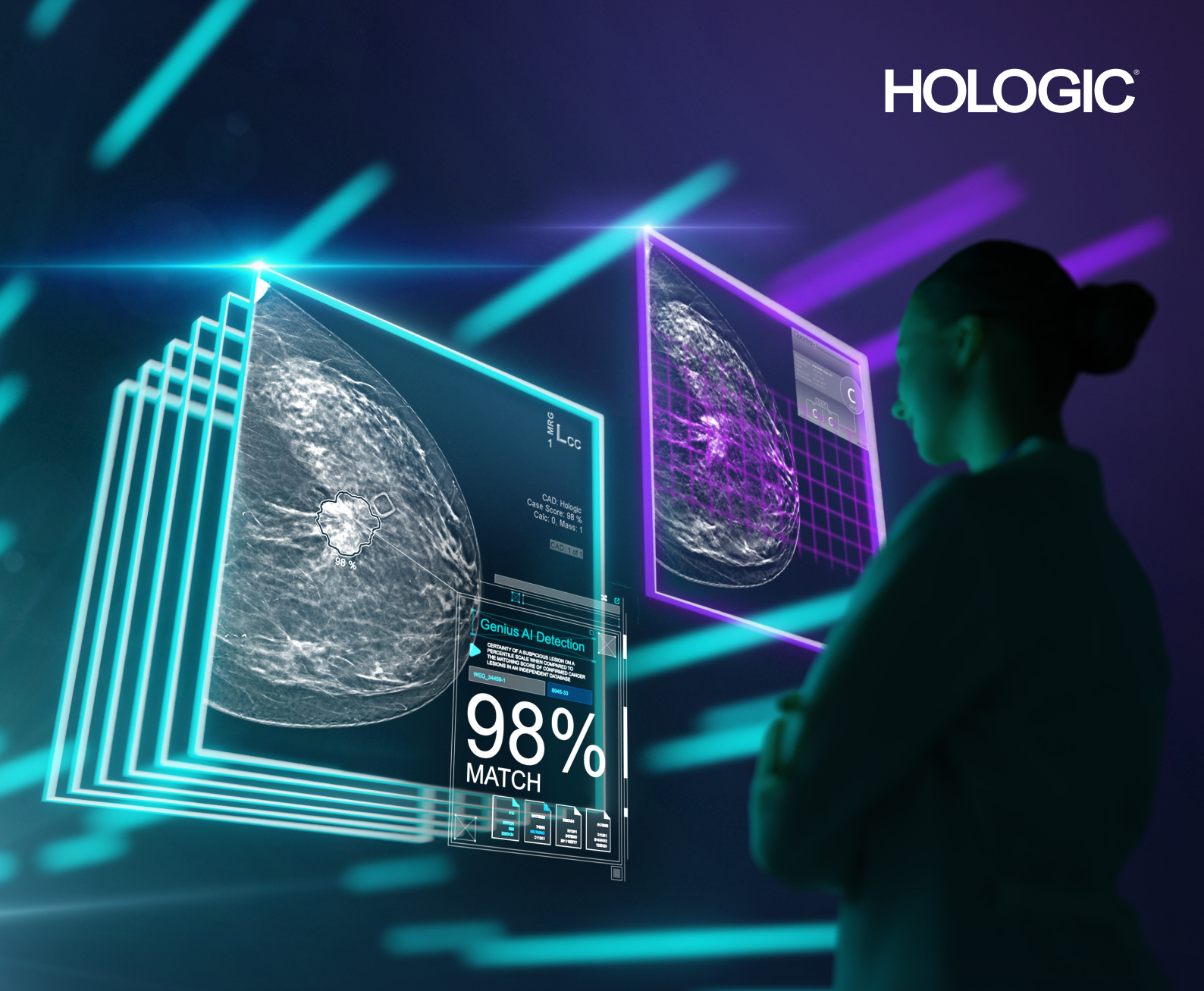

BJR
Breast Imaging & AI
Research papers from *BJR*
BIR
The British Institute of Radiology
HOLOGIC®

HOLOGIC®



Discover the integrated AI software solutions designed to improve early detection and diagnosis of breast cancer and help accelerate clinical workflow.

Future-proofed and incorporated in the 3Dimensions™ Mammography system enabling confident clinical decision making, providing better patient care.[1-7]

genius AI™ DETECTION

3DQuorum™ IMAGING TECHNOLOGY

Quantra™ BREAST DENSITY ASSESSMENT

www.3dimensionsmammography.eu

CE 2797 EC REP Hologic BV, Da Vincilaan 5, 1930 Zaventem, Belgium.

*Genius AI Detection Technology not CE marked. Not available for sale. Not for distribution.

References: 1. Data On File: DHM-06039 Rev 002 Bartoshevich J, Orefice T, Mansoor S, et.al Internal Study comparing Hologic's flat paddle to the SmartCurve paddle (18x24cm) (2017). 2. Friedewald, S.M., Rafferty, E.A., Rose, S.L., et al. Breast cancer screening using tomosynthesis in combination with digital mammography. JAMA. 2014 Jun 25;311(24):2499-507. 3. Zuckerman, S.P., Conant, E.F., Keller, B.M., et al. Implementation of Synthesized Two dimensional Mammography in a Population based Digital Breast Tomosynthesis Screening Program. Radiology. 2016 Dec;281(3):730- 736. 4. Skaane, P., Bandos, A., Eben, EB, et al. Two-view digital breast tomosynthesis screening with synthetically reconstructed projection images: comparison with digital breast tomosynthesis with full-field digital mammographic images. Radiology. 2014 Jun;271(3):655-63. 5. Bernardi, D., Macaskill, P., Pellegrini, M., et. al. Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study. Lancet Oncol. 2016 Aug;17(8):1105-13. 6. McDonald, E.S., Oustimov, A., Weinstein, S.P., et al. Effectiveness of Digital Breast Tomosynthesis Compared With Digital Mammography: Outcomes Analysis From 3 Years of Breast Cancer Screening. JAMA Oncol. 2016 Jun 1;2(6):737-43. 7. Rafferty, E.A., Durand, M.A., Conant, E.F., et al. Breast Cancer Screening Using Tomosynthesis and Digital Mammography in Dense and Nondense Breasts. JAMA. 2016 Apr 26;315(16):1784-6

Contents

Hologic has supported the production and distribution of this curated collection of independently produced articles originally published in *BJR* or *BJR*|*Open*. Hologic had no involvement with any of the articles prior to their inclusion in this collection.

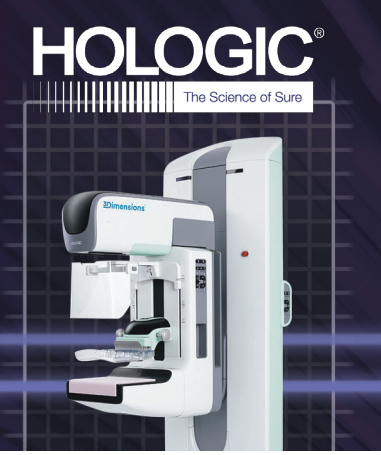
BJR

BJR is the international research journal of the British Institute of Radiology and is the oldest scientific journal in the field of radiology and related sciences.

An international, multi-disciplinary journal, *BJR* covers the clinical and technical aspects of medical imaging, radiotherapy, oncology, medical physics, radiobiology and the underpinning sciences. *BJR* is essential reading for radiologists, medical physicists, radiation oncologists, radiotherapists, radiographers and radiobiologists.

Breast imaging technologies that unlock the advantage of time

AI based solutions to enable you to achieve earlier and more accurate diagnosis and treatment



Pandemic provides 'turning point' for smarter breast cancer screening

The COVID-19 pandemic disrupted breast cancer screening across the UK; however, with the service getting back on track, there is a recognition that the pandemic could deliver the long-term legacy of stimulating smart breast screening technology.

"By the end of March this year, many units (though by no means all) will have brought screening intervals back to the recommended 36 months," says Dr Nisha Sharma, Director of Breast Screening and Clinical Lead for breast imaging from Leeds Teaching Hospitals, NHS Trust. "This amazing feat has been made possible by the good will of NHS screening staff who've gone to extraordinary lengths to reduce backlogs."

Not that the pandemic's impact on the screening service should be understated. In the UK, nearly one million breast screening appointments were delayed and, as a direct result in 2020 an estimated 10,725 breast cancer cases went undetected . Such disruptions can have tragic consequences. In England alone, it has been estimated that pandemic-related postponements of cancer diagnosis and treatment will result in approximately 18,000 additional cancer-related deaths .

Biggest threat is the shortage of radiologists

Even beyond the pandemic, threats to the breast screening service still exist, with by far the biggest issue being shortages of radiologists. According to the 2020 annual census from the Royal College of Radiologists (RCR), the UK radiologist workforce is short staffed by approximately 33% and needs nearly 2,000 more radiology consultants to ensure appropriate services. Furthermore, unless further steps are taken the RCR forecasts the shortfall will hit 3,600 by 2025 (equivalent to 44% of the workforce)[3].

Despite all the pandemic difficulties there are positive aspects which going forward could provide the 'turning point' in how breast screening services are delivered. Recent experiences have given everyone in the NHS an appetite to get involved with innovation and research to improve patient pathways. There's growing realization amongst clinicians that they need to start thinking outside the box by embracing smarter technology.

Smart technology to improve mammography workflow

Smarter technology, such as the incorporation of artificial intelligence (AI) solutions, could enable operational efficiencies providing the breast screening service with greater productivity using the workforce it already has. Going forward, smart technology won't perhaps allow for expansion, but it may enable clinicians to sustain services for longer.

With Hologic's commitment to improving lives, not only through its HGWHI, but also through its innovation across breast health, its latest technologies have the potential to improve mammography workflows, particularly alleviating pressures on image interpretation and reporting including reducing the Picture Archiving and Communication System (PACS) storage requirements for tomosynthesis images as well as helping radiologists prioritise patients, seek second opinions in real time, and turn around diagnoses for patients more quickly.

Hologic's 3Dimensions™ Mammography System is already aiding diagnosis in clinical practice where the use of tomosynthesis can separate out overlapping fibroglandular breast tissue to improve visualisation of abnormalities. While false positives lead to unnecessary recalls and further anxiety for the patient, an even more disturbing aspect is that 15 to 30% of breast cancers are not detected by standard 2D mammography[4].

Studies have shown that digital breast tomosynthesis (DBT), when compared to traditional 2D mammography, detects up to 65% more invasive breast cancers, and reduces patient recall by up to 40%[5,6]. However, the creation of 1 mm thick tomosynthesis slices, compared with 2D mammography, results in an increase in the number of images that radiologists need to review. With 2D mammography routinely there are four standard images taken (two views of each breast), but with DBT there can be between 150 and 300 images produced, leading to implications around the increased time needed to scroll through them and the contribution this may make to eye fatigue and concentration.

In response to such time challenges, Hologic has introduced a suite of imaging solutions powered by Genius AI technology. 3DQuorum™ Imaging Technology, uniquely reconstructs high-resolution tomosynthesis slices to create 6mm SmartSlices at the point of care[7]. When a Radiologist reads SmartSlices instead of 1mm tomosynthesis slices, the number of 3D™ images to review is reduced by 66%, leading to average interpretation time savings of one hour per day, based on eight hours of image interpretation time[8]. Such optimisation of time when utilising 3DQuorum™ and SmartSlices could allow radiologists to read 15% more cases per hour.

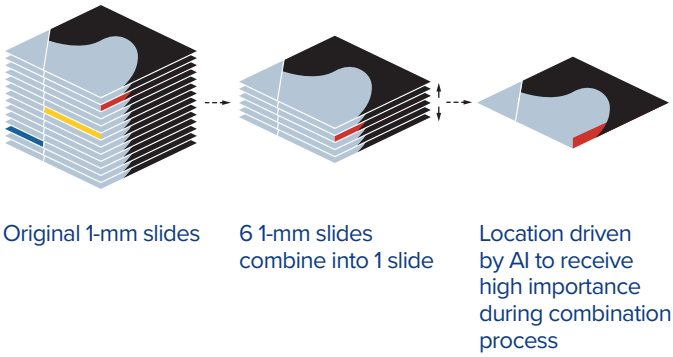


Figure 1: 2 Six 1-mm slices are combined into one 6-mm SmartSlice, giving extra weight to AI-located objects of interest

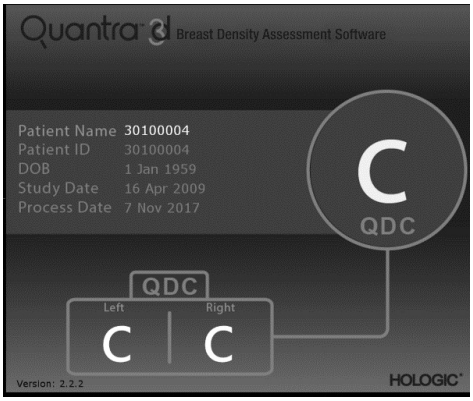
The addition of Genius AI™ Detection* deep learning technology assists radiologists in the detection of breast cancer by identifying suspicious lesions that could potentially be cancerous. The advanced artificial intelligence algorithm is trained using a high-volume image database and is intended to enhance the diagnostic accuracy of radiologists. Suspicious areas can be highlighted at the reading workstation for concurrent reading to aid interpretation. In addition to identifying potential cancers, Genius AI™ Detection software resulted in a difference of +9% in observed reader sensitivity for cancer cases and facilitated categorisation of high-risk cases according to 'priority', thereby enhancing workflow.

Studies have demonstrated that when Hologic's Genius AI™ Detection technology is used in combination with 3DQuorum™, radiologists can detect and diagnose breast cancer faster and more accurately[10].

Special role in women with dense breast tissue

Where new technologies may also come into their own is in risk stratification for women with dense breast tissue who are known to be at particularly high risk of developing breast cancer. According to American College of Radiology (ACR), Breast Imaging Reporting and Data System (Bi-RADS 5) criteria, used for reporting breast density, there are four categories of breast density: D1, almost all fatty tissue; D2, mostly fatty tissue with scattered areas of dense (fibrous and glandular) tissue; D3, mix of dense and fatty tissue; and D4, mostly dense tissue[9]. Patients in the D4 category, which Breast Cancer Now estimates affects over 700,000 women

in the UK[10], face the dual problem of increased density masking the detection of cancer as well as a 4.6-fold increased risk of developing breast cancer[11].



With such high risks, it would be of great value to identify these women who may benefit from more regular breast screening. In many countries, there is no official requirement for a quantitative assessment of breast density, leading to considerable intra and inter reader variability in qualitative assessments which leads to the possibility of unintended bias regarding screening decisions and reporting the density category[12, 13, 14].

Compatible with images acquired from all Hologic's Dimensions systems, including the 3Dimensions™ Mammography system, Quantra™ 2.2 breast density assessment software, a machine-learning algorithm analysing each patient's individual breast pattern and texture, provides objective assessments of breast density (using Bi-RADS 5 criteria).

Provision of a fully automated methodology that can be displayed on the acquisition workstation can offer an unbiased and reproducible option to assess and report breast density category. Hologic's Quantra software offers this fully automated choice, which is well-integrated into the existing mammographic workflow.

In tackling backlogs, AI machine learning and deep learning technologies could be used to analyse prior mammograms of women retrospectively and identify those at greatest risk. Such information might also allow screening intervals for lower risk women to be extended, thereby creating more efficient and targeted breast screening programmes.

Studies also show that using DBT as a screening tool for women with dense breasts delivers even greater benefits than in the general population. For women with dense breasts, the use of DBT, with or without synthesised 2D imaging, enables a significant increase in CDR (cancer detection rate) of both invasive and interval cancers in all density and age groups in comparison to conventional digital mammography. Furthermore, when using DBT, incremental recalls are less in high density screens and therefore a more density-tailored screening is beneficial[15, 16, 17, 18]. The European Commission has fully appreciated the benefits of using 3D mammography in women with dense breasts, issuing guidelines in August 2020, that women with previously detected high mammographic density should be screened with DBT[19].

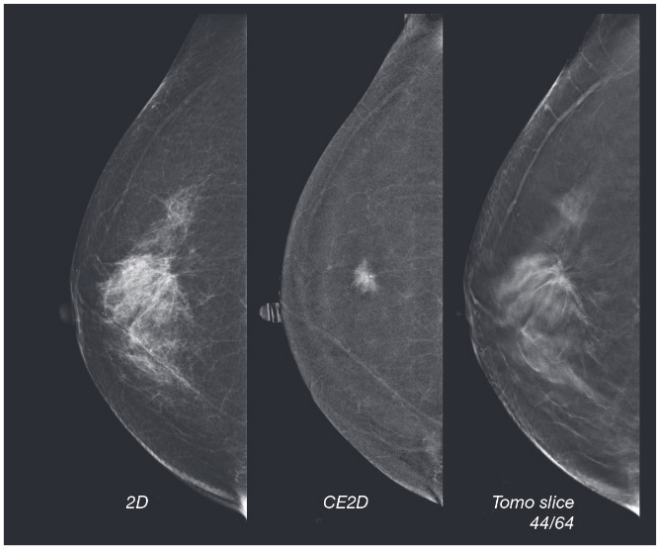


Figure 2: IDC enhances well on Contrast enhanced 2D image with spiculations seen on the tomo slice.

Multiple clinical studies have showed the benefit of screening mammography in reducing breast cancer [20, 21, 22]. The efficacy of screening depends upon the ability to identify cancers based on the differing absorption of x-rays in cancerous tissue compared to adipose and glandular tissue. Imaging a contrast agent identifies lesions and potential malignancies, providing additional information that may complement the results of mammography or tomosynthesis. Studies have also demonstrated that CESM may be a viable alternative to breast MRI as due to its clinical advantages, it can lower procedure costs and the procedure time is shorter. CESM could also be used as a method to triage and prioritise women who are most at risk, such as women with dense breasts, in a more cost-effective way[23].

Current state of play

Although 3D Mammography is widely available and used in the USA and Europe (including The Netherlands), in the UK, it is not used in routine breast screening, and is only for further assessment of recall patients and in the purely symptomatic setting.

The PROSPECTS trial, currently taking place in 10 NHS screening sites in England, will assess 100,000 women undergoing routine breast cancer screening, aged between 49 and 71 years. Patients are being randomised to 2D mammography or 2D mammography plus DBT. It is anticipated that the results from the PROSPECT's trial will provide information on numbers of cancers detected, size of cancers detected, false alarm recalls and cost effectiveness. This data will be used by the UK Breast Screening Committee to make decisions around whether DBT is appropriate for the UK setting.

If introduced, DBT together with other forms of smart technology, could not only transform breast screening services but also play an important role in workflow prioritisation, which could bring greater certainty and peace of mind to patients and clinicians, as well as helping address staff shortages. DBT and the use of AI has the potential to increase breast cancer screening capacity, by removing the need for review by two radiologists. Additionally, AI could effectively and efficiently highlight the areas that are of particular interest for the radiologist.

Based on a comparison with the average time taken to read a breast screening image, with AI less time is needed[10], improving the efficiency with which images are reviewed. This time saving could mean that radiologists are able to read more cases per day or could focus more of their time on the more complex cases or patients who are at higher risk.

Healthcare providers will also have to be ready for some serious investment beyond the actual technology, "Before introducing smart technology, organisations will have to ensure they have IT infrastructure in place that has the capacity to deal with bigger file sizes and ensure that systems are robust", cautions Tim Simpson, General Manager of UK & Ireland, Hologic.

Looking into the not so distant future, it is evident that AI is sure to revolutionise healthcare. There will be multiple benefits associated with the adoption of AI technology in breast imaging. Accelerating detection and increasing accuracy of breast cancer diagnosis, as well as enhancing clinical confidence, and improving clinical and operational workflow efficiencies, are areas where technology will support clinicians and patients alike.

Hologic is excited to be at the forefront of offering AI solutions which ultimately unlock the advantage of time for clinicians and patients, ensuring the best possible experiences and outcomes. Indeed, there is still much to be done, but the future of mammography with the inclusion of AI seems a promising one.

1. MISC-07567-EUR-EN Finsbury International Policy & Regulatory Advisers (2021). Impact of the COVID-19 pandemic on breast cancer screening in Italy, France, Spain and UK. Internal FIPRA report. Unpublished.

2. Lai A, Pasea L, Banerjee A, et al. Estimating excess mortality in people with cancer and multimorbidity in the COVID-19 emergency. BMJ Open 2020;10:e043828. doi:10.1136/bmjopen-2020-043828

3. The Royal College of Radiologists. Clinical radiology UK census 2020 report. London: The Royal College of Radiologists, 2021. Available from: https://www.rcr.ac.uk/publication/clinical-radiology-uk-workforce-census-2020-report

4. Gilbert FJ, Tucker L, GC Gilian M. Accuracy of Digital Breast Tomosynthesis for Depicting Breast cancer subgroups in a UK retrospective reading study. Radiology: Volume 277: Number 3, 2015

5. Friedewald SM, Rafferty EA, Rose SL, Durand MA, Plecha DM, Greenberg JS, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. JAMA. 2014 Jun 25;311(24):2499-507.

6. Rafferty E, Park J, Philpotts L, Poplack SP, Sumkin JH, Halpern EF, et al., Assessing Radiologist Performance Using Combined Digital Mammography and Breast Tomosynthesis Compared with Digital Mammography alone: Results of a Multicenter, Multireader trial. Radiology, 2013 Jan; 266(1):104-13. Epub 2012 Nov 20.

7. Hologic data on file. DHM-08611.

8. Hologic data on file. CSR-00116

9. D'Orsi CJ, Mendelson EB, Ikeda DM, et al. Breast Imaging Reporting and Data System: ACR BI-RADS—breast imaging atlas. Reston, VA: American College of Radiology, 2003

Read More: https://www.ajronline.org/doi/10.2214/AJR.12.10197

10. Breastcancernow.org. 2022. Good enough? Breast cancer in the UK. [online] Available at: https://breastcancernow.org/sites/default/files/good_enough_-_breast_cancer_in_the_uk_-_report.pdf#:~:text=Over%2050%2C000%20women%20and%20around%20350%20men%20are,men%20still%20die%20of%20breast%20cancer%20each%20year.> [Accessed 20 April 2022].

11. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev. 2006; 15 (6): 1159-69

12. Redondo A, Comas M, Macià F. Inter- and intraradiologist variability in the BI-RADS assessmentand breast density categories for screening mammograms.2012, Br J Radiol, Vol. 85(1019), pp.1465–1470.

13. Ciatto, S, Houssami N, Apruzzeseal A. Categorizing breast mammographic density: intra- andinterobserver reproducibility of BI-RADS density categories. 2005, Breast, Vol. 14, pp. 269-275

14. Timmers, JM, J. van Doorne-Nagtegaal H, L. M. Verbeek A. A dedicated BI-RADS training programme: Effect on the inter-observer variation among screening radiologist.. 2012, Radiology, Vol. 81, pp. 2184-2188.

15. Helge Østerås B, Martinsen ACT, Gullien R. Digital Mammography versus Breast Tomosynthesis: Impact of Breast Density on Diagnostic Performance in Population-based Screening. 2019. [Accessed 22 April 2022] (Online) https://pubs.rsna.org/doi/full/10.1148/radiol.2019190425#:~:text=Prospective%20(5%E2%80%9311)%20and,fatty%20and%20extremely%20dense%20breasts

16. Li T, Houssami N, Noguchi . Differential detection by breast density for digital breast tomosynthesis versus digital mammography population screening: a systematic review and meta-analysis. 2022 – [Accessed 22 April 2022] (Online) https://www.nature.com/articles/s41416-022-01790-x

17. Abdullah P, Alabousi M, Ramadan S. Synthetic 2D Mammography Versus Standard 2D Digital Mammography: A Diagnostic Test Accuracy Systematic Review and Meta-Analysis. Volume 217, Number 2, 2021. Available from: https://www.ajronline.org/doi/abs/10.2214/AJR.20.24204?journalCode=ajr

18. Lowry KP,Yates Coley R, Miglioretti DL. Screening Performance of Digital Breast Tomosynthesis vs Digital Mammography in Community Practice by Patient Age, Screening Round, and Breast Density. 2020. – [Accessed 22 April 2022] (Online) https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2768727

19. Healthcare-quality.jrc.ec.europa.eu. 2022. European Commission Initiative on Breast Cancer. [online] Available at: <https://healthcare-quality.jrc.ec.europa.eu/european-breast-cancer-guidelines/dense-breast> [Accessed 20 April 2022].

20. Tabar L, Vitak B, Chen HH, Yen MF, et al. Beyond randomized controlled trials: organized mammographic screening substantially reduces breast carcinoma mortality. Cancer 2001; 91:1724-1732

21. Tabar L, Vitak B, Chen TH, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. Radiology 2011; 260:658-663

22. Lehtimaki T, Lundin M, Linder N, et al. Long-term prognosis of breast cancer detected by mammography screening or other methods. Breast Cancer Res 2011; 13:R134

23. Smith, A. The Principles of Contrast Mammography. US: Hologic. 2014 – Reference 14 - Lobbes MB, Lalji U, Houwers J, et al. Contrast-enhanced spectral mammography in patients referred from the breast cancer screening programme. Eur Radiol. 2014 Jul;24(7):1668-76.

BJR

https://doi.org/10.1259/bjr.20200135

Received: 06 February 2020 | Revised: 02 August 2020 | Accepted: 10 August 2020

Cite this article as:
Lyburn ID, Pinder SE. Screening detects a myriad of breast disease – refining practice will increase effectiveness and reduce harm. *Br J Radiol* 2020; **93**: 20200135.

COMMENTARY

Screening detects a myriad of breast disease – refining practice will increase effectiveness and reduce harm

[1]IAIN D LYBURN, FRCR and [2]SARAH E PINDER, FRCPath

[1]Cheltenham Imaging Centre, Cobalt Medical Charity, Cheltenham, United Kingdom
[2]School of Cancer & Pharmaceutical Sciences, King's College, London, United Kingdom

Address correspondence to: Prof Iain D Lyburn
E-mail: *iain.lyburn@nhs.net*

ABSTRACT

For many individuals, the term 'cancer' equates to a disease that if untreated will progress, spread from the area initially affected and ultimately cause death. 'Breast cancer', however, is a diverse of range of pathological entities, incorporating indolent to fast-growing and aggressive lesions, with varying histological patterns, clinical presentations, treatment responses and outcomes. Screening for malignancy is based on the assumption that cancer has a gradual, orderly progression and that detecting lesions earlier in their natural history, and intervening, will reduce mortality. The natural history of epithelial atypia, ductal carcinoma *in situ* and even invasive breast cancer is poorly understood, but widely variable. We believe that population breast screening methodology needs to change to focus on diagnosis of lesions of greatest clinical relevance.

COMMENTARY

There is a growing understanding of the heterogeneity of breast cancer and its associated behaviour. Some lesions may never progress to 'cancer' and may potentially be managed with watchful waiting. Randomised clinical trials are underway investigating this approach; for instance, active surveillance trials for low risk ductal carcinoma *in situ* (DCIS).[1] With evolving systemic therapies and improved molecular tools, there is the potential to reduce treatment burden in lower-risk tumours.[2] Truly personalised medicine involves tailoring, not only clinical management with more intensive treatment for some cancers *vs* less intensive or no intervention for others, but potentially also dovetailing investigations for detection and for surveillance. Breast screening should aim to detect clinically relevant cancers, not lesions that would cause no harm during lifetime.

Cancer progression is variable. The wide range of factors that influence invasion and metastasis, including general physiological and nutritional status, co-morbidities, stromal and tumour characteristics, remain imperfectly understood. If a tumour develops slowly but is likely to progress if unchecked, early detection is likely to be beneficial. For tumours that develop rapidly or disseminate early, screen detection may not improve patient outcome. Screening detects a higher proportion of indolent disease, due to the inherent tendency to preferentially identify slower growing cancers because more rapidly growing cancers are more likely to present symptomatically between screens (interval cancers). Thus, some subtypes of invasive breast cancer (*e.g.* Grade 1 tubular cancers) are more often screen-detected than symptomatic in presentation.

Two-dimensional full field digital mammography (FFDM) is currently the imaging basis of breast screening. Integrated FFDM and digital breast tomosynthesis (DBT) improves breast-cancer detection.[3] Screening in centres in some countries now incorporates integrated FFDM and DBT. Mammography is an imperfect science. In women with dense fibroglandular tissue, the sensitivity and specificity are lower than in those with more radiolucent breasts. The efficiency, and possibly the effectiveness, of mammographic screening is also lower in users of hormone replacement therapy, in females with previous breast surgery, and those of lower body weight.[4] Partly for these reasons, but multifactorially, the reported estimates on the effect of mammography screening on breast cancer mortality reduction vary widely. Nevertheless, a review of randomised controlled mammography trials reported an estimated mortality reduction of 20% in females aged between 50 and 70 years old.[5]

There is, however, increasing debate on the importance of the negative impact of breast screening, including over diagnosis – the identification of tumours that would otherwise not become symptomatic within the female's lifetime - and subsequent over treatment of such lesions. Another negative impact of screening is that of recall for further assessment in those without malignancy ('false-positive recalls'); approximately, 4% of females screened in the UK are invited to attend for further investigation, with a wide range of recalls between centres.[6] There is evidence that increases in recall rates above-defined levels are almost exclusively associated with false-positive recalls with only a very small increase in detection of low/intermediate grade DCIS (*i.e.* not cancers likely to be life-threatening).[7] Under current diagnostic algorithms, recall leads to tissue sampling in approximately 50% of females. This yields specimens which are definitively benign or malignant in most cases but, in 5–9% of core biopsies,[6] not clearly either, *i.e.*a lesion of uncertain potential (B3).

The B3 category represents a heterogeneous group of lesions that have an increased risk of adjacent malignancy, but this ranges from a few percent to up to 40%, depending on the abnormality and the method of biopsy.[8] B3 lesions with epithelial atypia include atypical intraductal epithelial proliferation (AIDEP), lobular neoplasia (the combined term applied in core biopsy for atypical lobular hyperplasia (ALH) and lobular carcinoma *in situ*) or flat epithelial atypia (FEA). The histopathological diagnostic criteria for some of these include assessment of extent of the process, and thus require a volume of tissue which is often not possible on a 14G sample ('standard core'). There is an upgrade rate – chance of adjacent DCIS or invasive cancer – of about 40% for AIDEP in a 14G sample[8] which is less (about 20%) in vacuum biopsy (obtaining a larger sample). About 30% of cases of lobular neoplasia on 14G core biopsy will have adjacent DCIS or invasive cancer.[9] The upgrade rate for FEA is lower (11%[10]).In addition to the risk of there being contemporaneous adjacent cancer, the risk of subsequent cancer in females with atypical ductal hyperplasia or ALH is increased by three- to fourfold,[11] *i.e.* the risk of progression is low. B3 lesions are more common in screening than symptomatic practice but, despite guidance recommending vacuum-assisted excision for many,[12] 41% of females with a screen-detected B3 lesion in the UK in 2018 to 2019 underwent surgical excision.[13] There are no biological markers available which can be used to predict either the risk of adjacent or subsequent cancer and there is no global agreement on management or follow-up of such patients. The long-term benefits to the patient, and at population level, of identifying these lesions at screening is unclear.

DCIS is not one disease, as shown by different presentations, appearances (imaging and histological), biomarkers and genetics. Information on the natural history of DCIS is limited because current standard of care is surgical excision. Reported series of untreated DCIS are of modest numbers and most direct evidence is from series where histological reviews of disease were originally diagnosed as benign and therefore not completely excised (most often low grade).[14] Overall, between 14 and 53% of DCIS progresses to invasive cancer over a period of 10 or more years,[15] but not all DCIS is equal; low-grade DCIS has a slower rate of progression over a very long time (up to 40+ years). Significantly, whilst low-grade DCIS is associated with low-grade invasive cancer, high-grade DCIS tends to progress to Grade 2 and Grade 3 disease. At the population level, there is a negative correlation between DCIS detection rates and interval cancer (notably including Grade 3 cancer) rates, albeit in observational studies,[16] indicating that the screen detection of, at least some, DCIS is worthwhile.

Invasive breast cancer encompasses an even broader range of patterns than atypia and DCIS. Although about three-quarters are of no special type (ductal), there are many histological types (and subtypes of types) of invasive breast cancer. Some 'special types' of breast carcinoma, including tubular, tubulolobular, invasive cribriform and Grade 1 mucinous carcinomas, have a good prognosis with >80% 10-year-survival, whilst others have a poorer outcome.[17] Histological grade adds significant prognostic information to tumour type. However, no individual patient, imaging or pathological feature is very informative regarding the natural history of the lesion's origins. None are used alone for clinical management purposes; none are sufficiently good at identifying an excellent (or poor) group for recurrence or patient survival. Nevertheless, it is clear that the intrinsic biology of the cancer has a significant effect on long-term patient outcome – a small tumour may have a high metastatic potential and a large tumour may have low potential for dissemination. A greater understanding of more detailed biological factors will improve strategies for prevention and screening.

The identification of Grade 1 cancers less than 20 mm in size and Grade 2 and 3 cancers less than 10 mm in size at screening is likely to be beneficial, with a lower likelihood of developing metastatic disease from such lesions.[18] The significant negative association between screen-detected DCIS and the rate of invasive interval cancers suggests that detection and treatment of (at least some) DCIS is worthwhile in prevention of future invasive disease.[16] We believe that screening methodologies should concentrate on the identification of small high grade lesions (both DCIS and invasive), which are those most likely to influence patient outcome, rather than small low-grade tumours, which are those most likely to represent over diagnosis and subsequent over treatment.

Although associated with histological grade and type, genomics and other biomarkers provide additional information. Gene expression profiling categorises breast cancers into molecular subtypes: luminal A, luminal B, human epidermal growth factor receptor2 (HER2)–enriched, and basal-like, which have different patterns of disease, response to therapy and survival outcomes. The initial presentation of disease and subsequent metastatic spread are also influenced by molecular subtype.[19] For example, patients with luminal A and B cancers are more likely to develop metastases in the skeleton than are those with basal-like subtype tumours, who more frequently develop lung and brain metastases.[20] Breast carcinomas of differing molecular subtypes also show variation in response to therapies; such knowledge guides initial treatment planning and imaging follow-up. However, formal molecular genomic subtype analysis in UK day-to-day practice is currently not practical or cost-efficient and surrogate

immunohistochemical markers are utilised. Oestrogen receptor (ER), progesterone receptor (PR), and HER2 status are used to define surrogate molecular subtypes, but concordance with formal genetic analysis ranges from 41 to 100%.[21] Biomarkers may, however, be variably expressed within a cancer, *e.g* heterogeneous expression of Her2 and Her2 amplification is recognised, highlighting the presence of intratumoral heterogeneity.[22] However, driver genetic variations (*e.g.* in TP53, PIK3CA, PTEN, MYC and BRCA2) occur early in some cancers, and late in others, reflecting the complexity of tumour progression.[23] Although our understanding of the diagnostic and clinical implications of such intratumoral heterogeneity is imperfect, the presence of subclones within individual tumour may impact optimum approaches to imaging, tissue sampling and to patient treatment.

Given the heterogeneity of breast pathologies histologically and genomically and their variable behaviours and outcomes, more tailored approaches to detecting, classifying and managing screen detected entities are required. Artificial intelligence (AI) systems have been demonstrated to have the potential to be capable of surpassing human experts in breast cancer prediction on FFDM; using UK and USA data sets respectively, reductions of 5.7 and 1.2% in false positives, and 9.4 and 2.7% in false negatives have been demonstrated.[24] The AI system maintained non-inferior performance and reduced the workload of the second reader by 88%. Further AI system development and testing on data sets is likely to lead to clinical trials to improve the accuracy and efficiency of breast cancer screening. For instance, currently the morphological type of a mammographically indeterminate lesion does not appear to be correlated with cancer risk, but AI/machine learning could potentially extract imaging features to aid more refined categorisation.

The incidence of invasive breast cancer has risen since the early 1980s, especially those which are of less aggressive phenotype[25] with about 30% of screen detected invasive breast cancers being low risk by molecular profiling. We believe the future lies in risk-based screening, and identification of those for whom less screening is the best strategy as well as those who may potentially benefit from more frequent screening (*e.g.* based on their genetic risk).[26] The use of modalities other than mammography may aid detection of malignancy in dense breast tissue, identify biologically significant lesions and lead to developing more tailored screening regimens. Imaging with low-dose mammography, contrast-enhanced mammography, automated whole breast ultrasound, molecular imaging and/or MRI (including abbreviated protocols), could all contribute to breast screening programmes. The first 2 year screening round of the Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial, evaluating the use of supplemental MRI screening in females with extremely dense breast tissue and normal results on mammography, led to the diagnosis of significantly fewer interval cancers than mammography alone.[27] Using an abbreviated MRI protocol for breast cancer screening, the high diagnostic accuracy of full MRI protocols can be maintained, while the time and cost associated with traditional MRI examinations are minimised; in females with dense breasts undergoing screening, abbreviated breast MRI compared with DBT, has a significantly higher rate of invasive breast cancer detection.[28] Currently, in the UK and other countries, there are enhanced programmes for those at highest risk for lethal and rapidly growing cancers, for instance, carriers of BRCA1 and BRCA2 mutations but the efficacy of population-based screening will be also be improved by reducing the frequency of screens for those at lower risk.

In order to maximise the benefit of screening and tailoring regimens, high quality data capture is essential. Cancer registries should be enabled to allow integration of information including detailed tumour characteristics, treatment and outcome, to better understand tumour biology and prognostic significance. Many countries, including the UK, collate these at national level – analysis of 'bigger data' is likely to yield more fruitful results. Evolving models of stratified screening could be developed with lessons learnt and extrapolated to screening for other conditions.

In conclusion, to increase the effectiveness of breast screening, we should focus attention on methods for identification of lesions of greatest clinical consequence and adapt a more sophisticated, tailored approach to recall and to the range of pathological lesions, with subsequent reduced or enhanced intervention, potentially using various imaging modalities, as appropriate. At present, a large number of biopsies are carried out in a 'catch all' strategy for breast cancer, resulting in the diagnosis of a range of benign lesions, with or without epithelial atypia, and low-risk DCIS and low grade invasive disease, in addition to more clinically relevant cancers. The benefits of detection of this myriad of breast disease requires further research to increase our understanding of the relationship between screening methods and clinical outcome.

REFERENCES

1. Kanbayashi C, Thompson AM, Hwang E-SS, Partridge AH, Rea DW, Wesseling J, et al. The International collaboration of active surveillance trials for low-risk DCIS (loris, Lord, comet, LORETTA. *JCO* 2019; **37**(15_suppl): TPS603. doi: https://doi.org/10.1200/JCO.2019.37.15_suppl.TPS603
2. Esserman L, Gallant E, Alvarado M. Less is more: the evolving surgical approach to breast cancer. *Am Soc Clin Oncol Educ Book* 2016; **35**: e5–10. doi: https://doi.org/10.1200/EDBK_159060
3. Chong A, Weinstein SP, McDonald ES, Conant EF. Digital breast tomosynthesis: concepts and clinical practice. *Radiology* 2019; **292**: 1–14. doi: https://doi.org/10.1148/radiol.2019180760
4. Banks E, Reeves G, Beral V, Bull D, Crossley B, Simmonds M, et al. Influence of personal characteristics of individual women on sensitivity and specificity of mammography in the Million women study: cohort study. *BMJ* 2004; **329**: 477. doi: https://doi.org/10.1136/bmj.329.7464.477
5. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M.

The benefits and harms of breast cancer screening: an independent review. *Br J Cancer* 2013; **108**: 2205–40. doi: https://doi.org/10.1038/bjc.2013.177

6. https://files.digital.nhs.uk/0A/9D9F34/breast-screening-programme-eng-2018-19-report.pdfaccessed 02/02/2020
7. Blanks RG, Given-Wilson RM, Cohen SL, Patnick J, Alison RJ, Wallis MG. An analysis of 11.3 million screening tests examining the association between recall and cancer detection rates in the English NHS breast cancer screening programme. *Eur* 2019; **29**: 3812–9.
8. Rakha EA, Lee AHS, Jenkins JA, Murphy AE, Hamilton LJ, Ellis IO. Characterization and outcome of breast needle core biopsy diagnoses of lesions of uncertain malignant potential (B3) in abnormalities detected by mammographic screening. *Int J Cancer* 2011; **129**: 1417–24. doi: https://doi.org/10.1002/ijc.25801
9. Hussain M, Cunnick GH. Management of lobular carcinoma in-situ and atypical lobular hyperplasia of the breast--a review. *Eur J Surg Oncol* 2011; **37**: 279–89. doi: https://doi.org/10.1016/j.ejso.2011.01.009
10. Rudin AV, Hoskin TL, Fahy A, Farrell AM, Nassar A, Ghosh K, et al. Flat epithelial atypia on core biopsy and upgrade to cancer: a systematic review and meta-analysis. *Ann Surg Oncol* 2017; **24**: 3549–58. doi: https://doi.org/10.1245/s10434-017-6059-0
11. Dyrstad SW, Yan Y, Fowler AM, Colditz GA. Breast cancer risk associated with benign breast disease: systematic review and meta-analysis. *Breast Cancer Res Treat* 2015; **149**: 569–75. doi: https://doi.org/10.1007/s10549-014-3254-6
12. Pinder SE, Shaaban A, Deb R, Desai A, Gandhi A, Lee AHS, et al. Nhs breast screening multidisciplinary Working group guidelines for the diagnosis and management of breast lesions of uncertain malignant potential on core biopsy (B3 lesions. *Clin Radiol* 2018; **73**: 682–92. doi: https://doi.org/10.1016/j.crad.2018.04.004
13. https://associationofbreastsurgery.org.uk/media/64800/nhsbsp_abs_breast_screening_audit-2016-2017-v2-today.pdfaccessed 02/02/2020
14. Maxwell AJ, Clements K, Hilton B, Dodwell DJ, Evans A, Kearins O, et al. Risk factors for the development of invasive cancer in unresected ductal carcinoma in situ. *Eur J Surg Oncol* 2018; **44**: 429–35. doi: https://doi.org/10.1016/j.ejso.2017.12.007
15. Erbas B, Provenzano E, Armes J, Gertig D. The natural history of ductal carcinoma in situ of the breast: a review. *Breast Cancer Res Treat* 2006; **97**: 135–44. doi: https://doi.org/10.1007/s10549-005-9101-z
16. Duffy SW, Dibden A, Michalopoulos D, Offman J, Parmar D, Jenkins J, et al. Screen detection of ductal carcinoma in situ and subsequent incidence of invasive interval breast cancers: a retrospective population-based study. *Lancet Oncol* 2016; **17**: 109–14. doi: https://doi.org/10.1016/S1470-2045(15)00446-5
17. Pereira H, Pinder SE, Sibbering DM, Galea MH, Elston CW, Blamey RW, et al. Pathological prognostic factors in breast cancer. IV: should you be a typer or a grader? A comparative study of two histological prognostic features in operable breast carcinoma. *Histopathology* 1995; **27**: 219–26. doi: https://doi.org/10.1111/j.1365-2559.1995.tb00213.x
18. Evans AJ, Pinder SE, Burrell HC, Ellis IO, Wilson AR. Detecting which invasive cancers at mammographic screening saves lives? *J Med Screen* 2001; **8**: 86–90. doi: https://doi.org/10.1136/jms.8.2.86
19. Lam SW, Jimenez CR, Boven E. Breast cancer classification by proteomic technologies: current state of knowledge. *Cancer Treat Rev* 2014; **40**: 129–38. doi: https://doi.org/10.1016/j.ctrv.2013.06.006
20. Smid M, Wang Y, Zhang Y, Sieuwerts AM, Yu J, Klijn JGM, et al. Subtypes of breast cancer show preferential site of relapse. *Cancer Res* 2008; **68**: 3108–14. doi: https://doi.org/10.1158/0008-5472.CAN-07-5644
21. Guiu S, Michiels S, André F, Cortes J, Denkert C, Di Leo A, et al. Molecular subclasses of breast cancer: how do we define them? the IMPAKT 2012 Working group statement. *Ann Oncol* 2012; **23**: 2997–3006. doi: https://doi.org/10.1093/annonc/mds586
22. Joseph C, Papadaki A, Althobiti M, Alsaleem M, Aleskandarany MA, Rakha EA, et al. Breast cancer intratumour heterogeneity: current status and clinical implications. *Histopathology* 2018; **73**: 717–31. doi: https://doi.org/10.1111/his.13642
23. Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, Van Loo P, et al. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat Med* 2015; **21**: 751–9. doi: https://doi.org/10.1038/nm.3886
24. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature* 2020; **577**: 89–94. doi: https://doi.org/10.1038/s41586-019-1799-6
25. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. *N Engl J Med* 2012; **367**: 1998–2005. doi: https://doi.org/10.1056/NEJMoa1206809
26. Harkness EF, Astley SM, Evans DG. Risk-based breast cancer screening strategies in women. *Best Pract Res Clin Obstet Gynaecol* 2020; **65**: 3–17. doi: https://doi.org/10.1016/j.bpobgyn.2019.11.005
27. Bakker MF, de Lange SV, Pijnappel RM, Mann RM, Peeters PHM, Monninkhof EM, et al. Supplemental MRI screening for women with extremely dense breast tissue. *N Engl J Med* 2019; **381**: 2091–102. doi: https://doi.org/10.1056/NEJMoa1903986
28. Comstock CE, Gatsonis C, Newstead GM, Snyder BS, Gareen IF, Bergin JT, et al. Comparison of abbreviated breast MRI vs digital breast Tomosynthesis for breast cancer detection among women with dense breasts undergoing screening. *JAMA* 2020; **323**: 746. doi: https://doi.org/10.1001/jama.2020.0572

BJR|Open https://doi.org/10.1259/bjro.20210034

Received: 31 May 2021 | Revised: 29 July 2021 | Accepted: 30 July 2021

Cite this article as:
Neeter LM.F.H, Raat H.P.JF, Alcantara R, Robbe Q, Smidt ML, Wildberger JE, et al. Contrast-enhanced mammography: what the radiologist needs to know. *BJR Open* 2021; **3**: 20210034.

