

# The future of breast screening

A literature review of emerging technologies in breast cancer screening

Final report: 30 May 2018



Allen + Clarke has been independently certified as compliant with ISO9001:2015 Quality Management Systems



Document status:	Final			
Version and date:	Final 30 May 2018			
Author(s):	Stuart Beresford, Jacinta Cording, Anna Gribble,			
	Jack Haddow, Anna Scanlan, Esther White			
Filing Location:	W/Department of Health/BreastScreen			
	Australia/Horizon Scan			
Peer / technical	Dr Robyn Haisman-Welsh			
review:				
Verification that QA	Anna Gribble			
changes made:				
Proof read:	Anna Gribble, Esther White			
Formatting:	Jo Evans			
Final QA check and	Anna Gribble			
approved for release:				

# CONTENTS

KEY	TERMS	4
GUII	6	
KEY	FINDINGS	7
DAS	GHBOARD	18
INTF	RODUCTION	19
MET	THODOLOGY	21
BIOI	MARKERS	27
1.	BLOOD TESTS	28
2.	SALIVA TESTING	36
IMA	AGING MODALITIES	41
3.	AUTOMATED WHOLE BREAST ULTRASOUND	42
4.	CONTRAST ENHANCED MAMMOGRAPHY	50
5.	DIGITAL BREAST TOMOSYNTHESIS	58
6.	DUCTOSCOPY	67
7.	MAGNETIC RESONANCE IMAGING	71
8.	MICROWAVE IMAGING	87
9.	MOLECULAR BREAST IMAGING	94
10.	SPECTROSCOPY	101
11.	THERMOGRAPHY	107
12.	TOMOGRAPHY	114
REA	DING STRATEGIES	128
13.	COMPUTER AIDED DETECTION	129
14.	ARTIFICIAL INTELLIGENCE	146
15.	TELE-MAMMOGRAPHY	152
REFE	ERENCES	157

# **KEY TERMS**

Automated breast ultrasound
Artificial intelligence
Australian Health Ministers' Advisory Committee
Breast Imaging Reporting and Data System
BreastScreen Australia
Computer-aided detection
Computer-aided diagnosis
Cone-beam breast computer tomography
Craniocaudal (view)
Contrast detail mammography
Contrast-enhanced-cone-beam breast computer tomography
Contrast enhanced mammography
Contrast enhanced spectral mammography
Circulating tumour DNA
Circulating tumour cells
Computer tomography
Contrast enhanced mammography
Cancer detection rate
Ductal carcinoma in situ
Digital breast tomosynthesis
Digital mammography
Electrical impedance tomography
Fluorodeoxyglucose positron emission tomography
Full-field digital mammography
Hand-held ultrasound
Lobular carcinoma in situ
Mean glandular dose
Milligray
Millisievert
Microwave imaging
Molecular breast imaging

MLO	Mediolateral oblique (view)
MMU	Mobile Mammography Unit
MRI	Magnetic resonance imaging
NPV	Negative predictive value
ОСТ	Optical coherence tomography
PACS	Picture archiving and communication systems
PET	Positron emission tomography
PPV	Positive predictive value
RCT	Randomised control trial
s2DM	Synthesised two-dimensional mammogram

# **GUIDANCE ON READING THIS REPORT**

This report presents findings on two biomarkers, ten imaging modalities and three reading strategies that may have an application for the early detection of breast cancer in a populationbased screening environment. Each technology or technique is described in a chapter and, for each, we answer the following questions (where information is available):

- What is the technology or innovation?
- What stage of development or trial are the new tests at?
- What are its considered potential clinical value in five years? In 10 years?
- Does this innovation show high sensitivity and specificity?
- Is there evidence that this innovation is more acceptable to women (in general and by ethnic group) compared to digital mammography?
- What cost and safety findings have been reported?
- Does this technology reduce deaths due to breast cancer through early detection?
- Has this test been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?
- Has a national position statement been published about this innovation, and if so, what is the position? Is there consensus in position statements?

Sensitivity and specificity are only discussed for those technologies that have clinical application (i.e., not those that are only in early experimental phases).

Information responding to each question is presented in an evidence hierarchy (i.e., data from systematic reviews is presented first, followed by information from randomised controlled trials, prospective studies, retrospective studies, and grey literature.

The information about digital breast tomosynthesis (DBT) included in this horizon scan has been summarised from another literature review completed by *Allen + Clarke* on the role of DBT in screening healthy women for breast cancer. This work was completed concurrently to the horizon scan literature search and review.

# **KEY FINDINGS**

The Department of Health engaged *Allen + Clarke* to undertake a horizon scan literature review to identify new and emerging technologies likely to impact population-based breast cancer screening of asymptomatic women and the BreastScreen Australia (BSA) program. This report presents the findings from the literature review.

# Methodology

*Allen + Clarke* completed a systematic search of the Ovid Medline, CINAHL, Embase, ProQuest and Scopus databases as well as searches of health technology assessment and clinical trial databases. We used combinations of subject/index terms as appropriate to the search functionality of each database (eg, exploded term 'mammography' or exploded 'breast neoplasm' in combination with key words, or key words alone). Primary studies already incorporated into high-quality systematic or literature reviews were reviewed but not assessed unless additional relevant material not described in the review was included. The methodology underpinning the chapter on DBT is described in *Allen + Clarke*'s literature review on DBT in screening.

# Results

The technologies and innovations described in this report are:

- biomarkers (blood and saliva testing)
- innovations to existing imaging technologies that are already used in either breast cancer screening, assessment or diagnosis (such as mammographic techniques like contrast enhanced mammography and DBT, MRI or ultrasound)
- new imaging technologies (like molecular breast imaging, spectroscopy), and
- new ways of reading or interpreting images (computer-aided detection, artificial intelligence, and tele-mammography).

For two innovations reviewed – ductoscopy and thermography – there appears to be a clear acceptance that they will not have a role in future breast screening programs due to issues with relatively poor sensitivity and specificity compared to full-field digital mammography (FFDM) and/or low acceptability to women and health professionals.

The following technologies are likely to be of interest to breast screening programs: automated whole breast ultrasound, computer-aided detection/AI and DBT.

#### BIOMARKERS

#### **Blood testing**

- Five studies discussed blood testing for the early detection of breast cancer in asymptomatic women.
- The use of blood testing in breast cancer screening of asymptomatic women is still in the early stages, with much of the research focus being on identifying promising biomarkers that demonstrate sufficient sensitivity and specificity to support their use in further clinical testing. Research is most advanced concerning the use of microRNA in screening.
- There is no indication of the timeframe in which the full clinical potential of blood testing for breast cancer detection will be realised; however, preliminary results from primarily retrospective studies are promising and the technology is improving rapidly.
- Blood testing has not been incorporated into any national screening programs, nor are there any national position statements that have been released on their use in breast cancer screening.
- Current research is not sufficient to be able to identify whether blood biomarker testing is able to reduce deaths due to breast cancer through early detection.