REVIEW ARTICLE

Contrast-enhanced mammography: what the radiologist needs to know

[1,2]LIDEWIJ M.F.H. NEETER, [3]H.P.J. (FRANK) RAAT, [4]RODRIGO ALCANTARA, [2]QUIRIEN ROBBE, [1,5]MARJOLEIN L. SMIDT, [2]JOACHIM E. WILDBERGER and [1,2,6]MARC B.I. LOBBES

[1]GROW School for Oncology and Developmental Biology, Maastricht University, Maastricht, the Netherlands
[2]Department of Radiology and Nuclear Medicine, Maastricht University Medical Center, Maastricht, the Netherlands
[3]Department of Medical Imaging, Laurentius Hospital, Roermond, the Netherlands
[4]Department of Radiology, Hospital del Mar, Barcelona, Spain
[5]Department of Surgery, Maastricht University Medical Center, Maastricht, the Netherlands
[6]Department of Medical Imaging, Zuyderland Medical Center, Sittard-Geleen, the Netherlands

Address correspondence to: Lidewij M.F.H. Neeter
E-mail: *lidewij.neeter@mumc.nl*

ABSTRACT

Contrast-enhanced mammography (CEM) is a combination of standard mammography and iodinated contrast material administration. During the last decade, CEM has found its place in breast imaging protocols: after i.v. administration of iodinated contrast material, low-energy and high-energy images are retrieved in one acquisition using a dual-energy technique, and a recombined image is constructed enabling visualisation of areas of contrast uptake.

The increased incorporation of CEM into everyday clinical practice is reflected in the installation of dedicated equipment worldwide, the (commercial) availability of systems from different vendors, the number of CEM examinations performed, and the number of scientific articles published on the subject. It follows that ever more radiologists will be confronted with this technique, and thus be required to keep up to date with the latest developments in the field. Most importantly, radiologists must have sufficient knowledge on how to interpret CEM images and be acquainted with common artefacts and pitfalls.

This comprehensive review provides a practical overview of CEM technique, including CEM-guided biopsy; reading, interpretation and structured reporting of CEM images, including the accompanying learning curve, CEM artefacts and interpretation pitfalls; indications for CEM; disadvantages of CEM; and future developments.

INTRODUCTION

To date, full-field digital mammography (FFDM) remains the primary imaging tool in breast cancer imaging worldwide. FFDM plays a pivotal role in breast cancer detection in clinical practice as well as in screening programmes.[1] However, FFDM is less accurate in females with dense breast tissue.[2,3] To resolve this issue, many technologies have been proposed as adjuncts to FFDM, such as digital breast tomosynthesis (DBT), breast ultrasound (US), and breast magnetic resonance imaging (MRI). Contrast-enhanced mammography (CEM) – a combination of mammography and iodinated contrast material administration – is the latest addition, and has consistently been shown to increase diagnostic accuracy as compared to FFDM.[4–6] Unsurprisingly therefore, CEM is steadily gaining ground, as is reflected in the increasing numbers of CEM equipment, examinations, and published studies.[7] First commercially introduced in 2011, CEM is now being offered on five different systems by four vendors.[8,9] Even although system characteristics differ, all available systems use a similar approach and will therefore be uniformly referred to as CEM throughout this review.

A consequence of the growing popularity of CEM is that more and more radiologists will be confronted with this technique. Radiologists will be required to keep up to date with the latest developments in this field and to acquire sufficient knowledge of CEM image interpretation. Most importantly, radiologists need to become acquainted with artefacts commonly seen in CEM and consequent interpretation pitfalls.

The current comprehensive review gives a practical overview and recommendations for CEM technique, including CEM-guided biopsy; reading, interpretation and structured reporting of CEM images, including the accompanying

learning curve and an overview of CEM-specific artefacts and interpretation pitfalls; indications for CEM; disadvantages of CEM; and future developments.

CEM technique: Principles, image acquisition and patient handling

Small tumours depend on diffusion to acquire oxygen and nutrients for their growth. As the tumour expands, diffusion becomes insufficient. Parts of the tumour then become hypoxic, stimulating the release of vascular growth factors. The latter promote new blood vessel formation, ultimately creating vascularization of the tumour itself and providing access to the oxygen and nutrients required for further growth.[10] These rapidly formed new blood vessels are often 'leaky' to contrast agents. As a consequence, after intravascular administration some contrast agent will enter and 'enhance' the tumour interstitium. This can be exploited for diagnostic purposes, provided the proper imaging tool is used.[10] In CEM, iodinated contrast agents are used, usually at a concentration of 300–370 mg iodine/ml.[7]

Intravascular iodinated contrast administration will extend the room time of a typical CEM examination to 15–20 min, which is approximately twice the time required for a FFDM.[11–13] Contrast agent is administered through an i.v. catheter, usually placed in an antecubital vein, preferably using an automatic injector at rate 2–3 $ml\,s^{-1}$ and followed by a saline flush at the same flow rate. Before injecting the contrast agent, patency of vascular access is checked by a saline test bolus. Contrast dose is usually 1.5 mL/kg body weight, with a limit on maximum contrast volume (120 cc 300 mg iodine/mL at our institution). Contrast is preferably administered with an automatic injector at rate 2–3 $ml\,s^{-1}$, followed by a saline flush. Two minutes after contrast injection, the patient is positioned for mammographic imaging. It is recommended to preserve the intravenous access until 15 min after contrast administration, so as to enable prompt treatment of any late adverse reactions to the contrast injection.

It is not necessary to acquire mammographic images in a specific order. Optimally, image acquisition should take place between 2 and 10 min after contrast administration, as all studies have confirmed adequate diagnostic accuracy within this time window. Fortunately, this is more than sufficient for acquiring the standard four mammography views as well as any supplemental views that may be called for. In both FFDM and CEM, exposure time depends on breast size and settings used and generally varies between 4 and 10 s/view.[14,15] Each CEM view consists of one low-energy (LE) and one high-energy (HE) image, the additional exposure time is in the order of seconds per acquisition,[14] and breast compression is released in between image acquisitions.

A standard CEM examination consists of a craniocaudal (CC) and a mediolateral oblique (MLO) view of each breast, with supplemental views (such as spot compression view or rolled views) as requested by the radiologist. Vendors have developed varying strategies for dual-energy mammography, using different anode materials, filter materials, and image reconstruction algorithms for combining LE and HE images. A detailed overview of vendor system characteristics has recently been published by Jochelson and Lobbes[9]; an updated overview is given in Table 1.

CEM makes use of the photoelectric effect of iodine which enables highlighting areas of contrast uptake. The photoelectric effect itself depends on the energy of the X-ray beam and k-edge of the material. The absorption k-edge of iodine (33 keV) falls within the average range of the X-ray beam in mammography. Furthermore, iodine X-ray absorption, or mass attenuation coefficient, is higher than that of breast tissue (Figure 1).

During image acquisition, first the LE image is acquired using tube voltages varying between 26 and 30 kVp.[15–19] Even although iodinated contrast is already present within the breast at this point, the LE mean energy falls below the k-edge of iodine and, as several studies show, LE is equivalent to FFDM in terms of image quality.[20–22]

The HE image is acquired second. In HE image-acquisition, the X-ray beam ranges from 44 to 49 keV. A photoelectric effect occurs when an incoming 44–49 keV photon causes an electron from the k-shell of an iodine atom to eject, thereby increasing the attenuation of iodine. Because iodine contrast has 'leaked' into the tumour interstitium, the latter will be enhanced and the difference between tumour and breast tissue becomes more apparent.[23]

Although the HE image contains relevant information, this cannot be perceived by the human eye. The information is instead used in post-processing to construct the so-called recombined or iodine image showing areas of contrast uptake. The end-result of the imaging process is LE and recombined CEM images from both breasts in two standard views (see example in Figure 2; an overview of a standard image-acquisition protocol is presented in Figure 3).

Reading, interpretation and reporting CEM images

CEM learning curve

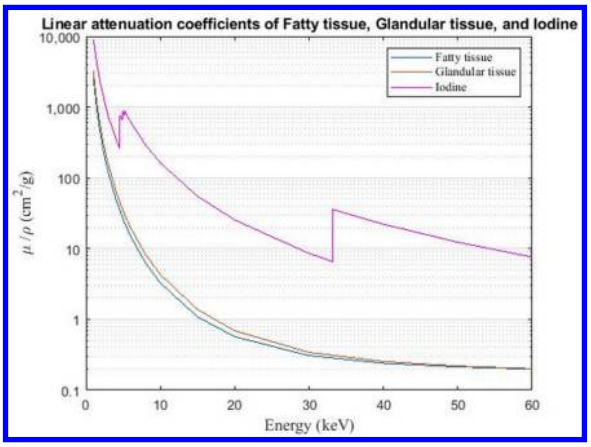
CEM is easy to learn, especially when readers have some experience with FFDM and MRI. This is supported by the results of the multi -reader study by Lalji et al,[24] in which seven radiologists and three residents assessed 199 cases (first LE images, followed by the complete CEM examination). Three levels of experience were distinguished: residents with marginal experience in CEM/FFDM; radiologists with at least two years' experience in CEM/FFDM; and radiologists with extensive experience in FFDM but none in CEM. Specificity and diagnostic performance increased significantly with CEM compared to FFDM regardless of level of experience. CEM sensitivity scores achieved by the residents (96.6%) and non-experienced CEM readers (95.9%) were similar to those of experienced readers (97.6%). These results suggest that novice CEM readers can reach a level equal to that of experienced radiologists.[24] This is supported by another study in which non-experienced high-school students, after a short introduction to breast cancer and CEM in general, evaluated the cases used in the study by Lalji et al. These students immediately reached a sensitivity of more than 80% in detecting breast cancers on recombined images.[25] This also implies that

Table 1. System characteristics of the five commercially available CEM systems Updated, from Jochelson and Lobbes[9].

	GE Healthcare Senographe Essential and Senobright	GE Healthcare Pristina and Senobright HD	Hologic Selenia Dimensions and 3Dimensoins I-View	Siemens Healthineers Mammomat Revelation Titanium CEM	Fujifilm Amulet Innovality CEDM
Low-energy acquisition					
Anode and filter material	Mo & Mo; Mo & Rh; Rh & Rh	Mo & Mo; Rh & Ag	W & Rh; W & Ag	W & Rh	W & Rh
Filter thickness (mm)	Mo, 0.03; Rh, 0.025	Mo, 0.03; Ag, 0.03	0.050	0.050	0.050
Tube voltage range (kV)	26–31	26–34	25–33	28–34	26–31
High-energy acquisition					
Anode and filter material	Mo & Al + Cu; Rh & Al + Cu	Mo & Cu; Rh & Cu	W & Cu	W & Ti	W & Al + Cu
Filter thickness (mm)	Al, 0.3; Cu, 0.3	0.25	0.3	1.0	Al, 0.7; Cu, 0.25
Tube voltage range (kV)	45–49	49	45–49	49	45–49
Complete CEM examination					
Mean glandular dose (mGy)	1.6–2.8	0.7–2.3	3.0	1.7	1.4
Total acquisition time (sec)	3–8	2–5	<2	15–22	5

Ag, silver; Al, aluminum; CEDM, contrast-enhanced digital mammography; Cu, copper; Mo, molybdenum; Rh, rhodium; Ti, titanium; W, tungsten.

Figure 1. Principle of iodine-based contrast enhancement. Mass attenuation coefficients of fatty tissue, glandular tissue, and iodine are shown. The iodine curve shows a steep elevation in attenuation at 33.2 keV, which is the k-edge of iodine. Differences in attenuation between breast tissue and iodinated contrast material are larger beyond the k-edge of iodine. Thus in high energy images (44–49 kVp), the differences in attenuation are larger than in low-energy images (26–30 kVp). Image processing of low- and high-energy images subsequently results in recombined images, showing contrast enhancement overlay.



semi-automatic software tools that are being developed might show steep learning curves (see 'Future developments').

It is not easy to determine how many CEM examinations must be read in order to be considered an experienced reader.[26] To the best of our knowledge, the only available study covering this specific topic is the one by Cheung et al, showing that a radiologist should read an average of 75 CEM examinations to reach a 90% probability of correct prediction.[27] Based on the above observations and the wide availability of CEM examinations, it is safe to assume that a minimum of 75 cases should be practised to acquire sufficient experience in clinical practice.

Hanging protocol

In practice, LE images are interpreted first to assess morphologic abnormalities, the recombined image being used for extra information.[24,27] This is the 'standard' hanging protocol proposed by the different vendors. However, alternative hanging protocols are feasible. To illustrate this, Van Geel et al compared CEM diagnostic accuracy using the 'standard' hanging protocol and an inverse hanging protocol (*i.e.,* first interpret the recombined image, followed by the LE image).[28] They found that sensitivity and specificity were equivalent between standard and inverse protocols, 98 and 99 versus 94% and 90%, respectively, but that the inverse hanging protocol led to an average decrease in reading time of 6.2 s/case. This was mainly due to shorter LE image evaluation in the inverse hanging protocol, average recombined image evaluation time remained similar.[28] Although time differences are small, they may become of interest in situations where large volumes of CEM examinations must be read, as is the case in screening programmes.

Figure 2. Contrast-enhanced mammographic images in a 67-year-old female recalled from the breast cancer screening program because of a new, spiculated mass in the right breast. A-D. Low energy images. E-H. Recombined images. Images were acquired of the right and left breast in craniocaudal (CC) and mediolateral oblique (MLO) views. The mass in the right breast is visible on low-energy images in both CC and MLO views (arrows in A and C). The recombined images of the right breast show enhancement of the lesion in both CC and MLO views (arrows in E and G). Histopathological results showed an invasive breast cancer of no special type, Grade 2, size 1.4 cm.

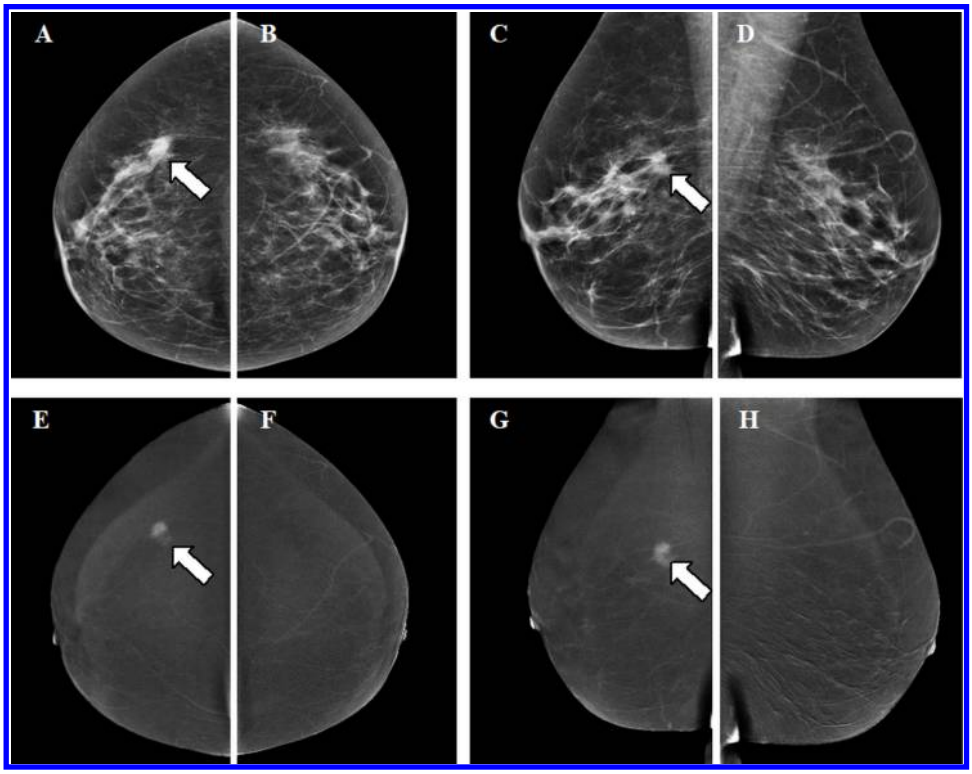
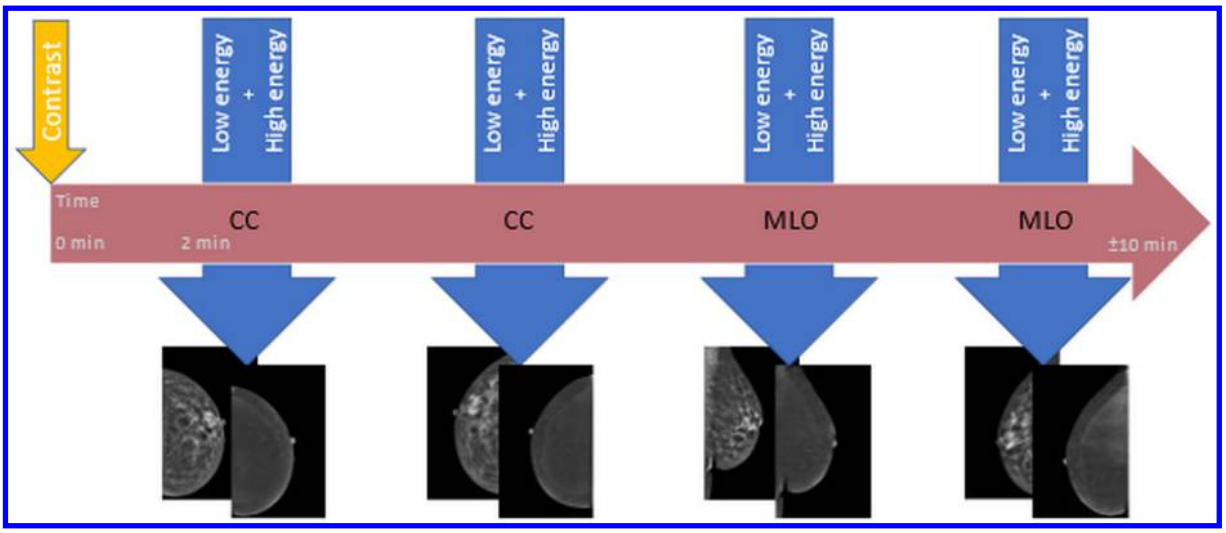


Figure 3. Diagram of image acquisition for contrast-enhanced mammography. The horizontal arrow represents the time window of 10 min in which a full (at least four views) contrast-enhanced mammography examination must be performed in order to be considered of diagnostic value. The iodine-based contrast agent is administered at time point zero (small vertical arrow), 2 min prior to the acquisition of the first view. Per view, one low energy and one high energy image are acquired within one compression (larger vertical arrows). The order of views may differ. After image processing, low energy and recombined images are retrieved for clinical assessment.



CEM artefacts

CEM can show artefacts, either related to the LE image or specific to the technique itself. Artefacts seen on the LE image are similar to those observed in FFDM and include air trapping, antiperspirant on the skin mimicking (micro)calcifications, and disruption of the X-ray beam by matter such as hair.[29,30] In general, these artefacts are well known and can be easily resolved by repeating image acquisition.

Some artefacts are specific to CEM and visible on the recombined image. An overview of these artefacts, their causes, and potential solutions, is provided in Table 2 (for artefact illustrations see Figures 4–6).[9,29–37]

Interpretation pitfalls

It is important to note that some lesions, such as invasive lobular carcinomas and mucinous carcinomas, are more difficult to detect using CEM. Van Nijnatten et al showed that invasive lobular carcinomas often show weak enhancement. On LE images such lesions appear as architectural distortions or asymmetries (instead of masses), rendering them difficult to spot on either type of CEM image.[38] Mucinous carcinomas contain large amounts of fluid and only limited numbers of vital tumour cells, and thus have limited blood supply (perfusion). As such, they only enhance slightly, or show rim enhancement, or sometimes show no enhancement at all.[24] Hence, the absence of enhancement in morphologically suspicious lesions cannot rule out breast cancer, and the CEM recombined image must therefore be seen as an adjunct to mammography rather than a replacement. Besides these two tumour types, cancers can be inherently difficult to detect due to their location in the mammographic field-of-view. CEM being a mammographic technique, some lesions may be overlooked in mammography blind spots, such as the medial part of the breast, the inframammary fold, the prepectoral zone, and the axillary tail.[24,39] Lesions in these areas are difficult to visualize in both FFDM and CEM, despite optimal breast positioning. If lesions are (partially) observed or suspected in these areas, breast MRI should be considered.

On the other hand, benign lesions can show enhancement on CEM, potentially resulting in false-positive findings. Common benign causes of enhancement are: fibroadenomas (Figure 4), atypical ductal hyperplasia, papilloma, infection or inflammation and radial scars.[24,40,41] Of the 128 benign lesions examined by Tsigginou et al, 37 showed enhancement on CEM (28.9%).[40] A similar percentage of enhanced benign lesions was seen by Deng et al. (12/44), and results suggest that the probability of a malignancy increases with stronger enhancement.[42] Although false-positive findings may lead to unnecessary biopsies or follow-up examinations, studies have shown that they occur less frequently in CEM than in FFDM.

Structural reporting of CEM examinations

LE images, being equivalent to FFDM, can be interpreted using the terminology suggested in the latest edition of the ACR BI-RADS lexicon.[43–45] To some extent, recombined images are comparable to standard MRI examinations, and therefore the use of standard MRI terminology is recommended when describing enhancement of lesions. For example, masses may be homogeneously or heterogeneously enhanced, or may show (irregular) rim enhancement. If no masses are observed, but instead architectural distortion or asymmetry is seen, the term 'non-mass enhancement' can be used in CEM reports, and the different characteristics described accordingly. However, some artefacts are specific to CEM and have acquired specific descriptions. For example, negative enhancement with or without a thin rim of enhancement also known as an 'eclipse sign', is the specific appearance of a cyst on CEM (Figure 5).[40,46,47] In addition, there are artefacts specific to recombined CEM images.

The amount of background parenchymal enhancement (BPE) in CEM can also be described as minimal, mild, moderate or marked enhancement, using terminology similar to that of MRI.[48,49] An increase in BPE is associated with increased odds for breast cancer.[48,49] The majority of patients showed to have minimal-to-mild BPE on CEM.[48,49] In a study by Sogani et al, three experienced breast imaging readers compared BPE levels between CEM and MRI showing agreements on BPE levels varying from moderate to substantial with $\kappa = 0.55$; $\kappa = 0.66$, and $\kappa = 0.67$.[49] Hence, interference of BPE is more or less comparable between the assessment of CEM and MRI.

At present, CEM is being considered for the ACR BI-RADS lexicon, and a comprehensive overview of structural reporting in CEM is expected to be available soon. Until that time, the recommendation is to keep LE and recombined image findings separate in the report, matching them where necessary, and to base the final BI-RADS classification on the complete CEM examination.[27]

Indications for CEM

The three most common indications for CEM are inconclusive findings, pre-operative staging, and response monitoring. Evidence of CEM efficacy in these settings, however, is mainly based on retrospective studies,[7] and proposed indications should be considered with this in mind. Current prospective trials such as the RACER and CMIST are ongoing and will provide scientific evidence for these indications.[50,51]

Inconclusive findings

One of the most studied aspects of CEM is its ability to act as 'problem solving' tool in the setting of inconclusive findings in conventional imaging, foremost a recall from the breast cancer screening programme. Despite low disease prevalence, CEM was shown to increase sensitivity, specificity, positive-predictive value (PPV), and negative-predictive value (NPV) in this population.[24,46]

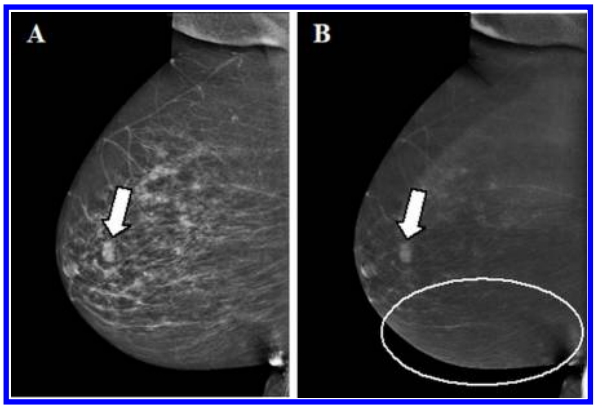
A feasibility study by Zuley et al suggests that CEM significantly reduces the false-positive rate (FPR) ($p = 0.017$) and significantly increases the true-positive rate (TPR) ($p = 0.019$) in BI-RADS 4 soft tissue lesions compared to FFDM/DBT.[52] Even in combination with ultrasound, the TPR of FFDM/DBT did not match that of CEM whilst the FPR significantly increased. Based on these results, CEM is likely to be more accurate than a FFDM/DBT/US combination. Moreover, supplemental US after negative CEM findings is questionable: the risk of finding false-positive lesions

Table 2. Overview of CEM specific artefacts

Artefacts	Cause	Appearance on recombined image	Solution
Ripple artefact	Slight motion of the breast between the LE and HE image acquisition. More often seen in increasing breast thickness and mainly in MLO view.	Thin black and white lines in a ripple-like structure (see circles in Figures 4 and 5).	Reduce movement of the patient during acquisition through patient instruction.
Rim artefact or "breast-in-breast"	Scattered radiation. In older systems: misalignment of the LE and HE anti scatter grids.	Double-breast contour in the form of a brighter breast edge-shape mimicking a "breast-in-breast" (see small arrows in Figure 6). This can be visible in both CC and MLO views.	Most commonly observed on CEM exams performed on first generation systems, less or not applicable in newer systems.
Skin line enhancement	Image filtration to equalize breast thickness.	The skin contour is partially highlighted (see larger arrow in Figure 6).	Most commonly observed on CEM exams performed on first generation systems, less or not applicable in newer systems. Can be easily dismissed if absence of any skin-related problems on the LE image.
Breast implants	Image distortion.	Poor recombined image quality with black or white areas surrounding implant.	Use other imaging modalities, such as MRI.
Axillary line artefact	Wrong usage of the small compression paddle.	Horizontal lines in the axillary region on the recombined image.	Use correct paddle size for large breasts.
Contrast splatter	Contamination of contrast on the skin (not a true technical artefact).	May mimic small (micro)calcifications on the LE image.	Prevent contamination by wearing gloves during contrast administration and/or washing hands before patient positioning. Distinction is easy: in contrast splatter, the corresponding lesion is extremely bright, whereas calcifications are black on the recombined image.
Skin lesions	Skin lesions such as haemangiomas showing enhancement, superimposed within the boundaries of the breast.	Mimics an intramammary enhancing lesion.	Check for noticeable skin lesions which might correspond to an enhancing lesion during breast positioning.

CC, cranio-caudal; CEM, contrast-enhanced mammography; HE, high-energy; LE, low-energy; MLO, mediolateral oblique; MRI, magnetic resonance imaging.

Figure 4. Enhancing fibroadenoma. A,B. Contrast-enhanced mammographic of right breast in mediolateral oblique view in a 63-year-old female recalled from screening because of a new ill-defined and partly obscured mass. A. Low-energy image showing the suspect mass (arrow in A). B. Corresponding recombined image in which the suspect lesion is showing enhancement (arrow in B). The lines visible in the caudal part of the breast (circle) are the result of slight motion between the low- and high-energy image acquisition, the ripple artefact. Histopathological results showed a classic fibroadenoma.



is increased without any real improvement in terms of cancer detection.[52]

The benefit of CEM in assessing suspicious breast calcifications is not as clear. A prospective study by Cheung et al in patients with screening recalls for suspicious microcalcifications found 88.9% sensitivity and 86.6% specificity.[53] In a similar study, Houben et al found a slight increase in diagnostic accuracy, with only 81.1% of ductal carcinoma *in situ* (DCIS) showing enhancement, but it was not sufficient to be of added value for clinical use in surgical treatment planning.[54] Considering these findings, it is currently not recommended to downgrade unenhanced calcifications to a lower BI-RADS classification. On the other hand, enhancement

Figure 5. Contrast-enhanced mammographic images in a 55-year-old female recalled from screening because of a new mass in the left breast. A. Low-energy image in mediolateral oblique view shows an ill-defined round mass (arrow in A). B. Corresponding recombined image. At the site of the suspect lesion a subtle 'eclipse sign' is visible, implicating a cyst (arrow in B). No screen-detected interval breast cancer has been reported in the 18-month follow-up period. The ripple artefact is also visible on the recombined image (circle).

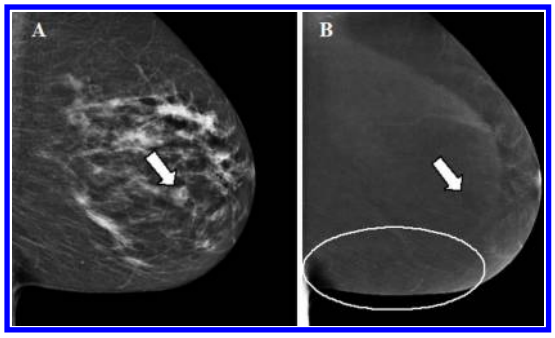
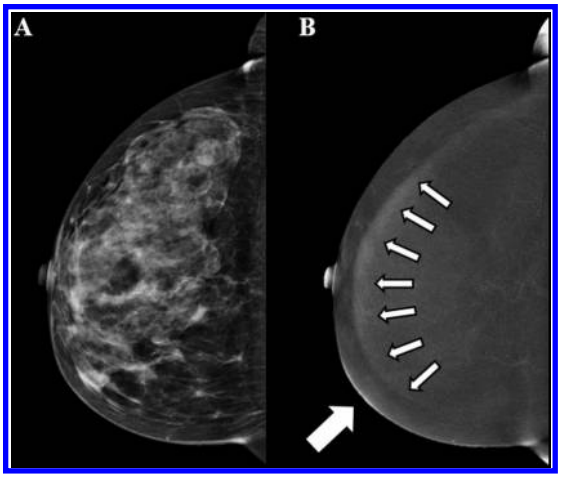


Figure 6. Contrast-enhanced mammographic images in craniocaudal view in a 63-year-old female. A. Low-energy image of the right breast. B. The rim artefact is shown in the recombined image (small arrows). In addition, the skin line enhancement artefact is visible in the anteromedial part of the breast (larger arrow). No suspicious findings were reported.



of calcifications may be sufficient grounds for an upgrade of the BI-RADS classification, but biopsy remains necessary.

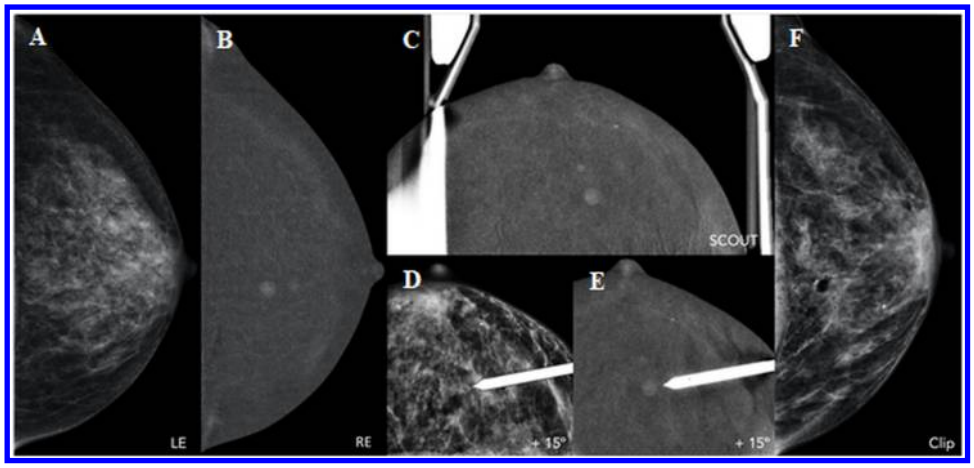
For patients with contraindications for MRI (claustrophobia, pacemaker, metallic implant), CEM is a good alternative; diagnostic performance appears to be comparable.[55–57] In a recent review by Xiang et al, pooled sensitivity was found to be 97% for both CEM and MRI, whereas accuracy and pooled specificity were higher for CEM: 98 and 66 versus 92% and 52%, respectively.[57] These pooled results may not be applicable to specific study populations. In a prospective study with BI-RADS 3–5 lesions comparing diagnostic performance of multiple breast imaging modalities including CEM, the best diagnostic performance was achieved using MRI.[58] Nevertheless, CEM performance makes it quite an acceptable alternative to breast MRI when the latter is not preferred or unavailable. However, relative strengths and weaknesses of each modality need to be investigated in more detail: in specific subpopulations and for diagnostic accuracy certainly, but also regarding cost-efficiency.

Pre-operative staging

Breast MRI is currently the reference standard for assessing tumour extent and presence of additional foci.[59] CEM has been evaluated as a tool for pre-operative staging and may provide a good alternative for MRI. CEM tends to slightly overestimate tumour size (in the order of mms'), while FFDM/LE and ultrasound tend to underestimate tumour size, compared to histological size.[55,60–62] Size measurements using CEM are comparable to those using MRI, and both are in concordance with or slightly overestimated compared with histological size.[58,59,61,63–65]

A single-centre retrospective study in the setting of preoperative breast staging (n = 326) found 93% sensitivity and 98% specificity for CEM. Furthermore, CEM led to a change in surgery type compared to conventional imaging in 18.4% of patients.[66] It is mostly symptomatic patients with palpable lesions who benefit

Figure 7. CEM-guided biopsy in a 61-year-old patient with palpable lesion in right breast (IDC, not shown) and additional contralateral (left breast) finding on diagnostic CEM. A,B. Low-energy (LE) and recombined image (RE) of left breast in craniocaudal view. There is a 6 mm mass enhancement at 12 o'clock, with no ultrasound correlation and not enough references on 2D/3D in order to favour a conventional mammographic-guided biopsy. C-E. The procedure of CEM-guided biopsy is similar to a standard stereotactic biopsy (one scout and a pair of angled stereotactic images) with the additional step of contrast media injection 2 min before compression and first imaging. Like a routine CEM, each acquisition is composed of one low-energy (LE) and one high-energy (HE) exposure. The inclusion of the enhancing lesion is confirmed with a recombined scout view (0°), followed by the two angled views. Another pair of stereotactic angled images (±15°) is sometimes acquired, previous to fire-forward, in order to confirm that the target was reached. F. Final CC view after clip placement. Histopathological results showed an invasive lobular carcinoma in the left breast.



from staging with CEM. In a study with 101 CEM-detected lesions, CEM led to 17 additional imaging and 12 additional biopsies, and the surgical treatment plan was changed for 20 patients.[39]

Response monitoring

Neoadjuvant chemotherapy (NAC) is increasingly used to treat locally advanced breast carcinomas. The aim of NAC is to reduce tumour size, thereby decreasing the need for mastectomy and/or lymph node dissection. In response monitoring, the tumour is usually assessed before, during and after treatment. Response to NAC is reflected in a decrease in tumour size as well as in changes in lesion enhancement. Before CEM, MRI was the most accurate imaging modality for tumour extent measurements and residual tumour evaluation.[67–69] However, initial results of studies on CEM in response monitoring are encouraging.

In a study by Iotti et al, 46 patients underwent both MRI and CEM before, during and after treatment. CEM better predicted the pathological complete response than MRI (Lin's coefficient 0.81 and 0.59, respectively); both imaging modalities underestimated the size of residual tumours, 4.1 mm on average for CEM and 7.5 mm on average for MRI.[59] In a similar study among 33 patients by Barra et al, Lin's coefficients of 0.7 and 0.4 were found, and residual tumour size was overestimated with an average of 8.0 mm for CEM and 18.0 mm for MRI.[70] Both studies suggest CEM to be more accurate than MRI in residual tumour evaluation.[59,70] A first systematic review and meta-analysis of CEM and MRI in response monitoring was recently published, including 6 CEM and 21 MRI studies. Pooled sensitivity for CEM was higher than that of MRI, 83vs 77%, whereas pooled specificities were equal, 82vs 82%.[71] Available data are limited, but so far CEM appears to be a good alternative to MRI in response monitoring.

CEM-guided biopsy

CEM-guided biopsy was developed to access enhancing lesions not seen on accompanying LE images or targeted US. It is a promising alternative to MRI-guided biopsy. The technique may be used to guide various interventional procedures of the breast, such as vacuum-assisted biopsy or excision (VAB or VAE), core needle biopsy, and pre-surgical wire localization. CEM-guided biopsy is based on the principle of (conventional) stereotactic procedures, using dual energy acquisition and i.v. administration of iodinated contrast media. Image acquisition is performed in a similar way to diagnostic CEM, including the 2-min wait after contrast administration (Figure 7). Inclusion of the enhancing lesion is confirmed with a recombined scout view (0°), after which a pair of dual-energy angled stereotactic images (±15°) is acquired with the objective indicated in each. Thus, the equipment automatically calculates the X, Y and Z coordinates allowing access to the target. Generally, enhancement will be visible for at least 5 to 7 min which is sufficient for target selection. After local anaesthesia, a needle is inserted into the breast until the limit point is reached, as defined by the support. Another pair of stereotactic angled images is sometimes acquired before the fire-forward to confirm that the objective was reached, or to redefine coordinates if it was not. Next, sampling is carried out with the vacuum system device. We recommend to extract at least 12 tissue samples in order to reduce sampling error. Lastly, it is crucial to mark the biopsy bed with a radiological marker, ideally using the same probe.

Disadvantages of CEM

CEM has two important disadvantages: the use of iodinated contrast agents and an increase in radiation dose. Potential benefits of CEM should always be weighed against these disadvantages.

Table 3. Overview of studies comparing radiation dose of CEM to FFDM on FDA approved, commercially available systems

	CEM					FFDM					Difference
Article	Patients (n)	Images (n) and VIEWS (n/n)	Mean breast thickness (mm) + (range)	Mean AGD (mGy) + (range)	System (vendor)	Patients (n)	Images (n) and VIEWS (n/n)	Mean breast thickness (mm) + (range)	Mean AGD (mGy) + (range)	System (vendor)	Dose ratio CEM / FFDM (%)
Badr *et al.* (2014)	104	391 CC/MLO (N/A)	56 (N/A)	2.65 (N/A)	Not reported	104	360 CC/MLO (N/A)	57 (N/A)	1.72 (N/A)	Not reported	54%
Jeukens *et al.* (2014)	47	195 CC/MLO (96/97)	58 (21-96)	2.80 (1.10–4.29)	Senographe Essential +SenoBright (GE Healthcare)	715	2782 CC/MLO (1238/1339)	56 (15-100)	1.55 (0.63–5.12)	Senographe Essential +SenoBright (GE Healthcare)	81%
James *et al* (2017)	173	174 single CC	63 (N/A)	3.0 (N/A)	Selenia Dimensions (Hologic)	6214	6215 single CC	47 (N/A)	1.8 (N/A	Selenia Dimensions (Hologic)	42%
Phillips *et al.* (2018)	45	180 CC/MLO (90/90)	56 (22-88)	2.49 (N/A)	Senographe Essential +SenoBright (GE Healthcare)	45	180 CC/MLO (90/90)	56 (N/A)	1. 1.40 (N/A) 2. 2.16 (N/A)	1. Senographe Essential (GE Healthcare) 2. Selenia Dimensions (Hologic)	1. 78% 2. 15%

AGD, average glandular dose; CC, cranio-caudal; CEM, contrast-enhanced mammography; FFDM, full-field digital mammography; MLO, mediolateral oblique.

Iodinated contrast material

Although the use of modern iodine-based contrast materials is considered safe, a possibility of mild, moderate or severe anaphylactoid reactions remains.[72] In a systematic review, Zanardo et al found the pooled rate of adverse reactions in CEM examinations to be 0.82%.[7] At our institution, we observed a 0.6% rate of adverse reactions in CEM examinations.[41] However, subjects with prior hypersensitivity reactions to any of the ingredients of iodinated contrast should be excluded from undergoing CEM, since breast MRI could be considered a safer alternative.

In addition to hypersensitivity reactions, iodinated contrast administration may cause post-contrast acute kidney injury. Patients at risk of acute kidney injury, such as those with renal insufficiency, incur a risk when undergoing CEM.[73] Since breast imaging never involves 'do or die' scenarios, alternative methods for diagnostics such as breast MRI should be used in such cases, in accordance with safety guidelines on the use of iodinated contrast material.[73,74]

Radiation dose

The first study on CEM radiation dose performed on a commercially available system approved by the U.S. Food and Drug Administration (FDA) (as opposed to a prototype or modified unit) was performed by Badr et al. They found a 54% higher mean radiation dose for CEM (2.65 mGy) than for FFDM (1.72 mGy).[75] Three other studies comparing CEM and FFDM radiation dose similarly found higher doses for CEM.[14,76,77]

Studies thus consistently find an increase in radiation dose for CEM, but the magnitude differs. This is presumably due to variation in system settings and different patient characteristics, projection views and breast thickness for example, may influence results. An overview of the various study characteristics is presented in Table 3.[14,75–77]

Although increased, CEM radiation dose remains within safe radiation dose limits according to the Mammography Quality Standards Act regulations (3.0 mGy per view).[78] The life-attributable risk (LAR) number for cancer incidence incurred by a complete CEM exam with four acquisitions at the age of 40 is 0.009% , and the LAR for cancer mortality is even lower, at 0.002%. These percentages drastically decrease with increasing age.[41,79] Nevertheless, the As Low As Reasonably Achievable (ALARA) principle is also applicable to CEM, meaning that risks should always be weighed against benefits.

Future developments

Continual technical developments are being explored to further advance CEM. These not only include technical hardware improvements but also advances in the post-processing algorithm, which may help to reduce CEM artefacts and improve image quality in general.

Enhancement plays an important role in the evaluation of CEM examinations, and there seems to be diagnostic information encompassed in the amount of enhancement than can be observed. For example, Lobbes et al found that grey values on recombined images were significantly higher for malignant

lesions than for benign lesions ($p = 0.002$) or cysts ($p < 0.001$).[80] Unfortunately, such differences cannot be accurately assessed through visual inspection and grey values of enhancement are difficult to quantify. Herein lies an opportunity for the use of artificial intelligence and radiomics.

Indeed, machine-learning algorithms with textural and morphological features are already able to distinguish benign lesions from malignancies with an accuracy of 90% (45/50).[81] Moreover, initial results from Marino et al reveal radiomics accuracies of 78 to 100% in differentiating between malignant lesions based on several tumour characteristics, such as (non-) invasiveness, three hormone receptor sensitivities (positive or negative), and tumour grade (Grades 1–3).[82] Finally, Wang et al created a radiomics monogram using 14 radiomics features and risk factors, and achieved an accuracy of 81% in predicting tumour response to NAC using CEM.[83] Currently, ongoing studies use deep learning algorithms to detect breast lesions on CEM and radiomics to subsequently classify them. The introduction of machine learning-based decision support tools for CEM appears to be only a matter of time.

CONCLUSION

Since its commercial introduction in 2011, CEM has been steadily incorporated as imaging tool in clinical practice. CEM is surprisingly easy to learn and confers logistic and diagnostic advantages over breast MRI. However, it is a relatively novel addition and future studies will certainly elaborate on its strengths and weaknesses, not only in terms of specific populations and diagnostic accuracy, but also in cost-effectiveness.

ACKNOWLEDGMENT

We wish to thank Estelle Nijssen for her assistance in editing the final manuscript.