#### Saliva testing

- Nine studies discussed saliva testing for the early detection of breast cancer in asymptomatic women.
- The use of saliva testing in breast cancer screening of asymptomatic women is still early stages, with much of the research focused on identifying promising biomarkers that are specific to breast cancer, and which can detect breast cancer with sufficient sensitivity and specificity to support their use in further clinical testing (including prospective studies and RCTs). Current research is limited to retrospective studies comparing samples from breast cancer patients with healthy controls.
- There is currently no indication of the timeframe in which the full clinical potential of saliva testing for breast cancer detection will be realised; however, preliminary results from retrospective studies are promising, with studies routinely finding sensitivity and specificity rates over 80%, or even over 90%, for a range of saliva biomarkers.
- Saliva testing has not been incorporated into any national screening programs, nor are there any national position statements that have been released on their use in breast cancer screening for asymptomatic women.
- Current research is not sufficiently advanced to identify whether saliva biomarker testing is able to reduce deaths due to breast cancer through early detection.

#### IMAGING MODALITIES

# Automated whole breast ultrasound

• Nineteen studies discussed automated breast ultrasound (ABUS) for the early detection of breast cancer in asymptomatic women.

- ABUS is already in clinical use for breast cancer screening in the United States and Canada. There is no evidence to suggest that ABUS has been implemented into a national screening program at this stage.
- The potential clinical value of ABUS in five to 10 years is not explicitly discussed in the literature, although findings have indicated its value as an adjunct screening tool for women with denser breasts.
- Many studies have found that using ABUS and FFDM leads to significantly higher cancer detection rates for women with denser breasts than using digital mammography alone.
- ABUS is likely to be highly acceptable to women as no safety issues have been identified and the examination process involves less discomfort than other modalities, including digital mammography.
- There is no evidence that ABUS reduces deaths due to breast cancer through early detection.
- ABUS has not been incorporated into any national breast screening programs, nor are there any national position statements that have been released on its use in breast cancer screening for asymptomatic women.

# Contrast enhanced mammography

- Twenty studies discussed contrast enhanced mammography for the early detection of breast cancer in asymptomatic women.
- The use of CEM in population-based breast cancer screening of asymptomatic women is seen to have some potential in clinical studies, and there are current clinical trials related to CEM, including with a focus on screening.
- There is currently no indication of the timeframe in which the full clinical potential of CEM for breast cancer screening will be realised; however, results from prospective clinical studies show that CEM reports higher percentages of sensitivity, specificity, positive predictive value and negative predictive value than conventional mammography.
- Current research is not sufficient to be able to identify whether CEM as a breast imaging tool is able to reduce deaths due to breast cancer through early detection.
- CEM has not been incorporated into any national breast screening programs, nor are there any national position statements that have been released on its use in breast cancer screening for asymptomatic women.

# Digital breast tomosynthesis

- The Department of Health commissioned *Allen + Clarke* to undertake a literature review on the use of DBT in screening. The results of this literature review are included in this report for completeness. Further detail on the methodology is contained in *Allen + Clarke*'s primary literature review document.
- There is strong evidence that both FFDM + DBT and DBT + s2DM provide superior performance for improved cancer detection and that DBT may be a more sensitive test than FFDM alone. The magnitude of improvement in CDR may be affected by reading

strategy (eg, double reading approaches result in higher detection rates compared to single reading).

- There is emerging evidence that, used as an adjunct screen to FFDM, DBT can reduce recall rates and false positives results compared to FFDM alone; however, some inconsistent results between large prospective trials are reported. These differences may reflect the already low rates of recall seen in some population-based screening programs in which the trials are embedded, or different reading or arbitration strategies. Further research investigating comparative performance of FFDM and DBT will help to unpick areas of uncertainty including the impact of double/single reading strategies and the impact of access to previous DBT images.
- While DBT improves breast cancer detection with (potentially) lower rates of recall than FFDM alone, there is insufficient evidence about the long-term mortality benefit to support the use of DBT alone as a primary screening test. Current research is not able to identify whether DBT as a breast imaging tool reduces deaths due to breast cancer through early detection.
- Few studies identified for this literature review investigated DBT alone compared to FFDM alone, although literature exploring different ways to integrate DBT into screening continues to develop quickly. Further research may also help determine which combination of approaches (FFDM + DBT, two-view DBT + s2DM, DBT<sub>MLO</sub> + FFDM, or some form of DBT alone) achieves the best balance between radiation dose, sensitivity and specificity.
- DBT has not yet been incorporated into any national screening programs; however, fully-paired trials embedded in population-based screening programs have been completed or are underway. In addition, DBT is available for private breast screening in several jurisdictions including Australia and New Zealand.

# Ductoscopy

- Two studies discussed ductoscopy for the early detection of breast cancer in asymptomatic women.
- The role of ductoscopy in breast cancer screening of asymptomatic populations is not clear as most of the research focuses on its role in the assessment and diagnosis of ductal carcinoma in-situ (DCIS). For example, recent research focuses on the use of ductoscopy with symptomatic patients that present with nipple discharge.
- There is no indication that ductoscopy will achieve clinical potential as a screening tool. There is little evidence to suggest ductoscopy is acceptable to women as a screening modality.
- Current research is not sufficient to be able to identify whether ductoscopy as a screening procedure can reduce deaths due to breast cancer through early detection.
- Ductoscopy has not been incorporated into any national breast screening programs, nor are there any national position statements that have been released on their use in breast cancer screening for asymptomatic women.

#### Magnetic resonance imaging

- Sixty-six studies discussed MRI for the early detection of breast cancer in asymptomatic women.
- The availability of MRI as a supplementary examination to mammography offers a clear clinical benefit in some situations, particularly women with a high risk for breast cancer, especially those with dense breast tissue. Compared to mammography, MRI is less specific but more sensitive to detect small tumours in subjects with high breast cancer risk. MRI is a non-invasive technique that gives extremely clear, detailed images of soft-tissue structures that other imaging techniques cannot achieve. Unlike mammography, MRI does not expose the tissue to ionising radiation and the contrasting agent used in MRI is less likely to produce an allergic reaction that may occur during the use of iodine-based substances in other imaging modalities.
- MRI has not been recommended for the general population due to high false-positive rates (which can lead to over-diagnosis with attendant cost and anxiety), high cost, time consumption, lack of adequate number of units, the need for experienced radiologists and lack of clinical utility. Some cancers, such as DCIS, are better detected by mammography than by MRI.
- MRI is expensive, and although not painful the patient must remain still during the examination, which was an issue for claustrophobic women. Data has also emerged indicating there can be accumulation of gadolinium in the brain in patients who have undergone multiple contrast-enhanced MRI studies.
- MRI has not been incorporated into any national breast screening programs. The American Cancer Society and the European Society of Breast Cancer Specialists have both released statements on the use of MRI in breast cancer screening.

# **Microwave imaging**

- Three studies discussed the use of microwave imaging for the early detection of breast cancer in asymptomatic women.
- The use of microwave imaging in breast cancer screening of asymptomatic women is still in the early stages. Much of the research focus is on identifying feasible microwave imaging systems that demonstrate sufficient sensitivity and detectability to support their use in further clinical testing. Research is most advanced in the use of ultrawideband frequency systems.
- There is currently no indication of the timeframe in which the full clinical potential of microwave imaging for breast cancer detection will be realised; however, preliminary results from clinical studies indicates that this technology has the potential for clinical use.
- Current research is not sufficient to be able to identify whether microwave imaging as a breast imaging tool is able to reduce deaths due to breast cancer through early detection.
- Microwave imaging has not been incorporated into any national breast screening programs, nor are there any national position statements that have been released on their use in breast cancer screening for asymptomatic women.

#### Molecular breast imaging

- Thirteen studies discussed molecular breast imaging for the early detection of breast cancer in asymptomatic women.
- Results from largely retrospective studies are promising regarding the effectiveness of MBI for screening purposes; however, more large-scale multi-centre prospective studies are required before conclusions can be drawn about the role of MBI in the early detection of breast cancer in asymptomatic women.
- There is currently no indication of the timeframe in which the full clinical potential of MBI for breast cancer detection will be realised. High sensitivity and moderate specificity rates have been found in existing research, with improvements in the detection of sub-centimetre lesions. Rates of cancer detection are improved when MBI is used as a supplement to screening mammography.
- Despite initial issues with radiation dosage, more recent MBI systems are demonstrating good detection ability with reduced effective radiation dosages (approximately 2.4 mSv). One study found that the cost of supplementing screening with MBI is higher per examination than for mammography alone, however combining MBI with mammography leads to a decrease in cost per cancer detected.
- There is growing evidence for the increased efficacy of MBI for the early detection of cancer for women with dense breasts (BIRADS 3 or 4) compared with screening mammography, although these results may underestimate the impact of MBI for women with more dense breasts due to categorisation issues.
- No specific information on the acceptability of MBI to women was identified; however, the average length of time for an MBI examination (40-45 minutes) is substantially longer than the length of time needed for a mammogram.
- MBI has not been incorporated into any national screening programs, nor are there any national position statements that have been released on its use in breast cancer screening for asymptomatic women.
- Current research is not sufficient to be able to identify whether MBI is able to reduce deaths due to breast cancer through early detection.

# Spectroscopy

- Two studies discussed the use of spectroscopy for the early detection of breast cancer in asymptomatic women
- The potential use of spectroscopic techniques in breast cancer screening of asymptomatic populations is not currently clear, with recent research focusing on developing and refining technology for clinical testing.
- There is currently no indication of the timeframe in which the full clinical potential of spectroscopic techniques for breast cancer detection of asymptomatic people will be realised; however, results from clinical studies indicates that optical mammography has the potential for clinical use in high risk groups. Research and development into spectroscopy for breast cancer detection is emerging and advancing.
- Current research is not sufficient to be able to identify whether spectroscopy as a screening procedure can reduce deaths due to breast cancer through early detection.

• Spectroscopic techniques have not been incorporated into any national breast screening programs, nor are there any national position statements that have been released on their use in breast cancer screening for asymptomatic women.

# Thermography

- Nineteen studies discussed thermography for the early detection of breast cancer in asymptomatic women.
- There is not a large body of evidence that supports the use of thermography as a tool for breast screening in asymptomatic women. Most studies use small sample sizes and the results vary significantly. Much of the research is focused on different methods and technologies for obtaining and interpreting thermographs rather than the actual use of thermography for a screening purpose.
- Current research is not sufficient to be able to identify whether thermography is able to reduce deaths due to breast cancer through early detection.
- Thermography has not been incorporated into any national screening programs. Importantly, national position statements discouraging thermography as a breast screening tool have been issued in the United States and Australia; however, these are often based on quite old literature and do not reflect the developments in the evidence that have occurred in more recent years.

# Tomography

• Eight studies discussed tomography for the early detection of breast cancer in asymptomatic women, across five different types of tomography.

# **Computer tomography (CT)**

- The development of breast CT systems for use in the early detection of breast cancer is still in its infancy, with only retrospective studies identified in the literature reviews.
- There is currently no indication of the timeframe in which the full clinical potential of breast CT for the early detection of breast cancer in asymptomatic women will be realised. Although breast CT scanning can visualise breast lesions and masses as well as or better than mammography, breast CT scanning performs worse in visualising microcalcifications. This may limit the incorporation of breast CT as a primary breast cancer screening modality in screening for asymptomatic women.
- Despite initial issues with radiation dose, more recent breast CT scanning systems are improving, achieving radiation dose levels comparable to conventional mammography. Additionally, cost is unlikely to be a barrier in the use of breast CT for the early detection of breast cancer.
- Although no primary studies were identified, one review of the breast CT literature stated that breast CT could play an important role in breast cancer screening for women with dense breasts.
- Findings suggest that breast CT is significantly more comfortable for women than mammography, primarily due to the lack of breast compression.
- Current research is not sufficient to be able to identify whether breast CT scanning is able to reduce deaths due to breast cancer through early detection.

• Breast CT has not been incorporated into any national screening programs, nor are there any national position statements that have been released on its use in breast cancer screening for asymptomatic women.

# Cone-beam breast computer tomography (CBBCT)

- The research assessing the use of CBBCT is still in the early stages, with research not yet advancing to clinical tests using asymptomatic samples.
- There is currently no indication of the timeframe in which the full clinical potential of CBBCT for breast cancer detection will be realised. That said, current results using symptomatic samples are promising in terms of its ability to improve on FFDM.
- Current research suggests that there is no statistically significant difference in radiation dose from CBBCT scans compared with FFDM. Findings related to the cost of CBBCT were not reported in the identified literature.
- No studies were identified that assessed the performance of CBBCT scans in asymptomatic women with dense breasts, however results from symptomatic samples suggest that contrast-enhanced CBBCT may improve scan sensitivity for women with dense breasts compared with FFDM.
- Reported patient comfort is higher for CBBCT than for mammography.
- Current research is not sufficient to be able to identify whether CBBCT scanning is able to reduce deaths due to breast cancer through early detection.
- CBBCT has not been incorporated into any national screening programs, nor are there any national position statements that have been released on its use in breast cancer screening for asymptomatic women.

#### **Positron emission tomography (PET)**

- Research into the use of PET for early breast cancer detection in asymptomatic has progressed to prospective clinical studies, however further research is still required on the safety and effectiveness of PET as a screening test.
- There are no clear timeframes regarding the clinical potential of PET as a screening test being reached; however, current research suggests that there may still be limitations with the ability of PET to detect small tumours.
- Relatively high levels of radiation continue to be an issue for PET systems, and cost is also identified as being potentially prohibitive to its incorporation into routine screening for asymptomatic women.
- No information was identified on the sensitivity and specificity of PET for asymptomatic women with dense breasts or women who have had breast surgery/augmentation compared with FFDM.
- No information was identified on the acceptability of PET for women compared with FFDM.
- Results from a nationwide fluorodeoxyglucose PET (FDG-PET) cancer screening program in Japan found that FDG-PET had a sensitivity of 84% in detecting breast cancer; this was not significantly different from rates found for mammography. Issues with radiation dosage and the cost of FDG-PET scans were noted as a barrier for incorporation into other screening programs.