CONFLICTS OF INTEREST

RA received institutional grant and consulting fee from GE Healthcare.
JW received institutional grants and speaker's fees from AGFA, Bayer Healthcare, Bard Medical, GE Healthcare, Optimed, Philips Healthcare, and Siemens Healthineers.
ML received several research grant and speaker's fees from GE Healthcare, Hologic, Bayer, and Guerbet.
The other authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

REFERENCES

1. Dibden A, Offman J, Duffy SW, Gabe R. Worldwide review and meta-analysis of cohort studies measuring the effect of mammography screening programmes on incidence-based breast cancer mortality. *Cancers* 2020; **12**: 976. doi: https://doi.org/10.3390/cancers12040976
2. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med* 2007; **356**: 227–36. doi: https://doi.org/10.1056/NEJMoa062790
3. Nazari SS, Mukherjee P. An overview of mammographic density and its association with breast cancer. *Breast Cancer* 2018; **25**: 259–67. doi: https://doi.org/10.1007/s12282-018-0857-5
4. Tagliafico AS, Bignotti B, Rossi F, Signori A, Sormani MP, Valdora F, et al. Diagnostic performance of contrast-enhanced spectral mammography: systematic review and meta-analysis. *Breast* 2016; **28**: 13–19. doi: https://doi.org/10.1016/j.breast.2016.04.008
5. Zhu X, Huang J-M, Zhang K, Xia L-J, Feng L, Yang P, et al. Diagnostic value of contrast-enhanced spectral mammography for screening breast cancer: systematic review and meta-analysis. *Clin Breast Cancer* 2018; **18**: e985–95. doi: https://doi.org/10.1016/j.clbc.2018.06.003
6. Suter MB, Pesapane F, Agazzi GM, Gagliardi T, Nigro O, Bozzini A. Diagnostic accuracy of contrast-enhanced spectral mammography for breast lesions : a systematic review and meta-analysis. *The Breast Elsevier Ltd* 2020; **53**: 8–17.
7. Zanardo M, Cozzi A, Trimboli RM, Labaj O, Monti CB, Schiaffino S, et al. Technique, protocols and adverse reactions for contrast-enhanced spectral mammography (CESM): a systematic review. *Insights Imaging* 2019; **10**: 76. doi: https://doi.org/10.1186/s13244-019-0756-0
8. Healthcare GE. GE Healthcare announces FDA 510(k) clearance of SenoBright Contrast Enhanced Spectral Mammography (CESM) for breast cancer diagnosis [Internet]. 2011. Available from: https://www.ge.com/news/press-releases/ge-healthcare-announces-fda-510k-clearance-senobright-contrast-enhanced-spectral [cited 2021 Feb 2].
9. Jochelson MS, Lobbes MBI. Contrast-Enhanced mammography: state of the art. *Radiology* 2021; **299**: 36–48. doi: https://doi.org/10.1148/radiol.2021201948
10. Cuenod CA, Fournier L, Balvay D, Guinebretière J-M. Tumor angiogenesis: pathophysiology and implications for contrast-enhanced MRI and CT assessment. *Abdom Imaging* 2006; **31**: 188–93. doi: https://doi.org/10.1007/s00261-005-0386-5
11. Berns EA, Hendrick RE, Solari M, Barke L, Reddy D, Wolfman J, et al. Digital and screen-film mammography: comparison of image acquisition and interpretation times. *AJR Am J Roentgenol* 2006; **187**: 38–41. doi: https://doi.org/10.2214/AJR.05.1397
12. Dromain C, Balleyguier C, Muller S, Mathieu M-C, Rochard F, Opolon P, et al. Evaluation of tumor angiogenesis of breast carcinoma using contrast-enhanced digital mammography. *AJR Am J Roentgenol* 2006; **187**: W528–37. doi: https://doi.org/10.2214/AJR.05.1944
13. Bernardi D, Ciatto S, Pellegrini M, Anesi V, Burlon S, Cauli E, et al. Application of breast tomosynthesis in screening: incremental effect on mammography acquisition and reading time. *Br J Radiol* 2012; **85**: e1174–8. doi: https://doi.org/10.1259/bjr/19385909
14. Jeukens CRLPN, Lalji UC, Meijer E, Bakija B, Theunissen R, Wildberger JE, et al. Radiation exposure of contrast-enhanced spectral mammography compared with full-field digital mammography. *Invest Radiol* 2014; **49**: 659–65. doi: https://doi.org/10.1097/RLI.0000000000000068
15. Daniaux M, De Zordo T, Santner W, Amort B, Koppelstätter F, Jaschke W, et al.

Dual-Energy contrast-enhanced spectral mammography (CESM. *Arch Gynecol Obstet* 2015; **292**: 739–47. doi: https://doi.org/10.1007/s00404-015-3693-2
16. Skarpathiotakis M, Yaffe MJ, Bloomquist AK, Rico D, Muller S, Rick A, et al. Development of contrast digital mammography. *Med Phys* 2002; **29**: 2419–26. doi: https://doi.org/10.1118/1.1510128
17. Jong RA, Yaffe MJ, Skarpathiotakis M, Shumak RS, Danjoux NM, Gunesekara A, et al. Contrast-Enhanced digital mammography: initial clinical experience. *Radiology* 2003; **228**: 842–50. doi: https://doi.org/10.1148/radiol.2283020961
18. Lewin JM, Isaacs PK, Vance V, Larke FJ. Dual-Energy contrast-enhanced digital subtraction mammography: feasibility. *Radiology* 2003; **229**: 261–8. doi: https://doi.org/10.1148/radiol.2291021276
19. Jochelson MS, Dershaw DD, Sung JS, Heerdt AS, Thornton C, Moskowitz CS, et al. Bilateral contrast-enhanced dual-energy digital mammography: feasibility and comparison with conventional digital mammography and MR imaging in women with known breast carcinoma. *Radiology* 2013; **266**: 743–51. doi: https://doi.org/10.1148/radiol.12121084
20. Francescone MA, Jochelson MS, Dershaw DD, Sung JS, Hughes MC, Zheng J, et al. Low energy mammogram obtained in contrast-enhanced digital mammography (CEDM) is comparable to routine full-field digital mammography (FFDM. *Eur J Radiol* 2014; **83**: 1350–5. doi: https://doi.org/10.1016/j.ejrad.2014.05.015
21. Blum KS, Antoch G, Mohrmann S, Obenauer S. Use of low-energy contrast-enhanced spectral mammography (CESM) as diagnostic mammography-proof of concept. *Radiography* 2015; **21**: 352–8. doi: https://doi.org/10.1016/j.radi.2015.02.005
22. Lalji UC, Jeukens CRLPN, Houben I, Nelemans PJ, van Engen RE, van Wylick E, et al. Evaluation of low-energy contrast-enhanced spectral mammography images by comparing them to full-field digital mammography using EUREF image quality criteria. *Eur Radiol* 2015; **25**: 2813–20. doi: https://doi.org/10.1007/s00330-015-3695-2
23. Jeukens C. Physics of contrast-enhanced mammography. In: Lobbes M, Jochelson M, eds.*Contrast-enhanced Mammogr Cham*. Switzerland: Springer; 2019. pp.. 23–39.
24. Lalji UC, Houben IPL, Prevos R, Gommers S, van Goethem M, Vanwetswinkel S, et al. Contrast-Enhanced spectral mammography in recalls from the Dutch breast cancer screening program: validation of results in a large multireader, multicase study. *Eur Radiol* 2016; **26**: 4371–9. doi: https://doi.org/10.1007/s00330-016-4336-0
25. van Nijnatten TJA, Smidt ML, Goorts B, Samiei S, Houben I, Kok EM, et al. Can high school students help to improve breast radiologists in detecting missed breast cancer lesions on full-field digital mammography? *J Cancer* 2019; **10**: 765–71. doi: https://doi.org/10.7150/jca.30494
26. Berg WA, Bandos AI, Zuley ML, Waheed UX. Training radiologists to interpret contrast-enhanced mammography: toward a standardized lexicon. *J Breast Imaging* 2021; **3**: 176–89. doi: https://doi.org/10.1093/jbi/wbaa115
27. Cheung Y-C, Lin Y-C, Wan Y-L, Yeow K-M, Huang P-C, Lo Y-F, , et al. Diagnostic performance of dual-energy contrast-enhanced subtracted mammography in dense breasts compared to mammography alone: interobserver blind-reading analysis. *Eur Radiol* 2014; **24**: 2394–403. doi: https://doi.org/10.1007/s00330-014-3271-1
28. van Geel K, Kok EM, Krol JP, Houben IPL, Thibault FE, Pijnappel RM, et al. Reversal of the hanging protocol of contrast enhanced mammography leads to similar diagnostic performance yet decreased reading times. *Eur J Radiol* 2019; **117**: 62–8. doi: https://doi.org/10.1016/j.ejrad.2019.05.013
29. Bhimani C, Li L, Liao L, Roth RG, Tinney E, Germaine P. Contrast-Enhanced spectral mammography: modality-specific artifacts and other factors which may interfere with image quality. *Acad Radiol* 2017; **24**: 89–94. doi: https://doi.org/10.1016/j.acra.2016.08.024
30. Nori J, Gill MK, Vignoli C, Bicchierai G, De Benedetto D, Di Naro F, et al. Artefacts in contrast enhanced digital mammography: how can they affect diagnostic image quality and confuse clinical diagnosis? *Insights Imaging* 2020; **11**: 16. doi: https://doi.org/10.1186/s13244-019-0811-x
31. Dromain C, Thibault F, Muller S, Rimareix F, Delaloge S, Tardivon A, et al. Dual-Energy contrast-enhanced digital mammography: initial clinical results. *Eur Radiol* 2011; **21**: 565–74. doi: https://doi.org/10.1007/s00330-010-1944-y
32. Yagil Y, Shalmon A, Rundstein A, Servadio Y, Halshtok O, Gotlieb M, et al. Challenges in contrast-enhanced spectral mammography interpretation: artefacts lexicon. *Clin Radiol* 2016; **71**: 450–7. doi: https://doi.org/10.1016/j.crad.2016.01.012
33. Phillips J, Fein-Zachary VJ, Slanetz PJ. Pearls and pitfalls of contrast-enhanced mammography. *J Breast Imaging* 2019; **1**: 64–72. doi: https://doi.org/10.1093/jbi/wby013
34. Lobbes M. Interpretation of contrast-enhanced mammography. In: Lobbes M, Jochelson M, eds.*Contrast-enhanced Mammogr Cham*. Switzerland: Springer; 2019. pp.. 61–75.
35. Gluskin J, Click M, Fleischman R, Dromain C, Morris EA, Jochelson MS. Contamination artifact that mimics in-situ carcinoma on contrast-enhanced digital mammography. *Eur J Radiol* 2017; **95**: 147–54. doi: https://doi.org/10.1016/j.ejrad.2017.08.002
36. Perry H, Phillips J, Dialani V, Slanetz PJ, Fein-Zachary VJ, Karimova EJ, Mehta TS, et al. Contrast-Enhanced mammography: a systematic guide to interpretation and reporting. *AJR Am J Roentgenol* 2019; **212**: 222–31. doi: https://doi.org/10.2214/AJR.17.19265
37. Lewis TC, Patel BK, Pizzitola VJ. Navigating contrast-enhanced digital mammography. *Appl Radiol* 2017; **46**: 21–8.
38. van Nijnatten TJA, Jochelson MS, Pinker K, Keating DM, Sung JS, Morrow M, et al. Differences in degree of lesion enhancement on CEM between ILC and IDC. *BJR|Open* 2019; **1**: 20180046. doi: https://doi.org/10.1259/bjro.20180046
39. Ali-Mucheru M, Pockaj B, Patel B, Pizzitola V, Wasif N, Stucky C-C, et al. Contrast-Enhanced digital mammography in the surgical management of breast cancer. *Ann Surg Oncol* 2016; **23**(Suppl 5): 649–55. doi: https://doi.org/10.1245/s10434-016-5567-7
40. Tsigginou A, Gkali C, Chalazonitis A, Feida E, Vlachos DE, Zagouri F, et al. Adding the power of iodinated contrast media to the credibility of mammography in breast cancer diagnosis. *Br J Radiol* 2016; **89**: 1067. doi: https://doi.org/10.1259/bjr.20160397
41. Houben IPL, Van de Voorde P, Jeukens CRLPN, Wildberger JE, Kooreman LF, Smidt ML, et al. Contrast-Enhanced spectral mammography as work-up tool in patients recalled from breast cancer screening has low risks and might hold clinical benefits. *Eur J Radiol* 2017; **94**: 31–7. doi: https://doi.org/10.1016/j.ejrad.2017.07.004
42. Deng C-Y, Juan Y-H, Cheung Y-C, Lin Y-C, Lo Y-F, Lin G, et al. Quantitative analysis of enhanced malignant and benign lesions on contrast-enhanced spectral mammography. *Br J Radiol* 2018; **91**: 20170605. doi: https://doi.org/10.1259/bjr.20170605
43. D'Orsi CJ, Sickles E, Mendelson E, Morris E. ACR BI-RADS atlas, breast imaging reporting and data system. *Reston, VA, USA* 2013;.
44. Knogler T, Homolka P, Hoernig M, Leithner R, Langs G, Waitzbauer M, et al. Application of BI-RADS descriptors in contrast-enhanced dual-energy mammography:

comparison with MRI. *Breast Care* 2017; **12**: 212–6. doi: https://doi.org/10.1159/000478899

45. Travieso-Aja MDM, Naranjo-Santana P, Fernández-Ruiz C, Severino-Rondón W, Maldonado-Saluzzi D, Rodríguez Rodríguez M, et al. Factors affecting the precision of lesion sizing with contrast-enhanced spectral mammography. *Clin Radiol* 2018; **73**: 296–303. doi: https://doi.org/10.1016/j.crad.2017.10.017
46. Lobbes MBI, Lalji U, Houwers J, Nijssen EC, Nelemans PJ, van Roozendaal L, et al. Contrast-Enhanced spectral mammography in patients referred from the breast cancer screening programme. *Eur Radiol* 2014; **24**: 1668–76. doi: https://doi.org/10.1007/s00330-014-3154-5
47. Li L, Roth R, Germaine P, Ren S, Lee M, Hunter K, et al. Contrast-Enhanced spectral mammography (CESM) versus breast magnetic resonance imaging (MRI): a retrospective comparison in 66 breast lesions. *Diagn Interv Imaging* 2017; **98**: 113–23. doi: https://doi.org/10.1016/j.diii.2016.08.013
48. Sorin V, Yagil Y, Shalmon A, Gotlieb M, Faermann R, Halshtok-Neiman O, et al. Background parenchymal enhancement at contrast-enhanced spectral mammography (CESM) as a breast cancer risk factor. *Acad Radiol* 2020; **27**: 1234–40. doi: https://doi.org/10.1016/j.acra.2019.10.034
49. Sogani J, Morris EA, Kaplan JB, D'Alessio D, Goldman D, Moskowitz CS, et al. Comparison of background parenchymal enhancement at contrast-enhanced spectral mammography and breast MR imaging. *Radiology* 2017; **282**: 63–73. doi: https://doi.org/10.1148/radiol.2016160284
50. Neeter LMFH, Houben IPL, Nelemans PJ, Van Nijnatten TJA, Pijnappel RM, Frotscher C, et al. Rapid access to contrast-enhanced spectral mammogRaphy in women recalled from breast cancer screening: the Racer trial study design. *Trials* 2019; **20**: 759. doi: https://doi.org/10.1186/s13063-019-3867-5
51. American College of RadiologyContrast Enhanced Mammography Imaging Screening Trial (CMIST) [Internet]. 2021. Available from: https://www.acr.org/Research/Clinical-Research/CMIST [cited 2021 May 10].
52. Zuley ML, Bandos AI, Abrams GS, Ganott MA, Gizienski T-A, Hakim CM, et al. Contrast enhanced digital mammography (CEDM) helps to safely reduce benign breast biopsies for low to moderately suspicious soft tissue lesions. *Acad Radiol* 2020; **27**: 969–76. doi: https://doi.org/10.1016/j.acra.2019.07.020
53. Cheung Y-C, Juan Y-H, Lin Y-C, Lo Y-F, Tsai H-P, Ueng S-H, et al. Dual-Energy contrast-enhanced spectral mammography: enhancement analysis on BI-RADS 4 Non-Mass microcalcifications in screened women. *PLoS One* 2016; **11**: e0162740. doi: https://doi.org/10.1371/journal.pone.0162740
54. Houben IP, Vanwetswinkel S, Kalia V, Thywissen T, Nelemans PJ, Heuts EM, et al. Contrast-Enhanced spectral mammography in the evaluation of breast suspicious calcifications: diagnostic accuracy and impact on surgical management. *Acta Radiol* 2019; **60**: 1110–7. doi: https://doi.org/10.1177/0284185118822639
55. Richter V, Hatterman V, Preibsch H, Bahrs SD, Hahn M, Nikolaou K, et al. Contrast-Enhanced spectral mammography in patients with MRI contraindications. *Acta Radiol* 2018; **59**: 798–805. doi: https://doi.org/10.1177/0284185117735561
56. Xing D, Lv Y, Sun B, Xie H, Dong J, Hao C, et al. Diagnostic value of contrast-enhanced spectral mammography in comparison to magnetic resonance imaging in breast lesions. *J Comput Assist Tomogr* 2019; **43**: 245–51. doi: https://doi.org/10.1097/RCT.0000000000000832
57. Xiang W, Rao H, Zhou L. A meta-analysis of contrast-enhanced spectral mammography versus MRI in the diagnosis of breast cancer. *Thorac Cancer* 2020; **11**: 1423–32. doi: https://doi.org/10.1111/1759-7714.13400
58. Petrillo A, Fusco R, Vallone P, Filice S, Granata V, Petrosino T, et al. Digital breast tomosynthesis and contrast-enhanced dual-energy digital mammography alone and in combination compared to 2D digital synthetized mammography and MR imaging in breast cancer detection and classification. *Breast J* 2020; **26**: 860–72. doi: https://doi.org/10.1111/tbj.13739
59. Iotti V, Ravaioli S, Vacondio R, Coriani C, Caffarri S, Sghedoni R, et al. Contrast-Enhanced spectral mammography in neoadjuvant chemotherapy monitoring: a comparison with breast magnetic resonance imaging. *Breast Cancer Res* 2017; **19**: 106. doi: https://doi.org/10.1186/s13058-017-0899-1
60. Blum KS, Rubbert C, Mathys B, Antoch G, Mohrmann S, Obenauer S. Use of contrast-enhanced spectral mammography for intramammary cancer staging: preliminary results. *Acad Radiol* 2014; **21**: 1363–9. doi: https://doi.org/10.1016/j.acra.2014.06.012
61. Tennant SL, James JJ, Cornford EJ, Chen Y, Burrell HC, Hamilton LJ, et al. Contrast-Enhanced spectral mammography improves diagnostic accuracy in the symptomatic setting. *Clin Radiol* 2016; **71**: 1148–55. doi: https://doi.org/10.1016/j.crad.2016.05.009
62. Patel BK, Garza SA, Eversman S, Lopez-Alvarez Y, Kosiorek H, Pockaj BA. Assessing tumor extent on contrast-enhanced spectral mammography versus full-field digital mammography and ultrasound. *Clin Imaging* 2017; **46**: 78–84. doi: https://doi.org/10.1016/j.clinimag.2017.07.001
63. Lobbes MBI, Lalji UC, Nelemans PJ, Houben I, Smidt ML, Heuts E, et al. The quality of tumor size assessment by contrast-enhanced spectral mammography and the benefit of additional breast MRI. *J Cancer* 2015; **6**: 144–50. doi: https://doi.org/10.7150/jca.10705
64. Travieso-Aja MDM, Maldonado-Saluzzi D, Naranjo-Santana P, Fernández-Ruiz C, Severino-Rondón W, Rodríguez Rodríguez M, et al. Diagnostic performance of contrast-enhanced dual-energy spectral mammography (CESM): a retrospective study involving 644 breast lesions. *Radiol Med* 2019; **124**: 1006–17. doi: https://doi.org/10.1007/s11547-019-01056-2
65. Lorek A, Steinhof-Radwańska K, Barczyk-Gutkowska A, Zarębski W, Boratyn-Nowicka A, Bobola A, et al. Retrospective comparison of contrast-enhanced spectral mammography with digital mammography in assessing tumor size in 668 cases of breast cancer. *Med Sci Monit* 2020; **26**: e926977: e926977: . doi: https://doi.org/10.12659/MSM.926977
66. Bicchierai G, Tonelli P, Piacenti A, De Benedetto D, Boeri C, Vanzi E, et al. Evaluation of contrast-enhanced digital mammography (CEDM) in the preoperative staging of breast cancer: large-scale single-center experience. *Breast J* 2020; **26**: 1276–83. doi: https://doi.org/10.1111/tbj.13766
67. Lobbes M, Prevos R, Smidt M. Response monitoring of breast cancer patientsreceiving neoadjuvant chemotherapy using breast MRI – a review of current knowledge. *J Cancer Ther Res* 2012; **1**: 34. doi: https://doi.org/10.7243/2049-7962-1-34
68. Lobbes MBI, Prevos R, Smidt M, Tjan-Heijnen VCG, van Goethem M, Schipper R, et al. The role of magnetic resonance imaging in assessing residual disease and pathologic complete response in breast cancer patients receiving neoadjuvant chemotherapy: a systematic review. *Insights Imaging* 2013; **4**: 163–75. doi: https://doi.org/10.1007/s13244-013-0219-y
69. Mann RM, Cho N, Moy L. Breast MRI: state of the art. *Radiology* 2019; **292**: 520–36. doi: https://doi.org/10.1148/radiol.2019182947
70. Barra FR, Sobrinho AB, Barra RR, Magalhães MT, Aguiar LR, de Albuquerque GFL, et al. Contrast-Enhanced mammography

(CEM) for detecting residual disease after neoadjuvant chemotherapy: a comparison with breast magnetic resonance imaging (MRI. *Biomed Res Int* 2018; **2018**8531916 doi: https://doi.org/10.1155/2018/8531916

71. Tang S, Xiang C, Yang Q. The diagnostic performance of CESM and CE-MRI in evaluating the pathological response to neoadjuvant therapy in breast cancer: a systematic review and meta-analysis. *Br J Radiol* 2020; **93**: 20200301. doi: https://doi.org/10.1259/bjr.20200301
72. Hunt CH, Hartman RP, Hesley GK. Frequency and severity of adverse effects of iodinated and gadolinium contrast materials: retrospective review of 456,930 doses. *AJR Am J Roentgenol* 2009; **193**: 1124–7. doi: https://doi.org/10.2214/AJR.09.2520
73. European Society of Urogenital RadiologyESUR Guidelines on Contrast Agents v10.0 [Internet]. 2018. Available from: http://www.esur.org/fileadmin/content/2019/ESUR_Guidelines_10.0_Final_Version.pdf [cited 2021 May 27].
74. National Institute for Health and Care Excellence Acute kidney injury: prevention. *detection and management (NICE Guideline NG148) [Internet]* 2019;. cited 2021 May 27.
75. Badr S, Laurent N, Régis C, Boulanger L, Lemaille S, Poncelet E. Dual-Energy contrast-enhanced digital mammography in routine clinical practice in 2013. *Diagn Interv Imaging* 2014; **95**: 245–58. doi: https://doi.org/10.1016/j.diii.2013.10.002
76. James JR, Pavlicek W, Hanson JA, Boltz TF, Patel BK. Breast radiation dose with CESM compared with 2D FFDM and 3D tomosynthesis mammography. *AJR Am J Roentgenol* 2017; **208**: 362–72. doi: https://doi.org/10.2214/AJR.16.16743
77. Phillips J, Mihai G, Hassonjee SE, Raj SD, Palmer MR, Brook A, et al. Comparative dose of contrast-enhanced spectral mammography (CESM), digital mammography, and digital breast tomosynthesis. *American Journal of Roentgenology* 2018; **211**: 839–46. doi: https://doi.org/10.2214/AJR.17.19036
78. FDAMammography Quality Standards Act Regulations [Internet]. 1999. Available from: https://www.fda.gov/radiation-emitting-products/regulations-mqsa/mammography-quality-standards-act-regulations [cited 2021 Apr 18].
79. National Research Council Health risks from exposure to low levels of ionizing radiation: BEIR VII phase 2. *National Academic Press* 2006;.
80. Lobbes MBI, Mulder HKP, Rousch M, Backes WH, Wildberger JE, Jeukens CRLPN. Quantification of enhancement in contrast-enhanced spectral mammography using a custom-made quantifier tool (I-STRIP): a proof-of-concept study. *Eur J Radiol* 2018; **106**: 114–21. doi: https://doi.org/10.1016/j.ejrad.2018.07.021
81. Patel BK, Ranjbar S, Wu T, Pockaj BA, Li J, Zhang N, et al. Computer-Aided diagnosis of contrast-enhanced spectral mammography: a feasibility study. *Eur J Radiol* 2018; **98**: 207–13. doi: https://doi.org/10.1016/j.ejrad.2017.11.024
82. Marino MA, Pinker K, Leithner D, Sung J, Avendano D, Morris EA, et al. Contrast-Enhanced mammography and radiomics analysis for noninvasive breast cancer characterization: initial results. *Mol Imaging Biol* 2020; **22**: 780–7. doi: https://doi.org/10.1007/s11307-019-01423-5
83. Wang Z, Lin F, Ma H, Shi Y, Dong J, Yang P, et al. Contrast-Enhanced spectral mammography-based radiomics nomogram for the prediction of neoadjuvant chemotherapy-insensitive breast cancers. *Front Oncol* 2021; **11**: 605230. doi: https://doi.org/10.3389/fonc.2021.605230

BJR https://doi.org/10.1259/bjr.20210779

Received: 25 June 2021 | Revised: 30 January 2022 | Accepted: 07 February 2022

Cite this article as:
Savaridas SL, Whelehan P, Warwick VR, Vinnicombe SJ, Evans AJ. Contrast-enhanced digital breast tomosythesis and breast MRI to monitor response to neoadjuvant chemotherapy: patient tolerance and preference. *Br J Radiol* 2022; **95**: 20210779.

FULL PAPER

Contrast-enhanced digital breast tomosythesis and breast MRI to monitor response to neoadjuvant chemotherapy: patient tolerance and preference

[1]SARAH L SAVARIDAS, MBChB, FRCR, [1]PATSY WHELEHAN, MSc, PgC, DCR, [1]VIOLET R WARWICK, MSc, PhD, [1,2]SARAH J VINNICOMBE, MBBS, BSc, MRCP, FRCR and [1]ANDREW J EVANS, MBChB, FRCR

[1]School of Medicine, University of Dundee, Ninewells Hospital & Medical School, Dundee, UK
[2]Gloucestershire Hospitals NHS Foundation Trust, Cheltenham, UK

Address correspondence to: Dr Sarah L Savaridas
E-mail: *s.savaridas@dundee.ac.uk*

Objective: Contrast-enhanced digital breast tomosynthesis (CE-DBT) is a novel imaging technique, combining contrast-enhanced spectral mammography and tomosynthesis. This may offer an alternative imaging technique to breast MRI for monitoring of response to neoadjuvant chemotherapy. This paper addresses patient experience and preference regarding the two techniques.
Methods: Conducted as part of a prospective pilot study; patients were asked to complete questionnaires pertaining to their experience of CE-DBT and MRI following pre-treatment and end-of-treatment imaging. Questionnaires consisted of eight questions answered on a categorical scale, two using a visual analogue scale (VAS), and a question to indicate preference of imaging technique. Statistical analysis was performed with Wilcoxon signed rank test and McNemar test for related samples using SPSS v. 25.
Results: 18 patients were enrolled in the pilot study. Matched CE-DBT and MRI questionnaires were completed after 22 patient episodes. Patient preference was indicated after 31 patient episodes. Overall, on 77% of occasions patients preferred CE-DBT with no difference between pre-treatment and end-of-treatment imaging. Overall experience (p = 0.008), non-breast pain (p = 0.046), anxiety measured using VAS (p = 0.003), and feeling of being put at ease by staff (p = 0.023) was better for CE-DBT. However, more breast pain was experienced during CE-DBT when measured on both VAS (p = 0.011) and categorical scale (p = 0.021).
Conclusion: Our paper suggests that patients prefer CE-DBT to MRI, adding further evidence in favour of contrast-enhanced mammographic techniques.
Advances in knowledge: Contrast mammographic techniques offer an alternative, more accessible imaging technique to breast MRI. Whilst other studies have addressed patient experience of contrast-enhanced spectral mammography, this is the first study to directly explore patient preference for CE-DBT over MRI in the setting of neoadjuvant chemotherapy, finding that overall, patients preferred CE-DBT despite the relatively long breast compression.

BACKGROUND

Breast cancer remains the most common malignancy in females in the European Union, accounting for 13.3% of all cancer diagnoses, and 7.3% of all cancer deaths.[1] With advances in oncological treatment, increasing numbers of females are receiving pre-surgical neoadjuvant chemotherapy (NACT). Initially used to downstage inoperable, locally advanced tumours, NACT is increasingly used to reduce extent for surgery, *e.g.* downgrading from mastectomy to breast conservation surgery.[2]

Imaging monitoring of treatment response is necessary throughout the course of NACT to gauge *in vivo* chemosensitivity and guide surgical planning. Contrast-enhanced MRI is considered the current gold-standard technique both for predicting complete pathological response and residual tumour size.[3–5] Unfortunately, due to pressures on imaging services, it can be difficult to obtain in a timely fashion. It is an expensive and time-consuming technique and for some patients, it is either contraindicated or poorly tolerated.[6–8]

Contrast-enhanced digital breast tomosynthesis (CE-DBT) is a novel technique which allows acquisition of contrast-enhanced spectral mammography (CESM) and digital breast tomosynthesis (DBT) images during the same breast

compression episode. CESM is a functional imaging technique which demonstrates the vascularity of breast lesions through dual energy subtraction. DBT is a pseudo-3D mammographic technique, which eliminates overlapping breast tissue and improves visibility of malignant structural features. DBT has demonstrated increased cancer detection rates, especially in dense breasts, when compared with full field digital mammography.[9] Published studies have demonstrated that CESM is at least as accurate as MRI for the detection of breast cancer[10,11] and early evidence suggests CESM may be comparable to MRI for monitoring patients treated with NACT.[12,13]

Emerging evidence of patient experience of CESM suggests an overall preference for CESM in place of MRI. Hobbs et al consider patient tolerance of CESM and MRI in the local pre-operative staging of breast cancer. They include feedback from 49 patients, with a significantly higher overall preference for CESM.[14] Phillips et al review patient preference and experience of CESM, MRI and digital mammography in the context of high risk screening, with 79% of patients indicating they would prefer CESM to MRI.[15] However, we can find no published studies on patient preference concerning contrast-enhanced mammographic techniques in the context of NACT. Furthermore, there appears to be no published evidence on patient experience with CE-DBT, which requires a longer period of compression in each position to allow the additional DBT acquisition. If CE-DBT is to replace MRI for some patients for monitoring response to NACT, it is essential to assess patient acceptability, specifically in this context.

METHODS

This research was conducted as part of the ethically approved prospective imaging study: CONtrast enhanced Digital breast tomosynthesis for monitoring Of Response to neoadjuvant chemotherapy (CONDOR). The aim of this pilot study was to compare the use of breast MRI with CE-DBT for monitoring tumour response to NACT. The results of the comparative accuracy of the two techniques will be published separately. Females aged over 18 years with symptomatic and screen-detected cancers undergoing NACT were eligible for inclusion. Exclusion criteria were contraindication to iodinated contrast, history of previous breast cancer surgery or implants, pregnancy and lactation.

Patients were imaged with both modalities prior to starting chemotherapy, at mid-treatment and at end-of-treatment, prior to surgery. Investigation of patient acceptability and preference was included in the aims of the study, and is the subject of this report. CE-DBT images were acquired using the commercially available Selenia Dimensions system (Hologic, Massachussetts). Imaging was commenced 3 min after intravenous (i.v.) administration of 1.5 mg/kg iodinated contrast agent (Omnipaque 300, GE Healthcare, Buckinghamshire), at rate of 2–3 ml/s. After 3 min, imaging in the CE-DBT unit was commenced. At pre-treatment, bilateral two-view (CC and MLO) CE-DBT was performed with delayed MLO of the index breast(s) acquired 9 min after injection. At mid- and end-of-treatment, only the breast(s) with malignancy were imaged. Breast MRI was performed on a Siemens 3T Prisma Fit scanner (Siemans Healthineers, Erlangen, Germany), using a dynamic contrast-enhanced protocol. The sequences included T1 2D axial high resolution, T2 axial turbo spin echo, diffusion sequences, T1 3D dynamic sequences (two pre-contrast and seven post-contrast) and a delayed T1 axial high resolution sequence, with a total scan time of approximately 40 min.

Participants were asked to complete questionnaires regarding their experience of CE-DBT and MRI, both following pre- and post-treatment imaging. The questions were identical on both questionnaires (Figure 1). There were eight questions with a 4-point categorical response scale. Two questions regarding anxiety and breast pain were assessed using a visual analogue scale (VAS); participants were asked to place a mark on a linear scale from 0 to 100. Finally, at both imaging time points, patients were asked to indicate a preference. To capture preference based on patient experience, rather than expectation of the test accuracy, the preference question was prefaced with 'assuming CE-DBT and MRI provided equivalent diagnostic information'. Several free-text boxes were provided to allow the participants to expand on responses.

Statistical analysis

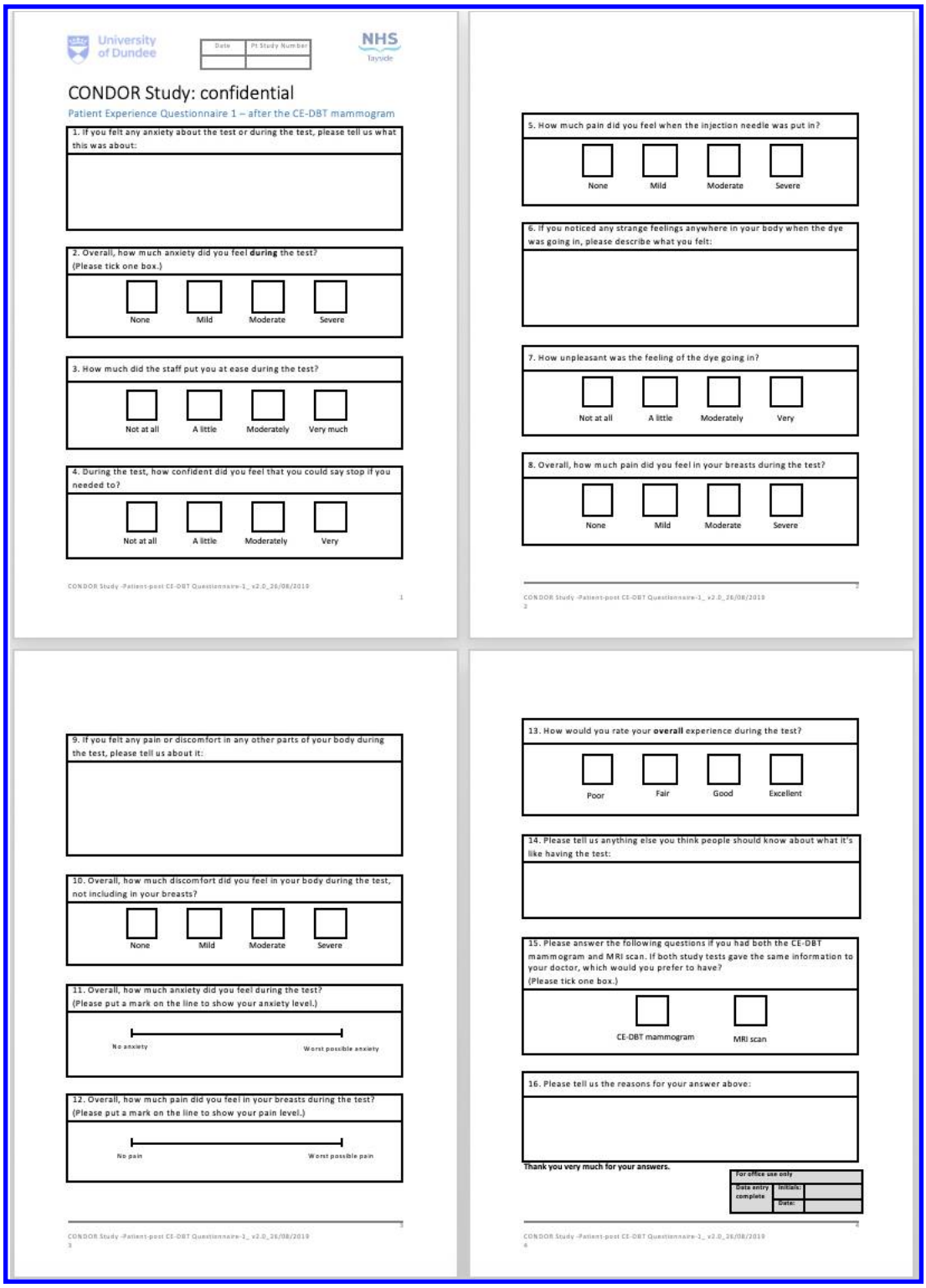
The Wilcoxon signed-rank test for related samples was used to assess for significant differences between the modalities on those questions using categorical response scales. Due to small sample size, it was not possible to analyse categorical data from the pre- and post-treatment questionnaires separately. Non-parametric VAS data were also analysed using a Wilcoxon signed-rank test as recommended by Heller et al.[16] In addition to the combined data set, subset analysis of pre-treatment breast pain and anxiety VAS data was performed. Binary outcome data were analysed using a McNemar test for related samples. The content of the free-text responses was summarised according to the subject matter and relative frequencies are shown in Table 1. All statistical analysis was performed on SPSS v. 25.0 (IBM, Armonk, NY).

RESULTS

Of the 31 patients eligible for the study, 10 declined. Reasons given were feeling overwhelmed (5)concerns regarding cannulation and/or contrast (2),extra travel (1)unknown (2) Three patients could not be recruited due to hospital logistics, leaving 18 patients who participated in the study. Average age of patients was 52.7 years (range 32–72 years). Average time between CE-DBT and MRI was 9.56 days (range 0–31 days) and 3.64 days (range 0–13 days) for pre- and post-treatment imaging respectively. On 11 (34%) occasions CE-DBT was performed first; the studies were performed on the same day on 7 (22%) occasions and MRI was performed first on 14 (44%) occasions. One patient withdrew at mid-treatment due to difficult intravenous access. Post-treatment MRI was omitted in two patients, as per standard care, as they only received four cycles of chemotherapy. Post-treatment CE-DBT was omitted for one patient who developed metastatic disease over the course of treatment. Therefore, 18 patients had pre-treatment CE-DBT and MRI and 14 had both CE-DBT and MRI post-treatment.

At pre-treatment, 17 (94%) and 14 (78%) patients completed questionnaires for CE-DBT and MRI respectively, with 13 (72%)

Figure 1. Patient questionnaire.



University of Dundee | Date | Pt Study Number | NHS Tayside

CONDOR Study: confidential

Patient Experience Questionnaire 1 – after the CE-DBT mammogram

1. If you felt any anxiety about the test or during the test, please tell us what this was about:

2. Overall, how much anxiety did you feel **during** the test? (Please tick one box.)
None | Mild | Moderate | Severe

3. How much did the staff put you at ease during the test?
Not at all | A little | Moderately | Very much

4. During the test, how confident did you feel that you could say stop if you needed to?
Not at all | A little | Moderately | Very

CONDOR Study -Patient-post CE-DBT Questionnaire-1_ v2.0_26/08/2019 1

5. How much pain did you feel when the injection needle was put in?
None | Mild | Moderate | Severe

6. If you noticed any strange feelings anywhere in your body when the dye was going in, please describe what you felt:

7. How unpleasant was the feeling of the dye going in?
Not at all | A little | Moderately | Very

8. Overall, how much pain did you feel in your breasts during the test?
None | Mild | Moderate | Severe

CONDOR Study -Patient-post CE-DBT Questionnaire-1_ v2.0_26/08/2019 2

9. If you felt any pain or discomfort in any other parts of your body during the test, please tell us about it:

10. Overall, how much discomfort did you feel in your body during the test, not including in your breasts?
None | Mild | Moderate | Severe

11. Overall, how much anxiety did you feel during the test? (Please put a mark on the line to show your anxiety level.)
No anxiety — Worst possible anxiety

12. Overall, how much pain did you feel in your breasts during the test? (Please put a mark on the line to show your pain level.)
No pain — Worst possible pain

CONDOR Study -Patient-post CE-DBT Questionnaire-1_ v2.0_26/08/2019 3

13. How would you rate your **overall** experience during the test?
Poor | Fair | Good | Excellent

14. Please tell us anything else you think people should know about what it's like having the test:

15. Please answer the following questions if you had both the CE-DBT mammogram and MRI scan. If both study tests gave the same information to your doctor, which would you prefer to have? (Please tick one box.)
CE-DBT mammogram | MRI scan

16. Please tell us the reasons for your answer above:

Thank you very much for your answers.

For office use only | Data entry complete | Initials: | Date:

CONDOR Study -Patient-post CE-DBT Questionnaire-1_ v2.0_26/08/2019 6

patients completing both. Post-treatment, 15 (94%) completed questionnaires following CE-DBT and 10 (67%) completed questionnaires following MRI, with 9 (64%) completing both. Thus, there were a total of 22 participant episodes completed by 16 patients, where all questionnaires were completed. As disproportionately more patients returned questionnaires following MRI, only matched questionnaires were used for further comparative statistical analysis. The content of free-text responses was reviewed for all completed questionnaires (CE-DBT $n = 32$, MRI $n = 24$). Free-texts were grouped according to content of the responses. For example, in response to the question regarding anxiety, responses classified as 'general anxiety'

Table 1. Summarised free-text responses from all completed questionnaire

Question	Free-text response (grouped)	True for CE-DBT (n = 32)	True for MRI (n = 24)
If you felt any anxiety about the test or during the test, please tell us what this was about:			
	General anxiety	3 (9%)	5 (21%)
	Breast pain / discomfort / compression	3 (9%)	0
	Cannulation / contrast	4 (13%)	0
	Being enclosed	0	3 (13%)
	Noise	0	1 (4%)
	Controlling breathing / Keeping still	0	2 (8%)
	Non-breast pain	0	1 (4%)
If you noticed any strange feelings anywhere in your body when the dye was going in, please describe what you felt:			
	Heat / flushing	23 (71%)	2 (8%)
	Cold sensation	0	6 (25%)
	Numbness	2 (6%)	0
	Need to urinate	12 (0.38)	3 (0.13)
	Strange taste	7 (0.22)	2 (0.08)
If you felt any pain or discomfort in any other parts of the body during the test, please tell us about it:			
	Leg (sciatic)	1 (3%)	1 (4%)
	Shoulders / upper limbs	0	7 (29%)
	Face / forehead	0	3 (13%)
Please tell us anything else you think people should know about what it's like having the test:			
	Noise	0	5 (20%)
	Headache / off balance	0	2 (8%)
	Wish to reassure other females / share positive experience	14 (44%)	5 (21%)
	Leaflet / more information	0	2 (8%)
	Non-breast pain	0	1 (4%)
	Sensation associated with contrast	5 (6%)	0
	Breast pain / discomfort	1 (3%)	0
Please tell us the reason for your answer above: [preference of technique]			
	Quicker technique	13 (41%)	0
	More comfortable	11 (34%)	4 (17%)
	More confidence in technique	1 (3%)	3 (13%)
	Less intimidating / more in control	3 (9%)	0
	Feeling unwell after MRI	2 (6%)	0

CE-DBT, contrast-enhanced digital breast tomosynthesis.

included "A little anxious. Mainly because it is the first time I have been in the breast screening clinic" [CE-DBT] and "I felt apprehensive beforehand" [MRI]. The free text responses to the question "*Please tell us anything else you think people should know about what it's like having the test*" were very varied. For example, many participants noted positive comments to reassure future patients regarding the imaging studies. These comments were classified as 'wish to reassure other females / share positive experience' and included comments such as "There is nothing to worry about" [CE-DBT] and 'It gets easier after the first one" [MRI]. The categorised responses are shown in Table 1.

Outcome data for questions answered with 4-point categorical response format is shown in Table 2; statistically significant results are given in bold.

Answers measured using VAS are displayed in Table 3.

Table 2. Patient experience from matched questionnaire, categorical data

	Rating	CE-DBT	MRI	*p*
Overall how much anxiety did you feel during the test?				
	None	14	9	
	Mild	6	10	
	Moderate	2	2	
	Severe	0	1	0.052
How much pain did you feel when the needle was put in?				
	None	11	10	
	Mild	11	12	
	Moderate	0	0	
	Severe	0	0	0.655
Overall, how much pain did you feel in your breasts during the test?				
	None	12	20	
	Mild	7	1	
	Moderate	1	1	
	Severe	1	0	**0.021**
Overall, how much discomfort did you feel in your body during the test, not including in your breasts?				
	None	16	11	
	Mild	5	8	
	Moderate	1	2	
	Severe	0	1	**0.046**
How much did the staff put you at ease during the test?				
	Very much	21	15	
	Moderately	1	5	
	A little	0	2	
	Not at all	0	0	**0.023**
During the test, how confident did you feel that you could say stop if you needed to?				
	Very much	22	19	
	Moderately	0	1	
	A little	0	1	
	Not at all	0	1	0.109
How unpleasant was the feeling of the dye going in?				
	Not at all	15	18	
	A little	6	3	
	Moderately	1	1	0.257
	Very much	0	0	
How would you rate your overall experience?				
	Excellent	14	8	
	Good	8	7	
	Fair	0	7	
	Poor	0	0	**0.008**

CE-DBT, contrast-enhanced digital breast tomosynthesis.

Table 3. Patient experience from matched questionnaires, questions answered using VAS

	CE-DBT (mean ± SD)	MRI (mean ± SD)	*p*
Anxiety (full data set)	6.45 ± 8.06	16.91 ± 20.77	**0.003**
Anxiety (pre-treatment only)	9.08 ± 9.31	22.92 ± 23.78	**0.015**
Breast pain (full dataset)	11.14 ± 18.60	3.86 ± 9.92	**0.011**
Breast pain (pre-treatment only)	14.23 ± 22.90	5.92 ± 12.54	**0.155**

CE-DBT, contrast-enhanced digital breast tomosynthesis; VAS, visual analogue scale.

Significant differences in favour of CE-DBT were seen for non-breast pain ($p = 0.046$), being put at ease by staff ($p = 0.023$) and overall experience ($p = 0.008$). Anxiety was lower for CE-DBT when measured using VAS ($p = 0.003$) and this remained significant for subset analysis of the pre-treatment data; there was no statistically significant difference when measured with the categorical scale. Breast pain was significantly higher with CE-DBT when measured with both the categorical scale ($p = 0.021$) and whole data set VAS ($p = 0.011$). While breast pain was higher for CE-DBT on subset analysis, this was not statistically significant. No statistically significant difference was seen between CE-DBT and MRI in patients' confidence that they could stop the test if needed.

Patient preference was recorded after 31 episodes, 16 following initial imaging, 15 following final imaging. One patient who selected both CE-DBT and MRI after final imaging this episode was excluded from further analysis. 11 patients recorded a preference at initial and final imaging. Results are displayed in Table 4.

Overall, on 77% of occasions patients preferred CE-DBT. Of the 11 patients who responded at both time points, there was no significant change in in the proportion preferring CE-DBT; 10 (91%) and 8 (82%) at initial and final imaging respectively, $p = 0.25$.