- Current research is not sufficient to be able to identify whether PET imaging is able to reduce deaths due to breast cancer through early detection of cancer in asymptomatic women.
- There have been no national position statements released regarding the use of PET for the early detection of breast cancer in asymptomatic women.

# **Optical coherence tomography (OCT)**

- The development of OCT systems for use in breast cancer screening is still in preclinical stages, with current systems only able to penetrate up to 2mm in most tissue. This means that most research into the use of OCT for cancer detection or diagnosis involves the use of either animal models or samples obtained via biopsies. Additionally, most of this research focuses on the use of OCT in treatment and diagnosis, rather than in screening.
- Current research is not sufficient to be able to identify whether OCT is able to reduce deaths due to breast cancer through early detection of cancer in asymptomatic women.

# **Electrical impedance tomography (EIT)**

- Research is still ongoing into the development of EIT systems that can image the breast with sufficient quality to accurately detect breast cancer. Some of these systems require breast compression for imaging, whereas others use a setup like breast CT, whereby subjects lay prone on an examination table with the breast hanging suspended through an opening in the table top. Because of the early stages of research in this area, there were no studies identified that used asymptomatic samples.
- Current research is not sufficient to be able to identify whether EIT is able to reduce deaths due to breast cancer through early detection of cancer in asymptomatic women.

#### **READING STRATEGIES**

#### **Computer-aided detection (CADe)**

- Fifty-seven articles discussed the use of CADe as a reading strategy for the early detection of breast cancer in asymptomatic women
- CADe, applied as a complementary technology to mammography interpretation, prompts the reader to lesions on the mammogram. The reader then needs to decide whether to recall CADe-prompted findings.
- Non-randomised studies have shown that CADe improves the sensitivity of a single reader, with an incremental cancer detection rate ranging between 1 and 19%. A limitation of CADe is that it also increases screening recall rate through a decrease in specificity. Studies have found that double-reading produces the same results as that for single reading with CADe, without the increased rate of recall. In terms of detecting lesion characteristics, good results have been observed for the detection of breast cancers presenting as microcalcifications. However, the rate of detection for circumscribed or spiculated masses and architectural distortions is not so positive.
- To expand the clinical value of CADe in the detection of breast cancer, several studies have focused on improving technical stages of the CADe process. Promising results have been observed, particularly regarding the application of deep learning systems, but further work needs to be done. While CADe could substitute the human second

reader in the future, depending on the first reader's experience, without improvements in its effectiveness (eg, a decrease in recall rate), CADe is unlikely to be a cost-effective alternative to double reading for mammography screening.

- Current research is not sufficient to be able to identify whether CADe is able to reduce deaths due to breast cancer through early detection.
- CADe has not been incorporated into any national breast screening programs, nor are there any national position statements that have been released on its use in breast cancer screening for asymptomatic women.

#### Artificial intelligence

- Fourteen studies discussed AI as a reading strategy for the early detection of breast cancer in asymptomatic women.
- The use of AI in breast cancer screening of asymptomatic women is still early stages but is closely linked to developments in CAD. While promising, machine learning was in its infancy with respect to demonstrating its utility in cancer screening.
- There is currently no indication of the timeframe in which the full clinical potential of AI for breast cancer detection will be realised; however, preliminary results for AI are promising.
- Current research is not sufficient to be able to identify whether AI testing is able to reduce deaths due to breast cancer through early detection.
- AI has not been incorporated into any national screening programs, nor are there any national position statements that have been released on their use in breast cancer screening for asymptomatic women.

#### **Tele-mammography**

- Four studies discussed tele-mammography's role in interpreting mammogram images to support the early detection of breast cancer in asymptomatic women.
- The use of tele-mammography appears to be at least moderately widespread, and seemingly has a range of benefits, including the ability to provide services to remote areas, and better utilisation of radiologists' time.
- Studies have found that there is no significant difference in screening outcomes between traditional mammography and tele-mammography technologies, with some even finding that consumer-grade image capture technologies (such as digital cameras) and visualisation devices (such as LCD screens) are suitable for performing tele-mammography.
- It is not clear whether tele-mammography has been incorporated into any national screening programs, or whether its use has been determined at sub-national level.

# Summary

From the literature reviewed, it appears that some imaging modalities reviewed for this horizon scan – namely ductoscopy and thermography – have no promising application in the early detection of breast cancer. In the case of ductoscopy, it may continue to have a role in the assessment and diagnosis of DCIS. Thermography appears to play a limited reliable role in breast cancer screening, assessment or diagnosis.

Other technologies are emerging. These may have a future in population screening, but it is too soon to determine the likelihood of this or the timeframe over which the technology will move from experimental testing to having a clear application to breast cancer screening. These technologies include blood testing, saliva testing, microwave imaging, molecular breast imaging, spectroscopy and tomography. In the case of biomarkers, further research focused on developing these innovative testing modalities to a suitable standard is needed before human or clinical trials can be undertaken. Further research to determine the role of the new imaging modalities in breast cancer detection and diagnosis is also required.

Other innovations covered in this horizon scan relate to improvements in imaging modalities already used in the detection, assessment and diagnosis of breast cancer. These imaging modalities are ABUS, CEM, and MRI. The technology underpinning these modalities is not so new and technologies have demonstrated clinical benefit. Therefore, this horizon scan has focused on innovations to improve their application to the early detection of breast cancer in asymptomatic women. For example, we have reviewed MRI-related innovations focused on improving the early detection of breast cancers for women who have a higher than average lifetime risk of developing breast cancer. Other breast cancer detection technologies (namely DBT) are already in being used in clinical practice for either the detection or in the assessment and diagnosis of breast cancer and have been included in this horizon scan for completeness.

Other innovations focus on the way in which screening modalities may be used to support a more personalised approach to breast cancer screening, reflecting the fact that women have different risk factors and that these can change over their lifetime. A move toward using different screening techniques at different times depending on the lifetime risk and individual circumstances could be informed by the innovations developing within the modalities discussed in this horizon scan. This horizon scan, however, did not consider any implementation issues associated with developing further innovations for in-practice techniques, nor did it consider the role of personalised screening programs based on access to a range of screening modalities.

The final group of innovations or technologies reviewed were reading strategies such as computer-aided detection, AI and tele-mammography. These strategies are concurrently both in development (i.e., creating and testing of new algorithms to support enhanced detection and developing deep learning/machine reading techniques) with some being used in clinical practice already. These strategies therefore sit between the other groupings, in terms of the type and breadth of research assessing their application to the early detection of breast cancer for asymptomatic women.

# DASHBOARD

Saliva tests

#### A state of play assessment reports the following ... Biomarkers Early R+D Early human In RCT with In clinical Used for No proposed (non-human clinical trials humans practice with diagnosis screening or prea screening and application clinical) setting assessment Blood tests Х х Х

Imaging modalities	Early R+D (non-human or pre- clinical)	Early human clinical trials	In RCT with humans	In clinical practice with a screening setting	Used for diagnosis and assessment	No proposed screening application
Automated whole breast ultrasound			х	х	х	
Contrast enhanced mammography				х	х	
DBT			x	x	х	
Ductoscopy						х
Magnetic resonance imaging			x	x	x	
Microwave imaging	х					
Molecular breast imaging		х		х		
Spectroscopy	х					
Thermography						x
Tomography	x	х				

Reading innovations	Early R+D (non-human or pre- clinical)	Early human clinical trials	In RCT with humans	In clinical practice with a screening setting	Used for diagnosis and assessment	No proposed screening application
Computer-aided detection	х			x	х	
Artificial intelligence	х				х	
Tele-mammography	х			x	x	

# **INTRODUCTION**

Breast cancer is one of the leading causes of death among women in Australia. Currently, there is no effective way to prevent the occurrence of breast cancer. Early detection is a first crucial step towards breast cancer diagnosis and treatment. Currently, evidence supports bilateral full field digital mammography (FFDM) as the standard test for the early breast cancer detection in asymptomatic women due to its low cost and accessibility. While screening with FFDM presents some limitations (for instance, it is less effective at detecting cancer in women aged under 40 years old or in women with more dense breasts and is less sensitive to small tumours (less than 1 mm, about 100,000 cells), it remains the most effective method for early detection of breast cancer as it provides high sensitivity for women with more fatty breasts and excellent performance on microcalcification detection.

New and emerging technologies to improve the early detection of breast cancer are constantly being developed, trialled and evaluated. As these new technologies develop and the evidencebase for these becomes sufficiently robust, effective screening tests that are comparable or superior to FFDM may develop.

New technologies that could shape the future of early detection of breast cancer could include:

- improvements to breast imaging technologies: ABUS, contrast enhanced mammography, DBT, ductoscopy, electrical impedance tomography and positron emission tomography (PET)
- imaging modalities that extend imaging of the breast: cone-beam breast computer tomography (CT) and molecular breast imaging
- new breast imaging modalities: radar-based microwave imaging, optical breast imaging and spectroscopy
- computer-aided diagnosis, remote/tele-radiology, machine reading or other AI applications
- biological tests to detect abnormal proteins produced by cancers
- risk assessment tools and tailored risk-based screening, and
- genomic testing or cancer gene fingerprinting.

R+D progression is reported in a variety of ways including peer-reviewed articles in scientific journals, health technology assessments, government papers, regulatory announcements, advice from professional bodies, product launches, media articles, websites and advertisements for breast screening products or services. Freely available information of variable quality can raise consumer expectations and concerns in a highly emotive subject such as early cancer detection. In Australia, for any new technological development to be considered for population-based breast screening or breast cancer imaging, it must meet the criteria set down in the Population Based Screening Framework (2016), endorsed by the Community Care and Population Principal Committee of Australian Health Ministers' Advisory Committee (AHMAC).

To keep up with new developments in health-related technologies, the Australian Government established the Australian Horizon Scanning Program, a collaborative Commonwealth and State initiative guided by the Health Policy Advisory Committee on Technology (HealthPACT). This program provides notice of significant new and emerging technologies to health departments in Australia and New Zealand and enables the exchange of information and evaluation of the potential impact of new technologies on their respective health systems.

One report relevant to newly emerging technologies for early breast cancer detection, and based on the Australian Horizon Scanning Program, has been published previously. In 2009, the Australia and New Zealand Horizon Scanning Network published the report *Emerging Technology Bulletin* 'New and emerging technologies for breast cancer detection'. This study reviewed international evidence on seven technologies for breast cancer detection: CT, PET, ultrasonography, thermography, electrical impedance tomography, molecular breast imaging (also known as scintimammography) and ductoscopy. In 2010, HealthPACT also completed a report on the development of molecular breast imaging (scintimammography).

# Scope

Building on the 2009 horizon scan (Australia and New Zealand Horizon Scanning Network, 2009), the 2018 <u>qualitative</u> horizon scan focuses on population-level screening. It will include:

1. Probable or possible changes to imaging technologies including technologies that are beginning to be used in clinical practice and those that are still in R+D:

Included technologies are: contrast enhanced mammography (CEM), ductoscopy, ABUS, magnetic resonance imaging, thermography, molecular breast imaging, cone-beam breast CT, position emission tomography, electrical impedance tomography, radar-based microwave imaging, and optical breast imaging

2. Probable or possible changes to reading/interpretation technologies

Included technologies are computer-aided diagnosis, remote/tele-radiology, machine reading or other AI applications

3. "Wildcard" changes such as biological screening tests which would have a significant impact on preferred screening modality (including blood, saliva or hair tests and genomic testing as a screening tool in its own right).

As the evidence allows, we will identify technologies using the following framework:

- May be of clinical value within a five-year timeframe
- May be of clinical value within a five- to ten-year timeframe, and
- May be of clinical value in more than 10 years.

DBT is the subject of a separate literature review (refer to *Allen + Clarke*'s literature review on the role of DBT in breast cancer screening) but we have included summarised findings from this literature review in the horizon scan for completeness. Risk assessment tools and genomic testing outside of a screening context are excluded.

# METHODOLOGY

#### Summary

This horizon scan provides a narrative, qualitative overview of research about promising imaging modalities, AI and key biomarkers that could potentially have an application in a population-based screening program for the early detection of breast cancer in asymptomatic women. No primary research or pooled analysis was undertaken.

The following databases were searched between 8 and 12 January 2018: EMBASE, Ovid Medline, CINAHL, ProQuest and Scopus. The range of key health technology assessment and clinical trials websites were also searched. Other technology or specific websites may also have been covered when searching for grey literature on aspects of the technology or innovation.

All returned citations and abstracts were assessed for relevance to the research questions and inclusion criteria. The same criteria were used to review the full-text and bibliographies of all articles proposed for inclusion.

No assessment of the overall quality of the evidence was undertaken.

Information about the methodology underpinning the chapter on DBT can be found in Allen + Clarke's literature review on the role of DBT in screening.

# **Objectives**

The Department of Health engaged *Allen + Clarke* to undertake a horizon scan (comprising a literature review and supplemented with key informant interviews) to identify new and emerging technologies likely to impact population-based breast cancer screening of asymptomatic women and the BSA program. This horizon scan:

"maintains a predictive or speculative focus, often based on low level evidence, and is aimed at informing policy and decision makers. It is not a definitive assessment of the safety, effectiveness, ethical considerations and cost effectiveness of a technology. In the context of a rapidly evolving technology, an horizon scanning report is a 'state of play' assessment that presents a trade-off between the value of early, uncertain information, versus the value of certain, but late information, that may be of limited value to policy and decision makers".