DISCUSSION

To our knowledge, this is the first study to assess patient experience of CE-DBT; furthermore, it is the first study to assess patient preference for any form of contrast-enhanced mammographic technique when used for assessing response to NACT.

It is beyond the scope of this paper to assess the relative accuracy of CE-DBT and MRI. This is the primary aim of the pilot study, and therefore patients were aware that the accuracy of CE-DBT for assessing response to NACT is currently unknown when they consented to join the study. This study has demonstrated that, assuming the test provided equivalent diagnostic information, the majority of patients preferred CE-DBT. This did not vary between pre- and post-treatment, suggesting that previous experience of the techniques did not influence attitude.

Table 4. Patient preference

Time point	Preference	
	CE-DBT (%)	MRI (%)
Pre-treatment	13 (81%)	3 (19%)
Post-treatment	10 (71%)	4 (29%)
Total	23 (77%)	7 (23%)

CE-DBT, contrast-enhanced digital breast tomosynthesis.

Overall experience was also significantly more positive for CE-DBT, with 64% reporting it as excellent, as opposed to only 36% reporting an excellent MRI experience. These findings are supported by previous studies that report a patient preference for CESM over MRI both in the setting of local staging and high risk screening. Similar to previous studies, the most common reasons for preference of CE-DBT or CESM were faster time and greater comfort.[14,15] Unlike the study by Hobbs et al, noise level was not cited as a reason for preference in our cohort, although five patients mentioned MRI-associated noise in the free-text boxes.

Consistent with a previous study,[14] anxiety was significantly higher in the MRI group when measured using a VAS ($p = 0.003$), and descriptively higher when measured using the categorical rating scale ($p = 0.052$). Specific anxiety related to MRI concerned the enclosed space, lying still for a prolonged period and the noise. Anxieties relating to cannulation and/or contrast administration were only recorded in the free text in relation to CE-DBT, not MRI. However, unlike the findings of Hobbs et al, no significant difference was demonstrated between modalities either in pain on cannulation or unpleasant sensations associated with contrast injection.[14] Sensations described varied between the two techniques, iodinated CE-DBT contrast more commonly associated with heat or flushing, the sensation of passing urine and odd taste, and gadolinium more commonly associated with a cold sensation.

Conversely, significantly more positive responses for CE-DBT were also seen in relation to being put at ease by staff. We suggest that the close proximity of staff during CE-DBT, enabling them to reassure patients, reduced the anxiety patients experienced.

It is accepted that breast pain relating to mammographic compression is a widely reported patient concern, and has been shown to be associated with non-re-attendance for mammographic screening.[17] Consistent with Hobbs et al, it is therefore perhaps not surprising that significantly more females experienced breast pain associated with CE-DBT than MRI (categorical $p = 0.021$, VAS $p = 0.011$).[14] However, it is reassuring that despite the increased compression time necessary to allow DBT acquisition in addition to CESM, compared to standard mammography, overall patient preference and experience remains in favour of CE-DBT. Unlike in previous studies, patients were also asked to

report on pain in the rest of the body experienced during both techniques; significantly more females experienced non-breast pain with MRI (p = 0.046). Pain was predominantly related to upper limb and pressure on the face / forehead experienced during MRI. This finding has not been previously reported and may offset the increased breast pain experienced with mammographic techniques.

The main limitation of the study is the small sample size of 18 patients. Because of this small sample size, it was necessary to pool responses for pre-treatment and end-of-treatment for statistical analysis of categorical data. Subset analysis of pre-treatment VAS data was performed. The majority of patients (10), were only included at one time point. However, the responses of six patients were included at both time points. As with most prospective trials, there is the possibility of selection bias, as patients opted into joining the study. Therefore, it is possible that the study cohort felt more positive towards new imaging techniques than the general population. Whilst these factors could potentially bias results, our findings are consistent with previous slightly larger studies.

Ideally, the order in which patients were imaged would have been randomised, but this was not possible logistically as the studies were booked according to availability and timing of chemotherapy cycles. Despite this the order in which patients had the two imaging techniques was fairly balanced, with MRI performed first on 44% occasions and CE-DBT on 34% occasions.

This study compared patient experience of CE-DBT to that of a 3 T MRI scanner, as opposed to a 1.5 T MRI scanner. It is possible that the negative experience of some patients associated with MRI may have been compounded by the higher field strength and the narrower bore of the magnet. However, whilst studies have demonstrated that patients experience symptoms such as vertigo/dizziness, headache and spinal pain more frequently with 3 T MRI,[18] in addition to comfort the primary reason for CE-DBT, preference was cited as the shorter study time. This factor would remain true irrespective of magnetic field strength. Therefore, whilst it is possible that the preference of CE-DBT was magnified by the higher field strength it is unlikely that it would alter overall preference.

It is also worth noting that the MRI protocol routinely used in this study is quite lengthy, as evaluation of imaging response to NACT is a research focus of the unit. A faster protocol might have increased patient acceptability.

Establishing patient acceptability is essential prior to any policy change. This pilot study has demonstrated promising results regarding patient experience of contrast mammographic techniques in the context of NACT. A large multicentre study is required to confirm these findings. Assessment of additional factors that may confound patient experience could be included, such as a body mass index, history of claustrophobia, degree of mammographic compression. Overall recruitment rate would likely be improved in centres routinely using CE-DBT at diagnosis as this would reduce the time pressure for patients to decide. Online or text-message based questionnaire versions may improve response rate.

CONCLUSION

Our results demonstrate an overall patient preference for CE-DBT over MRI, when used to monitor response to NACT. This finding supports the use of contrast-mammographic techniques as a potential alternative to breast MRI for an ever-increasing number of indications, assuming clinical effectiveness.

FUNDING

TENOVUS, British Society of Breast Radiology, Symposium Mammographicum.

REFERENCES

1. Cancer incidence and mortality in EU-27 countries EU Science Hub: European Commision; 2020 [updated 22/07/2020]. 2020. Available from: https://ec.europa.eu/jrc/en/news/2020-cancer-incidence-and-mortality-eu-27-countries
2. Petruolo O, Sevilimedu V, Montagna G, Le T, Morrow M, et al. How often does modern neoadjuvant chemotherapy downstage patients to breast-conserving surgery? *Ann Surg Oncol* 2021; **28**: 287–94. https://doi.org/10.1245/s10434-020-08593-5
3. Lobbes MBI, Prevos R, Smidt M, Tjan-Heijnen VCG, van Goethem M, et al. The role of magnetic resonance imaging in assessing residual disease and pathologic complete response in breast cancer patients receiving neoadjuvant chemotherapy: a systematic review. *Insights Imaging* 2013; **4**: 163–75. https://doi.org/10.1007/s13244-013-0219-y
4. Gu Y-L, Pan S-M, Ren J, Yang Z-X, Jiang G-Q. Role of magnetic resonance imaging in detection of pathologic complete remission in breast cancer patients treated with neoadjuvant chemotherapy: a meta-analysis. *Clin Breast Cancer* 2017; **17**: 245–55. https://doi.org/10.1016/j.clbc.2016.12.010
5. Murphy C, Mukaro V, Tobler R, Asher R, Gibbs E, et al. Evaluating the role of magnetic resonance imaging post-neoadjuvant therapy for breast cancer in the neonab trial. *Intern Med J* 2018; **48**: 699–705. https://doi.org/10.1111/imj.13617
6. Mann RM, Balleyguier C, Baltzer PA, Bick U, Colin C, et al. Breast mri: eusobi recommendations for women's information. *Eur Radiol* 2015; **25**: 3669–78. https://doi.org/10.1007/s00330-015-3807-z
7. Eshed I, Althoff CE, Hamm B, Hermann K-GA. Claustrophobia and premature termination of magnetic resonance imaging examinations. *J Magn Reson Imaging* 2007; **26**: 401–4. https://doi.org/10.1002/jmri.21012
8. Onega T, Tosteson ANA, Weiss J, Alford-Teaster J, Hubbard RA, et al. Costs of diagnostic and preoperative workup with and without breast mri in older women with a breast cancer diagnosis. *BMC Health Serv Res* 2016; **16**(1): 76. https://doi.org/10.1186/s12913-016-1317-6
9. Caumo F, Zorzi M, Brunelli S, Romanucci G, Rella R, et al. Digital breast tomosynthesis

with synthesized two-dimensional images versus full-field digital mammography for population screening: outcomes from the verona screening program. *Radiology* 2018; **287**: 37–46. https://doi.org/10.1148/radiol.2017170745

10. Łuczyńska E, Heinze-Paluchowska S, Hendrick E, Dyczek S, Ryś J, et al. Comparison between breast mri and contrast-enhanced spectral mammography. *Med Sci Monit* 2015; **21**: 1358–67. https://doi.org/10.12659/MSM.893018
11. Fallenberg EM, Dromain C, Diekmann F, Engelken F, Krohn M, et al. Contrast-enhanced spectral mammography versus mri: initial results in the detection of breast cancer and assessment of tumour size. *Eur Radiol* 2014; **24**: 256–64. https://doi.org/10.1007/s00330-013-3007-7
12. Patel BK, Hilal T, Covington M, Zhang N, Kosiorek HE, et al. Contrast-enhanced spectral mammography is comparable to mri in the assessment of residual breast cancer following neoadjuvant systemic therapy. *Ann Surg Oncol* 2018; **25**: 1350–56. https://doi.org/10.1245/s10434-018-6413-x
13. Iotti V, Ravaioli S, Vacondio R, Coriani C, Caffarri S, et al. Contrast-enhanced spectral mammography in neoadjuvant chemotherapy monitoring: a comparison with breast magnetic resonance imaging. *Breast Cancer Res* 2017; **19**: 106. https://doi.org/10.1186/s13058-017-0899-1
14. Hobbs MM, Taylor DB, Buzynski S, Peake RE. Contrast-enhanced spectral mammography (cesm) and contrast enhanced mri (cemri): patient preferences and tolerance. *J Med Imaging Radiat Oncol* 2015; **59**: 300–305. https://doi.org/10.1111/1754-9485.12296
15. Phillips J, Miller MM, Mehta TS, Fein-Zachary V, Nathanson A, et al. Contrast-enhanced spectral mammography (cesm) versus mri in the high-risk screening setting: patient preferences and attitudes. *Clin Imaging* 2017; **42**: 193–97. https://doi.org/10.1016/j.clinimag.2016.12.011
16. Heller GZ, Manuguerra M, Chow R. How to analyze the visual analogue scale: myths, truths and clinical relevance. *Scand J Pain* 2016; **13**: 67–75. https://doi.org/10.1016/j.sjpain.2016.06.012
17. Whelehan P, Evans A, Wells M, Macgillivray S. The effect of mammography pain on repeat participation in breast cancer screening: a systematic review. *Breast* 2013; **22**: 389–94. https://doi.org/10.1016/j.breast.2013.03.003
18. Weintraub MI, Khoury A, Cole SP. Biologic effects of 3 tesla (t) mr imaging comparing traditional 1.5 t and 0.6 t in 1023 consecutive outpatients. *J Neuroimaging* 2007; **17**: 241–45. https://doi.org/10.1111/j.1552-6569.2007.00118.x

BJR

https://doi.org/10.1259/bjr.20211013

Received: 31 August 2021 | Revised: 18 November 2021 | Accepted: 25 November 2021

Cite this article as:
Komarla R, Gilliland L, Piraner M, Seidel R, Clifford K, Kunjummen J. Imaging and pathologic features of non-calcified ductal carcinoma in situ: can sonography predict upgrade?. *Br J Radiol* 2021; **95**: 20211013.

FULL PAPER

Imaging and pathologic features of non-calcified ductal carcinoma in situ: can sonography predict upgrade?

ROHINI KOMARLA, MD, LAWRENCE GILLILAND, MD, MARIA PIRANER, MD, REBECCA SEIDEL, MD, KELLY CLIFFORD, MD and JEAN KUNJUMMEN, DO

Department of Radiology and Imaging Sciences, Emory University School of Medicine, Atlanta, Georgia

Address correspondence to: Dr Jean Kunjummen
E-mail: *jean.kunjummen@emory.edu*

Objective: The purpose of this study was to evaluate the imaging and pathologic features and upgrade rate of non-calcified ductal carcinoma *in situ* (NCDCIS). The study tested the hypothesis that lesions with sonographic findings have higher upgrade rate compared to lesions seen on mammography or MRI only.
Methods: This retrospective study included patients with ductal carcinoma *in situ* (DCIS) diagnosed by image-guided core breast biopsy from December 2009 to April 2018. Patients with microcalcifications on mammography or concurrent ipsilateral cancer on core biopsy were excluded. An upgrade was defined as surgical pathology showing microinvasive or invasive cancer.
Results: A total of 71 lesions constituted the study cohort. 62% of cases (44/71) had a mammographic finding, and 38% (27/71) of mammographically occult lesions had findings on either ultrasound, MRI, or both. Of the 67 cases that underwent sonography, a mass was noted in 56/67 (83.6%) cases and no sonographic correlate was identified in 11/67 (16.4%) cases. 21% (15/71) of lesions were upgraded on final surgical pathology. The upgrade rate of patients with sonographic correlate was 27% (15/56) *vs* with mammographic findings only was 0% (0/11).
Conclusion: DCIS should be considered in the differential diagnosis of architectural distortion, asymmetries, focal asymmetries, and masses, even in the absence of microcalcifications. NCDCIS diagnosed by ultrasound may be an independent risk factor for upgrade.
Advances in knowledge: Radiologists must be aware of imaging features of DCIS and consider increased upgrade rate when NCDCIS is diagnosed by ultrasound.

INTRODUCTION

Ductal carcinoma *in situ* (DCIS) is a non-invasive lesion, which is confined to the ductal structures with an intact basement membrane and therefore has a better prognosis than invasive disease. The incidence of DCIS has increased with the implementation of screening. According to the American Cancer Society, approximately 60,000 new cases of DCIS are diagnosed each year in the United States.[1] Microcalcifications are the hallmark of DCIS on mammography; however, other imaging findings such as masses, architectural distortions, focal asymmetries, developing asymmetries, and tubular ductal density have been described in the literature.[2–5] Non-calcified DCIS (NCDCIS) accounts for approximately 10–20% of DCIS cases.[2,6] The upgrade rate for DCIS to invasive disease reported in the literature ranges from 20 to 30%.[7,8] Less is known about the imaging features of non-calcified DCIS and whether NCDCIS portends upgrade at final surgical pathology. Upgrade rate is determined as the rate at which presumable pure DCIS patients have invasive cancer on final lumpectomy or mastectomy pathology. To our knowledge, there is only one published study which evaluated the upgrade rates of NCDCIS, concluding that upgrade risk is associated with older patient age and family history of breast cancer in a first-degree relative.[9]

The purpose of this study is to describe the spectrum of imaging and pathologic features of NCDCIS and to evaluate features associated with upgrade of NCDCIS, which may contribute to improving detection and understanding the variable biologic behavior of the disease. If a higher upgrade rate is associated with certain imaging features of NCDCIS, these may be considered an independent risk factor for upgrade and be useful in surgical planning given the range of treatment options. When DCIS is upgraded to invasive disease, invasive tumors may require additional treatment beyond standard DCIS treatment including sentinel lymph node biopsy and, furthermore, chemotherapy. Specifically,

this study evaluates whether there is a higher upgrade rate in NCDIS lesions with a sonographic correlate *vs* lesions seen only by mammography or MRI without a sonographic correlate.

METHODS AND MATERIALS

Study population

This Health Insurance Portability and Accountability Act compliant retrospective study was approved by the institutional review board with a waiver for informed consent. The pathology database was queried for patients who underwent core needle biopsies of masses, asymmetries, distortions, and non-mass enhancement seen on mammography, ultrasound or MRI yielding DCIS from December 2009 to April 2018. 290 patients underwent successful core needle biopsy showing DCIS during the study period. Patients with concurrent invasive or microinvasive cancer on core biopsy and patients with associated microcalcifications seen on mammography were excluded. 86 patients fulfilled criteria for the study. 15 cases did not have final pathology documented in the medical record and therefore the remaining 71 cases constitute our study cohort. All imaging findings were retrospectively reviewed by one of three dedicated breast radiologists (LG, JK, RS) with 8–10 years of experience. There was one case where ultrasound features could not be retrospectively reviewed due to poor quality and therefore sonographic features were categorized as unspecified, however, that case was not upgraded on final pathology. Patient characteristics such as age, race, clinical symptoms, personal history of breast cancer, and family history of breast cancer were recorded. Imaging findings, method of biopsy (modality, needle gauge, and number of cores), radiologic–pathologic concordance, and final surgical pathology results were also documented.

Mammography

Standard craniocaudal and mediolateral oblique two view mammography was performed using one of three mammography unit models (Lorad M3, Hologic, Bedford, MA; Senographe DMR, GE Healthcare, Milwaukee, WI; Selenia Dimensions, Hologic, Bedford, MA). Additional views were acquired, if necessary, per the interpreting radiologist. The mammographic lesions were evaluated according to the American College of Radiology BI-RADS lexicon.[10] Breast density and lesion morphologic characteristics including size were recorded.

Sonography

Real-time grayscale and color or power Doppler sonography was obtained using a 12.5-MHz linear array transducer (iU22, Philips Healthcare, Andover MA). Static ultrasound images were reviewed. Lesion size, shape, margin, echogenic pattern, and other sonographic features were documented. Lesions were categorized according to the BI-RADS lexicon.[10]

Magnetic resonance imaging

MRI was performed using 1.5 T scanner on either Magnetom Aera, Siemens Healthcare, Erlangen Germany or Signa, GE Healthcare, Milwaukee, WI with a dedicated breast coil. Patients were imaged in the prone position using standard contrast-enhanced technique in the axial and sagittal planes. Sequences for the breast MRI include: unenhanced axial T1, axial short-tau inversion recovery, dynamic contrast-enhanced T_1 weighted fat suppressed with subtraction followed by sagittal contrast enhanced T_1 weighted fat suppressed. Gadopentetatedimeglumin (Magnevist, Bayer HealthCare) or gadoteridol (ProHance, Bracco) was administered intravenously (0.1 mmoll^{-1} per kg of body weight). Background parenchymal enhancement was documented. Areas of abnormal enhancement were classified as mass or non-mass enhancement with defining morphologic features and distribution according to the BI-RADS lexicon.[10]

Core biopsy

Biopsies were performed using ultrasound, stereotactic, or MRI guidance. Modality, needle gauge, and number of cores obtained were documented. Stereotactic-guided biopsies were performed on a dedicated prone (LoRad DSM, Hologic) or upright (Affirm, Hologic, Bedford, MA or Senographe Essential Stereotaxy GE Healthcare, Milwaukee, WI) stereotactic biopsy systems. Stereotactic and MRI biopsies were performed using 9- gauge (Suros ATEC, Hologic) vacuum-assisted needle device. An average of 11 cores were obtained during stereotactic biopsies and 12 cores during MRI guidance. Ultrasound-guided core biopsies were performed using a 14-gauge spring-loaded needle (Achieve coaxial biopsy system, Care Fusion, IL or Bard Marquee core biopsy needle, AZ). An average of five cores were obtained during ultrasound-guided core biopsies. Following biopsy, a marker clip was placed at biopsy site in all cases and two-view post-procedural mammograms were obtained on all cases to document clip placement and accuracy of targeting.

Histopathological assessment

All pathology specimens were interpreted by a breast pathologist at the time the specimen was obtained. Pathologic specimens from the core biopsy and final surgical pathology were reviewed documenting DCIS subtype, grade, and receptor status when available. All results were reviewed by a breast radiologist to determine imaging concordance. An upgrade was defined as surgical excisional pathology showing microinvasive or invasive cancer.

Statistical analysis

χ^2 analysis was performed to determine statistical significance of results, specifically if NCDCIS with sonographic correlate had a higher upgrade rate than NCDCIS with mammographic or MRI findings only. Findings with a p-value of less than 0.05 were considered to be statistically significant. StatView v. 5.0 (SAS Institute, Cary, NC) was used to carry out the analyses.

RESULTS

Clinical findings

A total of 71 cases in 69 patients constituted our study cohort. In patients with more than one finding (two patients), each finding was documented individually. All patients were female with a mean age of 61 years (range 29–86 years). 38 of 71 cases (54%) were African-Americans, 26 (37%) were white females, 5 (7%) were Asians, 1 (1%) was Hispanic and 1 (1%) was unknown. Initial presentation was a palpable mass in 9/71 cases (13%), focal pain in 1/71 cases (1%), and ipsilateral bloody nipple discharge in 5/71 cases (7%). The remaining 56/71 (79%) cases were detected

by screening (mammographic and MRI). 22 of 71 cases occurred in patients with a personal history of breast cancer (31%) and 18/71 cases in patients with a family history of breast cancer in a first-degree relative (25%).

Imaging findings

Mammography

All patients underwent mammography within our institution (Emory Healthcare) or at an outside facility. Tomosynthesis was performed in 48 of 71 cases (68%). Breast density evaluation included: fatty in 3/71 (4%), scattered fibroglandular density in 31/71 (44%), heterogeneously dense in 33/71 (46%), and extremely dense in 4/71 (6%). Mammographic findings were identified in 44/71 cases (62%) and DCIS was mammographically occult in 27/71 cases (38%). Of the cases with no mammographic correlate, 18 had extremely dense or heterogeneously dense breast density and 9 had scattered fibroglandular density. These findings are summarized on Table 1. The predominant mammographic finding was focal asymmetry seen in 21 cases. Final BI-RADS assessment was given to each of the 71 lesions. BI-RADS 4 assessment was assigned in 66/71 (93%) of lesions and BI-RADS 5 assessment was assigned in 5/71 (7%) of lesions.

Sonography

Ultrasound was performed in all but four cases, in which a suspicious finding was seen on MRI, and these four cases proceeded directly to MR biopsy without a MRI-directed ultrasound. Of the remaining 67/71 cases that underwent sonography, a mass was noted in 56/67 cases and no sonographic correlate was identified in 11/67 cases. An ultrasound only finding (without mammographic correlate) was identified in six cases performed for bloody nipple discharge or palpable abnormality. In addition, one case was an incidental finding seen by ultrasound.

Masses were characterized based on morphologic features, specifically mass shape and margins. Sonographic features of masses are presented in Table 1. The predominant features were irregular shape (62%) and indistinct margins (47%). Of the 56 masses seen on ultrasound, 52/56 (93%) were solid, and 4/56 (7%) were complex solid and cystic masses.

MRI

MRI was performed in 46 cases: 30/46 (65%) to evaluate disease extent, 13/46 (28%) for high-risk screening, and 3/46 (7%) for evaluation for bloody nipple discharge. A total of 44 MRI findings were identified (two lesions were not detected by MRI, due to marked background parenchymal enhancement). Non-mass enhancement alone was the predominant finding in 28/46 (60.9%) cases. Of cases with non-mass enhancement, the distribution was as follows: focal 10/28 (36%), segmental 6/28 (21%), regional 5/28 (18%), linear 6/28 (21%), and diffuse 1/28 (4%). MRI features are displayed in Table 1.

Histopathological findings

Diagnosis was made by image-guided core biopsy using ultrasound guidance in 55/71 (78%), MRI guidance in 13/71 (18%), and stereotactic guidance in 3/71 (4%). All lesions demonstrated radiologic–pathologic concordance.

Table 1. Mammographic, sonographic, and MRI findings

Finding	Number n/N (%)
Mammography	
Architectural distortion	10/71 (14)
Asymmetry	2/71 (3)
Focal asymmetry	21/71 (30)
Mass	11/71 (15)
Negative	27/71(38)
Sonography	
Mass	56/67 (84)
Shape	
Irregular	35/56(62)
Oval	19/56 (34)
Round	1/56 (2)
Unspecified	1/56 (2)
Margins	
Circumscribed	9/56 (16)
Non-circumscribed	
Indistinct	26/56 (47)
Microlobulated	13/56 (23)
Angular	6/56 (11)
Spiculated	0/56 (0)
Unspecified	1/56 (2)
Orientation	
Non-parallel	28/56 (50)
Parallel	27/56 (48)
Unspecified	1/56 (2)
Negative	11/67 (16)
MRI	
NME	28/46 (61)
Mass	14/46 (31)
Mass with associated NME	2/46 (4)
Negative	2/46 (4)

NME, non-mass enhancement..

DCIS was identified on all core biopsy specimens with nuclear grade and receptor status documented. The predominant nuclear grade was intermediate, seen in 33 of 71 cases. Estrogen and Progesterone receptor positivity was present in 54/71 cases (76%).

All cases underwent surgical excision yielding a 21% upgrade rate. Histopathological features are presented in Table 2. Definitive pathology was obtained in the 71 patients, of whom 36/71 (51%) underwent a lumpectomy and 35/71 (49%) had mastectomies. Of the 15 patients (21%) who had an upgrade at surgery, 8/15 (53%) had lumpectomies and 7/15 (46.7%) had

Table 2. Histopathologic features of all 71 cases

	Number n/N (%)
DCIS nuclear grade	
Low	19/71 (27)
Intermediate	33/71 (46)
High	17/71 (24)
Unknown	2/71 (3)
Receptor status	
ER+/PR+	54/71 (76)
ER+/PR-	5/71 (7)
ER-/PR-	12/71 (17)
Final pathology:	
No upgrade	56/71 (79)
Upgrade	15/71 (21)
Microinvasive	6/15 (40)
Invasive	9/15 (60)

DCIS, ductal carcinoma *in situ*;ER, estrogen receptor; PR, progesterone receptor.

mastectomies. X^2 analysis $X^2 = 0.06$, $p < 0.05$ demonstrates that results are not statistically significant ($p = 1.0$). Of the 15 patients who got upgraded to invasive cancer, 6/15 patients (40%) had a palpable finding and the remaining 9/15 (60%) were detected by screening.

The upgrade rate of patients with sonographic correlate was 27% (15/56) *vs* with mammographic findings only was 0% (0/11). X^2 analysis $X^2 = 3.965$, $p < 0.05$ demonstrate the results are statistically significant ($p = 0.048$). Examples of imaging from four of the upgraded cases with sonographic findings are provided in Figures 1–4. Summarized imaging features and pathology results of the all upgraded cases are displayed in Table 3.

DISCUSSION

Approximately, 20% of diagnosed cases of breast cancer each year are *in-situ*. The hallmark of DCIS is microcalcifications. However, 10–20% of DCIS cases do not present with microcalcifications on mammography.[2,6] Therefore, the detection of NCDCIS can be challenging. Studies show that 50–80% of patients with NCDCIS are clinically symptomatic.[2,3,11] In our study, only 21% of patients reported clinical symptoms. This difference in observation is

Figure 1. A 63-year-old female with architectural distortion initially identified on screening mammography. Biopsy pathology was ductal carcinoma *in situ*, intermediate grade, and final pathology yielded invasive tubular carcinoma. (**A**) Right craniocaudal tomosynthesis mammogram image demonstrates an area of architectural distortion in the upper outer quadrant (white circle). (**B**) Grayscale sonographic image shows an irregular mixed echogenic mass with indistinct margins (white arrows).

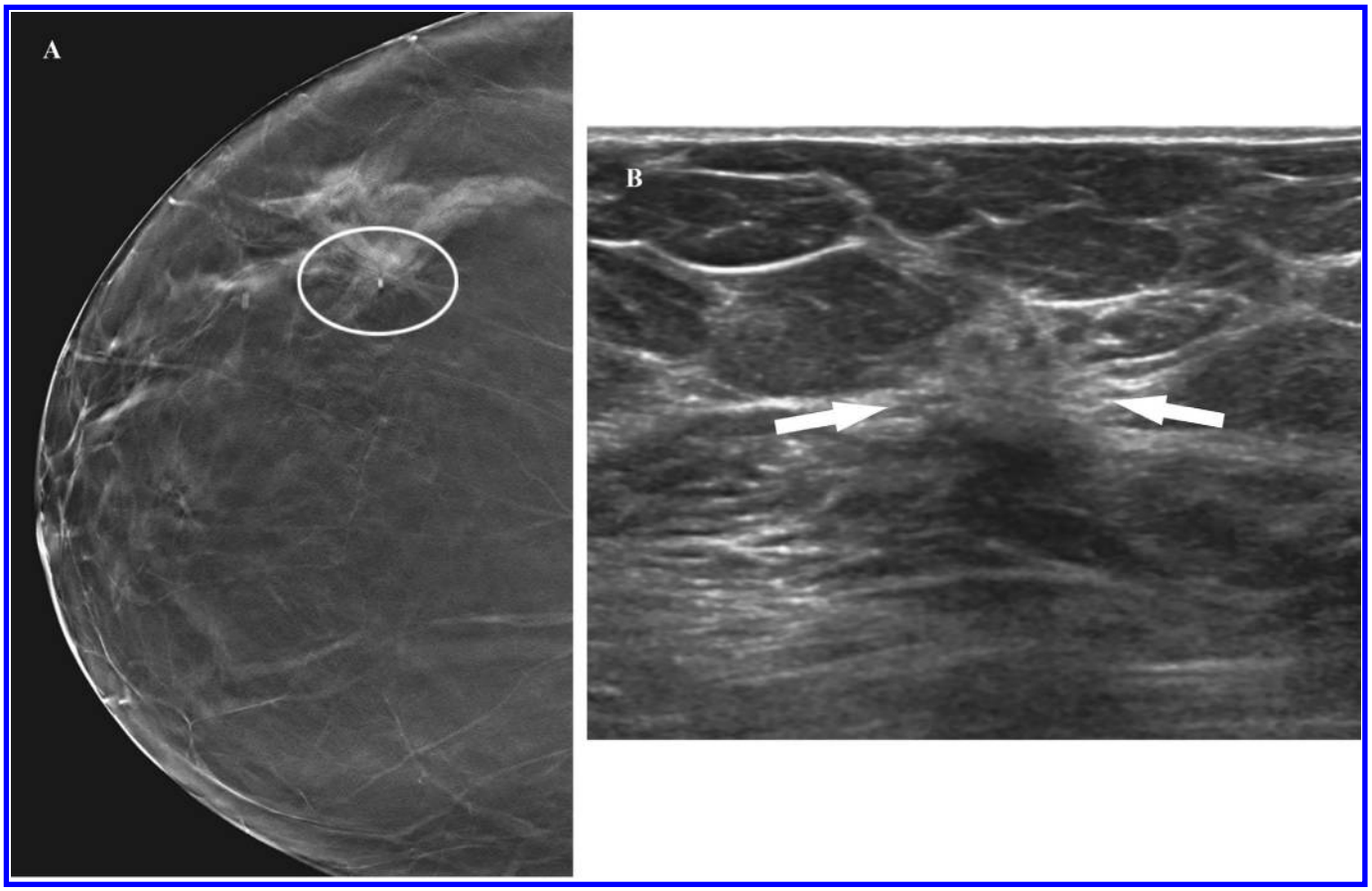
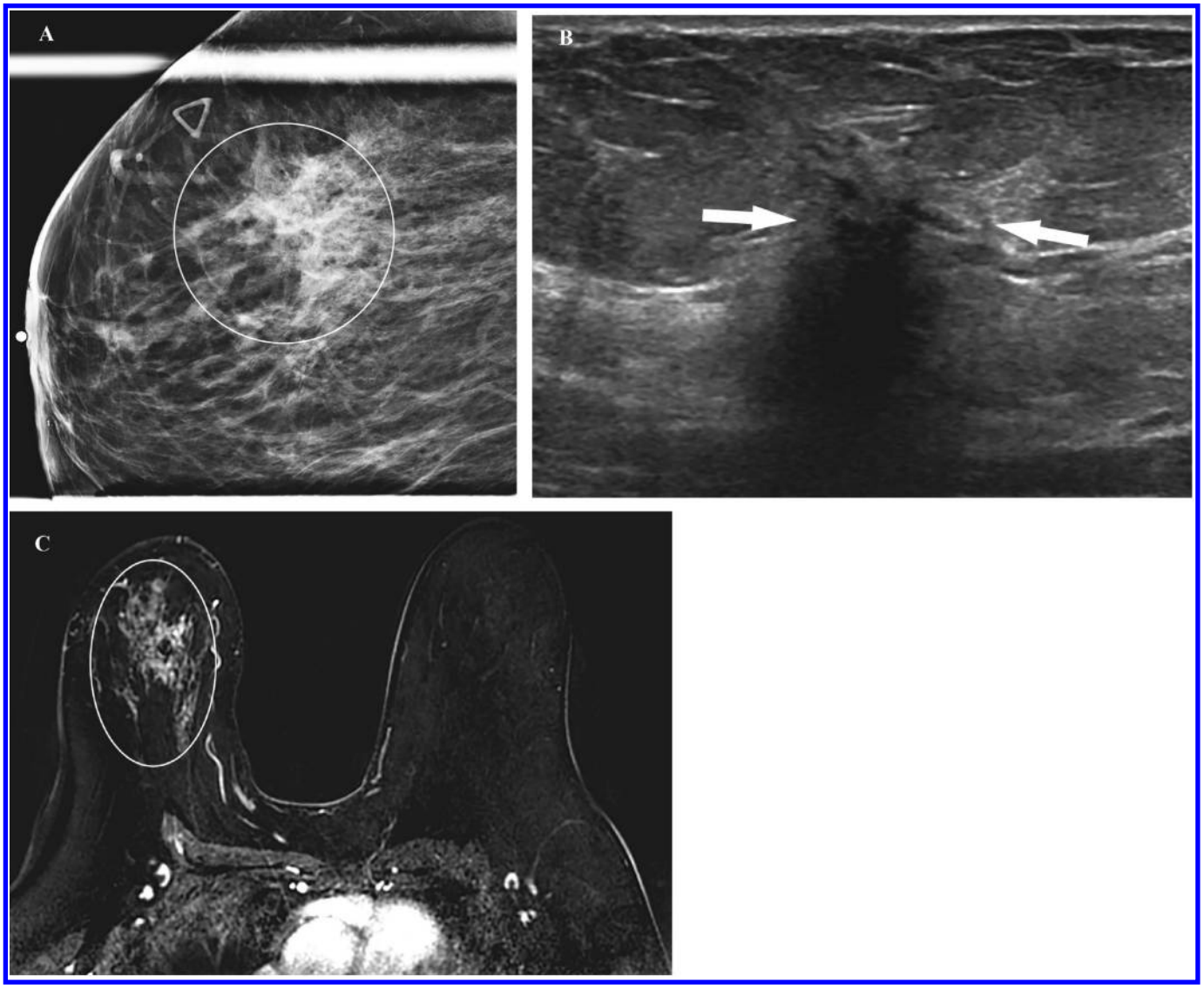


Figure 2. A 57 year-old female presenting with palpable lump in the right breast. Biopsy pathology was ductal carcinoma *in situ*, intermediate to high grade, and final pathology yielded microinvasive carcinoma. (**A**) Mediolateral spot compression mammogram of the right breast demonstrates a focal asymmetry (white circle). Triangular skin marker indicates a palpable abnormality is present at this site. (**B**) Grayscale sonographic image shows an irregular mass with indistinct margins at the 12 O'clock axis, site of palpable abnormality (white arrows). (**C**) Axial dynamic contrast-enhanced T_1 weighted subtraction image of the right breast shows regional asymmetric non-mass enhancement (white circle).

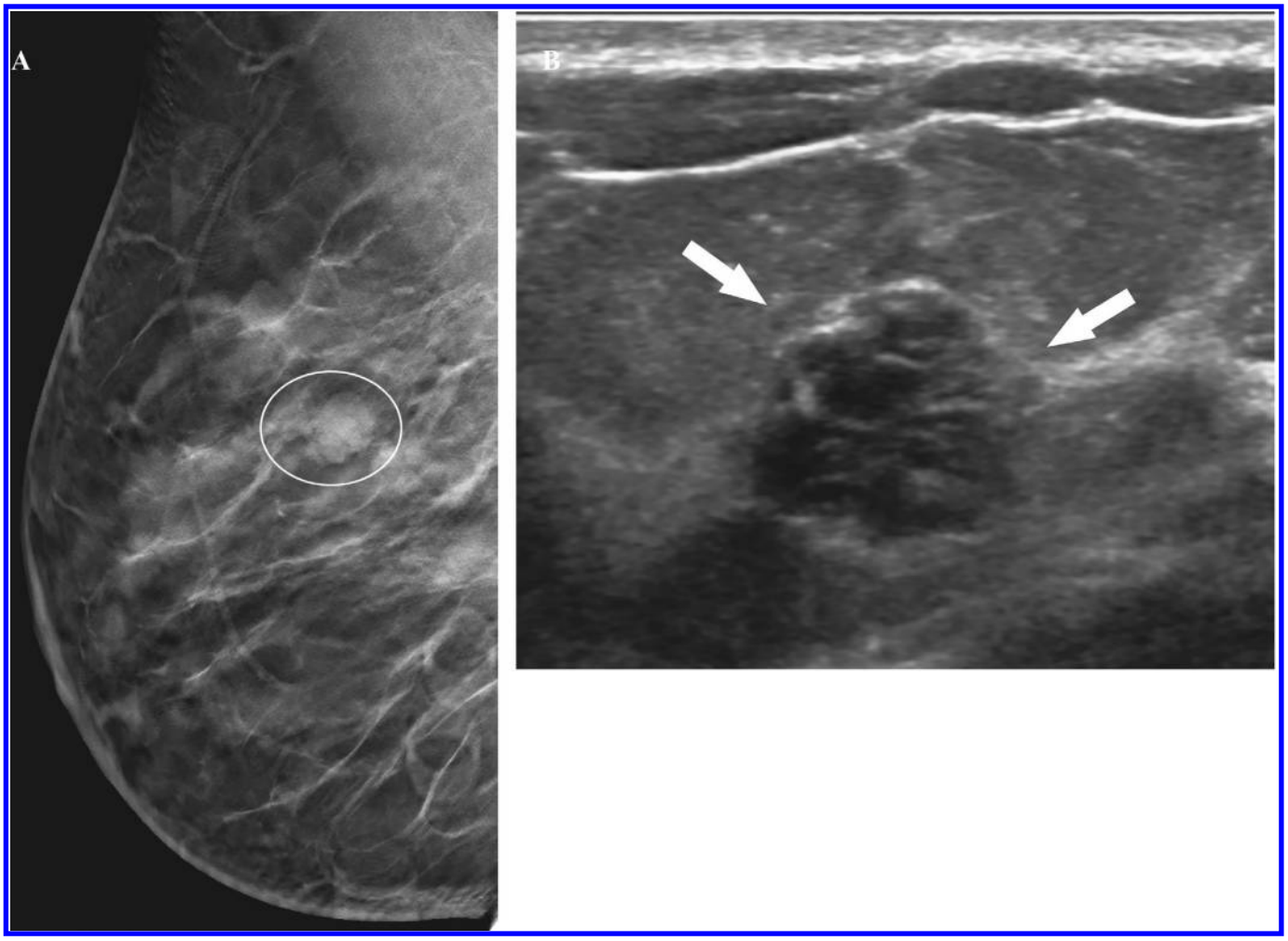


likely contributed to the detection of many cases with MRI, a modality not evaluated on prior studies. Two of the reference articles had a lower volume of cases, and all three studies were published more than 10 years prior.[2,3,11] Changes in technique and use of analog film likely contributed to differences in perception of findings on screening mammography.

Mammographic findings were seen in 62% (44/71) of cases and included architectural distortion, asymmetry, focal asymmetry, and mass, of which focal asymmetry was the most common finding at 48% of cases. This is similar to study by Cho et al[11] where focal asymmetries were the most common mammographic finding seen on NCDCIS. In our study, 38% (27/71) of cases were mammographically occult, which is higher than cases reported in the literature, ranging from 6 to 32%.[2,3,11] This may be due to the fact that many cases of NCDCIS were identified on MRI in our study. Of the cases with no mammographic correlate, 18/27 (67%) had extremely dense or heterogeneously dense breast density.

A mass was visualized by ultrasound in 84% (56/67) of NCDCIS cases in our study. This is consistent with other reports, in which the rate of detecting a mass as opposed to another type of lesion on sonography in DCIS ranged from 52 to 84%.[11–14] The most common sonographic appearance of NCDCIS, in our study, was an irregular mass with indistinct margins, which is

Figure 3. A 44 year-old female with oval mass initially identified on screening mammography. Biopsy pathology was ductal carcinoma *in situ*, intermediate grade, and final pathology yielded invasive mucinous carcinoma. (**A**) Right mediolateral tomosynthesis mammogram image demonstrates an oval mass in the upper outer quadrant (circle). (**B**) Sonographic image shows oval mass with microlobulated margins in the right breast 9 O'clock axis (white arrows).

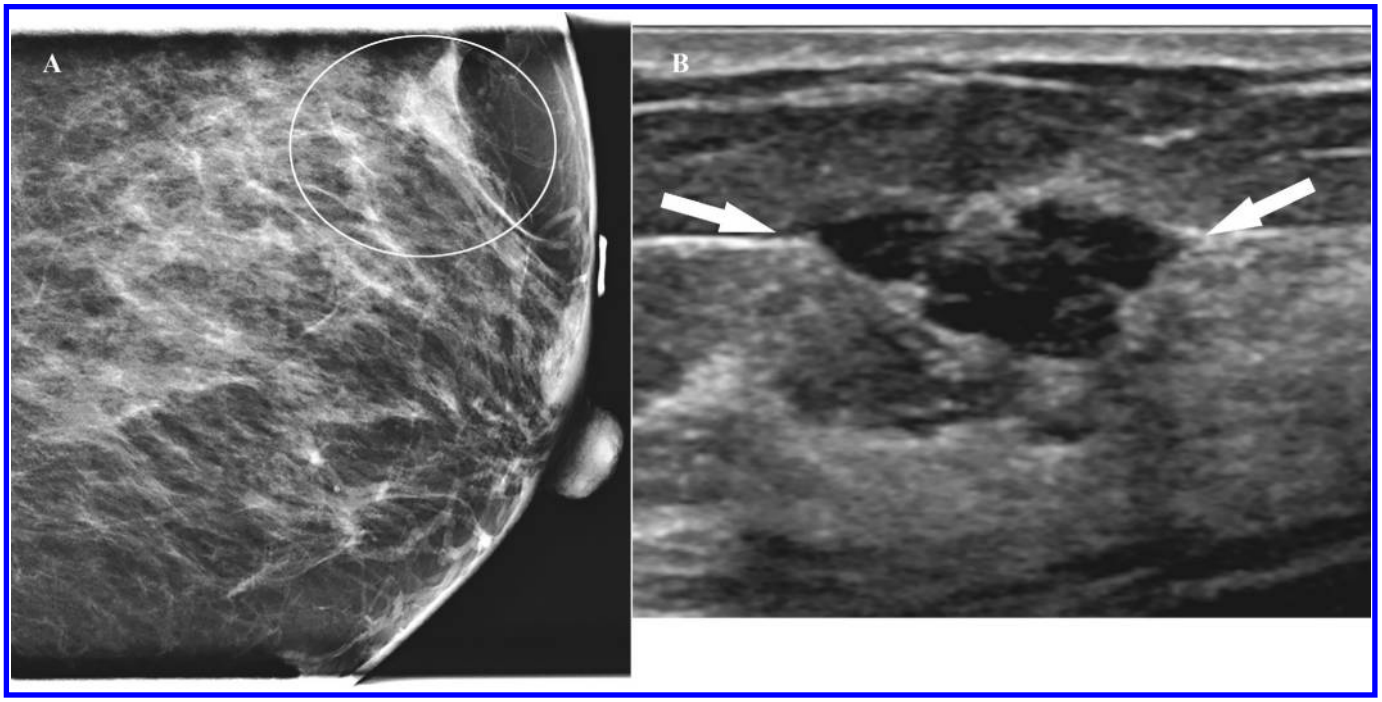


similar to prior studies.[3,12,13] NCDCIS can have variable imaging appearances and can present as a single or multiple hypoechoic masses. NCDCIS can have variable margins (microlobulated, indistinct, or circumscribed), variable shape, and appear as solid and cystic mass with variable internal vascularity.[15,16] However, sonographic appearance of NCDCIS remains non-specific. In our study, 34% (19/56) of sonographically detected NCDCIS had an oval shape and 16% (9/56) had circumscribed margins. This is similar to other studies in which sonographic features have been described as non-specific and can appear similar to benign processes.[11,15,16]

More recently, MRI features of DCIS have become better understood due to increased use of MRI both for screening in high-risk populations and for evaluation of disease extent in patients with newly diagnosed breast cancer. DCIS has varied appearance on MRI.[17–19] In our study, MRI was performed in 46 women yielding 44 findings, with non-mass enhancement (NME) being the most common MRI finding in 61%. This is similar to study by Scott-Moncrieff et al where 60% of non-calcified DCIS presented as NME.[18] Most common distribution of NME in our study was focal, in 36% of cases. Segmental distribution of NME was noted to be the most common for DCIS in prior studies 33%[17] and 14–77%[19] with focal distribution seen in 16%[17] and 16–33%.[19]

Van Lujit et al studied the distribution of grades of DCIS in females participants in a large volume screening program and in 4232 women with a diagnosis of DCIS, the distribution was as follows: 17.7% low grade, 31.4% intermediate grade, and 50.9% high grade.[20] In our study, 76% of non-calcified DCIS was predominantly low and intermediate grade and were ER/PR positive, which may suggest that NCDCIS has a better prognosis. This is in line with results seen by Kim et al.[3] Estrogen and progesterone are common biologic markers in breast cancer, which can predict the efficacy of patients' response to hormonal therapy. Hormonal receptor positivity suggests non-comedonecrosis and better prognosis.[21]

Figure 4. A 55 year-old female with a history of invasive ductal carcinoma status post lumpectomy with negative surgical margins 5 years prior presented for palpable abnormality at the lumpectomy site. Biopsy pathology was ductal carcinoma *in situ*, high grade, with final pathology yielding microinvasive ductal carcinoma. (**A**) Right mediolateral mammographic view demonstrates stable post-lumpectomy change (circle) compared to prior years. Triangular skin marker indicates a palpable abnormality is present at this site. (**B**) Sonographic image with color flow shows an irregular, hypoechoic mass with indistinct margins with vascularity (arrows).