The report provides the literature review component of *Allen + Clarke*'s work.

# **Research questions**

Main Question	Supplementary questions
What innovations to <u>existing imaging</u> <u>technologies</u> OR <u>reading/interpretation</u> <u>technologies</u> can improve early detection of breast cancer in asymptomatic women?	<ul> <li>What is the innovation?</li> <li>What stage of development or trial is this innovation at?</li> <li>What is its considered potential clinical value in five years? In 10 years?</li> <li>What cost and safety findings have been reported?</li> <li>Does this technology reduce deaths due to breast cancer through early detection?</li> <li>FOR IMAGING TECHNOLOGIES: Does this innovation show high sensitivity and specificity for women with dense breasts and women who have had breast surgery/augmentation compared to digital mammography?</li> <li>FOR INTERPRETATION/READING TECHNOLOGIES: Does this innovation show higher sensitivity and specificity for women with dense breasts and women who have had breast surgery/augmentation compared to digital mammography?</li> <li>FOR INTERPRETATION/READING TECHNOLOGIES: Does this innovation show higher sensitivity and specificity for women with dense breasts and women who have had breast surgery/augmentation compared to single human view?</li> <li>Is there evidence that this innovation is more acceptable to women (in general and by ethnic group) compared to digital mammography?</li> <li>Has this innovation been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?</li> <li>Has a national position statement been published about this innovation, and if so, what is the position? Is there consensus in position statements?</li> </ul>
Are there new or novel <u>imaging</u> or <u>reading/interpretation</u> <u>technologies</u> that result in early detection of breast cancer in asymptomatic women?	<ul> <li>What is the new technology?</li> <li>What stage of development or trial is this new technology at?</li> <li>What is its considered potential clinical value in five years? In 10 years?</li> <li>What cost and safety findings have been reported?</li> <li>Has this technology been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?</li> <li>Has a national position statement been published about this technology, and if so, what is the position? Is there consensus in position statements?</li> </ul>
Are there new or novel biological tests that result in early detection of breast cancer in asymptomatic women?	<ul> <li>What is the new test?</li> <li>Does this technology reduce deaths due to breast cancer through early detection?</li> <li>What stage of development or trial is this new test at?</li> <li>What is its considered potential clinical value in five years? In 10 years?</li> <li>What cost and safety findings have been reported?</li> <li>Has this test been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?</li> <li>Has a national position statement been published about this test, and if so, what is the position? Is there consensus in position statements?</li> </ul>

# Literature search

The following databases were searched between 8 and 12 January 2018:

- AustHealth
- Australian Medical Index
- Australian Public Affairs Information Service (APAIS) Health
- Cinahl
- Cochrane Library including, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, the Cochrane Central Register of Controlled Trials (CENTRAL), the Health Technology Assessment Database, the NHS Economic Evaluation Database
- Current Contents
- Embase
- OVID Medline (and Pre-Medline)
- ProceedingsFirst
- ProQuest
- PsycInfo
- Web of Science Science Citation Index Expanded.

Supplemental searches of the following websites searched were conducted in January and February 2018:

- <u>Australian Clinical Trials Registry www.actr.org.au/default.aspx</u>
- Breast Cancer Research Institute of Australia: www.breastcancertrials.org.au/
- Breast Cancer Network Australia <u>www.bcna.org.au</u>
- <u>Clinical Trials Registry www.clinicaltrials.gov</u>
- <u>Current Controlled Trials metaRegister www.controlled-trials.com/</u>
- <u>Health Technology Assessment International www.htai.org</u>
- International Network for Agencies for Health Technology Assessment
   www.inahta.org/
- <u>Medicines and Healthcare Products Regulatory Agency (UK)</u> <u>www.mhra.gov.uk/index.htm</u>
- <u>National Library of Medicine Health Services/Technology Assessment Text</u> <u>www.ncbi.nlm.nih.gov/books</u>
- <u>National Library of Medicine Locator Plus database www.locatorplus.gov</u>
- <u>National Institute for Health Research UK HTA programme</u> www.journalslibrary.nihr.ac.uk/programmes/hta/
- <u>New York Academy of Medicine Grey Literature Report</u> <u>www.nyam.org/library/grey.shtml</u>

- <u>Trip database www.tripdatabase.com</u>
- <u>U.K. National Research Register www.portal.nihr.ac.uk/Pages/NRRArchive.aspx</u>
- <u>US Food and Drug Administration, Center for Devices and Radiological Health.</u> <u>www.fda.gov/cdrh/databases.html</u>
- Websites of Specialty Organisations (dependent on technology topic area), and
- <u>National Breast Cancer Foundation</u>.

To complete a systematic search, we used combinations of subject/index terms where appropriate (eg, exploded term 'mammography' or exploded 'breast neoplasm') in combination with key words, or key words alone depending on the search functionality of each database or website (eg, main searches included 'tomography' PLUS 'breast cancer' PLUS 'screen\*' in the title or abstract).

The following limits were applied on all searches:

- a date criterion (1 January 2010 31 December 2017 or 2010 onwards)
- English language, and
- study type restrictions (where available and appropriate, we restricted returns from research databases to peer-reviewed systematic reviews, literature reviews, RCT, observational studies and clinical trials).

As 'human' was not used as a limiter, some animal studies were returned for the search terms. These were excluded once identified. Duplicate citations and a small number of false hits/inaccurate returns were removed before all initial returned citations and abstracts were reviewed for relevance to the main research questions.

Material was excluded if it:

- did not relate to a test that could be used in a population-based screening program for the early detection of breast cancer in asymptomatic women
- compared potential screening performance to tests other than FFDM
- focused on a study population other than asymptomatic women
- related to surveillance, diagnosis or treatment.

To determine if this first search retrieved the correct range of available research, a validation process was completed using recent systematic or literature reviews relevant to the technology (where these were available).

From this first sweep, full texts for all proposed inclusions were retrieved and reviewed for relevance to the research questions, inclusion criteria and documented PICOT criteria. A critical appraisal of study design (to determine overall quality) was completed and the bibliography of each included article was reviewed to identify other relevant research that may be of interest.

The citation review process is described at the beginning of each chapter.

# **Interpretation notes**

We describe the body of research for each technology/innovation in each chapter of this report. In summary, limited literature sources and/or grey material were identified for some of the technologies and innovations reviewed as part of this horizon scan. For emerging technologies (biomarkers and some new imaging techniques), the research is in its infancy and is focused on developing tools with adequate sensitivity and specificity (as is the case for blood and saliva testing and some of the imaging techniques like molecular breast imaging and spectroscopy).