In our study, at surgical excision, 15/71 (21%) of NCDCIS diagnosed on core biopsy were upgraded to microinvasive (40%) or invasive disease (60%), similar to the known 20–30% upgrade rate of DCIS reported in the literature.[7,8] 14 of 15 cases were intermediate- or high-grade disease with 1 case of low-grade disease upgraded to microinvasive disease at final pathology. All upgraded cases manifested as masses on ultrasound. Mammographic findings included focal asymmetry, architectural distortion, mass, and asymmetry. Upgrade cases presented either as a mass or NME on MRI. In our study, NCDCIS with ultrasound correlate had a higher upgrade rate (27%) compared to NCDCIS seen on mammogram or MRI only without a sonographic correlate, which was statistically significant ($p < 0.05$). This information can help radiologists determine if axillary ultrasound or MRI may be appropriate for evaluation of disease extent in cases where upgrade is more likely based on grade and sonographic findings. Although there is no randomized control data, the National Comprehensive Cancer Network recommends that a sentinel node biopsy should only be done if a patient is undergoing a mastectomy and in patients where the anatomic location of the DCIS and initial surgery will interfere with mapping a sentinel node should additional surgery be needed.[22] Given the higher upgrade rate of non-calcified DCIS with a sonographic finding, this information may be valuable in deciding if a sentinel node biopsy should be done. Additionally, given that some institutions are moving toward less invasive treatment of DCIS, knowledge that patients with sonographic findings have an increased upgrade rate could be of benefit if a surgeon is considering imaging and clinical surveillance of DCIS rather than excision.

Limitations of our study include its retrospective design and single institution experience. Also, the patient population is from a large academic institution with subspecialized radiologists, pathologists, and surgeons, and therefore the results may not apply to all practices. Another limitation is that ultrasound-guided biopsies at our institutions are performed with fewer samples and smaller gauge spring-loaded devices compared to stereotactic, tomosynthesis, and MRI-guided biopsies, which result in differences in sampling. Reviewing upgrade rates on NCDCIS at facilities that utilize larger gauge vacuum assisted devices for ultrasound-guided core needle biopsies would be helpful. Another limitation of our study is the small sample size, which reflects the rarity of NCDCIS. Reports like ours, which whittle down to subsets of an already rare disorder, are even more limited in number and generalizability and will require a multi-institutional effort to attain clinically significant sample sizes.

CONCLUSION

DCIS should be considered in the differential diagnosis of architectural distortion, asymmetries, focal asymmetries, and masses even in the absence of microcalcifications. Mammography typically shows a focal asymmetry in NCDCIS. The sonographic appearance most frequently observed is a single irregularly shaped mass with indistinct margins, and most common MRI

Table 3. Upgraded cases: findings and pathology

Case	Mammo	Ultrasound	MRI	Bx	Bx Pathology, Grade	Surgical Pathology
1	Negative	Irregular mass with indistinct margins	Not done	Ultrasound	DCIS with comedonecrosis, high	Microinvasive ductal carcinoma
2	Focal asymmetry	Oval mass with microlobulated margins	Focal NME	Ultrasound	Micropapillary DCIS with focal necrosis, intermediate	Microinvasive ductal carcinoma
3	Architectural distortion	Irregular mass with indistinct margins	Not done	Ultrasound	Solid and cribiform DCIS with focal necrosis, intermediate	Invasive tubular carcinoma
4	Focal asymmetry	Irregular mass with indistinct margins	Diffuse NME	Ultrasound	DCIS, high	Microinvasive carcinoma
5	Negative	Irregular mass with indistinct margins	Regional NME	Ultrasound	Micropapillary DCIS, high	Microinvasive ductal carcinoma
6	Focal asymmetry	Oval mass with microlobulated margins	Not done	Ultrasound	Solid and cribiform DCIS, intermediate	Invasive ductal carcinoma
7	Focal asymmetry	Irregular mass with indistinct margins	Not done	Ultrasound	Cribiform DCIS, low	Microinvasive ductal carcinoma
8	Focal asymmetry	Irregular mass with microlobulated margins	Regional NME	Ultrasound	Cribiform DCIS, intermediate	Invasive ductal carcinoma
9	Architectural distortion	Irregular mass with indistinct margins	Irregular mass with spiculated margins	Ultrasound	Cribiform and solid DCIS, high	Microinvasive DCIS
10	Round mass	Irregular mass with indistinct margins	Not done	Ultrasound	Cribiform DCIS, intermediate	Invasive ductal carcinoma
11	Asymmetry	Oval mass with indistinct margins	Not done	Ultrasound	Solid and cribiform DCIS with comedonecrosis, intermediate	Invasive ductal carcinoma
12	Negative	Oval mass with circumscribed margins	Oval mass	Ultrasound	Solid and micropapillary DCIS, intermediate	Foci of invasive ductal carcinoma
13	Focal asymmetry	Irregular mass with indistinct margins	Not done	Ultrasound	Comedo and cribiform DCIS, high	Invasive ductal carcinoma
14	Oval mass	Oval mass with circumscribed margins	Irregular mass	Ultrasound	Cribiform and solid DCIS with comedonecrosis and papilloma, intermediate	Invasive ductal carcinoma
15	Oval mass	Oval mass with microlobulated margins	Not done	Ultrasound	DCIS with mucinous features, intermediate	Invasive mucinous carcinoma

Bx, biopsy;DCIS, ductal carcinoma in-situ; NME, non-mass enhancement DCIS.

finding is NME. Given the increased upgrade rate observed, sonographic findings in NCDCIS should be considered a risk factor for upgrade and be considered in surgical planning. Further studies are needed to confirm the correlation between sonographic findings and upgrade rate in patients with NCDCIS due to our small sample size and single institution data.

REFERENCES

1. Breast Cancer Facts and Figures 2017-2018. American Cancer Society. 2017. Available from: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/breast-cancer-facts-and-figures-2017-2018.pdf [Accessed 1/13/2020].
2. Ikeda DM, Andersson I. Ductal carcinoma in situ: atypical mammographic appearances. *Radiology* 1989; **172**: 661–6. doi: https://doi.org/10.1148/radiology.172.3.2549563
3. Kim JH, Ko ES, Kim DY, Han H, Sohn J-H, Choe DH. Noncalcified ductal carcinoma in situ: imaging and histologic findings in 36 tumors. *J Ultrasound Med* 2009; **28**: 903–10. doi: https://doi.org/10.7863/jum.2009.28.7.903
4. Mun HS, Shine HJ, Kim HH, Cha JH. Screening-detected calcified and non-calcified ductal carcinoma in situ: differences in the imaging and histopathologic features. *Clin Radiol* 2013; **2012**: 27–35. doi: https://doi.org/10.1016/j.crad.2012.09.003
5. Scudder JM, Parikh J. Imaging features of noncalcified DCIS. *Breast Cancer Res* 2012; **14**(S1): 37. doi: https://doi.org/10.1186/bcr3292
6. Stomper PC, Connolly JL, Meyer JE, Harris JR. Clinically occult ductal carcinoma in situ detected with mammography: analysis of 100 cases with radiologic-pathologic correlation. *Radiology* 1989; **172**: 235–41. doi: https://doi.org/10.1148/radiology.172.1.2544922
7. Hogue J-C, Morais L, Provencher L, Desbiens C, Poirier B, Poirier Éric, et al. Characteristics associated with upgrading to invasiveness after surgery of a DCIS diagnosed using percutaneous biopsy. *Anticancer Res* 2014; **34**: 1183–91.
8. Sim YT, Litherland J, Lindsay E, Hendry P, Brauer K, Dobson H, et al. Upgrade of ductal carcinoma in situ on core biopsies to invasive disease at final surgery: a retrospective review across the Scottish breast screening programme. *Clin Radiol* 2015; **70**: 502–6. doi: https://doi.org/10.1016/j.crad.2014.12.019
9. Lamb LR, Kim G, Oseni TO, Bahl M. Noncalcified ductal carcinoma in situ (DCIS): rate and predictors of upgrade to invasive carcinoma. *Acad Radiol* 2021; **28**: e71–6. doi: https://doi.org/10.1016/j.acra.2020.02.011
10. D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA. ACR BI-RADS® Atlas. In: *Breast Imaging Reporting and Data System*. 5th ed. Reston: VA, American College of Radiology; 2013.
11. Cho KR, Seo BK, Kim CH, Whang KW, Kim YH, Kim BH, et al. Non-calcified ductal carcinoma in situ: ultrasound and mammographic findings correlated with histological findings. *Yonsei Med J* 2008; **49**: 103–10. doi: https://doi.org/10.3349/ymj.2008.49.1.103
12. Scoggins ME, Fox PS, Kuerer HM, Rauch GM, Benveniste AP, Park YM, et al. Correlation between sonographic findings and clinicopathologic and biologic features of pure ductal carcinoma in situ in 691 patients. *AJR Am J Roentgenol* 2015; **204**: 878–88. doi: https://doi.org/10.2214/AJR.13.12221
13. Park J-S, Park Y-M, Kim E-K, Kim S-J, Han S-S, Lee S-J, et al. Sonographic findings of high-grade and non-high-grade ductal carcinoma in situ of the breast. *J Ultrasound Med* 2010; **29**: 1687–97. doi: https://doi.org/10.7863/jum.2010.29.12.1687
14. Shin HJ, Kim HH, Kim SM, Kwon GY, Gong G, Cho OK. Screening-detected and symptomatic ductal carcinoma in situ: differences in the sonographic and pathologic features. *AJR Am J Roentgenol* 2008; **190**: 516–25. doi: https://doi.org/10.2214/AJR.07.2206
15. Moon WK, Myung JS, Lee YJ, Park IA, Noh D-Y, Im J-G. US of ductal carcinoma in situ. *Radiographics* 2002; **22**: 269–81. doi: https://doi.org/10.1148/radiographics.22.2.g02mr16269
16. Wang LC, Sullivan M, Du H, Feldman MI, Mendelson EB. US appearance of ductal carcinoma in situ. *Radiographics* 2013; **33**: 213–28. doi: https://doi.org/10.1148/rg.331125092
17. Jansen SA, Newstead GM, Abe H, Shimauchi A, Schmidt RA, Karczmar GS. Pure ductal carcinoma in situ: kinetic and morphologic Mr characteristics compared with mammographic appearance and nuclear grade. *Radiology* 2007; **245**: 684–91. doi: https://doi.org/10.1148/radiol.2453062061
18. Scott-Moncrieff A, Sullivan ME, Mendelson EB, Wang L. MR imaging appearance of noncalcified and calcified DCIS. *Breast J* 2018; **24**: 343–9. doi: https://doi.org/10.1111/tbj.12948
19. Greenwood HI, Heller SL, Kim S, Sigmund EE, Shaylor SD, Moy L. Ductal carcinoma in situ of the breasts: review of MRI imaging features. *Radiographics* 2013; **33**: 1569–88. doi: https://doi.org/10.1148/rg.336125055
20. van Luijt PA, Heijnsdijk EAM, Fracheboud J, Overbeek LIH, Broeders MJM, Wesseling J, et al. The distribution of ductal carcinoma in situ (DCIS) grade in 4232 women and its impact on overdiagnosis in breast cancer screening. *Breast Cancer Res* 2016; **18**: 47. doi: https://doi.org/10.1186/s13058-016-0705-5
21. Karayiannakis AJ, Bastounis EA, Chatzigianni EB, Makri GG, Alexiou D, Karamanakos P. Immunohistochemical detection of oestrogen receptors in ductal carcinoma in situ of the breast. *Eur J Surg Oncol* 1996; **22**: 578–82. doi: https://doi.org/10.1016/S0748-7983(96)92242-8
22. National Comprehensive Cancer Network. Breast Cancer (Version 6. 2020). Available from: https://www.nccn.org/professionals/physician_gls/pdf/breast.pdf [Accessed October 1 2020].

BJR

https://doi.org/10.1259/bjr.20201105

Received: 09 September 2020 | Revised: 07 December 2020 | Accepted: 17 December 2020

Cite this article as:
Whelehan P, Ali K, Vinnicombe S, Ball G, Cox J, Farry P, et al. Digital breast tomosynthesis: sensitivity for cancer in younger symptomatic women. *Br J Radiol* 2021; **94**: 20201105.

FULL PAPER

Digital breast tomosynthesis: sensitivity for cancer in younger symptomatic women

[1,2]PATSY WHELEHAN, DCR, MSc, [1]KULSAM ALI, PhD, [3]SARAH VINNICOMBE, FRCR, [4]GRAHAM BALL, PhD, [5]JULIE COX, FRCR, [6]PAUL FARRY, FRCR, [7]MAGGIE JENKIN, MSc, [8]KEITH LOWRY, FRCR, [9]STUART A MCINTOSH, FRCS, [1]RACHEL NUTT, PhD, [10]RACHEL OEPPEN, FRCR, [6]MICHAEL REILLY, FRCR, [10]MICHAELA STAHNKE, FRCR, [7]JIM STEEL, FRCR, [2]YEE TING SIM, FRCR, [1]VIOLET WARWICK, PhD, [11]LOUISE WILKINSON, FRCR, [4]DIMITRIOS ZAFEIRIS, PhD and [1]ANDREW J EVANS, FRCR

[1]School of Medicine, University of Dundee, Mailbox 4, Ninewells Hospital & Medical School, Dundee, UK
[2]NHS Tayside, Dundee, UK
[3]Gloucestershire Hospitals NHS Foundation Trust, Cheltenham, UK
[4]Nottingham Trent University, Nottingham, UK & Intelligent OMICS Ltd, Nottingham, UK
[5]South Tyneside and Sunderland NHS Foundation Trust, Sunderland, UK
[6]Western Health and Social Care Trust, Londonderry, UK
[7]Derriford Hospital, Plymouth, UK
[8]Belfast Health and Social Care Trust, Belfast, UK
[9]Queen's University, Belfast, UK
[10]University Hospital Southampton NHS Foundation Trust, Southampton, UK
[11]Oxford University Hospitals NHS Foundation Trust, Oxford, UK

Address correspondence to: Ms Patsy Whelehan
E-mail: *p.j.whelehan@dundee.ac.uk*

Objective: Full-field digital mammography (FFDM) has limited sensitivity for cancer in younger women with denser breasts. Digital breast tomosynthesis (DBT) can reduce the risk of cancer being obscured by overlying tissue. The primary study aim was to compare the sensitivity of FFDM, DBT and FFDM-plus-DBT in women under 60 years old with clinical suspicion of breast cancer.

Methods: This multicentre study recruited 446 patients from UK breast clinics. Participants underwent both standard FFDM and DBT. A blinded retrospective multireader study involving 12 readers and 300 mammograms (152 malignant and 148 benign cases) was conducted.

Results: Sensitivity for cancer was 86.6% with FFDM [95% CI (85.2–88.0%)], 89.1% with DBT [95% CI (88.2–90%)], and 91.7% with FFDM+DBT [95% CI (90.7–92.6%)]. In the densest breasts, the maximum sensitivity increment with FFDM +DBT over FFDM alone was 10.3%, varying by density measurement method. Overall specificity was 81.4% with FFDM [95% CI (80.5–82.3%)], 84.6% with DBT [95% CI (83.9–85.3%)], and 79.6% with FFDM +DBT [95% CI (79.0–80.2%)]. No differences were detected in accuracy of tumour measurement in unifocal cases.

Conclusions: Where available, DBT merits first-line use in the under 60 age group in symptomatic breast clinics, particularly in women known to have very dense breasts.

Advances in knowledge: This study is one of very few to address the accuracy of DBT in symptomatic rather than screening patients. It quantifies the diagnostic gains of DBT in direct comparison with standard digital mammography, supporting informed decisions on appropriate use of DBT in this population.

INTRODUCTION

Full-field digital mammography (FFDM) has limited sensitivity for breast cancer in younger females with denser breasts. Subgroup analysis in the DMIST (Diagnostic Performance of Digital *vs* Film Mammography for Breast-Cancer Screening) trial showed that the sensitivity of FFDM in females under 50 with dense breasts was only 59%.[1] Because DMIST was a screening trial, the cancers would have been smaller than those found in a symptomatic population. Lower mammographic sensitivity has been demonstrated in younger females presenting symptomatically in earlier studies with film-screen mammography: 67% on average in females under 60 years *vs* 87% in those aged 60–70.[2] Although evidence on sensitivity rates of FFDM in symptomatic populations is limited, a study in Germany has demonstrated that young age and dense breasts remain risk factors for false-negative mammography in symptomatic females in the digital era.[3]

Digital breast tomosynthesis (DBT) has the potential to alleviate the problem of cancers being masked on FFDM

by the dense breast tissue which is characteristic of younger breasts, because the technology partially separates overlapping structures.[4] It has been shown in a sample of patients with dense breasts and either screen-detected or symptomatically presenting lesions that DBT has a sensitivity of about 88% – a 10% increment over FFDM.[5] Sensitivity and other diagnostic performance parameters have rarely been compared in exclusively symptomatic patient samples. Two such studies have now been published but both involved only the Hologic Selenia Dimensions equipment (Hologic Inc., Marlborough, MA).[6,7] The study by Bian and colleagues, in females with dense breasts, found that sensitivity increased from 58.8% with FFDM to 68.1% with DBT, although no statistical test of this difference is reported.[6] In their sample of symptomatic patients not selected by breast density, Tang and colleagues found statistically significant improvements in sensitivity with FFDM plus DBT compared to FFDM alone, which they reported separately for each of two radiologists.[7] The sensitivity increments were in the order of 20%, with little change in specificity. Because DBT technology differs significantly between vendors, results from single-vendor studies, such as these two, are not necessarily generalisable to other equipment.

The aim of our multicentre study was primarily to compare the sensitivity for breast cancer of DBT, FFDM, and the two combined, using the Siemens Mammomat Inspiration unit (Siemens Healthcare GmbH, Erlangen, Germany) in females aged under 60 years presenting with symptoms or signs of possible breast cancer. Secondary aims were to compare specificity, differential sensitivity according to mammographic breast density and breast cancer type, and to compare accuracy for assessing tumour size.

METHODS AND MATERIALS

Approvals

The study was approved by the National Health Service (NHS) Research Ethics Service and received management approval in all participating institutions. The study was registered on a public database: ClinicalTrials.gov; NCT01241981.

Patients

Patients were recruited from specialist breast multidisciplinary clinics in five UK hospitals, to which they had been referred for investigation of breast symptoms. They were eligible if female, aged under 60 years, if they had an abnormality which the clinician performing physical examination graded as having a greater than 20% likelihood of malignancy, and if they were referred for and agreed to mammography. Patients classified as normal or benign on clinical examination ("P"-score one or two on a scale of 1–5) were excluded. The purpose was to achieve the requirements of the power calculation to detect a difference in sensitivity, while avoiding excessive recruitment overall. The upper age limit was informed by previous research on the sensitivity of mammography in symptomatic females of different age groups.[2] Patients aged over 25 but below the local age threshold for mammography to be used as a first-line imaging procedure (usually 40 years) but in whom ultrasound examination gave sufficient cause for suspicion to justify mammography were also eligible, irrespective of clinical suspicion.

Patients were excluded if they lacked capacity to give informed consent, were pregnant or lactating, or if they had obvious locally advanced breast cancer or severe co-morbidities expected to preclude surgical treatment. During the recruitment period, the DBT function was not cleared by the manufacturer for use on patients with breast implants. In some clinics, eligible patients were not approached because of logistical issues, *e.g.* equipment breakdown, no radiologist with DBT reporting training available in the clinic, or no-one available to take written informed consent. It was not feasible to keep records of patients who met the eligibility criteria but were not approached.

Following written informed consent, all participants underwent a combined examination consisting of bilateral FFDM and DBT on a Siemens Mammomat Inspiration unit. Both standard care imaging and DBT findings were taken into account in the real-time diagnostic triple assessment process.

Sample size

A power calculation for a χ^2 variance test was performed using Statistica v. 8 (StatSoft Inc., Tulsa, OK), assuming a population variance of 0.2. To detect a 25% reduction in the FFDM occult rate, from an expected 20% for mammography to 15% using DBT or FFDM +DBT, with a statistical power of 0.8, it was calculated that 150 participants with cancer were required. The numbers used reflected an element of uncertainty regarding the variance of the sample. The size of the difference to be detected was chosen based on the chief investigator's professional judgement on the level of benefit required to influence clinical practice. In order to include the required number of participants with cancer, 446 participants were recruited in total, of whom 154 had cancer.

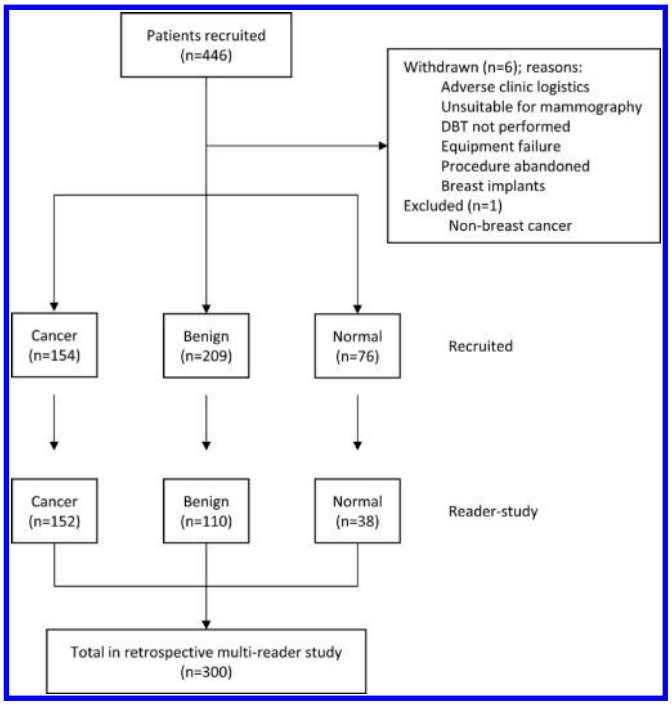
Retrospective multireader study

The retrospective reading exercise which is the subject of this manuscript included all the recruited cancer cases except for two in which we could not retrieve the full imaging data set from the recruiting site (n = 152). Randomised selection of normal and benign cases was undertaken to provide a total of 300 cases for inclusion in the reader study. Further details of the sample are provided in Figure 1 and in the Results section text. Randomised assignment of the 300 cases into batches of 50 was undertaken, which resulted in similar distributions per batch of patient age, and cancer, benign and normal cases.

The FFDM-only, DBT-only, and FFDM + DBT images for each batch of 50 cases were separately packaged with viewing software, and each batch of 50 cases was assigned to 2 readers from a pool of 12. Thus, each case was read twice under each of the three conditions (300 cases × 3 conditions × 2 readers = 1800 exam-reads in total). No reader read the same case twice. All readers read a total of three batches, one each of FFDM-only, DBT-only and FFDM + DBT. Allocation of specific batches to readers was randomised, as was the order in which they read their FFDM, DBT and FFDM + DBT batches.

All readers were trained and clinically experienced with Siemens DBT. 11 were consultant radiologists and 1 was a radiographer. Radiographers in the UK are able to undertake mammography

Figure 1. Participation flowchart. DBT, digital breast tomosynthesis.



interpretation and reporting, subject to recognised additional training and terms of employment.

Data collection in the reader study was facilitated by a computer-based tool designed for observer studies in mammography and tomosynthesis (Medical Extensible Viewer - "MedXViewer" - National Co-ordinating Centre for Physics in Mammography, Guildford, UK). MedXViewer enabled display of the images on five mega-pixel mammography reporting monitors with on-screen tools for the readers to mark regions of interest (ROIs), measure lesions where applicable, and describe abnormalities. In line with real-life practice, readers were provided with information on clinical presentation and patient age. They were instructed to ignore inconsequential benign radiological features that they would pass over in the clinic. For significant lesions, readers recorded a suspicion score according to the UK 1–5 scale, where 1 is normal and 5 is malignant.[8] Scores of 3–5 were considered malignant in the analysis of sensitivity for cancer (sometimes known as "complete" sensitivity).

In the FFDM reading condition, we also asked readers to provide a BI-RADS® (5th edition) breast density score: a: The breasts are almost entirely fatty; b: There are scattered areas of fibroglandular density; c: The breasts are heterogeneously dense, which may obscure small masses; d: The breasts are extremely dense, which lowers the sensitivity of mammography[9], and to use an on-screen 0–100 mm visual analogue scale (VAS) to assign an area-based percentage mammographic density to the mammogram, based on their impression of all images in the examination. The FFDM images were also subjected to software assessment of percent volumetric breast density using Volpara® Data Manager™ software (Volpara Solutions Ltd., Wellington, New Zealand), algorithm v. 1.5.0. The value used for analysis was the mean of the per-image output values for the images in the FFDM examination of the non-cancer-bearing breast. Bilateral cancer cases (n = 5) and participants with cancer with only one breast examined (n = 2) were therefore excluded from this subanalysis. Volpara data were missing for six cases because the raw DICOM images required for software processing were unobtainable.

Readers measured lesion size using an on-screen ruler. For analysis of the relative accuracy of malignant lesion measurements in patients with unifocal cancer, only the FFDM alone and DBT alone reading conditions were included. Reader measurements were compared to the histopathological whole tumour diameter (WTD). Patients treated with neoadjuvant systemic therapy were excluded from the disease-extent analyses.

Ground truth

The ground truth was established from the results of triple assessment (clinical examination, medical imaging and histopathological examination as applicable). Using the MedXViewer software, the mammograms for each case were annotated and the ground truth recorded (malignant, benign or normal) by one of two senior consultant radiologists from the pool of readers. They were provided with both the FFDM and DBT images and the triple assessment information to enable them to identify and classify the lesions. They marked each lesion by a generously sized freehand ROI on each view where it was visible, on the two modalities. If a malignant lesion known to be present was occult on FFDM and DBT, they marked its location based on the information available from ultrasound, MRI and histopathology findings. There were three such occult cases. When subsequently participating in the reader study, the two radiologists were only assigned cases on which they had not performed ground-truth marking.

The ground truth data and all the reader data were combined and exported from MedXViewer to a spreadsheet for analysis. Each lesion was assigned a unique identifier by MedXViewer, incorporating lesion-matching across different mammographic projection images. The readers' marks and interpretations captured by the software were automatically compared to the ground truth marks and diagnoses. Thus, the software recorded whether a reader had successfully detected a lesion and correctly identified it as malignant or benign. To score a true positive, the reader mark had to be within the corresponding generously-sized ROI applied at ground-truth marking.

Analytical and statistical methods

The performance of the modalities was based on sensitivity and specificity and the plotting of receiver operator characteristic curves (true positives *vs* false positives). The analyses were conducted at the per-breast level. In order to determine population variation a Monte Carlo subsampling approach was applied to the data, where the population was sampled 20 times for a randomised subset of 30–50% of the data set depending on the size of the data. The sensitivity and specificity of the results were then calculated for each of these Monte Carlo derived subsets. The variance in sensitivity and specificity and the confidence

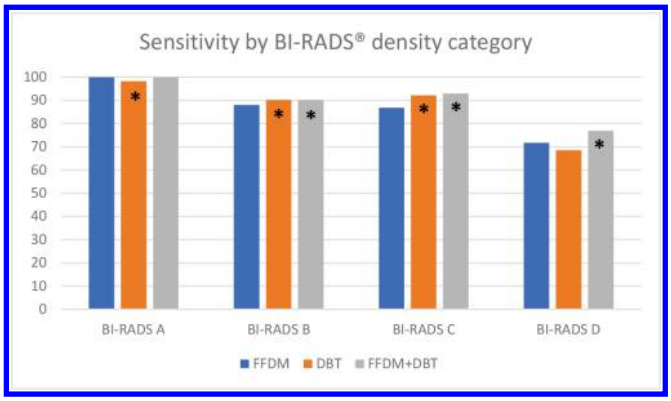
Table 1. Characteristics of patients included in the multireader study

	Patients: *n* = 300 Patients with cancer: *n* = 152 Breasts with cancer: *n* = 157
Mean patient age (range)	47 (24–60)
Mass as dominant radiological feature in malignancies	140/157 (89%)
Unifocal tumours	134/157 (85%)
Multifocal tumours	23/157 (15%)
Mean tumour size, unifocal breast cancers (range)	32 mm (5–95 mm)
Median tumour size, unifocal breast cancers	25 mm
DCIS	2/157 (1%)
Invasive (ductal) no special type, of which	127/157 (81%)
Grade 1	9/127
Grade 2	52/127
Grade 3	66/127
ILC	19/157 (12%)
Mixed ductal/lobular	2/157 (1%)
Other invasive carcinoma (Mucinous *n* = 3, one each tubular, micropapillary, Metaplastic, malignant phyllodes)	7/157 (4%)

DCIS, ductal carcinoma in situ; ILC, invasive lobular carcinoma.

intervals were calculated from this population of subsets. The same values were used to plot receiver operating characteristic (ROC) curves and calculate the area under the curve (AUC) for each simulation, utilising the ROCR package in R.[10] Significance between approaches was tested using a paired two-way Student's *t*-test on the Log normalised values.

Figure 2. Sensitivity (%) according to BI-RADS® density category. (Values significantly different to FFDM at $p < .05$ are denoted by asterisks). DBT, digital breast tomosynthesis; FFDM, full-field digital mammography.



RESULTS

Sample description

446 patients were recruited between March 2011 and April 2016. Figure 1 provides a recruitment flowchart and Table 1 shows the characteristics of the cases included in the retrospective multireader study.

Overall sensitivity

Sensitivity for breast cancer was 86.6% with FFDM [95% CI (85.2–88.0%)], 89.1% with DBT [95% CI (88.2–90.0%)], and 91.7% with FFDM + DBT [95% CI (90.7–92.6%)]. Comparing the values by *t*-test, the differences in sensitivity for cancer between modalities were statistically significant - FFDM *vs* DBT: $p = .004$; DBT *vs* FFDM + DBT: $p < .001$; FFDM *vs* FFDM + DBT: $p < .001$.

In the reader study, there were four cases picked up by FFDM but not by either reader with DBT. The features were as follows: ill-defined mass, $n = 2$; well-defined mass, $n = 1$; lobulated mass with associated calcifications, $n = 1$. There were eight cases picked up by DBT but not by either reader with FFDM. The features were as follows: spiculated mass, $n = 5$; well-defined mass, $n = 1$; ill-defined mass, $n = 1$; ill-defined mass with associated calcifications, $n = 1$.

Sensitivity according to mammographic density

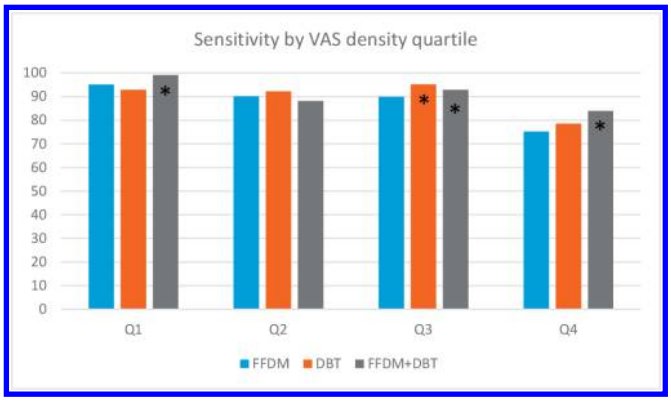
For each mammogram, there were two reader classifications using the BI-RADS® four-category density system.[9] The two BI-RADS® values per patient were applied to all observer-readings. The following distribution of BI-RADS® density categories was seen ($n = 157$ breasts with cancer x 2 BI-RADS® reads; total $n = 314$): category A (almost entirely fatty), $n = 23$ (7%); category B (scattered areas of fibroglandular density), $n = 132$ (42%); category C (heterogeneously dense), $n = 128$ (41%); category D (extremely dense), $n = 31$ (10%). Agreement between the readers on the BI-RADS® category for each patient was 62%. Variations in percentage cancer sensitivity according to BI-RADS® density category are shown in Figure 2.

For the 0–100 VAS values (observers' assessments of percentage dense area estimated for the mammogram overall), the mean of the two readers' scores was used and was applied to all breasts for the analysis. The data were divided into quartiles and the ranges for each quartile were as follows: – Q1: 4–31, Q2: 32–41, Q3: 42–63, Q4: 64–86. Variations in percentage cancer sensitivity according to VAS density are shown in Figure 3.

For volumetric percentage breast density assessed by Volpara® software, the mean of the per-image values for each patient (non-cancer-bearing breast only) was used for analysis, and the data were divided into quartiles. Ranges within the quartiles were as follows: Q1: 2.37–4.87, Q2: 4.91–7.03, Q3: 7.18–13.09, Q4: 13.15–39.05. Patients with bilateral cancer were excluded ($n = 5$) and Volpara® data were unavailable for six patients. Variations in percentage cancer sensitivity according to Volpara® density are shown in Figure 4.

In summary, decreased sensitivity with increasing breast density was less marked with DBT than with FFDM. By all three density

Figure 3. Sensitivity (%) according to VAS percent density quartile. (Values significantly different to FFDM at $p < .05$ are denoted by asterisks). DBT, digital breast tomosynthesis; FFDM, full-field digital mammography; VAS, visual analogue scale



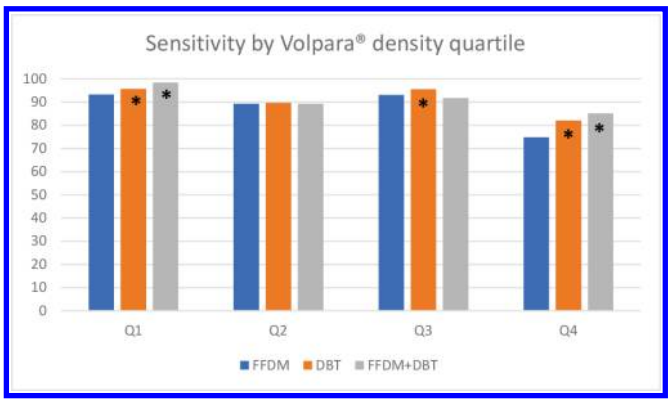
measures, FFDM + DBT was more sensitive than FFDM in the most dense category, whereas the advantage of DBT alone was most apparent in the third most dense category. Only the automated density assessment method (Volpara®) showed a statistically significant sensitivity increment in the most dense breasts for DBT alone: DBT 82.0 *vs* 74.8% for FFDM, $p < .001$.

The largest subgroup benefit detected in the study was the 10.3% sensitivity increment seen in the densest breasts according to the Volpara® measurement (85.1% with FFDM +DBT *vs* 74.8% with FFDM alone, $p < .001$).

Sensitivity in different tumour types

Analysing sensitivity separately for the invasive lobular (ILC) and combined non-lobular invasive cancers revealed no statistically significant differences between modalities in the lobular group: FFDM: 84.1% [95% CI (80.4–87.8%)]; DBT: 85.6% (82.0–89.1%); FFDM + DBT: 87.7% (84.9–90.4%). *t*-test results were: FFDM *vs* DBT: $p = .55$; FFDM *vs* FFDM + DBT: $p = .11$; DBT *vs* FFDM + DBT: $p = .33$.

Figure 4. Sensitivity (%) according to Volpara® density quartiles. (Values significantly different to FFDM at $p < .05$ are denoted by asterisks). DBT, digital breast tomosynthesis; FFDM, full-field digital mammography



The results for non-lobular invasive cancer sensitivity, which were overwhelmingly the larger group (136 breasts *vs* 19 breasts), closely reflect the overall results: FFDM: 86.3% [95% CI (85.6–87.1%)]; DBT: 89.4% (88.7–90.2%); FFDM + DBT: 90.7% (90.0–91.4%). *t*-test results were: FFDM *vs* DBT: $p < .001$; FFDM *vs* FFDM + DBT: $p < .001$; DBT *vs* FFDM + DBT: $p = .01$.

Specificity

Specificity was 81.4% with FFDM [95% CI (80.5–82.3%)], 84.6% with DBT [95% CI (83.9–85.3%)], and 79.6% with FFDM +DBT [95% CI (79.0–80.2%)]. Differences were statistically significant by *t*-test at: FFDM *vs* DBT: $p < .001$; FFDM *vs* FFDM +DBT: $p = .003$; DBT *vs* FFDM +DBT: $p < .001$. Of note, in the subgroup with the highest sensitivity gain using the FFDM +DBT modality (*i.e.* a 10% sensitivity increment in cases with breast density in the highest Volpara® quartile), there was no specificity penalty: 87.4% with FFDM [95% CI (85.8–89.0%)], and 87.3% with FFDM +DBT (85.5–89.2%), $p = .94$.

Receiver operating characteristic analysis

The AUC for FFDM was 0.90; for DBT it was 0.92; for FFDM +DBT it was 0.92.

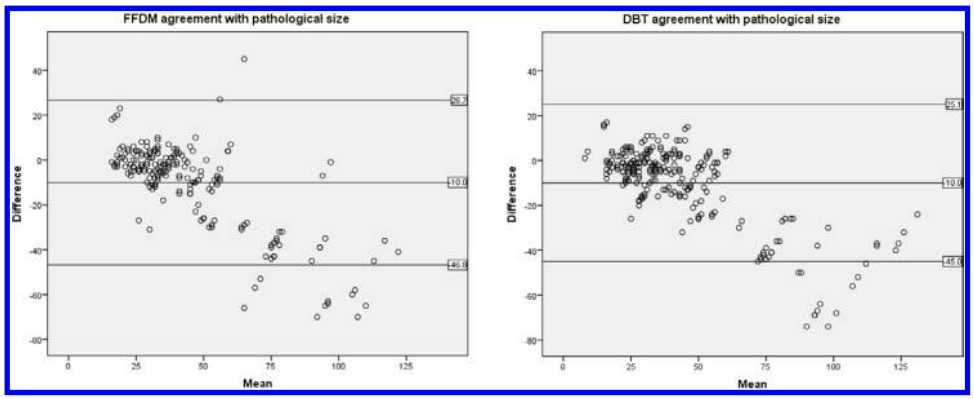
Assessment of tumour size in unifocal cancer cases

There were 214 reader measurements of unifocal malignant lesions not treated with neoadjuvant systemic therapy under the FFDM condition and 260 under the DBT condition. (The difference in numbers reflects the higher sensitivity of DBT.) Absolute agreement between reader measurements and histopathology measurements by intraclass correlation coefficient was 0.41 for FFDM [95% CI (0.25–0.60)] and 0.55 for DBT [95% CI (0.28–0.70)]. The rate of overestimation of histopathological WTD by more than 5 mm was 10.3% with FFDM and 8.5% with DBT ($p = .82$). The rate of underestimation by more than 5 mm was 47.2% with FFDM and 46.2% with DBT ($p = .50$). The mean discrepancy between the readers' measurements and the histopathological WTD was identical for the two modalities – a 10 mm underestimation – and Bland–Altman 95% limits of agreement were very similar at +26.7 to −46.8 for FFDM and +25.1 to −45.0 for DBT. Please see Figure 5 for Bland–Altman plots. Common to both modalities was a tendency for greater underestimation with increasing lesion size. ILC and lesions which included radiological microcalcifications were over-represented in the top quintile of size underestimation.

DISCUSSION

Our study contributes to the limited body of evidence evaluating the effectiveness of modern mammographic imaging in patients presenting with symptoms of possible breast cancer. While it can be argued that the importance of any single imaging modality is limited in the context of multimodality diagnostic breast clinics, even triple assessment does not completely eliminate false-negative findings in symptomatic patients,[11,12] so the sensitivity performance of each individual element still matters. Furthermore, if mammography is negative in the presence of malignant clinical or sonographic findings, or if there is size discrepancy between assessment methods, there can be a tendency to resort

Figure 5. Bland–Altman plots for agreement between imaging tumour size and final histopathological size (unifocal only). DBT, digital breast tomosynthesis; FFDM, full-field digital mammography.



to an expensive MRI scan, despite limited evidence supporting MRI for diagnostic problem-solving.[13]

Overall sensitivity for cancer was high with all three modalities in our study (FFDM: 86.6% - FFDM +DBT: 91.7%). Given that the sensitivity of FFDM was so high, it is unsurprising that the overall gains from adding DBT were clinically relatively modest. A recent study in the symptomatic setting was conducted in China in a sample where 149 of 197 participants had BI-RADS® C or D density.[7] That study, using Hologic Selenia Dimensions equipment, did not compare FFDM with DBT alone but found that sensitivity for FFDM in their whole sample was 72% – much lower than in our sample – and for FFDM and DBT combined was 91% – similar to our value. Information on tumour size was not provided in the publication. Diagnostic studies with mixed samples of screen-detected and symptomatic lesions, using a prototype GE tomosynthesis device (GE Healthcare, Chicago, IL), have also shown lower FFDM sensitivity and higher sensitivity gains with DBT than were seen in our study.[14,15]

Differential sensitivity by breast density in our study varied according to method of assessing density but the overall pattern is for the sensitivity gains from DBT to be more apparent in denser breasts. Again, however, our FFDM performance compares favourably with published values. In a recent study[6] comparing FFDM and DBT using Hologic Selenia Dimensions equipment in a symptomatic population with dense breasts (BI-RADS® C or D), sensitivity for cancer was considerably lower with both modalities than for females with dense breasts in our study, at 59% for FFDM (*vs* a mean sensitivity in our BI-RADS® C and D cases of 79%) and 68% (*vs* 80%) for DBT. The mean tumour size of 23 mm in that study, compared to 32 mm in ours, may help explain the generally lower sensitivity.

In ILC, descriptively FFDM +DBT gave a 3.6% increment over FFDM alone but there were no statistically significant differences in sensitivity for ILC between modalities, possibly because there were only 19 cases of ILC in our study. A previous larger multireader study including screen-detected and symptomatic cases of ILC, using Hologic Selenia Dimensions equipment, found a statistically significant 15% sensitivity increment with DBT (85% *vs* 70%).[16] It has also been shown that reader ratings of lesion conspicuity in ILC are higher with DBT than FFDM.[17]

Specificity in our study was about 3% higher for DBT *vs* FFDM, with a 5% drop in overall specificity for FFDM + DBT. We think the lower specificity with the combined modalities is most likely just a function of having two tests instead of one. Tang and colleagues[7] found no difference in specificity for FFDM *vs* FFDM +DBT, with generally lower specificity than ours, at 72 and 71% respectively. Bian and colleagues[6] achieved higher specificity for both FFDM and DBT alone, rising to 95% for DBT, which may further explain the relatively low sensitivity values in that study. DBT has been shown to improve specificity in screening studies[18] but maximising specificity is less important in the symptomatic triple assessment clinic than in the screening of well females, especially in females with clinical suspicion of cancer as in this study.

We detected only very small differences in AUC values (0.90 for FFDM and 0.92 for DBT), similar to the study by Tang and colleagues,[7] which demonstrated an improvement from 0.85 to 0.9.

Accurate estimation of tumour extent is important in guiding therapeutic decision-making. Our study detected little descriptive improvement and no statistically significant improvement in the accuracy of measuring the size of unifocal cancers with DBT compared to FFDM. Conversely, several earlier studies have found DBT size assessment to be more accurate than FFDM,[19,20] although it has also been shown that there is a greater risk of overestimation of tumour size with tomosynthesis,[21,22] which was not our experience in this study. Our finding that underestimation of tumour size in ILC persists with DBT is in line with previous work,[17] but that study also included only a small number of ILC cases.

Study strengths and limitations

Study strengths included the use of multiple centres and multiple readers, and the strict blinding of readers between modalities. The inclusion of multiple measures of breast density was also a strength. It could be considered a limitation that the images were

read under simulated rather than real-life practice conditions, but that approach was necessary in order to conduct a robustly blinded study. Synthesised 2D images which can be used in place of standard FFDM were not available at the time of image acquisition for our study. We did not follow up the patients to ascertain false-negative triple assessment cases, therefore the study assesses the relative sensitivity of the modalities. Because ours was a study of patients presenting with suspicious clinical symptoms, it does not add to the evidence base on the clinical utility of DBT in DCIS. Like others of its kind, ours was a single-vendor study and results may not be generalisable to other vendors' equipment.

CONCLUSIONS

FFDM and DBT in combination provided a small but statistically significant improvement in sensitivity for cancer in our sample of younger symptomatic patients, from 86.6 to 91.7% overall.

The greatest improvements in sensitivity, over FFDM alone, were seen with the combined modality in the densest breasts (an increment of 9% when density was measured by human-assigned area-based percentage, and of 10% when density was measured by Volpara® software).

The overall sensitivity improvement with combined FFDM and DBT was at the cost of a small reduction in specificity, from 81.4 to 79.6%.

No advantage was seen for assessment of unifocal tumour size.

Although our study has not shown FFDM to be sufficiently inferior to mandate the replacement or supplementation of FFDM with DBT for all younger females in the symptomatic clinic, where it is available it does merit first-line use in the under 60 age group, particularly in females who are known to have very dense breasts. If breast density is not known in advance from prior mammography, DBT could be performed after negative FFDM in females with dense breasts, rather than in combination at the outset.

The benefits of DBT should be weighed against the additional radiation dose, acquisition time, reading time and data storage costs. The contribution of DBT to triple assessment in symptomatic females with dense breasts needs to be reassessed in comparison with the performance of other potential diagnostic tests such as increasingly available contrast-enhanced mammography.