Other imaging technologies and reading strategies (including DBT, MRI or ultrasound) are already be used in screening environments as either a primary screening tool (DBT) or as a supplemental screening examination following FFDM or instead of FFDM (MRI and ultrasound). The body of evidence discussing these techniques in a screening setting is deep and we relied less heavily on grey literature. Also, where a technology is in use in a screening setting (and usually it is used for groups of women who have high lifetime risk of cancer), we have focused our approach to reporting the evidence base to the innovations that are coming (rather than its current use as an adjunct or supplemental screen). For example, we provide contextual material about the use of ultrasound in the early detection of breast cancer in asymptomatic women screening but have focused our review on the innovations arising from the development and testing of ABUS. Further information about supplemental screening tests for women with more dense breasts is discussed in *Allen + Clarke*'s literature review on breast density and mammography (*Allen + Clarke*, unpublished).

Some other technologies like tomography have a clinical application in the diagnosis of breast cancer and treatment management. Again, deep bodies of evidence underpin the clinical use of the technology; however, literature may have much less focus on the technology's impact on the early detection of breast cancer in asymptomatic women. In these instances, we have focused only on the literature discussing the technology's potential application to a screening environment.

This horizon scan does not separately review the literature regarding the use of genomic screening as an early cancer detection modality. This is because current genomic screening literature is primarily focused on the identification of genetic risk predictors for the development of breast cancer, rather than the actual detection of malignancy in asymptomatic women. The use of genomic screening is therefore in identifying high-risk women who could potentially benefit from individualised breast cancer screening using technologies that detect the presence of breast cancer, such as FFDM or other modalities that are reviewed in this horizon scan. This falls outside of the scope of the current scan, which was focused on early breast cancer detection for asymptomatic women in the general population (i.e., healthy women at average risk of developing breast cancer).

Although genomic screening falls outside the scope of the current horizon scan, it is not a reflection on the potential importance of genomic screening for informing future decisions around population-level screening. The search for genetic risk predictors – often conducted through genome-wide association studies involving population-level data – will help to identify individuals who are at a higher risk of developing breast cancer. This in turn could potentially inform the development of individualised screening (and treatment) approaches for certain sub-populations, thereby improving the early detection of breast cancer among those at highest risk of developing malignancies. Genomics is therefore an area that should be considered in future decisions regarding individualised breast cancer screening programs for asymptomatic women.

The following biomarkers, imaging modalities and reading strategies are included in this horizon scan:

#### **Biomarkers**

- Blood tests
- Saliva tests

#### **Imaging modalities**

- Automated whole breast ultrasound
- Contrast enhanced mammography
- Digital breast tomosynthesis
- Ductoscopy
- Magnetic resonance imaging
- Microwave imaging
- Molecular breast imaging
- Spectroscopy
- Thermography
- Tomography

#### **Reading strategies**

- Artificial intelligence
- Computer-aided detection
- Tele-mammography

# BIOMARKERS

This section covers:

- Blood tests, and
- Saliva testing.

# 1. BLOOD TESTS

# Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

The use of blood testing for the early detection of breast cancer in asymptomatic women was not covered in the 2009 report.

# **1.1.** Blood testing as a breast cancer screening test

Most biological processes of cancer result in the release of biomarkers that are present in the blood (Cree, 2011). These biomarkers are easily and inexpensively measured in bodily fluids (mainly in serum or plasma), with results being quickly available to clinicians and patients (Holdenrieder, Pagliaro, Morgenstern, & Dayyani, 2016). Furthermore, the tests are not overly invasive and do not increase patient risk, as they do not require biopsies or involve radiation (Hyun, Kim, Gwak, & Jung, 2016). This makes blood biomarkers a promising alternative screening modality for the early detection of breast cancer in asymptomatic women.

There are currently no identified blood biomarkers that have sufficient sensitivity and specificity to support their use in existing breast cancer screening programs for asymptomatic women (Cree et al., 2017); however, there are several different blood biomarkers that show promise. A significant amount of current research is focused on developing more specific and sensitive tests using these markers. Three biomarkers were studied in the identified literature:

- 1. circulating microRNAs
- 2. circulating tumour DNA (ctDNA), and
- 3. circulating tumour cells (CTCs).

This section of the horizon scan presents findings related to these three biomarkers.

# **1.2.** Summary of key findings

- The use of blood testing in breast cancer screening of asymptomatic women is still in the early stages, with much of the research focus being on identifying promising biomarkers that demonstrate sufficient sensitivity and specificity to support their use in further clinical testing. Research is most advanced concerning the use of microRNA in screening.
- There is no indication of the timeframe in which the full clinical potential of blood testing for breast cancer detection will be realised; however, preliminary results from primarily retrospective studies are promising and the technology is improving rapidly.
- Blood testing has not been incorporated into any national screening programs, nor are there any national position statements that have been released on their use in breast cancer screening.
- Current research is not sufficient to be able to identify whether blood biomarker testing is able to reduce deaths due to breast cancer through early detection.

# 1.3. Literature search results (number of studies returned)

From the literature search, a total of 22 articles related to blood testing were identified. Abstract contents were then reviewed and 20 of these articles were excluded for numerous reasons, including: relating to risk assessment or diagnosis rather than cancer detection; involving only non-human subjects or symptomatic participants; or relating to cancer generally rather than breast cancer specifically. One further article was subsequently excluded because it could not be located (Weigel & Dowsett, 2010). The reference lists of both included and excluded articles were reviewed, which led to the inclusion of a further four studies.

Because of the relatively early stage of research in this area, many of the more recent articles do not include asymptomatic samples or participants. Additionally, much of the current research focus for blood testing relates to breast cancer diagnosis or treatment, rather than to early detection in asymptomatic women. We therefore included the most recent articles we could find on each of the biomarkers, meaning that the only article identified for CTC testing was published before the 2010 cut-off for the literature search (eg, Mostert, Sleijfer, Foekens, & Gratama, 2009). Furthermore, studies already incorporated into an existing systematic or literature review were not separately assessed for the current scan.

A total of five articles were then reviewed to answer the key research questions in relation to the use of blood testing for breast cancer screening in asymptomatic women.

Systematic and/or literature reviews ctDNA: One study (Cheuk, Shin, & Kwong, 2017) CTCs: One study (Mostert et al., 2009) RCTs None identified Prospective studies microRNAs: One study (Godfrey et al., 2013) Retrospective studies ctDNA: One study (Kloten et al., 2013) microRNAs: One study (Ng et al., 2013) MicroRNAs: One study (Ng et al., 2013) MicroRNAs: One study (Ng et al., 2013)

# **1.4.** Study findings and discussion

# 1.4.1. What stage of development or trial are the new tests at?

# **Circulating tumour DNA (ctDNA)**

ctDNA is a circulating cell free DNA that has been linked to the presence of numerous types of cancer, including breast cancer (Cree et al., 2017). Because ctDNA is produced as a result of cell

death it can also appear because of the presence of a number of other health conditions including myocardial infarction, pregnancy, and serious infections or inflammatory conditions (Cree et al., 2017). High levels of ctDNA are therefore not specific to cancer and may reduce specificity when used in breast cancer screening of asymptomatic women. Research is currently focused on identifying ctDNA markers that are specific to breast cancer, thereby improving its potential to be used in screening.