ACKNOWLEDGEMENTS

We gratefully acknowledge the patients who participated in this study. A large number of administrative, clinical and trials staff at the participating sites supported the study; their invaluable contributions are gratefully acknowledged. We thank Dr Melissa Hill of Volpara Health Technologies Ltd. for helpful comments on the draft manuscript, and Dr Mark Halling-Brown and colleagues at the National Co-ordinating Centre for Physics in Mammography for providing and tailoring the MedXViewer software and for associated data management support.

DECLARATIONS OF INTEREST

Eighteen of 19 co-authors have no conflict of interest to declare. One author (SV) provides paid expert feedback to Siemens Healthcare GmbH on product design.

FUNDING

The equipment upgrade required for the study was funded, for the lead site only, by Ninewells Cancer Campaign. Further small grants to support the running of the study were provided by Tenovus, the Sir John Fisher Foundation, and a local fundraiser for the lead site, Mrs Fiona Edwards. None of the funders had input to the design or reporting of the study. The study was supported in-kind by the University of Dundee, by the host NHS organisations, and by the NHS clinical research support services and networks in England, Scotland and Northern Ireland.

REFERENCES

1. Pisano ED, Hendrick RE, Yaffe MJ, Baum JK, Acharyya S, Cormack JB, et al. Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. *Radiology* 2008; **246**: 376–83. doi: https://doi.org/10.1148/radiol.2461070200
2. Sibbering DM, Burrell HC, Evans AJ, Yeoman LJ, Wilson ARM, Robertson JFR, et al. Mammographic sensitivity in women under 50 years presenting symptomatically with breast cancer. *The Breast* 1995; **4**: 127–9. doi: https://doi.org/10.1016/0960-9776(95)90008-X
3. Häberle L, Fasching PA, Brehm B, Heusinger K, Jud SM, Loehberg CR, et al. Mammographic density is the main correlate of tumors detected on ultrasound but not on mammography. *Int J Cancer* 2016; **139**: 1967–74. doi: https://doi.org/10.1002/ijc.30261
4. Hooley RJ, Durand MA, Philpotts LE. Advances in digital breast tomosynthesis. *AJR Am J Roentgenol* 2017; **208**: 256–66. doi: https://doi.org/10.2214/AJR.16.17127
5. Chae EY, Kim HH, Cha JH, Shin HJ, Choi WJ. Detection and characterization of breast lesions in a selective diagnostic population: diagnostic accuracy study for comparison between one-view digital breast tomosynthesis and two-view full-field digital mammography. *Br J Radiol* 2016; **89**: 20150743. doi: https://doi.org/10.1259/bjr.20150743
6. Bian T, Lin Q, Cui C, Li L, Qi C, Fei J, et al. Digital breast tomosynthesis: a new diagnostic method for Mass-Like lesions in dense breasts. *Breast J* 2016; **22**: 535–40. doi: https://doi.org/10.1111/tbj.12622
7. Tang W, Hu F-X, Zhu H, Wang Q-F, Gu Y-J, Peng W-J. Digital breast tomosynthesis plus mammography, magnetic resonance imaging plus mammography and mammography alone: a comparison of diagnostic performance in symptomatic women. *Clin Hemorheol Microcirc* 2017; **66**: 105–16. doi: https://doi.org/10.3233/CH-16242
8. Maxwell AJ, Ridley NT, Rubin G, Wallis MG, Gilbert FJ, Michell MJ, et al. The Royal College of Radiologists Breast Group breast imaging classification. *Clin Radiol* 2009; **64**:

624–7. doi: https://doi.org/10.1016/j.crad.2009.01.010
9. D'Orsi C, Sickles E, Mendelson E, Morris E, Bassett L. *ACR BI-RADS atlas*. 5th ed. Reston, VA: American College of Radiology; 2013.
10. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. *Bioinformatics* 2005; **21**: 3940–1. doi: https://doi.org/10.1093/bioinformatics/bti623
11. Britton P, Duffy SW, Sinnatamby R, Wallis MG, Barter S, Gaskarth M, et al. One-stop diagnostic breast clinics: how often are breast cancers missed? *Br J Cancer* 2009; **100**: 1873–8. doi: https://doi.org/10.1038/sj.bjc.6605082
12. Coolen A, Leunen K, Menten J, van Steenbergen W, Neven P. False-Negative tests in breast cancer management. *Neth J Med* 2011; **69**: 324–9.
13. Lehman CD, Lee AY, Lee CI. Imaging management of palpable breast abnormalities. *AJR Am J Roentgenol* 2014; **203**: 1142–53. doi: https://doi.org/10.2214/AJR.14.12725
14. Chan H-P, Helvie MA, Hadjiiski L, Jeffries DO, Klein KA, Neal CH, et al. Characterization of breast masses in digital breast Tomosynthesis and digital mammograms. *Acad Radiol* 2017; **24**: 1372–9. doi: https://doi.org/10.1016/j.acra.2017.04.016
15. Chae EY, Kim HH, Cha JH, Shin HJ, Choi WJ. Detection and characterization of breast lesions in a selective diagnostic population: diagnostic accuracy study for comparison between one-view digital breast tomosynthesis and two-view full-field digital mammography. *Br J Radiol* 2016; **89**: 20150743. doi: https://doi.org/10.1259/bjr.20150743
16. Mariscotti G, Durando M, Houssami N, Zuiani C, Martincich L, Londero V, et al. Digital breast tomosynthesis as an adjunct to digital mammography for detecting and characterising invasive lobular cancers: a multi-reader study. *Clin Radiol* 2016; **71**: 889–95. doi: https://doi.org/10.1016/j.crad.2016.04.004
17. Chamming's F, Kao E, Aldis A, Ferré R, Omeroglu A, Reinhold C, et al. Imaging features and conspicuity of invasive lobular carcinomas on digital breast tomosynthesis. *Br J Radiol* 2017; **90**: 20170128. doi: https://doi.org/10.1259/bjr.20170128
18. Houssami N, Macaskill P, Marinovich ML, Hunter KE. Breast cancer screening using tomosynthesis or mammography: a meta-analysis of cancer detection and recall. *JNCI J Natl Cancer Inst [Internet]* 2018; **110**: 942–9.
19. Luparia A, Mariscotti G, Durando M, Ciatto S, Bosco D, Campanino PP, et al. Accuracy of tumour size assessment in the preoperative staging of breast cancer: comparison of digital mammography, tomosynthesis, ultrasound and MRI. *Radiol Med* 2013; **118**: 1119–36. doi: https://doi.org/10.1007/s11547-013-0941-z
20. Helal MH, Mansour SM, Zaglol M, Salaleldin LA, Nada OM, Haggag MA. Staging of breast cancer and the advanced applications of digital mammogram: what the physician needs to know? *Br J Radiol* 2017; **90**: 20160717. doi: https://doi.org/10.1259/bjr.20160717
21. Mercier J, Kwiatkowski F, Abrial C, Boussion V, Dieu-de Fraissinette V, Marraoui W, et al. The role of tomosynthesis in breast cancer staging in 75 patients. *Diagn Interv Imaging* 2015; **96**: 27–35. doi: https://doi.org/10.1016/j.diii.2014.06.010
22. Marinovich ML, Bernardi D, Macaskill P, Ventriglia A, Sabatino V, Houssami N. Agreement between digital breast tomosynthesis and pathologic tumour size for staging breast cancer, and comparison with standard mammography. *Breast* 2019; **43**: 59–66. doi: https://doi.org/10.1016/j.breast.2018.11.001

BJR

https://doi.org/10.1259/bjr.20201166

Received: 27 September 2020 | Revised: 29 November 2020 | Accepted: 05 December 2020

Cite this article as:
Chen ST, Covelli J, Okamoto S, Daniel BL, DeMartini WB, Ikeda DM. Clumped vs non-clumped internal enhancement patterns in linear non-mass enhancement on breast MRI. *Br J Radiol* 2020; **94**: 20201166.

FULL PAPER

Clumped vs non-clumped internal enhancement patterns in linear non-mass enhancement on breast MRI

[1]SHU TIAN CHEN, MD, [2]JAMES COVELLI, MD, [3]SATOKO OKAMOTO, MD, PhD, [2]BRUCE L DANIEL, MD, [2]WENDY B DEMARTINI, MD and [2]DEBRA M IKEDA, MD

[1]Department of Diagnostic Radiology, Chang-Gung Memorial Hospital, Chiayi, Taiwan
[2]Department of Radiology, Stanford University School of Medicine, California, United States
[3]Department of Radiology, St. Marianna University School of Medicine, Kawasaki, Japan

Address correspondence to: Ms Shu Tian Chen
E-mail: *sugarcan99@hotmail.com*

Objective: To compare positive predictive values (PPVs) of clumped *vs* non-clumped (homogenous and heterogeneous) internal enhancement on MRI detected linear non-mass enhancement (NME) on MRI-guided vacuum-assisted breast biopsy (MRI-VABB).
Methods: With IRB (Institutional Review Board) approval, we retrospectively reviewed 598 lesions undergoing MRI-VABB from January 2015 to April 2018 that showed linear NME. We reviewed the electronic medical records for MRI-VABB pathology, any subsequent surgery and clinical follow-up. The X^2 test was performed for univariate analysis.
Results: There were 120/598 (20%) linear NME MRI-VABB lesions with clumped (52/120, 43%) *vs* non-clumped (68/120, 57%) internal enhancement, average size 1.8 cm (range 0.6–7.6 cm). On MRI-VABB, cancer was identified in 22/120 (18%) lesions, ductal carcinoma *in situ* (DCIS) was found in 18/22 (82%) and invasive cancer in 4 (18%). 3/31 (10%) high-risk lesions upgraded to DCIS at surgery, for a total of 25/120 (21%) malignancies. Malignancy was found in 12/52 (23%) clumped lesions and in 13/68 (19%) of non-clumped lesions that showed heterogeneous (5/13, 38%) or homogenous (8/13, 62%) internal enhancement. The PPV of linear NME with clumped internal enhancement (23.1%) was not significantly different from the PPV of non-clumped linear NME (19.1%) (p = 0.597). The PPV of linear NME lesions <1 cm (33.3%) was not significantly different from the PPV of lesions ≥1 cm (18.6%) (p = 0.157).
Conclusions: Linear NME showed malignancy in 21% of our series. Linear NME with clumped or non-clumped internal enhancement patterns, regardless of lesion size, might need to undergo MRI-VABB in appropriate populations.
Advances in knowledge: Evaluation of linear NME lesions on breast MRI focuses especially on internal enhancement pattern.

INTRODUCTION

Clinical indications for contrast-enhanced MRI of the breast include screening of high-risk populations, evaluating extent of disease in patients with known breast cancer, or workup of abnormal imaging on mammography or ultrasound.[1–3] Although breast MRI is the most sensitive imaging modality for detecting breast cancer,[4] it has lower specificity for breast cancer due to overlap of features of benign and malignant lesions.[5,6] Because MR-guided biopsies are invasive, time-consuming and expensive, it is important to define criteria for suspicious breast masses needing biopsy.

The fifth edition of the Breast Imaging and Reporting Data System classifies abnormal enhancement into three types: focus, mass, and nonmass enhancement (NME).[7] Lesion type, abnormality morphology, and kinetics were used to discriminate benign from malignant lesions.[8] Linear distribution is of the BI-RADS distribution descriptors of NME and can be subclassified into internal enhancement patterns of clumped, homogeneous, heterogeneous, and clustered ring.[7] Among NME lesions, segmental and clumped linear enhancement patterns are seen more frequently as suspicious characteristics for malignancy.[9] However, the positive predictive values (PPVs) of linear NME varied among the previous studies, and the distribution and internal characteristics are often interwoven.[9–12]

Currently, there are few studies on linear NME lesions focusing only on internal enhancement pattern and research is required to identify the associations with malignancy on this topic. We aim to compare the PPVs of clumped internal

enhancement characteristics versus non-clumped (homogeneous, heterogeneous, and clustered ring) on linear NME on MRI-guided vacuum-assisted breast biopsy (VABB).

METHODS AND MATERIALS

Study population

This Health Insurance Portability and Accountability Act (HIPAA) compliant single institution study was approved by our institutional review board and granted a waiver of informed consent. We retrospectively reviewed the imaging, pathology, and radiology reports of breast lesions in patients who underwent 9-gauge MRI-guided VABB at our institution from January 2015 to April 2018.

MRI acquisition

MRI was performed in the prone position with a 16-channel dedicated breast coil on a 3T system (GE Healthcare, Milwaukee, WI, USA). Each study included an axial T1W 3D gradient echo, axial T2W fat-suppressed images (CUBE), axial diffusion-weighted 2D EPI images with $B = 0$ and $B = 600$, and axial multiphase centrically encoded 3D T1W SPGR (Spoiled gradient recalled echo) images with Dixon separation of fat and water into separate images: a pre-contrast high-resolution image, 14 view shared rapid dynamic contrast-enhanced (DCE) images (15 sec each) with bolus power injection of 0.1 mmol/kg of gadobutrol (Gadovist, Bayer Health Care, Berlin, Germany) and a 20-ml saline flush 15 s after start of the acquisition, four high-resolution post-contrast phases (150 s each). Multiplanar reformation, thin slab maximal intensity projection (MIP) images and rotating MIP images were created from the first high-resolution post-contrast "peak" phase. Kinetic enhancement curves were performed by placing a region of interest on suspicious enhancing lesions using Aegis (Sentinelle Medical Inc., Toronto, Canada) or by DynaCAD (Invivo, Gainesville, FL, USA) system.

MRI-guided vacuum-assisted breast biopsies (VABB) were obtained on either a 1.5T or 3T scanner (Signa, GE Healthcare, Milwaukee, WI, USA) with 16-channel dedicated breast table top Sentinelle breast coils (Invivo, Gainesville, FL, USA) and open grid. Targeting was obtained using computer software (DynaCAD, Invivo, Gainesville, FL, USA, or Aegis, Sentinelle Medical, Inc, Toronto, Canada). Contrast enhancement of the target was confirmed after a bolus i.v. injection 0.1 mmol/kg of gadolinium contrast. All biopsies were performed with a 9-gauge VABB system (ATEC; Hologic Inc., Bedford, MA, USA), followed by placement of a titanium or MRI compatible stainless steel marker at the site of biopsy. For all patients, a post-biopsy mammogram was performed to confirm the position of the clip relative to the biopsy site.

MRI interpretation and data acquisition

All breast MRI studies were reported according to the fifth edition of the ACR BI-RADS MRI lexicon[7] and were initially interpreted by one of the seven fellowship trained MQSA-certified radiologists in the breast imaging department. Two breast-imaging radiologists (JC and SO) reviewed the MR images on a GE workstation to confirm that lesions undergoing MRI-guided VABB were comprised of linear non-mass enhancement, and noting that only lesions that were given a BI-RADS category four or five were biopsied. Lesion and patient characteristics were recorded in a HIPAA compliant data base including: examination indication, patient age, personal history of breast cancer, lesion size, internal enhancement characteristics, kinetic features prompting biopsy, MRI-guided breast VABB core biopsy results, and subsequent surgery results or >1 year imaging follow-up after biopsy. Lesions were separated into clumped and non-clumped (which include heterogeneous, homogeneous, clustered ring) internal enhancement characteristic categories. Enhancement kinetic curve descriptions were based on BI-RADS terminology of initial phase (slow, medium, and fast) and delayed phase (persistent, plateau, and washout), and recorded the most suspicious kinetic curve (initial, delay, or both), and recorded as described on the initial report by the original interpreting radiologist.

Statistical analysis

Statistical Package for Social Sciences (SPSS) for Windows, Statistics v.16.0. (SPSS Inc., Chicago, USA) was used to perform statistical analyses. The PPV of each internal enhancement characteristic was first calculated using standard averages and percentages. Exact 95% CI were calculated for each PPV. The X^2 test was used for univariate analysis. *P*-values less than 0.05 were considered statistically significant. Inter-observer agreement was examined using the intraclass correlation coefficient (ICC).

RESULTS

Study populations

There were 598 MRI-VABB biopsies in the study period of which 120/598 (20%) showed linear NME, which comprise the study group. The ICC showed a good inter-observer reliability for the interpretation of linear NME on breast MRI (average measure of the ICC = 0.85). The indications of the breast MRI were screening of high-risk populations 60/120 (50%), staging 37/120 (30.8%), workup of abnormal imaging on mammography or ultrasound 22/120 (18.3%), and treatment-response assessment after neoadjuvant chemotherapy 1/120 (0.9%). Of these 120 linear NME lesions, there were 25 (20.8%) cancers, four invasive ductal carcinoma (IDC) and 21 ductal carcinoma *in situ* (DCIS) (Figure 1) (Table 1), and which will be described in detail subsequently. The mean patient age was 54.1 years old (ranged 27–82 years). The average number of core biopsies per lesion was nine cores (range: 4–17 cores). 53/120 (44%) NME lesions were operated upon. Of the remaining 67/120 (54%) NME not operated upon, two benign findings found at VABB were lost to follow up, and 65 (64 benign, one high-risk findings at VABB) showed no suspicious imaging and/or clinical features on follow up. The mean interval between VABB and imaging follow-up is 1.7 years.

Lesion characteristics on MR imaging

The average size of the linear NME lesions was 1.8 cm (range 0.6–7.6 cm). Of these, 52/120 (43%) had clumped internal enhancement characteristics and 68/120 (57%) were non-clumped. Non-clumped internal characteristics were heterogeneous, homogenous, and clustered ring, but there were no lesions with clustered ring internal enhancement in this population. Of the 68 lesions with non-clumped internal enhancement characteristics, 35/68 (51%) showed heterogeneous internal

Figure 1. The pathological result of the 120 linear non-mass enhancement (NME) lesions undergoing vacuum-assisted breast biopsy (VABB). ALH, Atypical lobular hyperplasia; ADH, Atypical ductal hyperplasia; DCIS, Ductal carcinoma *in situ*; IDC, Invasive ductal carcinoma; LCIS, Lobular carcinoma *in situ*; RS/CSL, Radial scar/complex sclerosing lesion.

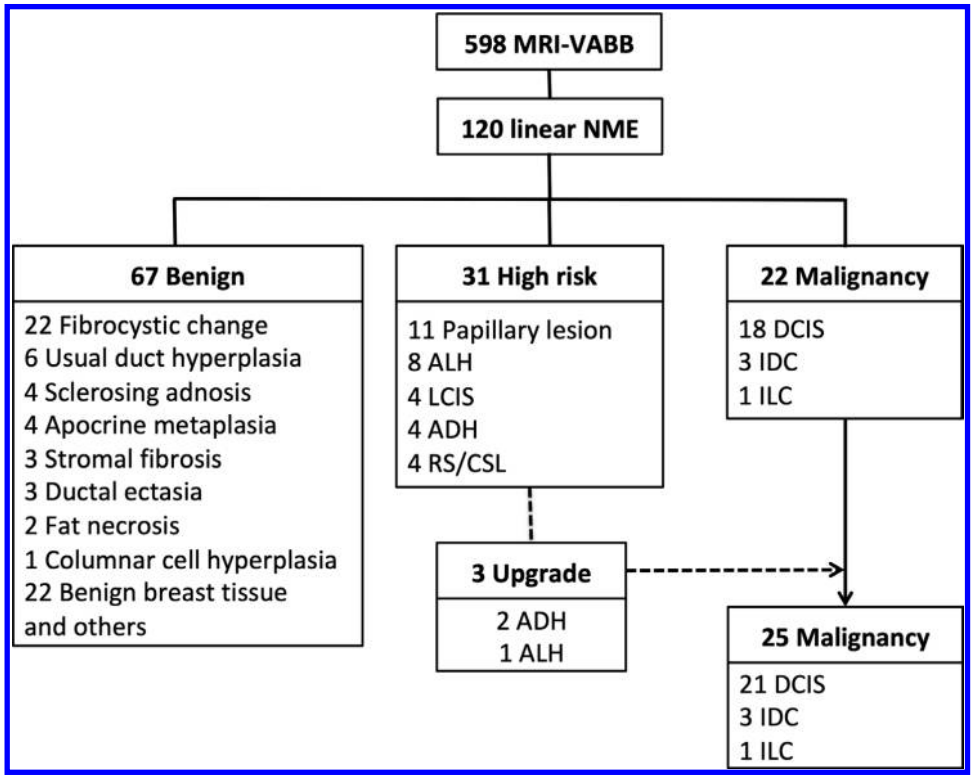
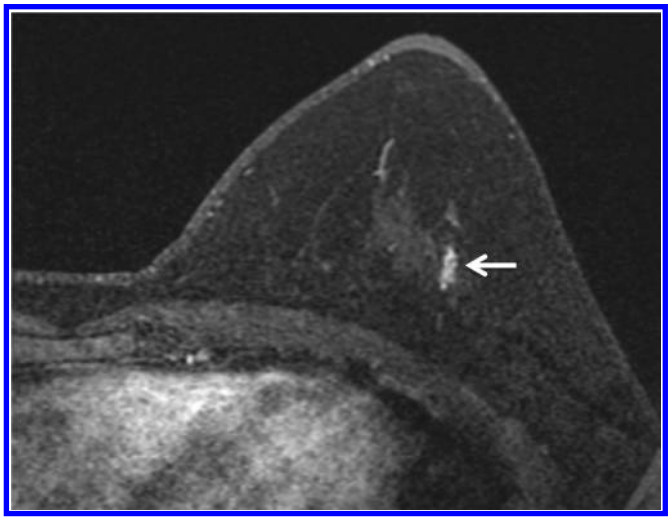


Table 1. Internal Enhancement Patterns and Kinetic Curve in 120 findings of Linear Nonmass Enhancement on Contrast-Enhanced Breast MRI

Descriptor	Lesions N (%)	No. of cancer	PPV	*p* value	Histological diagnosis	
					Invasive	DCIS
All linear NME	120 (100)	25	20.8	NA	4/25 (16.0)	21/25 (84.0)
Internal enhancement						
Clumped	52 (43.3)	12	23.1	0.522	1	11
Homogeneous	33 (27.5)	8	24.2		3	5
Heterogeneous	35 (29.2)	5	14.3		0	5
Kinetic curve						
Early[a]	67 (100)	10		0.243	1/10 (10)	9/10 (90)
Fast	54 (80.6)	9	16.7		0	9
Medium	6 (9.0)	1	16.7		1	0
Slow	7 (10.4)	0	0		0	0
Delayed[b]	85 (100)	16		0.754	2/16 (12.5)	14/16 (87.5)
Washout	33 (38.8)	7	21.2		0	7
Plateau	24 (28.2)	4	16.7		2	2
Persistent	28 (33.0)	5	17.9		0	5

DCIS, Ductal carcinoma *in situ*; NA, Not applicable.
[a]Unknown Early Kinetic Phase *n* = 53.
[b]Unknown Delayed Kinetic Phase *n* = 35.

Figure 2. 47-year-old female underwent staging MRI, axial T_1-weighted contrast-enhanced image showed a 1.4 cm non-mass enhancement (NME) linear distribution with non-clumped (heterogeneous) internal enhancement (arrow). MRI-guided vacuum-assisted breast biopsy showed lobular carcinoma *in situ* and atypical lobular hyperplasia.



enhancement (Figure 2) and 33/68 (49%) showed homogeneous internal enhancement (Figure 3).

Initial and delayed kinetic data were reported in 67 and 85 lesions, respectively. The initial phase was fast in 54/67 (81%) lesions, medium in 6/67 (9%) lesions, and slow in 7/67 (10%) lesion. The delayed phased showing washout 33/85 (39%), plateau 24/85 (28%), and persistent 28/85 (33%) kinetics (Table 1).

Figure 3. 46-year-old female underwent staging MRI for right invasive ductal carcinoma. Sagittal maximum intensity projection showed a 1.7 cm non-mass enhancement (NME) linear distribution with non-clumped (homogeneous) internal enhancement (arrow). The pathology result was intermediate grade ductal carcinoma *in situ*.

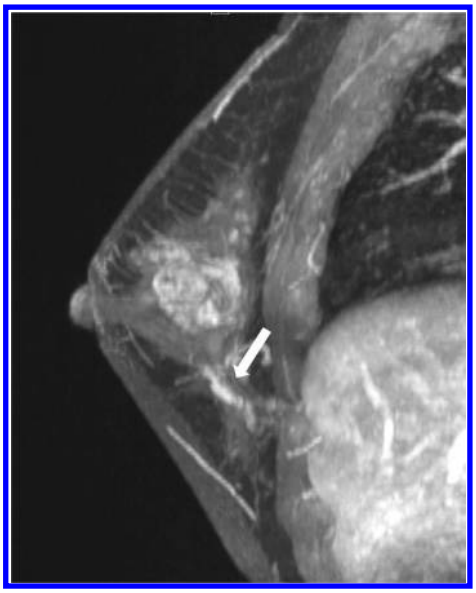
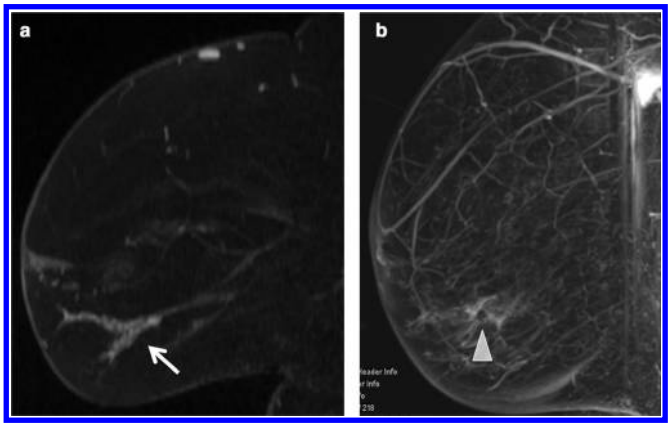


Figure 4. 79-year-old female underwent screening MRI, sagittal T_1-weighted contrast-enhanced image (A) showed a 2.6 cm non-mass enhancement (NME) linear distribution with clumped internal enhancement (arrow). Maximum intensity projection (MIP) of post MRI-guided vacuum-assisted breast biopsy (B). Susceptibility artefact from the biopsy clip is present (arrowhead). The pathology result was atypical ductal hyperplasia and excision yielded to intermediate grade ductal carcinoma *in situ*.



On MRI-VABB, there were 67/120 (56%) benign lesions, 22/120 (18%) malignancies (18 DCIS, 3 IDC, one invasive lobular carcinoma), 31/120 (21%) were high-risk lesions. Of the high-risk lesions, 3/31 (17%) upgraded to DCIS at surgery (Figure 4), for a total of 25/120 (21%) malignancies. Malignancy was found in 12/52 (23%) lesions with clumped internal enhancement and in 13/68 (19%) of lesions with non-clumped internal enhancement characteristics. Of the linear non-clumped lesions with malignancy, 5/13 (38%) had heterogeneous internal enhancement, and 8/13 (62%) had homogeneous internal enhancement. Kinetic data of the 25 patients with biopsy proven malignancy were evaluated and initial and delayed kinetic data were available in 10 and 16 lesions. The largest number of these malignant lesions showed fast initial Phase 9/10 (90%) followed by washout delayed Phase 7/16 (44%) (Table 1).

Positive predict value of linear NME

Overall, the PPV of NME with linear distribution was 20.8% (25/120; 95% CI, 13.5–28.1). The PPV of linear NME with clumped internal enhancement (23.1%; 95% CI, 11.6–34.6) was not significantly different from the PPV of non-clumped linear NME (19.1%; 95% CI, 9.8–28.4) (p = 0.597). The PPV of linear NME lesions <1 cm (33.3%; 95% CI, 11.5–55.1) was not significantly different from the PPV of lesions ≥1 cm (18.6%; 95% CI, 11.1–26.2) (p = 0.157). There was no association with malignancy between clumped and non-clumped NME in combination with size or suspicious kinetic data (Table 2).

DISCUSSION

The fifth edition of the ACR BI-RADS lexicon made specific changes to the distribution and internal enhancement patterns of NME.[13] With regard to distribution, the word "ductal" was replaced with "linear" to describe enhancement arrayed in a line or a line that branches.[7] Our study showed that linear NME on breast MRI may represent malignancies such as IDC, invasive

Table 2. PPVs for Malignancy in Linear Distributed NME on Contrast-enhanced MRI according to lesion characteristics and kinetic curves

Descriptors	Benign lesions (n)	Malignant lesions (n)	PPV (%)[a]	*p* value
Clumped	40	12	23.1 (11.6, 34.6)	0.597
Non-clumped	55	13	19.1 (9.8, 28.4)	
NME <1 cm	12	6	33.3 (11.5, 55.1)	0.157
NME ≥ 1 cm	83	19	18.6 (11.1, 26.2)	
Clumped linear <1 cm	6	2	25.0 (0.0, 55.0)	0.544
Clumped linear ≥1 cm	35	9	20.5 (8.6, 32.4)	
Non-clumped linear <1 cm	7	3	30.0 (1.6, 58.4)	0.336
Non-clumped linear ≥1 cm	47	11	19.0 (8.9, 29.1)	
Clumped and fast	17	2	10.5 (0.0, 24.3)	0.412
Non-clumped and fast	29	6	17.1 (4.6, 29.6)	
Clumped and washout	10	2	16.7 (0.0, 37.8)	0.494
Non-clumped and washout	16	5	23.8 (5.6,42.0)	

[a]Numbers in parentheses are 95% confidence intervals.

lobular carcinoma, DCIS or various benign breast processes as previous studies.[14–18] The PPVs of linearly distributed NME show range from 11 to 67% in the literature.[9,10,12,14,19] In our study of BIRADS four or five linear NME lesions leading to MRI-guided VABB, we found that linear NME had a PPV of 20.8%. As a result of this variability, studies of additional features of linear NME were performed to determine if specificity could be improved. Machida et al studied the PPVs of branching verses non-branching linear NME, showing that a branching pattern was a significantly stronger predictor of malignancy than was the linear pattern.[20] Combining these features may be used to improve the diagnosis of linear NME lesions.

To our knowledge, this is the only study that looked specifically at the PPVs of the internal enhancement characteristics of linear NME. Our results show no significant difference in PPVs of clumped linear non-mass enhancement (23.1%) *vs* non-clumped (19.1%) – heterogeneous (14.3%) or homogenous (24.2%) –- internal enhancement characteristics. In review of previous studies, most of currently published articles evaluating the PPVs of malignancy for linear distribution and associated internal enhancement were performed before the publication of the updated lexicon (Table 3). The clumped enhancement had been mixed with ductal or branching pattern, which may explain the wide range of PPV of overall clumped lesions in these studies,

Table 3. Review of Literature on Malignancy *vs* Benign Pathology in Linear Non-Mass Enhancement (NME) and Stratified by Internal Enhancement Characteristics

Study		Total No. of Biopsied NMEs	No. of linear/ Total No. of NMEs (%)	No. of Malignant Linear NMEs/Total linear NME(%)	No. of Malignant clumped/Total clumped NME (%)	No. of Malignant linear clumped/ Total linear clumped (%)
Lead Author	Year					
Liberman et al[10]	2002	40	21/40 (53)	5/21 (24)	9/22 (41)	5/16 (31)
Liberman et al[12]	2003	150	88/150 (59)	23/88 (26)	NA	18/52 (35)
Morakkabati et al[11]	2005	38	NA	2/10 (20)	NA	10/17 (59)[a]
Tozaki et al[9]	2006	30	8/30 (27)	2/6 (33)	7/8 (88)	3/4 (75)
Sakamoto et al[19]	2008	102	9/102 (9)[b]	1/9 (11)	2/10 (20)	NA
Uematsu et al[14]	2012	124	9/124 (7)	6/9 (67)	26/32 (81)	NA
Ballesio et al[15]	2014	94	19/94 (20)	12/19 (63)	9/14 (64)	NA
Chikarmane et al[21]	2017	144	38/144 (26)	12/38 (32)	11/33 (33)	NA
Our Study	2019	598	120/598 (20)	25/120 (21)	NA	12/52 (23)

NA, Not available.
[a]Including both linear and segmental distribution.
[b]Assessed by Linear-ductal pattern.

ranging from 20 to 88%. A few studies have performed a detailed analysis of linear clumped enhancement. Liberman et al reported that the PPV of linear clumped lesion ranged 31–35%.[10,12] A higher PPV of clumped enhancement was demonstrated by Morakkabati et al (PPV = 59%); however, in that study, NME included segmental and linear distributions and may have had an inherent bias toward malignancy. Tozaki and Fukuda[9] found extremely high PPV for linear clumped NME, with malignancy found in 3 of 4 (75%) lesions; however, the study had small numbers of both total biopsied NME and clumped lesions. Our results reflect the results of the highest number of total biopsied linear distribution (n = 120) and linear clumped lesions (n = 52) of all studies published to date (Table 3).

We found that the PPV of linear NME <1 cm (33.3%) was not significantly different the PPV of lesions ≥ 1 cm (18.6%), comparable to the result by Gutierrez et al that size was not a significant predictor of malignancy for NME.[22] This finding counters the conclusions of Machida et al who concluded that NME lesions with a linear pattern that are smaller than 1 cm can be managed with follow-up.[20] One possible reason for this difference is that our population was different than that of the Machida et al in that we only studied patients who underwent MRI-guided biopsy and all patients were at very high risk for breast cancer. This suggests, in the appropriate population (as in 80% of our population are either high-risk screening or staging for known breast cancer), linear NME might need to be biopsied rather than observed even if the lesion size is less than 1 cm.

Studies have shown that most common malignant causes of NME are DCIS and diffuse invasive breast cancers, particularly lobular cancer, but also occasionally ductal cancers.[23,24] In particular, DCIS most commonly manifests as NME, less frequently as a mass or a focus.[25] Rosen et al showed that pure DCIS lesions show NME in 59% of cases, whereas 14% present as an enhancing mass, 14% show no enhancement, and 12% as a focus.[26] It has been suggested that DCIS appears as NME because DCIS typically follows the ductal system, which means the most frequent enhancement was within a segmental or linear (ductal) distribution, and internal enhancement is usually clumped.[10,11,27,28] Our results supported these findings demonstrated in previous studies in that 11/21 (52%) of our linear NME DCIS showed clumped internal enhancement, higher than either heterogeneous (5/21, 24%) or homogeneous (5/21, 24%) internal enhancement if each of the non-clumped patterns were compared alone. Overall, our results showed most malignant linear distributed NME were DCIS (21/25, 84%), and many fewer were invasive cancer (3/25 were IDC, and only one invasive lobular carcinoma).

The classically suspicious kinetic curve for a lesion detected at MR imaging includes fast early enhancement and/or delayed washout, or plateau-type enhancement.[29,30] Kinetic information has proved to be helpful in the differential diagnosis of MR imaging – detected mass lesions.[29] However, the variability of kinetic curves in both benign and malignant NME has led a number of investigators to consider kinetic features unreliable for a diagnosis of malignancy, particularly for DCIS.[31–33] This is because kinetic curve is influenced by several pathological factors, including the extend and pattern of vascularization, vessel permeability, cellularity, interstitial pressure, and the fraction of the extracellular space.[32] Jansen et al evaluated 852 MRI detected lesions (552 mass, 261 NME, and 39 focus) and reported that in NME, only washout parameters may be relevant.[24] In our study, fast-washout pattern was the most frequent descriptor for malignant pathology in linear NME, and most are DCIS. In DCIS, a fast-washout pattern may associate with a high density of ducts, an abundance of blood vessels, and a high degree of inflammatory cell infiltration.[33] However, none of the kinetic features were significant predictors of carcinoma when compared to the entire study population, which was similar to Liberman et al[12] study

Our study has some limitations. First, it is a retrospective study and we examined a high-risk patient population and included only lesions referred for biopsy, which may lead to selection bias. Second, there was no clustered ring enhancement (CRE) in our study. CRE can be difficult to differentiate from other internal enhancement patterns and may be present with either clumped or heterogeneous NME. This challenge raises the question of whether CRE should be assessed differently, such as assessed as categorically present or not present, to reflect clinical practice, and possibly be listed as an associated feature in the BI-RADS atlas.[21]

CONCLUSIONS

Linear NME showed malignancy in 21% of our series. Our recommendations would be, in specific patient groups, such as screening in females at high-risk for breast cancer or staging for known breast cancer, linear NME with clumped or non-clumped internal enhancement patterns might need to undergo MRI-VABB regardless of lesion size or type of kinetic curve.

REFERENCES

1. DeMartini W, Lehman C. A review of current evidence-based clinical applications for breast magnetic resonance imaging. *Top Magn Reson Imaging* 2008; **19**: 143–50. doi: https://doi.org/10.1097/RMR.0b013e31818a40a5
2. Landercasper J, Linebarger JH. Contemporary breast imaging and concordance assessment: a surgical perspective. *Surg Clin North Am* 2011; **91**: 33–58. doi: https://doi.org/10.1016/j.suc.2010.10.003
3. Heywang-Köbrunner SH, Hacker A, Sedlacek S. Magnetic resonance imaging: the evolution of breast imaging. *Breast* 2013; **22 Suppl 2**(Suppl 2): S77–82. doi: https://doi.org/10.1016/j.breast.2013.07.014
4. Harms SE, Flamig DP, Hesley KL, Meiches MD, Jensen RA, Evans WP, et al. Mr imaging of the breast with rotating delivery of excitation off resonance: clinical experience with pathologic correlation. *Radiology* 1993; **187**: 493–501. doi: https://doi.org/10.1148/radiology.187.2.8475297
5. Peters NHGM, Borel Rinkes IHM, Zuithoff NPA, Mali WPTM, Moons KGM, Peeters

PHM. Meta-Analysis of Mr imaging in the diagnosis of breast lesions. *Radiology* 2008; **246**: 116–24. doi: https://doi.org/10.1148/radiol.2461061298

6. Kristoffersen Wiberg M, Aspelin P, Perbeck L, Boné B. Value of Mr imaging in clinical evaluation of breast lesions. *Acta Radiol* 2002; **43**: 275–81. doi: https://doi.org/10.1034/j.1600-0455.2002.430308.x
7. Sickles EA, D'Orsi CJ, Bassett LW, Appleton CM, Berg WA, Burnside ES. ACR BI-RADS® atlas, breast imaging reporting and data system. *Reston, VA: American College of Radiology* 2013;: 39–48.
8. Schnall MD, Blume J, Bluemke DA, DeAngelis GA, DeBruhl N, Harms S, et al. Diagnostic architectural and dynamic features at breast MR imaging: multicenter study. *Radiology* 2006; **238**: 42–53. doi: https://doi.org/10.1148/radiol.2381042117
9. Tozaki M, Fukuda K. High-spatial-resolution MRI of non-masslike breast lesions: interpretation model based on BI-RADS MRI descriptors. *AJR Am J Roentgenol* 2006; **187**: 330–7. doi: https://doi.org/10.2214/AJR.05.0998
10. Liberman L, Morris EA, Lee MJ-Y, Kaplan JB, LaTrenta LR, Menell JH, et al. Breast lesions detected on MR imaging: features and positive predictive value. *AJR Am J Roentgenol* 2002; **179**: 171–8. doi: https://doi.org/10.2214/ajr.179.1.1790171
11. Morakkabati-Spitz N, Leutner C, Schild H, Traeber F, Kuhl C. Diagnostic usefulness of segmental and linear enhancement in dynamic breast MRI. *Eur Radiol* 2005; **15**: 2010–7. doi: https://doi.org/10.1007/s00330-005-2755-4
12. Liberman L, Morris EA, Dershaw DD, Abramson AF, Tan LK. Ductal enhancement on MR imaging of the breast. *AJR Am J Roentgenol* 2003; **181**: 519–25. doi: https://doi.org/10.2214/ajr.181.2.1810519
13. Edwards SD, Lipson JA, Ikeda DM, Lee JM. Updates and revisions to the BI-RADS magnetic resonance imaging lexicon. *Magn Reson Imaging Clin N Am* 2013; **21**: 483–93. doi: https://doi.org/10.1016/j.mric.2013.02.005
14. Uematsu T, Kasami M. High-spatial-resolution 3-T breast MRI of nonmasslike enhancement lesions: an analysis of their features as significant predictors of malignancy. *AJR Am J Roentgenol* 2012; **198**: 1223–30. doi: https://doi.org/10.2214/AJR.11.7350
15. Ballesio L, Di Pastena F, Gigli S, D'ambrosio I, Aceti A, Pontico M, et al. Non mass-like enhancement categories detected by breast MRI and histological findings. *Eur Rev Med Pharmacol Sci* 2014; **18**: 910–7.
16. Hahn SY, Han B-K, Ko EY, Shin JH, Hwang J-Y, Nam M. Mr features to suggest microinvasive ductal carcinoma of the breast: can it be differentiated from pure DCIS? *Acta Radiol* 2013; **54**: 742–8. doi: https://doi.org/10.1177/0284185113484640
17. Scoggins M, Krishnamurthy S, Santiago L, Yang W. Lobular carcinoma in situ of the breast: clinical, radiological, and pathological correlation. *Acad Radiol* 2013; **20**: 463–70. doi: https://doi.org/10.1016/j.acra.2012.08.020
18. Tan H, Li R, Peng W, Liu H, Gu Y, Shen X. Radiological and clinical features of adult non-puerperal mastitis. *Br J Radiol* 2013; **86**: 20120657. doi: https://doi.org/10.1259/bjr.20120657
19. Sakamoto N, Tozaki M, Higa K, Tsunoda Y, Ogawa T, Abe S, et al. Categorization of non-mass-like breast lesions detected by MRI. *Breast Cancer* 2008; **15**: 241–6. doi: https://doi.org/10.1007/s12282-007-0028-6
20. Machida Y, Tozaki M, Shimauchi A, Yoshida T. Two distinct types of linear distribution in Nonmass enhancement at breast MR imaging: difference in positive predictive value between linear and branching patterns. *Radiology* 2015; **276**: 686–94. doi: https://doi.org/10.1148/radiol.2015141775
21. Chikarmane SA, Michaels AY, Giess CS. Revisiting Nonmass enhancement in breast MRI: analysis of outcomes and follow-up using the updated BI-RADS atlas. *AJR Am J Roentgenol* 2017; **209**: 1178–84. doi: https://doi.org/10.2214/AJR.17.18086
22. Gutierrez RL, DeMartini WB, Eby PR, Kurland BF, Peacock S, Lehman CD. Bi-Rads lesion characteristics predict likelihood of malignancy in breast MRI for masses but not for nonmasslike enhancement. *AJR Am J Roentgenol* 2009; **193**: 994–1000. doi: https://doi.org/10.2214/AJR.08.1983
23. Giess CS, Raza S, Birdwell RL. Patterns of nonmasslike enhancement at screening breast MR imaging of high-risk premenopausal women. *Radiographics* 2013; **33**: 1343–60. doi: https://doi.org/10.1148/rg.335125185
24. Jansen SA, Shimauchi A, Zak L, Fan X, Karczmar GS, Newstead GM. The diverse pathology and kinetics of mass, nonmass, and focus enhancement on MR imaging of the breast. *J Magn Reson Imaging* 2011; **33**: 1382–9. doi: https://doi.org/10.1002/jmri.22567
25. Greenwood HI, Heller SL, Kim S, Sigmund EE, Shaylor SD, Moy L. Ductal carcinoma in situ of the breasts: review of Mr imaging features. *Radiographics* 2013; **33**: 1569–88. doi: https://doi.org/10.1148/rg.336125055
26. Rosen EL, Smith-Foley SA, DeMartini WB, Eby PR, Peacock S, Lehman CD. Bi-Rads MRI enhancement characteristics of ductal carcinoma in situ. *Breast J* 2007; **13**: 545–50. doi: https://doi.org/10.1111/j.1524-4741.2007.00513.x
27. 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**: 356–78.
28. Kim J-A, Son EJ, Youk JH, Kim E-K, Kim MJ, Kwak JY, et al. Mri findings of pure ductal carcinoma in situ: kinetic characteristics compared according to lesion type and histopathologic factors. *AJR Am J Roentgenol* 2011; **196**: 1450–6. doi: https://doi.org/10.2214/AJR.10.5027
29. Kuhl CK, Mielcareck P, Klaschik S, Leutner C, Wardelmann E, Gieseke J, et al. Dynamic breast MR imaging: are signal intensity time course data useful for differential diagnosis of enhancing lesions? *Radiology* 1999; **211**: 101–10. doi: https://doi.org/10.1148/radiology.211.1.r99ap38101
30. Bluemke DA, Gatsonis CA, Chen MH, DeAngelis GA, DeBruhl N, Harms S, et al. Magnetic resonance imaging of the breast prior to biopsy. *JAMA* 2004; **292**: 2735–42. doi: https://doi.org/10.1001/jama.292.22.2735
31. Newell D, Nie K, Chen J-H, Hsu C-C, Yu HJ, Nalcioglu O, et al. Selection of diagnostic features on breast MRI to differentiate between malignant and benign lesions using computer-aided diagnosis: differences in lesions presenting as mass and non-mass-like enhancement. *Eur Radiol* 2010; **20**: 771–81. doi: https://doi.org/10.1007/s00330-009-1616-y
32. Kuhl CK, Schild HH. Dynamic image interpretation of MRI of the breast. *J Magn Reson Imaging* 2000; **12**: 965–74. doi: https://doi.org/10.1002/1522-2586(200012)12:6<965::AID-JMRI23>3.0.CO;2-1
33. Menell JH, Morris EA, Dershaw DD, Abramson AF, Brogi E, Liberman L, et al. Determination of the presence and extent of pure ductal carcinoma in situ by mammography and magnetic resonance imaging. *Breast J* 2005; **11**: 382–90. doi: https://doi.org/10.1111/j.1075-122X.2005.00121.x

BJR

https://doi.org/10.1259/bjr.20180801

Received: 17 September 2018 | Revised: 18 November 2018 | Accepted: 24 November 2018

Cite this article as:
Choudhery S, Axmacher J, Conners AL, Geske J, Brandt K. Masses in the era of screening tomosynthesis: Is diagnostic ultrasound sufficient?. *Br J Radiol* 2019; **92**: 20180801.

FULL PAPER

Masses in the era of screening tomosynthesis: Is diagnostic ultrasound sufficient?

[1]SADIA CHOUDHERY, MD, [1]JESSICA AXMACHER, MD, [1]AMY LYNN CONNERS, MD, [2]JENNIFER GESKE, MS and [1]KATHY BRANDT, MD

[1]Department of Radiology, Mayo Clinic, Rochester MN, US
[2]Biomedical Statistics and Informatics, Mayo Clinic, Rochester MN, US

Address correspondence to: Dr Sadia Choudhery
E-mail: *choudhery.sadia@mayo.edu*

Objectives: The purpose of this study is to compare diagnostic outcomes of digital breast tomosynthesis screen-detected masses worked up with mammography first with those evaluated with diagnostic ultrasound initially.
Methods: All masses recalled from screening digital breast tomosynthesis between July 1, 2017 and December 31, 2017 that were sent either to diagnostic mammography or ultrasound were compared. Size, shape, margins, visibility on ultrasound, diagnostic assessment and pathology of all masses along with breast density were evaluated.
Results: 102/212 digital breast tomosynthesis screen-detected masses were worked up with diagnostic mammography initially and 110/212 were worked up with ultrasound directly. There was no significant difference in ultrasound visibility of masses sent to diagnostic mammography first with those sent to ultrasound first (p = 0.42). 4 (4%) masses sent to mammogram first and 2 (2%) masses sent to ultrasound first were not visualized. There was a significant difference in size between masses that were visualized under ultrasound versus those that were not (p = 0.01), when masses in both groups were assessed cumulatively.
Conclusions: 98% of digital breast tomosynthesis screen-detected masses sent to ultrasound directly were adequately assessed without diagnostic mammography.
Advances in knowledge: There is potential for avoiding a diagnostic mammogram for evaluation of majority of digital breast tomosynthesis screen-detected masses.

INTRODUCTION

In the era of digital breast tomosynthesis (DBT), the need for diagnostic mammography (MG) before a diagnostic ultrasound for masses recalled from screening tomosynthesis has been questioned.[1–3] Historically, most masses recalled from two-dimensional (2D) screening mammography underwent diagnostic mammography prior to ultrasound.[4,5] In this setting, diagnostic mammography views have been shown to increase the specificity of mammography by improving margin assessment, determining lesion location, and confirming persistence of the screen-detected mass.[5,6] In comparison to 2D imaging, DBT allows better differentiation of true findings from superimposition of fibroglandular tissue, increases mass margin visibility, and improves location assessment.[7–9] In addition, studies have shown that DBT has similar accuracy as routine diagnostic mammography for non-calcified findings,[10] is comparable to spot compression mammography for characterizing masses as benign or malignant,[1] and is equivalent or better than spot compression mammography for evaluating findings recalled from 2D screening mammography.[2,3,11]

Currently, institutions and practitioners vary in how the work-up of masses recalled from screening tomosynthesis is performed, with some opting for ultrasound first and others performing diagnostic mammography before ultrasound. There are no clear American College of Radiology practice guidelines for work-up of DBT-detected masses. The purpose of this study is to compare outcomes of masses recalled from screening DBT worked-up initially with diagnostic mammography with those first evaluated with diagnostic ultrasound.

METHODS AND MATERIALS

Study subjects, imaging technique, and interpretation

Our Institutional Review Board approved this retrospective Health Insurance Portability and Accountability Act-compliant study. Informed consent was waived.

We performed a retrospective review of our mammography reporting system for all screen-detected masses from July 1, 2017, from the time of our conversion to

all three-dimensional screening, to December 31, 2017. All non-calcified masses without associated features of architectural distortion or skin thickening and with follow-up atMayo Clinic Rochester, MN were included in this study. A mass was defined, according to the fifth edition of the Breast Imaging-Reporting and Data System (BI-RADS), as a three-dimensional lesion seen on two views with complete or partial convex border.[12] Lesions that did not fit the BI-RADS definition of a mass were excluded.

Mammographic screening and diagnostic DBT examinations were performed using Hologic Selenia Dimensions digital breast tomosynthesis (Hologic, Bedford, MA) with synthesized 2D (C-View^{TM}) images. Screening exams included bilateral DBT mediolateral oblique (MLO) and craniocaudal (CC) views with C-View, and diagnostic exams included CC and MLO DBT spot compression views with C-View of the mass and a 2D full-field digital MG ML view of the breast. Ultrasound exams were performed by breast imaging specialized sonographers and/or breast imaging radiologists using GE Logiq E9 (General Electric Healthcare, Milwaukee, WI).

All exams were prospectively interpreted by 1 of 18 breast imaging radiologists with 1–5 years of tomosynthesis experience and 1–29 years of practice in a tertiary academic setting. Recalled masses were either worked-up with diagnostic MG first or sent directly to diagnostic ultrasound at the discretion of the screening radiologist as specified on the screening mammogram report. The diagnostic work-up was completed in one visit by any 1 of the 18 breast imaging radiologists and resulted in a final BI-RADS assessment. Masses sent to diagnostic MG first were either given a final assessment based on the MG or, if directed by the diagnostic radiologist, were further evaluated with diagnostic ultrasound before a final combined assessment was established. Masses sent directly to diagnostic ultrasound were either given a final assessment based on the ultrasound or were further worked-up with diagnostic MG after the ultrasound before a final combined assessment was determined. Patients given BI-RADS, one or two were returned to routine screening mammography. Patients with BI-RADS three were recommended to have a 6 month follow-up imaging exam, and patients with BI-RADS four or five masses were recommended to have core needle biopsies. The final diagnostic BI-RADS assessment was recorded for each mass. Pathology was recorded from core needle biopsies of masses that underwent sampling. The BI-RADS density assessment from the screening study was also recorded for each patient as dense, comprising the BI-RADS categories of heterogeneously dense and extremely dense, and non-dense, comprising the BI-RADS categories of almost entirely fatty and scattered areas of fibroglandular density.

Data collection and statistical analysis

Two breast imaging radiologists, not blinded to the final outcomes, reviewed the screening DBT to determine if the recalled mass met the BI-RADS definition used in this study. In addition, each radiologist assessed and recorded the shape, size, margins, and density of each mass seen on the screening DBT exam. χ^2 or Fisher's exact tests were used to compare categorical characteristics by MG or ultrasound first, and *t*-tests were used to compare mass size by MG or ultrasound first. Descriptive statistics are provided as frequencies (N) and percent (%) or mean and standard deviations (SDs). Statistical analyses were conducted using SAS (v. 9.4; Cary, NC).

RESULTS

2072/16,916 patients were recalled during our study period, which included 214 patients with 226 masses. Of these 226 masses, 14 masses were excluded on retrospective review since they did not fit the BI-RADS definition of a mass. 102/212 (48%) masses were sent to diagnostic MG first whereas 110/212 (52%) masses were sent to ultrasound initially. There was no significant difference in the size of the masses ($p = 0.31$) or density of the breasts ($p = 0.74$) in the two groups (Table 1). There was a significant difference in the shape of masses in these two groups ($p < 0.0001$) (Table 1) with 65 (64%) oval, 21 (21%) round, and 16 (16%) irregular masses in the MG first group in comparison to 96 (87%) oval, 12 (11%) round, and 2 (2%) irregular masses in the ultrasound first group. Similarly, there was a significant difference in the margins of masses in these two groups ($p < 0.0001$) with 66 (65%) circumscribed, 12 (12%) indistinct, 12 (12%) obscured, and 12 (12%) spiculated masses in the MG first group in comparison to 97 (88%) circumscribed, 8 (7%) indistinct, and 5 (5%) obscured masses in the ultrasound first group. There was also a significant difference in the density ($p = 0.039$) of masses, BI-RADS assessments ($p = 0.0002$), and pathology of masses ($p = 0.011$) in the two groups (Table 1). There were 55 (54%) BI-RADS 1 and 2, 11 (11%) BI-RADS 3, and 36 (35%) BI-RADS 4 and 5 masses in the MG first group compared to 85 (77%) BI-RADS 1 and 2, 7 (6%) BI-RADS 3, and 18 (16%) BI-RADS 4 and 5 masses in the US first group ($p = 0.0002$). There were 22 (61%) benign and 14 (39%) malignant lesions in the MG first group compared to 14 (94%) benign and 1 (6%) malignant lesions in the ultrasound first group ($p = 0.011$). Cysts were the most common benign lesion in both groups, 39/80 (49%) and 74/104 (71%) in the MG and ultrasound first groups respectively. Invasive ductal carcinoma (IDC) was the most common malignant pathology in the MG first group with 8/14 (57%) malignant masses representing IDC. The only malignant mass in the ultrasound first group was also IDC. Positive-predictive values from screening mammography (PPV1) were 14 and 0.9% for MG and US first groups respectively and from biopsy recommendation (PPV2)/from biopsy performed (PPV3) were 39 and 6% for MG and US first groups respectively.

94/102 (92%) masses sent to MG first were subsequently worked up with ultrasound and no ultrasound was performed for the remaining eight masses (Figure 1). Of these eight masses that did not undergo ultrasound evaluation, three were felt to be stable benign lymph nodes, three were mammographically stable masses, and two were thought to represent overlapping tissue by the radiologist interpreting the diagnostic mammogram. Of the 94 masses sent to subsequent ultrasound in this group, 90 (96%) had a sonographic correlate, and 4 (4%) were not visualized with ultrasound (Table 2). 3/4 masses that were not visualized by ultrasound were round, 1/4 was oval, 3/4 had circumscribed margins and 1/4 had obscured margins on the screening DBT. All 4 (100%) occurred in females with non-dense breasts. In

Table 1. Characteristics of masses worked up with diagnostic MG versus diagnostic ultrasound first

		MG first	Ultrasound first	*p*-value
		n = 102	*n* = 110	
	Category	*n* (%) or mean ± SD	*n* (%) or mean ± SD	
Breast density	Non-dense	87 (87.9)	95 (86.4)	0.74
	Dense	12 (12.1)	15 (13.6)	
Size (cm)[a]		0.9 ± 0.5	1.0 ± 0.6	0.31
Shape[a]	Irregular	16 (15.7)	2 (1.8)	<0.0001
	Oval	65 (63.7)	96 (87.3)	
	Round	21 (20.6)	12 (10.9)	
Margins[a]	Circumscribed	66 (64.7)	97 (88.2)	<0.0001
	Indistinct	12 (11.8)	8 (7.3)	
	Obscured	12 (11.8)	5 (4.6)	
	Spiculated	12 (11.8)	0 (0)	
Mass density[a]	High density	17 (16.7)	7 (6.4)	0.039
	Equal density	79 (77.5)	90 (81.8)	
	Low density	3 (2.9)	10 (90.9)	
	Fat-containing	3 (2.9)	3 (2.7)	
Final BI-RADS	1/2	55 (53.9)	85 (77.3)	0.0002
	3	11 (10.8)	7 (6.4)	
	4/5	36 (35.3)	18 (16.4)	
Pathology[b]	Benign	22 (61.1)	17 (94.4)	0.011
	Malignant	14 (38.9)	1 (5.6)	

MG, mammography; SD, standard deviation.
[a]Mass characteristics are based on findings on screening digital breast tomosynthesis (DBT)
[b]Pathology is for masses that underwent core needle biopsy

the final assessment, 3/4 of these recalled "masses" were felt to represent overlapping fibroglandular tissue. These lesions looked like masses on the screening exam but did not definitely persist on diagnostic MG and an ultrasound was done for confirming the absence of a true lesion. 1/4 of the masses persisted on diagnostic mammography and was not seen under subsequent diagnostic US so a stereotactic biopsy was performed with resulting pathology of apocrine cysts with stromal fibrosis, which was considered concordant with the imaging findings.

108/110 (98%) of masses sent to US first had a sonographic correlate, and 2/110 (2%) did not (Figures 1–3). The two masses not seen on US were subsequently sent to a diagnostic mammogram. Both of these masses were oval in shape, one had circumscribed and one had obscured margins on the screening DBT (Table 2). Both masses occurred in females with non-dense breasts. The mass with the obscured margins effaced on subsequent diagnostic mammography and was thought to represent overlapping fibroglandular tissue. The mass with circumscribed margins persisted on subsequent diagnostic MG so a stereotactic core biopsy was performed with pathology revealing fibrocystic changes, which was considered concordant with imaging findings.

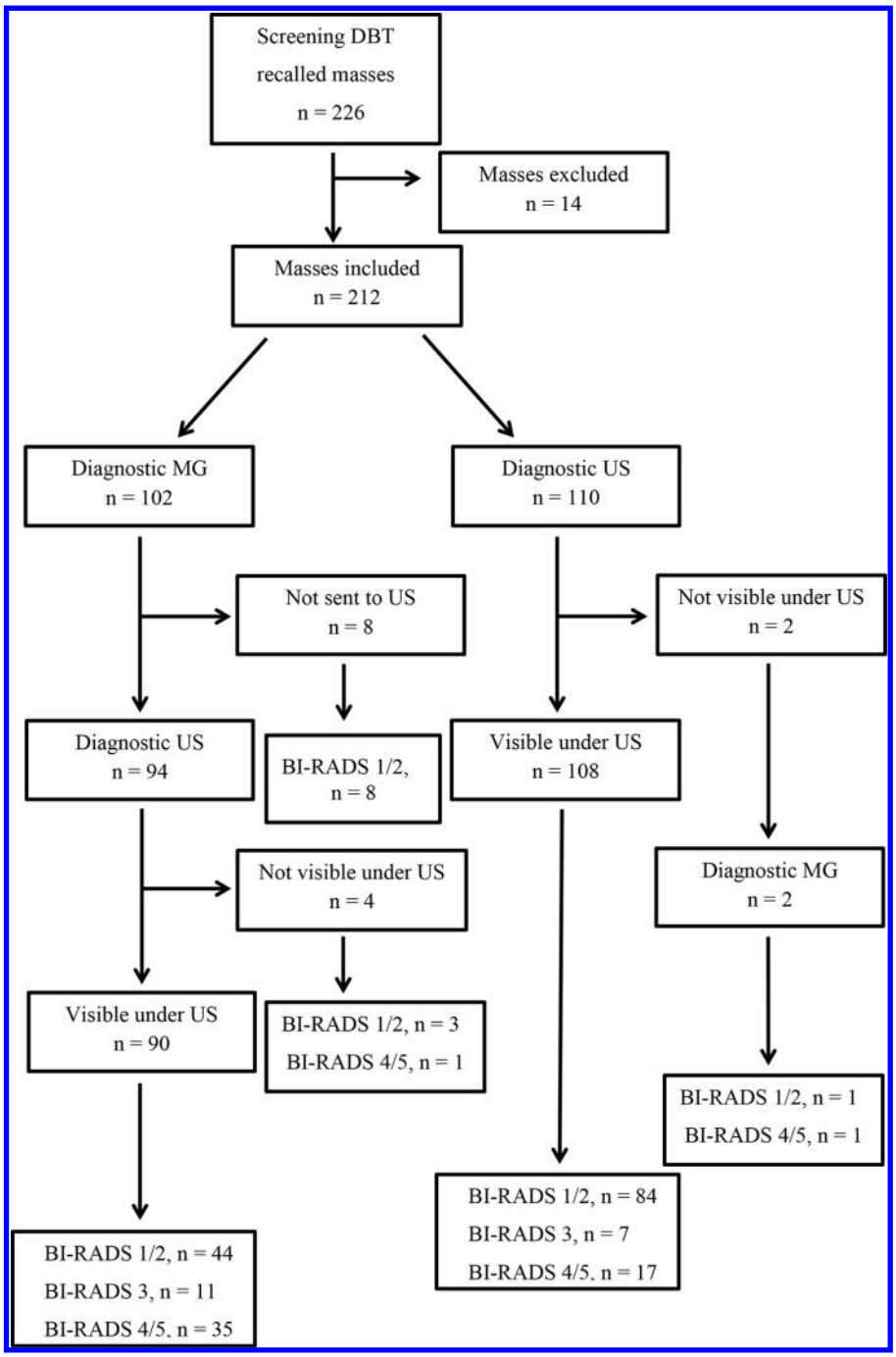
Overall, 198/204 (97%) of DBT-screen detected masses were found to have a sonographic correlate in our study. There was no significant difference in the US visibility of masses sent to MG first compared to those sent initially to US ($p = 0.42$). There was a significant difference in the mean size between masses that were seen on US (0.9 ± 0.5 cm) versus those that were not seen (0.5 ± 0.4 cm) ($p = 0.01$), when assessing the US visibility of masses in the MG first and US first groups cumulatively.

DISCUSSION

Our study shows that diagnostic mammography can be eliminated for working up majority of DBT screen-detected masses, defined as two-view lesions with complete or partial convex borders. 92% of masses sent to MG first were further evaluated with a diagnostic US after the MG, and 96% of these masses were successfully visualized by US. In comparison, 98% of masses sent to US first were satisfactorily assessed with US alone and 2% were further evaluated with a diagnostic mammogram after the US.

Studies have shown that DBT screening increases cancer detection rate,[13–17] reduces recall rate,[14,15,18,19] and improves diagnostic accuracy.[18,20] DBT also has advantages in the diagnostic setting. Several studies have shown that diagnostic DBT is

Figure 1. Flow chart demonstrating the diagnostic pathways of masses recalled from screening digital breast tomosynthesis in our study. DBT, digital breast tomosynthesis; MG, mammography.



comparable to spot compression mammography.[1–3,10,11] Our study is one of the first to evaluate the impact of screening DBT on the subsequent diagnostic work-up of recalled findings. Since DBT improves differentiation of true findings from superimposition of fibroglandular tissue, mass margin visibility, and location assessment[7–9] and is equivalent to routine mammography in the diagnostic setting,[1–3,10,11] it is not surprising that most DBT screen-detected masses in the US first group in our study did not require a diagnostic mammogram. Additionally, the US visibility of masses in the MG first and the US first groups did not vary significantly ($p = 0.42$).

In our study, we found no significant differences in the size of the masses or the density of the breasts in our MG first and US first groups. However, there was a difference in the shape, margins, and pathology of the masses between the two groups. Masses sent to US first were more likely to be oval (87.3% *vs* 63.7%), circumscribed (88.2% *vs* 64.7%), and benign (94.6% *vs*

Table 2. Characteristics of masses visible and not visible on ultrasound in the diagnostic MG and diagnostic ultrasound first groups

		MG first		Ultrasound first	
		Not visible on ultrasound	Visible on ultrasound	Not visible on ultrasound	Visible on ultrasound
		n = 4	*n* = 90	*n* = 2	*n* = 108
	Category	*n* (%) or mean ± SD		*n* (%) or mean ± SD	
Breast density	Non-dense	4 (100)	75 (86.2)	2 (100)	93 (86.1)
	Dense		12 (13.8)		15 (13.9)
Size (cm)[a]		0.4 ± 0.1	0.9 ± 0.5	0.9 ± 0.6	1.0 ± 0.6
Shape[a]	Irregular		16 (17.8)		2 (1.9)
	Oval	1 (25)	57 (63.3)	2 (100)	94 (87.0)
	Round	3 (75)	17 (18.9)		12 (11.1)
Margins[a]	Circumscribed	3 (75)	58 (64.4)	1 (50)	96 (88.9)
	Indistinct		10 (11.1)		8 (7.4)
	Obscured	1 (25)	10 (11.1)	1 (50)	4 (3.7)
	Spiculated		12 (13.3)		
Mass density[a]	High density	1 (25)	16 (17.8)		7 (6.5)
	Equal density	3 (75)	69 (76.7)		90 (83.3)
	Low density		3 (33.3)	2 (100)	8 (7.4)
	Fat-containing		2 (2.2)		3 (2.8)
Final BI-RADS	1/2	3 (75)	44 (48.9)	1 (50)	84 (77.8)
	3		11 (12.2)		7 (6.5)
	4/5	1 (25)	35 (38.9)	1 (50)	17 (15.7)
Pathology[b]	Benign	1 (100)	21 (60)	1 (100)	16 (94.1)
	Malignant		14 (40)		1 (5.9)

MG, mammography; SD, standard deviation.
[a]Mass characteristics are based on findings on screening digital breast tomosynthesis (DBT).
[b]Pathology is for masses that underwent core needle biopsy.

79.2%) than masses sent to MG first. Cysts were more common among the benign lesions in the US first group than the MG first group (71% *vs* 49%). This suggests that radiologists reading the screening DBT exam were more inclined to send benign-appearing masses to ultrasound first. We conjecture that this may have been because the benign appearing masses were likely to be cysts which are typically seen well and definitively assessed by US. For the more suspicious appearing masses, it is likely that the radiologists wished to assess margins and confirm the finding, as has been done historically, with diagnostic MG. However, studies have shown that margin assessment with ultrasound alone can accurately predict the benign or malignant nature of a mass,[21,22] and diagnostic MG before US may not be necessary for margin assessment. It should be noted that the majority of the masses in our study that were eventually not seen by US in both groups had benign features (3/4 round, 1/4 oval, and 3/4 circumscribed in the MG first group versus 2/2 oval and 1/2 circumscribed in the US first group). This suggests that mammographic shape and margins may not significantly affect US visibility. However, size may be a factor as the masses not seen by US were significantly smaller (mean size of 0.5 ± 0.4 cm) than those seen under US (mean size of 0.9 ± 0.5 cm), when assessing the US visibility of masses in the MG first and US first groups cumulatively. Additionally, all masses that were not visible by US in both groups occurred in non-dense breasts. Hence, size and breast density may impact the US visibility of masses, but our conclusions are limited because only six masses in both groups combined were not visible under US, and some did not persist as masses on diagnostic mammography.

In our study, eight masses sent to the MG first group were not further worked up with US. Of these eight masses, three were felt to be lymph nodes present on prior exams, three were other mammographically stable masses, and two were thought to represent overlapping tissue. On retrospective review by the study radiologists, the three lymph nodes and three masses were mammographically stable, and some radiologists might not have recalled these masses on the initial screening exam. If these had gone to US first, they would likely have been seen, but correlation with the current and prior screening mammograms would have been needed to establish their stability and arrive at the final benign assessments for these lesions. Therefore, it is critical

Figure 2. A 64-year-old female recalled from screening DBT for an oval, circumscribed mass in the right breast (circles) and sent to ultrasound directly. (A, B) MLO (A) and CC (B) synthesized 2D (C-View™) screening mammographic views. (C, D) enlarged MLO (C) and CC (D) DBT images. (E) Ultrasound demonstrates an oval, anechoic mass with posterior acoustic enhancement, consistent with a cyst. CC, cranio caudal; DBT, digital breast tomosynthesis; MLO, medio lateral oblique.

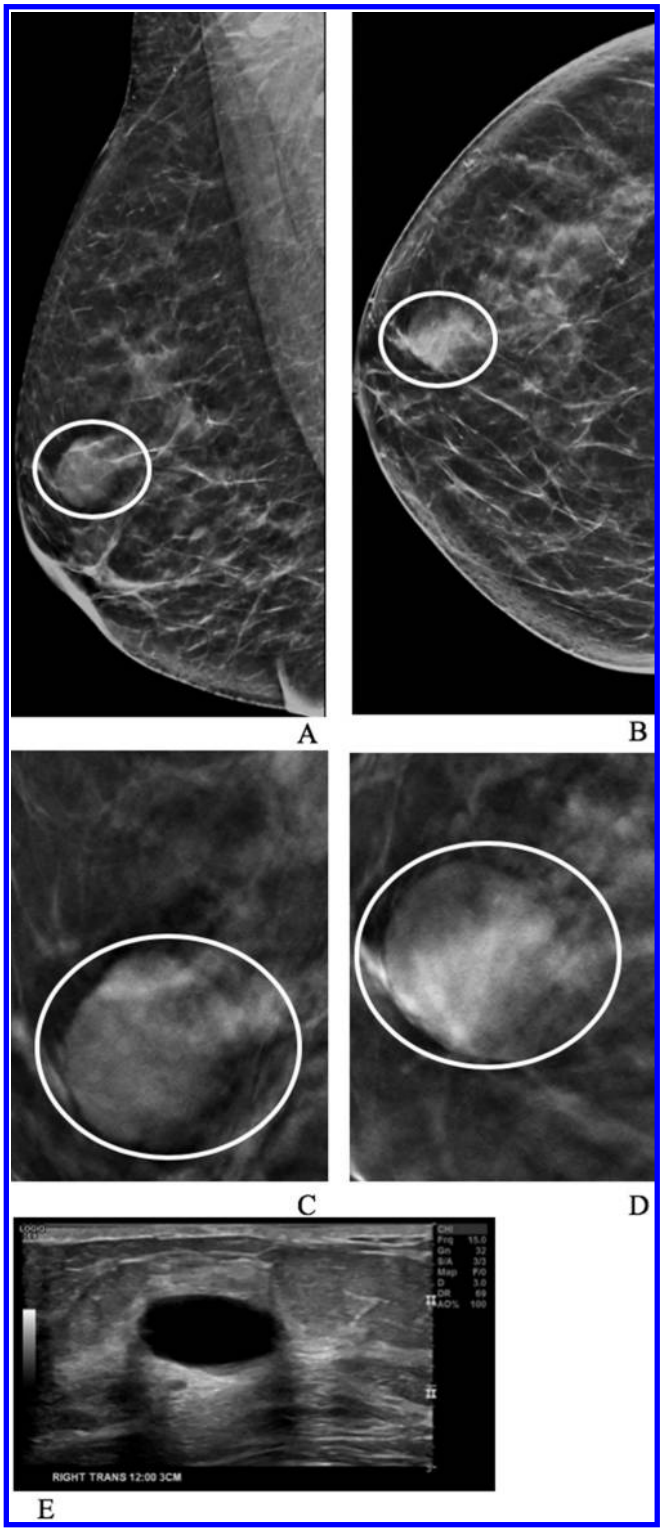
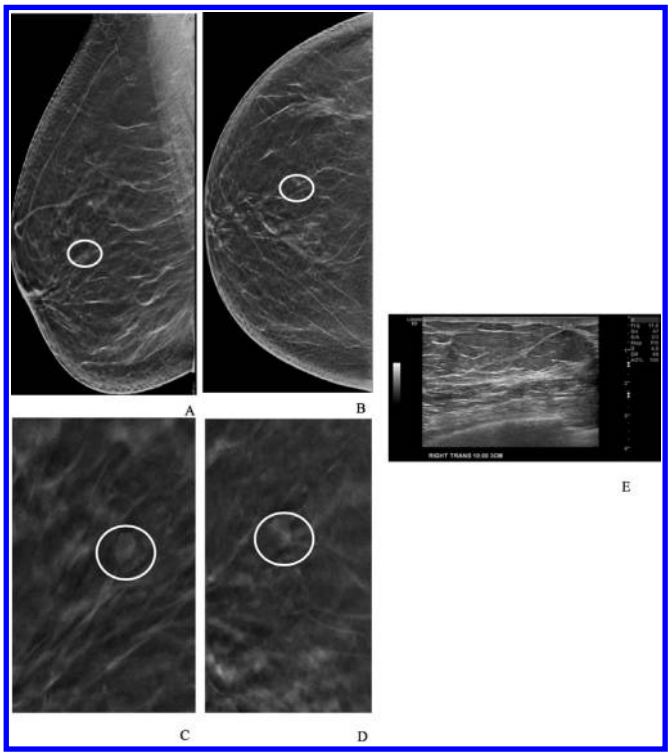


Figure 3. A 45-year-old female recalled from screening DBT for an oval, circumscribed mass in the right breast (circles) and sent to ultrasound first. (A, B) MLO (A) and CC (B) 2D screening mammographic views. (C,D) magnified MLO (C) and CC (D) DBT images. (E) Ultrasound does not show an abnormality to correspond to the mammographic finding. A diagnostic mammogram was subsequently performed and the finding persisted. A stereotactic biopsy was recommended and revealed fibrocystic changes. 2D, two-dimensional; CC, cranio caudal; DBT, digital breast tomosynthesis; MLO, medio lateral oblique.



to note that a DBT screen-detected mass sent directly to diagnostic US cannot be interpreted alone and correlation with other studies, particularly prior mammograms, is critical as in the rest of diagnostic breast imaging.

Additionally, if the two masses representing overlapping tissue in this MG first group went to US first, no US findings would have been detected and a subsequent diagnostic mammogram with spot compression views would have been needed to establish a final assessment. This was the case for one mass in our US first group. Hence, if a mass sent to US directly from DBT screening is not seen, a diagnostic mammogram should be subsequently performed as the mass may represent superimposition of normal tissue. However, this was rare in our study with only 2/110 (2%) of the masses in our US first group not being visualized and needing diagnostic MG after US.

Our study has a few limitations. It is a retrospective, single-institution study. We had 18 readers who read the screening DBT, and there was likely variability in a finding being defined as a mass versus a focal asymmetry by the reader. We did not respectively review findings called focal asymmetries to see if they qualified

as masses. There was a selection bias in which masses went to MG or ultrasound first as the decision was at the discretion of the screening radiologist. We did not account for the years of training or expertise of the interpreting radiologist. The two study radiologists were not completely blinded to the final outcomes of the patients. We are currently in an early stage of DBT screening, which may explain the low positive predictive value of biopsies in this study, and these findings may not be generalizable to practices with multiple years of DBT screening experience. In addition, our practice is a tertiary academic center, and our findings may not be generalizable to other practice settings. Most of our diagnostic ultrasounds are performed initially by breast imaging specialized sonographers with subsequent scanning by the breast imaging radiologist. This practice pattern is not universal and could impact the success rate of mass visibility. Also, due to our recent conversion to universal DBT screening, 2 year follow-up was not obtained in this study. Lastly, we only evaluated masses and did not assess asymmetries and focal asymmetries in this study. Studies evaluating all types of lesions recalled from screening DBT are needed to further establish optimal diagnostic pathways in the era of increasing DBT utilization.

In summary, our study indicates that the majority (98%) of masses, as defined by the BI-RADS, recalled from screening DBT can be adequately assessed with a diagnostic ultrasound alone. When a recalled mass is not seen by ultrasound, a diagnostic mammogram should be subsequently performed to provide a final assessment. This indicates a possibility to forego diagnostic mammography for work-up of majority of DBT screen-detected masses. Eliminating diagnostic MG can decrease cost and radiation to patients and increase the diagnostic workflow efficiency for radiology practices.

ACKNOWLEDGEMENTS

We would like to acknowledge Sonia Watson's help in preparation of this manuscript.

REFERENCES

1. Noroozian M, Hadjiiski L, Rahnama-Moghadam S, Klein KA, Jeffries DO, Pinsky RW, et al. Digital breast tomosynthesis is comparable to mammographic spot views for mass characterization. *Radiology* 2012; **262**: 61–8. doi: https://doi.org/10.1148/radiol.11101763
2. Tagliafico A, Astengo D, Cavagnetto F, Rosasco R, Rescinito G, Monetti F, et al. One-to-one comparison between digital spot compression view and digital breast tomosynthesis. *Eur Radiol* 2012; **22**: 539–44. doi: https://doi.org/10.1007/s00330-011-2305-1
3. Ni Mhuircheartaigh N, Coffey L, Fleming H, O' Doherty A, McNally S. With the advent of tromosynthesis in the workup of mammographic abnormality, is spot compression mammography now obsolete? An initial clinical experience. *Breast J* 2017; **23**: 509–18. doi: https://doi.org/10.1111/tbj.12787
4. Faulk RM, Sickles EA. Efficacy of spot compression-magnification and tangential views in mammographic evaluation of palpable breast masses. *Radiology* 1992; **185**: 87–90. doi: https://doi.org/10.1148/radiology.185.1.1523339
5. Berkowitz JE, Gatewood OM, Gayler BW. Equivocal mammographic findings: evaluation with spot compression. *Radiology* 1989; **171**: 369–71. doi: https://doi.org/10.1148/radiology.171.2.2704800
6. Sickles EA. Combining spot-compression and other special views to maximize mammographic information. *Radiology* 1989; **173**: 571. doi: https://doi.org/10.1148/radiology.173.2.2798895
7. Peppard HR, Nicholson BE, Rochman CM, Merchant JK, Mayo RC, Harvey JA. Digital breast tomosynthesis in the diagnostic setting: indications and clinical applications. *Radiographics* 2015; **35**: 975–90. doi: https://doi.org/10.1148/rg.2015140204
8. Chan HP, Helvie MA, Hadjiiski L, Jeffries DO, Klein KA, Neal CH, et al. Characterization of breast masses in digital breast tomosynthesis and digital mammograms: An observer performance study. *Acad Radiol* 2017; **24**: 1372–9. doi: https://doi.org/10.1016/j.acra.2017.04.016
9. Friedewald SM, Young VA, Gupta D. Lesion localization using the scroll bar on tomosynthesis: Why doesn't it always work? *Clin Imaging* 2018; **47**: 57–64. doi: https://doi.org/10.1016/j.clinimag.2017.07.019
10. Brandt KR, Craig DA, Hoskins TL, Henrichsen TL, Bendel EC, Brandt SR, et al. Can digital breast tomosynthesis replace conventional diagnostic mammography views for screening recalls without calcifications? A comparison study in a simulated clinical setting. *AJR Am J Roentgenol* 2013; **200**: 291–8. doi: https://doi.org/10.2214/AJR.12.8881
11. Whelehan P, Heywang-Köbrunner SH, Vinnicombe SJ, Hacker A, Jänsch A, Hapca A, et al. Clinical performance of Siemens digital breast tomosynthesis versus standard supplementary mammography for the assessment of screen-detected soft-tissue abnormalities: a multi-reader study. *Clin Radiol* 2017; **72**: 95.e9–95.e15. doi: https://doi.org/10.1016/j.crad.2016.08.011
12. Sickles E, D'Orsi C, Bassett L. ACR BI-RADS Mammography. ACR BI-RADS Atlas. In: *Breast Imaging Reporting and Data System*. Reston, VA: American College of Radiology; 2013.
13. Skaane P, Bandos AI, Gullien R, Eben EB, Ekseth U, Haakenaasen U, et al. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. *Radiology* 2013; **267**: 47–56. doi: https://doi.org/10.1148/radiol.12121373
14. Haas BM, Kalra V, Geisel J, Raghu M, Durand M, Philpotts LE. Comparison of tomosynthesis plus digital mammography and digital mammography alone for breast cancer screening. *Radiology* 2013; **269**: 694–700. doi: https://doi.org/10.1148/radiol.13130307
15. Sharpe RE, Venkataraman S, Phillips J, Dialani V, Fein-Zachary VJ, Prakash S, et al. Increased cancer detection rate and variations in the recall rate resulting from implementation of 3D digital breast tomosynthesis into a population-based screening program. *Radiology* 2016; **280**: 981. doi: https://doi.org/10.1148/radiol.2016164018
16. Hofvind S, Hovda T, Holen ÅS, Lee CI, Albertsen J, Bjørndal H, et al. Digital breast tomosynthesis and synthetic 2D mammography versus digital mammography: evaluation in a population-based screening program. *Radiology* 2018;

287: 787–94. doi: https://doi.org/10.1148/radiol.2018171361

17. Ciatto S, Houssami N, Bernardi D, Caumo F, Pellegrini M, Brunelli S, et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. *Lancet Oncol* 2013; **14**: 583–9. doi: https://doi.org/10.1016/S1470-2045(13)70134-7
18. Rafferty EA, Park JM, Philpotts LE, Poplack SP, Sumkin JH, Halpern EF, et al. Assessing radiologist performance using combined digital mammography and breast tomosynthesis compared with digital mammography alone: results of a multicenter, multireader trial. *Radiology* 2013; **266**: 104–13. doi: https://doi.org/10.1148/radiol.12120674
19. Durand MA, Haas BM, Yao X, Geisel JL, Raghu M, Hooley RJ, et al. Early clinical experience with digital breast tomosynthesis for screening mammography. *Radiology* 2015; **274**: 85–92. doi: https://doi.org/10.1148/radiol.14131319
20. Rafferty EA, Park JM, Philpotts LE, Poplack SP, Sumkin JH, Halpern EF, et al. Diagnostic accuracy and recall rates for digital mammography and digital mammography combined with one-view and two-view tomosynthesis: results of an enriched reader study. *AJR Am J Roentgenol* 2014; **202**: 273–81. doi: https://doi.org/10.2214/AJR.13.11240
21. Rahbar G, Sie AC, Hansen GC, Prince JS, Melany ML, Reynolds HE, et al. Benign versus malignant solid breast masses: US differentiation. *Radiology* 1999; **213**: 889–94. doi: https://doi.org/10.1148/radiology.213.3.r99dc20889
22. Hong AS, Rosen EL, Soo MS, Baker JA. BI-RADS for sonography: positive and negative predictive values of sonographic features. *AJR Am J Roentgenol* 2005; **184**: 1260–5. doi: https://doi.org/10.2214/ajr.184.4.01841260

BJR

Received: 15 April 2017 | Revised: 9 June 2017 | Accepted: 26 June 2017

https://doi.org/10.1259/bjr.20170269

Cite this article as:
Giannini V, Mazzetti S, Marmo A, Montemurro F, Regge D, Martincich L. A computer-aided diagnosis (CAD) scheme for pretreatment prediction of pathological response to neoadjuvant therapy using dynamic contrast-enhanced MRI texture features. *Br J Radiol* 2017; **90**: 20170269.

FULL PAPER

A computer-aided diagnosis (CAD) scheme for pretreatment prediction of pathological response to neoadjuvant therapy using dynamic contrast-enhanced MRI texture features

[1,2]VALENTINA GIANNINI, PhD, [1,2]SIMONE MAZZETTI, PhD, [2]AGNESE MARMO, MSc, [3]FILIPPO MONTEMURRO, MD, [1,2]DANIELE REGGE, MD, Prof and [2]LAURA MARTINCICH, MD

[1]Department of Surgical Sciences, University of Torino, Turin, Italy
[2]Department of Radiology, Candiolo Cancer Institute, Torino, Italy
[3]Department of Breast Cancer, Candiolo Cancer Institute, Candiolo, Italy

Address correspondence to: Dr Valentina Giannini
E-mail: *valentina.giannini@ircc.it*

Objective: To assess whether a computer-aided, diagnosis (CAD) system can predict pathological Complete Response (pCR) to neoadjuvant chemotherapy (NAC) prior to treatment using texture features.
Methods: Response to treatment of 44 patients was defined according to the histopatology of resected tumour and extracted axillary nodes in two ways: (a) pCR+ (Smith's Grade = 5) *vs* pCR− (Smith's Grade < 5); (b) pCRN+ (pCR+ and absence of residual lymph node metastases) *vs* pCRN−. A CAD system was developed to: (i) segment the breasts; (ii) register the DCE-MRI sequence; (iii) detect the lesion and (iv) extract 27 3D texture features. The role of individual texture features, multiparametric models and Bayesian classifiers in predicting patients' response to NAC were evaluated.
Results: A cross-validated Bayesian classifier fed with 6 features was able to predict pCR with a specificity of 72% and a sensitivity of 67%. Conversely, 2 features were used by the Bayesian classifier to predict pCRN, obtaining a sensitivity of 69% and a specificity of 61%.
Conclusion: A CAD scheme, that extracts texture features from an automatically segmented 3D mask of the tumour, could predict pathological response to NAC. Additional research should be performed to validate these promising results on a larger cohort of patients and using different classification strategies.
Advances in knowledge: This is the first study assessing the role of an automatic CAD system in predicting the pathological response to NAC before treatment. Fully automatic methods represent the backbone of standardized analysis and may help in timely managing patients candidate to NAC.

INTRODUCTION

Neoadjuvant chemotherapy (NAC) has a leading role in the preoperative treatment of patients with large breast lesions,[1,2] thanks to its clinical advantages. First, it allows a downstage of the tumour so that more conservative therapies could be proposed instead of mastectomy.[3,4] In addition, an improved survival rate has been reported for patients achieving pathological Complete Response (pCR) after the treatment.[5] Moreover, within this therapy, it is possible to monitor the treatment response by measuring the "*in vivo*" tumour changes during and after NAC. It has been demonstrated that NAC could lead to a pCR in up to 30% of patients with breast cancers.[6,7] However, the rate of response to NAC therapy is limited and dependent on the subtypes of breast cancers.[8–15] As a consequence, ineffective chemotherapy could be stopped and unnecessary toxicity for the patient could be avoided.[5] Indeed, early identification of treatment response would be of key importance to improve patients' management, since it would enable the use of alternative, potentially more effective therapies tailored to individual patient.

Magnetic Resonance Imaging (MRI) is currently used in clinical practice to assess the response at the end of NAC.[16] In several studies the variation of morpho-functional features provided by MRI before and during the course of NAC has been demonstrated to be potential "surrogate" biomarkers in the early discrimination between

responder and non-responder patients at the end of the treatment.[16–19] However, the suboptimal reproducibility of these features represents the main limitation for its real application in the daily clinical practice.

Recently, the predictive value of quantitative biomarkers based on textural characteristics of the image has been exploited. Textural analysis has gained wide applications in medical image analysis[20–25] for its ability to characterize the spatial dependence of grey-levels using high order statistics. In particular, it has been proven useful to detect and characterize breast lesions and, more recently, has shown promises in predicting tumour response to therapy.[15,26–29] However, most of previous studies evaluated quantitative parameters acquired after the completion of, at least, one cycle of chemotherapy, therefore their findings could not be used to redirect treatment regimens for patient who are not likely to respond to NAC.[29] Moreover, the majority of previous studies did not use as reference standard the pathologic response after surgery, which is the time point better associated with the final prognosis of the patient.[26] To the best of our knowledge, only three previous studies exploited the role of two-dimensional textural analysis of Dynamic Contrast Enhanced (DCE)-MRI obtained prior to treatment to predict pathological response to NAC.[26,28,29] In particular, Teruel et al[26] on a dataset of 58 patients showed that 4 individual texture features were significantly correlated with pCR, with an area under the receiver characteristics curve (ROC) curve (AUC) of 0.68. Michoux et al[28] and Golden et al[29] developed multiparametric classifiers to predict pathological non-response to therapy using 69 and 60 patients, respectively. The first reached a predictive accuracy of 68%, while the latter obtained an AUC equal to 0.68 in predicting pathological Complete Response. These previous results were promising, however further investigations are needed to better generalize these findings, *i.e.* including larger groups of tumour subtypes, to exploit the whole tumour characteristics by using a three-dimensional (3D) approach, and to standardize the methods, by developing automatic computer-aided diagnosis (CAD) scheme to segment the tumours.

In this scenario, the objective of this proof-of-concept study is to assess the feasibility of a CAD scheme able to automatically extract quantitative 3D biomarkers and classify each patient according to the likelihood of pCR to NAC, by considering also the different immunohistochemical subtypes.

METHODS AND MATERIALS

Patients

Patients were retrospectively included from a prospective single-centre observational study performed at our institution between November 2010 and February 2014, having the following inclusion criteria: (a) age between 18 and 65 years; (b) presence of imaging-guided core-biopsy proven Stage II/III operable breast cancer (T > 3 cm) or inoperable locally-advanced breast cancer and (c) unifocal or multiple masses at baseline MRI. This study was conducted in compliance with the ethical regulatory issues of our Institution and patients were asked to provide written informed consent before entering the study. All patients were evaluated by our multidisciplinary team clinic before and after the completion of NAC. Enrolled subjects underwent 4 cycles of treatment based on a combination of doxorubicin 50 mg m^{-2} bolus i.v. followed by paclitaxel 175 mg m^{-2} as a 3 h i.v. infusion. In patients with baseline left ventricular ejection fraction <55%, doxorubicin was omitted and monochemotherapy with paclitaxel 225 mg m^{-2} as a 3 h i.v. infusion was administered. Cycles were repeated every 21 days if absolute neutrophil count ≥ 1500 μl^{-1} and platelets ≥ 100,000 μl^{-1}, and otherwise delayed until resolution of hematologic toxicity.[17] Women with HER2-positive breast cancer received also trastuzumab. All patients underwent surgery, that was performed between 14 and 35 days after the completion of NAC.

Immunoistochemical analysis and pathological tumour response

Immunoistochemical (IHC) analysis was performed on specimens from imaging-guided core-biopsies. Positivity for Estrogen Receptor (ER) and Progesteron Receptor (PgR) status was defined as immunostaining in ≥1% of invasive tumour cells, while Ki67 was considered positive when expressed by more than 14% of tumour cells.[30] HER2 status was assessed according to ASCO/CAP Guideline recommendation.[31] Positivity was defined as 3+ score by IHC in >30% of invasive tumour cells using the HercepTest (Dako, Glostrup, Denmark). Equivocal cases at IHC (2+ score or 3+ in ≤30% of invasive tumour cells), were subjected to fluorescence *in situ* hybridization analysis. A ratio of HER2 gene signals to chromosome 17 signals of more than 2.2 was used as a cut-off to define HER2 gene amplification. Tumours were divided into Luminal A (ER-positive and Ki67 <14% and HER2-negative), Luminal B (ER-positive and Ki67 ≥14% and either HER2-positive or HER2-negative), HER2-enriched (ER-negative and HER2-positive) and triple-negative (ER-negative and HER2-negative).[30] The histopathological tumour response was evaluated using a five-point assessment scheme described by Smith et al[32] Grade 1, some alteration to individual malignant cells but no reduction in overall numbers as compared with the pre-treatment core biopsy; Grade 2, a mild loss of invasive tumour cells but overall cellularity still high; Grade 3, a considerable reduction in tumour cells up to an estimated 90% loss; Grade 4, a marked disappearance of invasive tumour cells such that only small clusters of widely dispersed cells could be detected; and Grade 5, no invasive tumour cells identifiable in the sections from the site of the previous tumour, that is, only *in situ* disease or tumour stroma remained. Grade 5 response was deemed to represent a pCR of the primary cancer. Pathological Complete Response at axillary level was classified as absence of residual invasive tumour in the lymph nodes.

MRI protocol

MRI examination was carried on before the first cycle of chemotherapy (baseline), within 2 weeks from the second cycle (intermediate) and after the completion of the planned treatment, within 1 week before surgery. MRI was acquired with a 1.5T equipment and dedicated phased-array 8-channel coil (HDx Signa Excite, GE HealthCare Milwaukee, WI), with the patient in the prone position and following the recommended technical requirements for breast imaging.[16] In particular, the DCE-MRI study was performed using a fat-sat 3D fast spoiled gradient-echo sequence (VIBRANT®,

General Electric, Milwaukee, WI) having slice thickness = 2.6 mm; acquisition matrix = 416 × 416, and flip angle = 10°. A total of six scans were acquired for each study: one baseline, 4 contrast-enhanced frames with 90 s time resolution, and one delayed frame acquired 8 min after i.v. contrast agent administration (Multihance, Bracco Imaging, Milan, Italy). Contrast-enhanced study was started simultaneously with the bolus injection of 0.1 mmol kg^{-1} of gadolinium chelate, infused in the antecubital vein by power injector, at a rate of 2 ml s^{-1} and followed by a saline flush. 33 patients were acquired along the axial plane with repetition time/echo time (TR/TE) = 5.4/2.6 ms and pixel size = 0.39 mm^2, while 11 patients were acquired using a sagittal sequence with TR/TE = 4.8/1.9 ms and pixel size = 0.22 mm^2.

Image analysis

The lesion segmentation method is based on a fully automatic algorithm previously developed[33] that consists of different steps. First, the sagittal volumes are converted into axial images and conveniently resampled (upsampled along the *x*-axis and downsampled along the *z*-axis). Then, an elastic registration is performed to align the enhanced DCE-MRI frames to the unenhanced one, thus correcting for misalignments due to patient motions. Once all datasets are registered, the breasts are segmented using the algorithm developed by Giannini et al[34] conveniently adapted to the fat-sat 3D fast spoiled gradient-echo sequence. Finally, the tumour is automatically segmented on the subtracted mean intensity projection image over time normalized by the contrast enhancement of the mammary vessels. This normalization has been demonstrated useful to cope with the significant variations of signal intensities between patients due to different scanners, coils, acquisition modalities, types and amounts of contrast agent injected, patients' physiology, and other external factors, and to provide a reliable automatic segmentation.[33] Since it would be possible that the automatic algorithm produces some false positive (FP), an experienced radiologist (more than 20 years of experience in interpreting breast MRI) selected, for each patient, the true positive among the segmented areas. In cases of multifocal disease, the largest tumour was selected as the index tumour and taken into consideration for the subsequent steps.

Features extraction

Twenty-seven 3D textural features were extracted from the subtracted post-contrast first frame of the pre-NAC (baseline) DCE-MRI studies. In particular, 17 features were derived from the grey-level co-occurrence matrices (GLCM),[35] and 10 were computed from the grey-level run length method (GLRLM).[36] The GLCM is a tabulation of how often different combinations of pixel brightness values (*i.e.* grey levels) occur between neighbouring voxels in an image along a given direction. Therefore, the GLCM allows the calculation of second order statistics, *i.e.* describing the relationship between groups of pixels in the image. Conversely, in the GLRLM method, each element $GLRLM_\theta(i,j)$ represents the number of occurrences of the *j* adjacent elements with grey level *i* calculated in direction θ.

Before extracting texture parameters, we first equalized the region of interest (ROI) histogram by rescaling into 256 bins the signal intensities within each ROI between the first and the 99th percentile. Then, to take into account the contribution of all voxels adjacent to the reference one, the GLCMs and the GLRLMs were generated for each of the 13 directions characterizing a 3D image. In the case of the GLCM calculation, a distance of one voxel was chosen. Finally, the 13 matrices were averaged to enable the method to be rotationally invariant to the distribution of texture. Finally, the following 17 features were obtained from GLCMs: contrast,[35] correlation1,[35] correlation2,[37] energy,[35] entropy,[35] homogeneity,[35] sum variance,[35] sum entropy,[35] sum average,[35] difference variance,[35] difference entropy,[35] information measure of correlation1,[35] information measure of correlation2,[35] cluster prominence,[38] cluster shade,[39] dissimilarity[40] and maximum probability.[41] Moreover, the following 10 features were computed form the GLRLMs: short run emphasis,[41] long run emphasis,[41] grey level distribution,[41] run length distribution,[41] low grey level runs emphasis,[42] high grey level runs emphasis,[42] short run low grey level emphasis,[43] short run high grey level emphasis,[43] long run low grey level emphasis[43] and long run high grey level emphasis.[43] All texture features were computed using an in-house software implemented using C++ and the ITK libraries.[44]

Statistical analysis

Response to treatment was dichotomized as following:

(a) at breast level: pCR+ (Smith's Grade = 5) *vs* pCR− (Smith's Grade < 5);
(b) at both breast and axillary level: pCRN+ (Smith's Grade = 5 plus either complete absence of residual nodal metastases or presence of nodal micrometastasis) *vs* pCRN− (Smith's Grade < 5 and any other status of axillary lymph node metastases).

Tumour subtypes and immunoistochemical characteristics were classified as previously described and differences between pCR+ and pCR−, and between pCRN+ and pCRN− tumours were assessed using the Fisher's exact mid-P test.

Age and tumour size were expressed as median with interquartile ranges in parentheses and their association with NAC response was evaluated using the Mann-Whitney test.

The relationship between outcome (pCR and pCRN) and texture features was explored by two approaches: mono-parametric and multiparametric. We first evaluated the performance of each 3D texture parameter in predicting the pathological response to therapy at breast with or without nodal level, using the ROC curve. AUC, sensitivity and specificity at the best cut-off were computed. The best cut-off is the one that maximizes the Youden index, which is the cut-point of the ROC curve that optimizes the biomarker's differentiating ability when equal weight is given to sensitivity and specificity.[45] A p-value < 0.05 was considered as indicating a AUC significantly greater than 0.5.

Analyses were performed with a statistical software (MedCalc Statistical Software version 17.4, Ostend, Belgium).

Afterwards, we combined the features into two different multiparametric classifiers. For both classifiers, we performed a

feature selection step to discard uninformative characteristics in order to prevent over-fitting, speed up the learning process as well as improve the model's interpretability. The first classifier was the logistic regression model, in which features were selected using the backward regression method. This method consists in entering all the variables in the model and sequentially removing (one-at-a-time) those that are non-significant ($p > 0.20$) for the model (*i.e.* having the largest p).

Subsequently, a Bayesian classifier was tested. In this case, the classical approach referred to as the "filter approach" was used to perform feature selection. This method consists in first ranking all features based on a criterion independent of the classifier, and then, selecting features from this rank list by setting a threshold which accounted for the classifier performance. Two different ranking methods were used: the Fisher (F)-score method[46] and the value of the AUC of the individual features. To avoid to arbitrarily select a threshold on the number of features, we used a previously published method,[47] in which the first n features were extracted from the sorted list and the classification performance achieved with this feature subset was computed. The classification performance was thus derived as a function of the number of the n first-ranked features. Performance was measured as the AUC, accuracy, sensitivity and specificity derived from a leave-one-out cross-validation.[48] Leave-one-out approach involves training on all but one case, testing the classification on the left out patient, and repeating the procedure until each case has been tested individually. Accuracy, sensitivity and specificity were then estimated and used to identify the set of features that yielded best predictive models.

RESULTS

44 patients were included in the dataset. Patients and lesions characteristics are reported in Table 1. Age, tumour size and subtypes were not different between pCR+ and pCR− and between pCRN+ and pCRN−.

ER positive and PgR positive tumours were less represented in the pCR+ and pCRN+ groups, while no differences were observed between Ki67 positive and negative tumours. pCR+ were significantly more represented in HER2 positive group and in tumours with both negative ER and PgR. pCRN+ were significantly more represented in tumours with both negative ER and

Table 1. Patient characteristics and receptor status stratified according tumour response to neoadjuvant chemoterapy

	All ($n = 44$)	pCR+ ($n = 15$)	pCR− ($n = 29$)	p-value	pCRN+ ($n = 13$)	pCRN− ($n = 31$)	p-value
Age	46 (39–53)	46 (40–53)	47 (38–52)	0.9802[a]	46 (40–54)	47 (38–53)	0.8773[a]
Size	37.5 (30–50)	38 (30–50)	36 (30–50)	0.7369[a]	38 (30–52)	36 (30–49)	0.4965[a]
Histological type							
IDC	39	13 (33.3%)	26 (66.7%)	0.8234[b]	13 (33.3%)	26 (66.7%)	0.2223[b]
ILC	3	1 (33.3%)	2 (66.7%)	0.7701[b]	0	3 (100%)	0.3739[b]
Mucinous cancer	1	0	1 (100%)	0.6705[b]	0	1 (100%)	0.6477[b]
Squamous cancer	1	1 (100%)	0	0.1705[b]	0	1 (100%)	0.6477[b]
Immunohistochemical							
ER positivity	29	4 (13.8%)	25 (86.2%)	0.0002 [b]	4 (13.8%)	25 (86.2%)	0.0027[b]
PgR positivity	27	4 (14.8%)	23 (85.2%)	0.0013[b]	4 (14.8%)	23 (85.2%)	0.0114[b]
ER– & PgR–	14	10 (71.4%)	4 (28.6%)	0.0002 [b]	8 (57.1%)	6 (42.9%)	0.0076[b]
Ki67 > 14%	38	14 (36.8%)	24 (63.2%)	0.0939 [b]	12 (31.6%)	26 (68.4%)	0.4959[b]
HER2 positivity	16	9 (56.2%)	7 (43.8%)	0.0122[b]	7 (43.8%)	9 (56.2%)	0.1307[b]
Subtypes							
Luminal A	4	0	4 (100%)	0.1899[b]	0	4 (100%)	0.1865[b]
Luminal B/HER2–	18	2 (11.1%)	16 (88.9%)	0.0144[b]	2 (11.1%)	16 (88.9%)	0.0313[b]
Luminal B/HER2+	5	2 (40.0%)	3 (60.0%)	0.8273[b]	2 (40.0%)	3 (60.0%)	0.4619[b]
Triple negative	4	3 (75.0%)	1 (25.0%)	0.0509[b]	3 (75.0%)	1 (25.0%)	0.0379[b]
HER2-enriched	10	7 (70.0%)	3 (30.0%)	0.0039[b]	5 (50.0%)	5 (50.0%)	0.0871[b]

pCR+ Smith's Grade = 5; pCR−, Smith's Grade < 5; pCRN+, pCR+ plus either complete absence of residual nodal metastases or presence of nodal micrometastasis; pCRN−, Smith's Grade < 5 and any other status of axillary lymph node metastases. IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; ER, estrogen receptor; PgR, progesterone receptor and HER2, epidermal growth factor receptor 2. Age and tumour size are expressed as median with interquartile ranges in parentheses, while other measurements are expressed as counts with percentages in parenthesis.

[a]p-value of the Mann-Whitney test.

[b]p-value of the Fisher's exact mid-P test.

Table 2. Performance of individual parameters in predicting pCR, measured from statistically significant ROC curves with cut-offs determined on the basis of Youden index

Variable	AUC	SE	Cut-off	Sensitivity (%)	Specificity (%)
Contrast	0.722	0.0851	>3385	46.7	93.1
Correlation	0.715	0.0848	$\leq 1.465 \times 10^{-4}$	60.0	82.8
Sum variance	0.674	0.0813	>7,4770	86.7	51.7
Difference variance	0.699	0.0874	$\leq 2.46 \times 10^{-5}$	46.7	89.7
Difference entropy	0.713	0.0859	>4.751	60.0	82.8
LRE	0.676	0.0806	>1.247	80	58.6
LRHGE	0.708	0.0777	>2,4213	93.3	55.2

AUC, area under the ROC curve; SE,standard error;LRE, long run emphasis, LRHGE, low run high grey level emphasis; pCR+ Smith's Grade = 5; pCR−, Smith's Grade < 5; pCRN+, pCR+ plus either complete absence of residual nodal metastases or presence of nodal micrometastasis; pCRN−, Smith's Grade < 5 and any other status of axillary lymph node metastases.

PgR. According to tumour subtypes, pCR+ were significantly more represented in the HER2-enriched group, while pCRN+ were significantly more represented in the triple negative group. Both pCR+ and pCRN+ were less represented in patients with luminal B/HER2− tumours.

Mono-parametric approach

When individual parameters were compared with pCR outcome, we found 7 statistically significant parameters with AUC greater than 0.5 (Table 2). The feature with the highest AUC was contrast with a sensitivity and specificity at the best cut-off equal to 46.7 and 93.1%, respectively. Two examples of the image processing pipeline are shown in Figures 1 and 2. Contrast has been found statistically higher for pCR+ tumours (Figure 1) than for the pCR− (Figure 2), while correlation showed lower values for pCR+ tumours. Considering texture features from GLRLM, we obtained that higher long run emphasis and higher low run high grey level emphasis were correlated with better response to therapy. The 3D parameters correlated ($p < 0.05$) with the pCRN outcome were cluster shade, sum variance, long run emphasis and higher low run high grey level emphasis (Table 3). The highest AUC was reached using low run high grey level emphasis, and was equal to 0.747, with a 100% sensitivity and 54.8% specificity at the best cut-off point.

Multi-parametric approach

Using the logistic regression classifier, we obtained a model to predict pCR response in which two parameters were kept: sum

Figure 1. 60 mm invasive ductal carcinoma in a 74 year-old female. (a) Result of the breast segmentation algorithm superimposed to the first contrast-enhanced image subtracted to the precontrast one. (b) Normalized maximum intensity projection over time image of the breast region. (c) Tumour segmentation obtained by the computer-aided diagnosis scheme superimposed to the maximum intensity projection over time image. Once the segmentation has been obtained, the radiologist selected the tumour to exclude false positive findings (red box). (d) Three-dimensional render of the mask of the tumour multiplied for the subtracted first contrast-enhanced frame. The 2 most discriminative features of both grey-level co-occurrence matrices and grey-level run length method algorithm are reported for this tumour. Pathological response Grade = 3/5, estrogen receptor status = 99%, progesterone receptor status = 13% and Ki67 = 11%, HER2 status = negative.

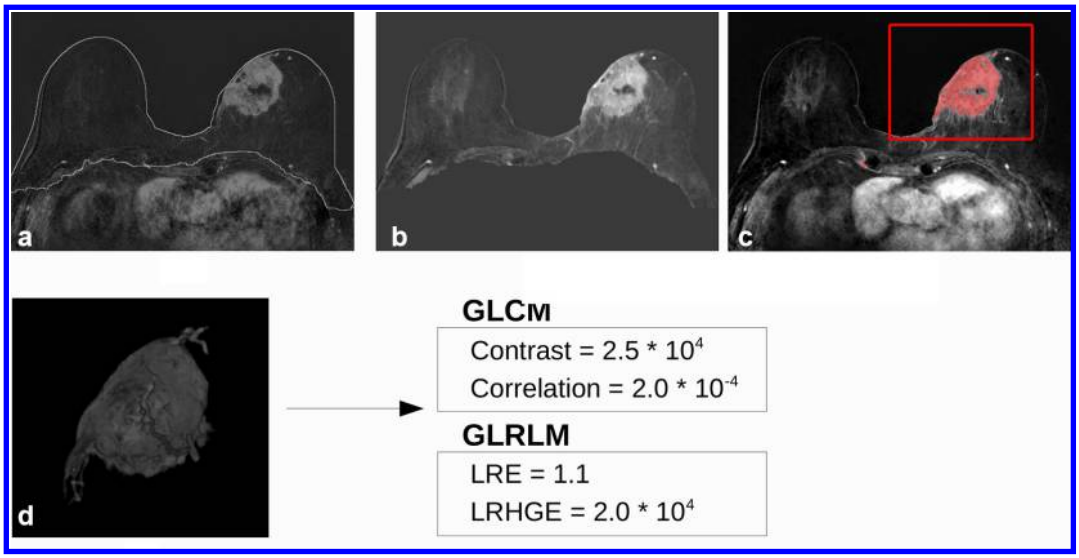
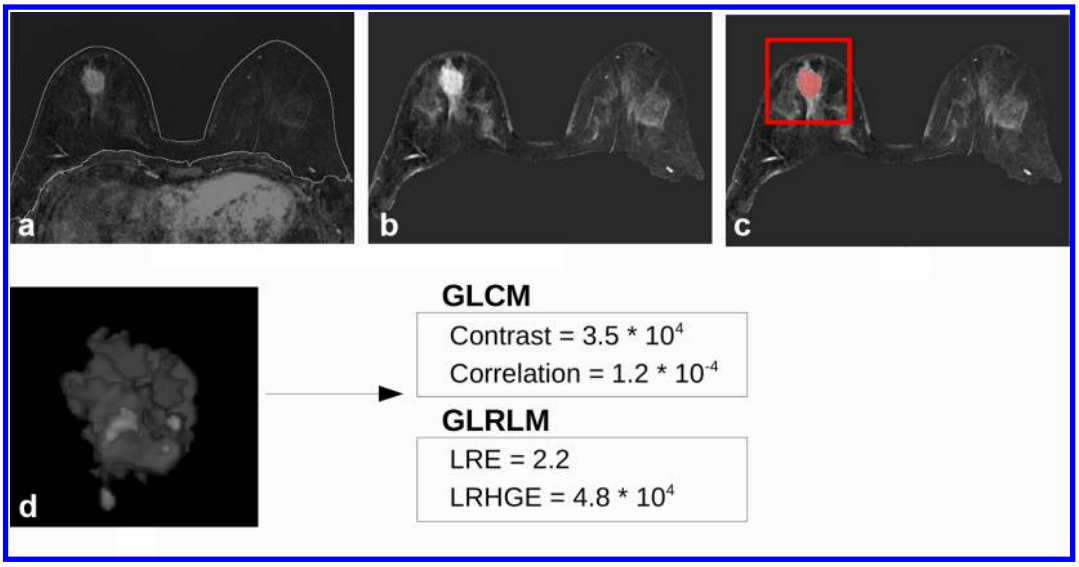


Figure 2. 26 mm invasive ductal carcinoma in a 43 year-old female. (a) Result of the breast segmentation algorithm superimposed to the first contrastenhanced image subtracted to the precontrast one. (b) Normalized maximum intensity projection over time of the breast region. (c) Tumour segmentation obtained by the CAD scheme superimposed to the maximum intensity projection over time image. Once the segmentation has been obtained, the radiologist selected the tumour to exclude false positive findings (red box). (d) Three-dimensional render of the mask of the tumour multiplied for the subtracted first contrastenhanced frame. The 2 most discriminative features of both GLCM and GLRLM algorithm are reported for this tumour. Pathological Complete Response (5/5), estrogen receptor status = 20%, progesterone receptor = status 25%, Ki67 = 30%, HER2 status = positive.



variance (p = 0.04) and difference entropy (p = 0.01). The AUC of the model was 0.795 (95% CI [0.647–0.902]), with a sensitivity and a specificity at the best cut-off (0.36) of 80 and 69%, respectively (Figure 3). When the logistic regression classifier was used to predict pCR with a lymph nodal response, *i.e.* pCRN, three texture features were maintained in the model: cluster shade (p = 0.04), long run emphasis (p = 0.11) and low run high grey level emphais (p = 0.19). The AUC of the model was equal to 0.764 (95% CI [0.612–0.879]), with a sensitivity of 46% and a specificity of 100% at the best cut-off (0.53).

When predicting pCR with the Bayesian classifier the best accuracy (70%) has been obtained when the first 6 features ranked by the F-score (cluster shade, correlation, contrast, difference entropy, correlation 1 and difference variance) were fed into the classifier. Specificity and sensitivity were equal to 72 and 67%, respectively. The Bayesian classifier was also tested to predict the pCR associated with a lymph node response, *i.e.* pCRN. In this case, the best results have been obtained when the features were ranked according to the AUC value of each individual feature. When the first 2 features (low run high grey level emphasis and long run emphasis) were included, the highest sensitivity has been reached (69%) with a specificity of 61% and an accuracy of 64%.

DISCUSSION

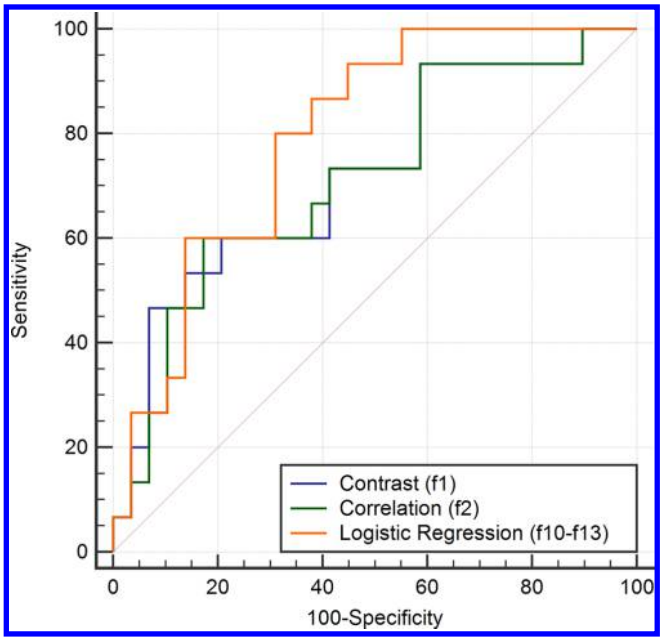
In this study we developed and tested a CAD scheme that aims to predict pathological response to NAC based on 3D texture features extracted from an automatic segmentation of the tumour at baseline MRI examination. We demonstrated that some individual texture features can discriminate between responder and non-responders at breast and axillary level before NAC. When analysing pathologic response at breast level (pCR+ *vs* pCR–),

Table 3. Performance of individual parameters in predicting pCRN, measured from statistically significant ROC curves with cut-offs determined on the basis of Youden index

Variable	AUC	SE	Cut-off	Sensitivity (%)	Specificity (%)
Cluster shade	0.685	0.0842	≤7039	92.3	41.9
Sum variance	0.687	0.0817	>74770	92.3	51.6
LRE	0.712	0.0790	>1.258	84.6	61.3
LRHGE	0.747	0.0731	>24213	100.0	54.8

AUC, area under the ROC curve; SE, standard error;LRE, long run emphasis, LRHGE, low run high grey level emphasis; pCR+ Smith's Grade = 5; pCR–, Smith's Grade < 5; pCRN+, pCR+ plus either complete absence of residual nodal metastases or presence of nodal micrometastasis; pCRN–, Smith's Grade < 5 and any other status of axillary lymph node metastases.

Figure 3. Receiver operating characteristic (ROC) curve of the logistic regression classifier compared with the ROC curves of the most discriminative individual features in predicting complete pathological response.



seven parameters reached statistical significance, with very different diagnostic performances (three parameters reached high sensitivity, while the other four had good specificity). When considering the prediction of pathological response at both breast and axillary level (pCRN+ *vs* pCRN−), four parameters were able to discriminate, before the treatment, between patients achieving pCRN+ and subjects obtaining pCRN−, and all of them reached high sensitivity, with poor specificity. These results demonstrated that DCE-MRI could be used as a non-invasive examination to predict at the fair level which patients will likely respond to NAC.

Previously, only Michoux et al[28] and Teruel et al[26] assessed the role of individual texture features in predicting pathological response to NAC. Michoux et al[28] demonstrated that the inverse difference moment, which is a measure of the local homogeneity of the grey levels, was inversely correlated to the NAC response, and that it showed the highest AUC (0.711) in predicting NAC response. Analogously, we obtained that higher contrast, which is inversely correlated to homogeneity, was an index of a better response to NAC (AUC = 0.722). This behaviour might be explained by the higher vascularity which characterizes tumours that are more likely to respond to NAC, *i.e.* pCR+.[49] Indeed, in higher vascularized tumours the general shape of the blood vessels is altered and deformed, becoming very rough and resulting in increased irregularity. Thus, when a contrast agent is injected, if the tumoural region shows a higher vascularization, there is enhanced intensity of blood vessels in the images and a corresponding increase of contrast values.[50] In addition, in our dataset we showed that patients with negative steroid receptor status were more likely to achieve pCR+ and it has been demonstrated that VEGF+ phenotype is more frequently associated with a negative steroid receptor status.[51] Therefore, higher contrast in pCR+ patients might be also explained by the fact that most pCR+ patients had a negative steroid receptor status. On the other side, compared with the study of Michoux et al[28] we did not obtain a significant AUC when considering the homogeneity of the ROI, *per se*. The reason might be twofold. First, Michoux et al[28] derived the texture features from a 2D ROI, defined as the largest region of contiguous pixels with the same behaviour in amplitude and wash-in of the signal intensity *vs* time curve, rather than using a 3D segmentation of the tumour. Therefore, they did not evaluate differences within different slices of the tumour that might affect the homogeneity of the ROI. Second, their dataset included only patients with invasive ductal carcinoma, therefore homogeneity could be biased by the choice of a single subtype of cancer. Teruel et al[26] showed that sum variance and sum entropy were the most predictive parameters for NAC response. Despite the great variations in imaging protocols and patient's cohort, we obtained similar results when considering sum variance and difference entropy, demonstrating the great potential of texture features as standard and robust quantitative biomarkers.

One of the main advantages of our work relies on the fact that we used the pathological response after surgery as reference standard, which represents the best time point better associated with the final prognosis of the patient, also considering the evolution of the patients from their clinical response to their final pathological outcome.[26]

A second important result of our study is that we demonstrated that it could be feasible to develop robust models to early predict the response to NAC at both breast and axillary level by combining different texture features in a multiparametric approach. In particular, we have shown that a two-parameter logistic regression classifier can predict pCR with higher accuracy than the mono-parametric approach (0.795 *vs* 0.722). More importantly, we demonstrated that a cross-validated Bayesian classifier can reach an accuracy of 70% in predicting pCR. This result could appear similar to that obtained with some individual features, however comparison is flawed as the cross-validation approach could not be performed in the mono-parametric analysis. Cross-validation is necessary to get an unbiased estimate of the predictive accuracy of the model in patients that were not used to train the classifier and to assess the relevance of the working hypothesis and its clinical applicability. Only two previous studies developed cross-validated classifier to predict pre-treatment pathological response to NAC based on texture features. Michoux et al[28] obtained an accuracy of 68% with a specificity of 62% and sensitivity of 84% in predicting non-responder patients using a *k*-means classifier. Similarly, Golden et al[29] reported that the use of 31 GLCM-derived features prior to treatment was able to predict pCR with an AUC of 0.68 for patient with triple-negative breast cancers. In our work we reached a slightly higher accuracy, and we also took advantages of two innovative approaches that could overcome some limitations of previous studies.

First, our dataset comprised both invasive ductal carcinoma and invasive lobular carcinoma, which is more representative of the

day clinical practice, and it comprises different tumour subtypes. Second, we developed a fully automatic CAD scheme able to segment the whole tumour and provide a 3D mask which is less operator dependent.[52] To the best of our knowledge, there are no studies that extracted pre-treatment texture features from an automatically segmented 3D mask of the tumour. Fully automatic lesion detection could represent the backbone of standardized analysis and may contribute to the introduction of quantitative biomarkers for a timely management of patients candidate to NAC. Moreover, the automatic segmentation is also able to reduce the post-processing time for the radiologist. Another strength of our work relies on the fact that we tried to define some imaging biomarkers that were not only related to the pathological response at breast level, but also to the overall loco-regional response, because of its relevant clinical implications in terms of adjuvant treatment and prognosis. Indeed, it has been shown that the achievement of pCR at both breast and axillary node levels is associated with improved long term clinical outcomes.

There are some limitations of our work. First, this is a retrospective study based on a limited number of patients, therefore a second study with a larger dataset should be performed to validate our results. However, in this proof-of-concept study we obtained promising results from texture features extracted from an automatic 3D segmentation of the tumour, and this could set the basis for the development of computer-assisted prediction solution for breast MRI. Second, in our study a specific timepoint corresponding to the enhancement peak on intensity-time curves has been evaluated, based on the findings of previously published works. Indeed, Ahmed et al[15] evaluated the correlation between pre-NAC texture features and the clinical response to NAC at different timepoints, and they found significant differences occurring at 1 and 2 min after contrast injection. Further tests on late timepoints should be conducted to evaluate whether different timepoints could better predict pathological response to NAC at breast and nodal level. A possible further limitation could be the inclusion of mass tumour only, due to the fact that in our institution tumour response in patients with non-mass cancers is monitored by clinical examinations and conventional imaging. However, our clinical work-up aims to combine clinical evidences, resource optimization and tailored treatment. It has been demonstrated that accuracy of MRI in the assessment of tumour response is higher in mass lesions than in diffuse cancers[53] and that the likelihood of conservative treatment after NAC is lower in patients with non-mass lesions at baseline.[54] As a consequence, we prefer to reserve a costly examination to those patients who more frequently benefit from both MRI and conservative surgical treatment after NAC.

In conclusion, in this work we demonstrated that a CAD scheme, that extract texture features from an automatically segmented 3D mask of the tumour, could help in predicting pathological response to NAC at both breast and axillary level, which is a response more relevant in terms of adjuvant treatment and prognosis. From a clinical point of view, such methods should, ideally, obtain a very high negative predictive value (*i.e.* ≥ 90%), thus achieving a twofold advantage for patients: (a) an early modification of the treatment for those patients that are not likely responding, (b) a reduction of toxicity due to unnecessary treatments. In this study, we reached a negative predictive value of 81% when predicting pCR+ patients, therefore additional research should be performed to increase the performance of our method. In general, we would like to develop and test other statistical classifiers (*e.g.* support vector machine) and/or unsupervised algorithms, *i.e.* *k*-means or hierarchical clustering to improve classification performance. Besides, it would be interesting to combine texture features with dynamic and pharmacokinetics modelling, thus adding others functional information. Finally, it would be necessary to validate our promising results on a larger prospective cohort of patients and using different imaging acquisition protocols.

However, findings of this work might help in developing scheme that will help to better select patients eligible for NAC, thus avoiding unnecessary treatment if the regimen is predicted to be unsuccessful.

REFERENCES

1. Kaufmann M, von Minckwitz G, Smith R, Valero V, Gianni L, Eiermann W, et al. International expert panel on the use of primary (preoperative) systemic treatment of operable breast cancer: review and recommendations. *J Clin Oncol* 2003; **21**: 2600–8. doi: https://doi.org/10.1200/JCO.2003.01.136
2. Heys SD, Hutcheon AW, Sarkar TK, Ogston KN, Miller ID, Payne S, et al. Neoadjuvant docetaxel in breast cancer: 3-year survival results from the aberdeen trial. *Clin Breast Cancer* 2002; **3**(Suppl. 2): S69–S74. doi: https://doi.org/10.3816/CBC.2002.s.015
3. Mieog JS, van der Hage JA, van de Velde CJ. Preoperative chemotherapy for women with operable breast cancer. *Cochrane Database Syst Rev* 2007; **18**: CD005002.
4. Fisher B, Bryant J, Wolmark N, Mamounas E, Brown A, Fisher ER, et al. Effect of preoperative chemotherapy on the outcome of women with operable breast cancer. *J Clin Oncol* 19 19; **16**: 2672–85; **16**: 2672–85. doi: https://doi.org/10.1200/JCO.1998.16.8.2672
5. Kim H, Kim HH, Park JS, Shin HJ, Cha JH, Chae EY. Prediction of pathological complete response of breast cancer patients undergoing neoadjuvant chemotherapy: usefulness of breast MRI computer-aided detection. *Br J Radiol* 2015; **88**: 20150143.
6. Press MF, Sauter G, Buyse M, Bernstein L, Guzman R, Santiago A, et al. Alteration of topoisomerase II-alpha gene in human breast cancer: association with responsiveness to anthracycline-based chemotherapy. *J Clin Oncol* 2011; **29**: 859–67. doi: https://doi.org/10.1200/JCO.2009.27.5644
7. Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast

cancer. *Lancet* 2003; **362**: 362–9. doi: https://doi.org/10.1016/S0140-6736(03)14023-8

8. Barbi GP, Marroni P, Bruzzi P, Nicolò G, Paganuzzi M, Ferrara GB. Correlation between steroid hormone receptors and prognostic factors in human breast Cancer. *Oncology* 1987; **44**: 265–9. doi: https://doi.org/10.1159/000226492
9. von Minckwitz G, Sinn HP, Raab G, Loibl S, Blohmer JU, Eidtmann H, et al. Clinical response after two cycles compared to HER2, Ki-67, p53, and bcl-2 in independently predicting a pathological complete response after preoperative chemotherapy in patients with operable carcinoma of the breast. *Breast Cancer Res* 2008; **10**: R30. doi: https://doi.org/10.1186/bcr1989
10. Esserman L, Kaplan E, Partridge S, Tripathy D, Rugo H, Park J, et al. MRI phenotype is associated with response to doxorubicin and cyclophosphamide neoadjuvant chemotherapy in stage III breast cancer. *Ann Surg Oncol* 2001; **8**: 549–59. doi: https://doi.org/10.1007/s10434-001-0549-8
11. Nishimura R, Osako T, Okumura Y, Hayashi M, Arima N. Clinical significance of Ki-67 in neoadjuvant chemotherapy for primary breast cancer as a predictor for chemosensitivity and for prognosis. *Breast Cancer* 2010; **17**: 269–75. doi: https://doi.org/10.1007/s12282-009-0161-5
12. Fangberget A, Nilsen LB, Hole KH, Holmen MM, Engebraaten O, Naume B, et al. Neoadjuvant chemotherapy in breast cancer-response evaluation and prediction of response to treatment using dynamic contrast-enhanced and diffusion-weighted MR imaging. *Eur Radiol* 2011; **21**: 1188–99. doi: https://doi.org/10.1007/s00330-010-2020-3
13. Kuerer HM, Newman LA, Smith TL, Ames FC, Hunt KK, Dhingra K, et al. Clinical course of breast Cancer patients with complete pathologic primary tumor and axillary lymph node response to doxorubicin-based neoadjuvant chemotherapy. *J Clin Oncol* 1999; **17**: 460–9. doi: https://doi.org/10.1200/JCO.1999.17.2.460
14. Gralow JR, Burstein HJ, Wood W, Hortobagyi GN, Gianni L, von Minckwitz G, et al. Preoperative therapy in invasive breast cancer: pathologic assessment and systemic therapy issues in operable disease. *J Clin Oncol* 2008; **26**: 814–9. doi: https://doi.org/10.1200/JCO.2007.15.3510
15. Ahmed A, Gibbs P, Pickles M, Turnbull L. Texture analysis in assessment and prediction of chemotherapy response in breast cancer. *J Magn Reson Imaging* 2013; **38**: 89–101. doi: https://doi.org/10.1002/jmri.23971
16. Sardanelli F, Boetes C, Borisch B, Decker T, Federico M, Gilbert FJ, et al. Magnetic resonance imaging of the breast: recommendations from the EUSOMA working group. *Eur J Cancer* 2010; **46**: 1296–316. doi: https://doi.org/10.1016/j.ejca.2010.02.015
17. Martincich L, Montemurro F, De Rosa G, Marra V, Ponzone R, Cirillo S, et al. Monitoring response to primary chemotherapy in breast cancer using dynamic contrast-enhanced magnetic resonance imaging. *Breast Cancer Res Treat* 2004; **83**: 67–76. doi: https://doi.org/10.1023/B:BREA.0000010700.11092.f4
18. Tardivon AA, Ollivier L, El Khoury C, Thibault F. Monitoring therapeutic efficacy in breast carcinomas. *Eur Radiol* 2006; **16**: 2549–58. doi: https://doi.org/10.1007/s00330-006-0317-z
19. Bufi E, Belli P, Costantini M, Cipriani A, Di Matteo M, Bonatesta A, et al. Role of the apparent diffusion coefficient in the prediction of response to neoadjuvant chemotherapy in patients with locally advanced breast cancer. *Clin Breast Cancer* 2015; **15**: 370–80. doi: https://doi.org/10.1016/j.clbc.2015.02.002
20. Chen W, Giger ML, Li H, Bick U, Newstead GM. Volumetric texture analysis of breast lesions on contrast-enhanced magnetic resonance images. *Magn Reson Med* 2007; **58**: 562–71. doi: https://doi.org/10.1002/mrm.21347
21. Chan HP, Wei D, Helvie MA, Sahiner B, Adler DD, Goodsitt MM, et al. Computer-aided classification of mammographic masses and normal tissue: linear discriminant analysis in texture feature space. *Phys Med Biol* 1995; **40**: 857–76. doi: https://doi.org/10.1088/0031-9155/40/5/010
22. Li H, Giger ML, Olopade OI, Margolis A, Lan L, Chinander MR. Computerized texture analysis of mammographic parenchymal patterns of digitized mammograms. *Acad Radiol* 2005; **12**: 863–73. doi: https://doi.org/10.1016/j.acra.2005.03.069
23. Gibbs P, Turnbull LW. Textural analysis of contrast-enhanced MR images of the breast. *Magn Reson Med* 2003; **50**: 92–8. doi: https://doi.org/10.1002/mrm.10496
24. Antel SB, Collins DL, Bernasconi N, Andermann F, Shinghal R, Kearney RE, et al. Automated detection of focal cortical dysplasia lesions using computational models of their MRI characteristics and texture analysis. *Neuroimage* 2003; **19**: 1748–59. doi: https://doi.org/10.1016/S1053-8119(03)00226-X
25. Vignati A, Mazzetti S, Giannini V, Russo F, Bollito E, Porpiglia F, et al. Texture features on T2-weighted magnetic resonance imaging: new potential biomarkers for prostate cancer aggressiveness. *Phys Med Biol* 2015; **60**: 2685–701. doi: https://doi.org/10.1088/0031-9155/60/7/2685
26. Teruel JR, Heldahl MG, Goa PE, Pickles M, Lundgren S, Bathen TF, et al. Dynamic contrast-enhanced MRI texture analysis for pretreatment prediction of clinical and pathological response to neoadjuvant chemotherapy in patients with locally advanced breast cancer. *NMR Biomed* 2014; **27**: 887–96. doi: https://doi.org/10.1002/nbm.3132
27. Aghaei F, Tan M, Hollingsworth AB, Qian W, Liu H, Zheng B. Computer-aided breast MR image feature analysis for prediction of tumor response to chemotherapy. *Med Phys* 2015; **42**: 6520–8. doi: https://doi.org/10.1118/1.4933198
28. Michoux N, Van den Broeck S, Lacoste L, Fellah L, Galant C, Berlière M, et al. Texture analysis on MR images helps predicting non-response to NAC in breast cancer. *BMC Cancer* 2015; **15**: 574. doi: https://doi.org/10.1186/s12885-015-1563-8
29. Golden DI, Lipson JA, Telli ML, Ford JM, Rubin DL. Qualitative and quantitative image-based biomarkers of therapeutic response in triple-negative breast cancer. *AMIA Jt Summits Transl Sci Proc* 2013; **2013**: 62.
30. Martincich L, Deantoni V, Bertotto I, Redana S, Kubatzki F, Sarotto I, et al. Correlations between diffusion-weighted imaging and breast cancer biomarkers. *Eur Radiol* 2012; **22**: 1519–28. doi: https://doi.org/10.1007/s00330-012-2403-8
31. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, et al.American Society of Clinical Oncology/College of American Pathologists guideline recommendations for Human Epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* 2007; **25**: 118–45. doi: https://doi.org/10.1200/JCO.2006.09.2775
32. Smith IC, Heys SD, Hutcheon AW, Miller ID, Payne S, Gilbert FJ, et al. Neoadjuvant chemotherapy in breast Cancer: significantly enhanced response with docetaxel. *J Clin Oncol* 2002; **20**: 1456–66. doi: https://doi.org/10.1200/JCO.2002.20.6.1456
33. Vignati A, Giannini V, De Luca M, Morra L, Persano D, Carbonaro LA, et al. Performance of a fully automatic lesion detection system for breast DCE-MRI. *J Magn Reson Imaging* 2011; **34**: 1341–51. doi: https://doi.org/10.1002/jmri.22680

34. Giannini V, Vignati A, Morra L, Persano D, Brizzi D, Carbonaro L, et al. A fully automatic algorithm for segmentation of the breasts in DCE-MR images. *Conf Proc IEEE Eng Med Biol Soc* 2010; **2010**: 3146–9. doi: https://doi.org/10.1109/IEMBS.2010.5627191
35. Haralick RM, Shanmugam K, Dinstein Its'Hak. Textural features for image classification. *IEEE Trans Syst Man Cybern* 1973; **SMC-3**: 610–21. doi: https://doi.org/10.1109/TSMC.1973.4309314
36. Conners RW, Harlow CA. A theoretical comparison of texture algorithms. *IEEE Trans Pattern Anal Mach Intell* 1980; **2**: 204–22. doi: https://doi.org/10.1109/TPAMI.1980.4767008
37. Clausi DA. An analysis of co-occurrence texture statistics as a function of grey level quantization. *Can J Remote Sens* 2002; **28**: 45–62. doi: https://doi.org/10.5589/m02-004
38. Nystuen JA, Garcia FW. Sea ice classification using SAR backscatter statistics. *IEEE Trans Geosci Remote Sens* 1992; **30**: 502–9. doi: https://doi.org/10.1109/36.142928
39. Unser M. Sum and difference histograms for texture classification. *IEEE Trans Pattern Anal Mach Intell* 1986; **8**: 118–25. doi: https://doi.org/10.1109/TPAMI.1986.4767760
40. Barber DG, LeDrew EF. SAR sea ice discrimination using texture statistics: a multivariate approach. *Photogramm Eng Remote Sensing* 1991; **57**: 385–95.
41. Shokr ME. Evaluation of second-order texture parameters for sea ice classification from radar images. *J Geophys Res* 1991; **96**: 10625–40. doi: https://doi.org/10.1029/91JC00693
42. Galloway MM. Texture analysis using gray level run lengths. *Comp Graph Image Process* 1975; **4**: 172–9. doi: https://doi.org/10.1016/S0146-664X(75)80008-6
43. Dasarathy BV, Holder EB. Image characterizations based on joint gray level—run length distributions. *Pattern Recognit Lett* 1991; **12**: 497–502. doi: https://doi.org/10.1016/0167-8655(91)80014-2
44. Johnson HJ, McCormick M, Ibanez L. The ITK software guide. 3rd ed. Kitware, Inc; 2013. http://www.itk.org/ItkSoftwareGuide.pdf
45. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950; **3**: 32–5. doi: https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3
46. Duda RO, Hart PE, Stork DG. *Pattern classification*; 2000.
47. Niaf E, Rouvière O, Mège-Lechevallier F, Bratan F, Lartizien C. Computer-aided diagnosis of prostate cancer in the peripheral zone using multiparametric MRI. *Phys Med Biol* 2012; **57**: 3833–51. doi: https://doi.org/10.1088/0031-9155/57/12/3833
48. Baumann K. Cross-validation as the objective function for variable-selection techniques. *TrAC Trends in Analytical Chemistry* 2003; **22**: 395–406. doi: https://doi.org/10.1016/S0165-9936(03)00607-1
49. Kuo WH, Chen CN, Hsieh FJ, Shyu MK, Chang LY, Lee PH, et al. Vascularity change and tumor response to neoadjuvant chemotherapy for advanced breast cancer. *Ultrasound Med Biol* 2008; **34**: 857–66. doi: https://doi.org/10.1016/j.ultrasmedbio.2007.11.011
50. Raja JV, Khan M, Ramachandra VK, Al-Kadi O. Texture analysis of CT images in the characterization of oral cancers involving buccal mucosa. *Dentomaxillofac Radiol* 2012; **41**: 475–80. doi: https://doi.org/10.1259/dmfr/83345935
51. Coradini D, Pellizzaro C, Veneroni S, Ventura L, Daidone MG. Infiltrating ductal and lobular breast carcinomas are characterised by different interrelationships among markers related to angiogenesis and hormone dependence. *Br J Cancer* 2002; **87**: 1105–11. doi: https://doi.org/10.1038/sj.bjc.6600556
52. Liney GP, Gibbs P, Hayes C, Leach MO, Turnbull LW. Dynamic contrast-enhanced MRI in the differentiation of breast tumors: user-defined versus semi-automated region-of-interest analysis. *J Magn Reson Imaging* 1999; **10**: 945–9. doi: https://doi.org/10.1002/(SICI)1522-2586(199912)10:6<945::AID-JMRI6>3.0.CO;2-I
53. Loo CE, Straver ME, Rodenhuis S, Muller SH, Wesseling J, Vrancken Peeters MJ, et al. Magnetic resonance imaging response monitoring of breast cancer during neoadjuvant chemotherapy: relevance of breast cancer subtype. *J Clin Oncol* 2011; **29**: 660–6. doi: https://doi.org/10.1200/JCO.2010.31.1258
54. Price ER, Wong J, Mukhtar R, Hylton N, Esserman LJ. How to use magnetic resonance imaging following neoadjuvant chemotherapy in locally advanced breast Cancer. *World J Clin Cases* 2015; **3**: 607–13. doi: https://doi.org/10.12998/wjcc.v3.i7.607