Blood testing based on ctDNA is still at the stage of identifying biomarkers with sufficient specificity and sensitivity to have a screening application. Screening blood tests based on ctDNA involve the identification of methylation in a panel of tumour-suppressor genes. A panel of genes is required because no single gene has been found to be methylated in every breast cancer specimen (Kloten et al., 2013). However, there is currently no scientific consensus as to which genes should be included in the panel used in testing. Identifying genes that provide adequate levels of specificity and sensitivity in distinguishing between individuals with breast cancer and those without is currently the focus of a large majority of literature in this area.

# Circulating tumour cells (CTCs)

CTCs are defined as cells shed from either the primary tumour or its metastases, that are circulating in the periphery blood system. There have been numerous attempts at developing measures of CTCs in the blood but there is not one specific feature of CTCs that reliably differentiates them from regular blood cells (Mostert et al., 2009). CTCs are also rare events, meaning that they are present only at extremely low levels in the blood. These issues affect the sensitivity and specifity of CTCs in detecting early breast cancer, and prohibit the adoption of CTC detection into current screening programs.

Research looking at the potential utility of CTC measurement in screening for breast cancer is still in concept testing stages. As mentioned above, there is no one feature of CTCs that allow them to be reliably differentiated from normal blood cells. Also, different histological and molecular types of tumours produce different CTC features. Current research is focused on identifying the optimal array of CTC markers to produce sufficient sensitivity and specificity in screening. We were only able to identify one study that assessed the use of CTCs in breast cancer screening. Mostert et al.'s 2009 study also discussed the use of CTCs in cancer screening more broadly.

# **Circulating microRNAs**

Small non-protein-coding RNAs (microRNAs) play important roles in the formation of cancer, with levels of microRNAs being found to be elevated in individuals with breast cancer (Ng et al., 2013). MicroRNAs are present in both human plasma and serum, with the presence of tumours altering their expression. The stability of microRNA levels in samples of plasma and serum make them a potentially useful test marker in screening for breast cancer (Godfrey et al., 2013).

Although there is not much existing research on the application of microRNA testing for use in breast cancer screening, the research is slightly more advanced than for cfDNA or CTC tests. Prospective field studies have begun on the use of microRNA testing in breast cancer screening for asymptomatic women, with the results from one study published (Godfrey et al., 2013) and another study currently being conducted (Giordano, Gallo, Petracci, Chiorino, & Segnan, 2017). This study is expected to complete data collection in March 2018. These studies will provide

more reliable information about the utility and challenges of implementing microRNA testing into screening programs at the population level than is available from retrospective studies.

# 1.4.2. What are their considered potential clinical value in five years? In 10 years?

The identified literature did not discuss timeframes in which potential clinical value may be realised, as the testing of biomarkers as part of screening programs is still in its relative infancy. That said, two papers noted the speed at which this technology has been developing over the past decade for the screening of cancers, including prostate cancer and breast cancer (Holdenrieder et al., 2016; Volik, Alcaide, Morin, & Collins, 2016).

Although we were not able to determine an indicative timeframe for the full potential of these tests to be realised, there have been a growing number of studies assessing the potential specificity and sensitivity of the biomarkers for detecting breast cancer in asymptomatic women. In order to replace the current standard breast cancer screening test (FFDM), blood biomarkers need to be able to achieve a sensitivity of at least 70% and a specificity of at least 85% (Kloten et al., 2013). Findings from the identified literature are outlined below for each biomarker, separated by the type of study methodology utilised.

# ctDNA

# Systematic and/or literature reviews

One systematic review on 14 studies assessing the utility of ctDNA testing in breast cancer detection was identified (Cheuk et al., 2017); however, the focus of this review was partially on breast cancer diagnosis rather than on early detection. Furthermore, the review did not include specificity and/or sensitivity rates found in each of the studies. The reviewed studies all used retrospective case-controlled designs to assess the performance of ctDNA testing in breast cancer. The studies included in the systematic review are provided in Blood Test Table 1 (overleaf).

Individual studies have been accessed where possible, and sensitivity and specificity rates for samples from breast cancer patients compared with healthy controls have been included, where available. Overall, sensitivity ranged from 62-95%, and specificity ranged from 87-100%. These results are promising in suggesting that ctDNA testing may be sufficiently specific and sensitive in the early detection of cancer in asymptomatic women to provide a viable alternative to FFDM.

Despite the limitations of the review, there were several findings that provide insights into how specificity and sensitivity might be improved in ctDNA testing. The authors noted that results from the 14 studies promote the use of a multigene panel in ctDNA testing (used in seven studies) rather than single gene testing (used in seven studies), as multigene panels produced higher sensitivity and specificity during cancer detection. Also, the sensitivity of ctDNA testing was found to be lower in plasma or serum samples than in primary tumour tissue samples, with the use of plasma or serum samples decreasing sensitivity by 9 to 15 percentage points in two studies.

The review discussed the limitations of existing ctDNA research, including relatively small sample sizes and the need to increase the number of studies including healthy controls so that the utility of ctDNA for use in breast cancer screening for asymptomatic women can be further investigated. They also noted the need for standardised guidelines for conducting research in

this area so that studies can be pooled and meta-analysed to overcome limitations posed by small sample sizes.

These results are supportive of the utility of this gene panel for breast cancer screening blood tests, with specificity and sensitivity rates higher than that commonly found for digital mammography. Although promising, it is important that these results are replicated in other studies, including studies that directly compare the performance of ctDNA testing and FFDM.

Blood Test Table 1: Studies included in Cheuk et al. (2017) systematic review

Note: Retrieved from Cheuk et al. (2017).

Author	Year	Sample size	Type of specimen	Single or multigene	Sensitivity	Specificity
Chimonidou et al.	2013	Control, 60; cancer, 114	Plasma	Single	Not provided	Not provided
Chimonidou et al.	2013	Control, 37; cancer, 27 (discovery), 46 (metastatic), 123 (validation)	Plasma	Single	Not provided	Not provided
Dulaimi et al.	2004	Control, 20; benign, 8; cancer, 34	Serum	3 genes	Not accessible	Not accessible
Fu et al.	2015	Benign, 60; cancer, 155	Plasma	Single	No control	No control
Guerrero-Preston et al.	2014	Control, 20 (discovery), 86 (validation); cancer, 20 (discovery), 154 (validation)	Plasma	5 genes	94%	87%
Hagrass et al.	2014	Benign, 100; cancer, 120	Serum	Single	No control	No control
Hoque et al.	2006	Control, 76 (discovery), 38 (validation); cancer, 93 (discovery), 47 (validation)	Plasma	4 genes	62%	87%
Martínez-Galán et al.	2008	Control, 74; benign, 34; cancer, 106; post-operational, 60	Serum	5 genes	81%	88%
Ng et al.	2011	Control, 60 (discovery), 20 (validation); breast cancer, 60 (discovery), 38 (validation); gastric cancer, 45 (discovery), 20 (validation)	Plasma	Single	87%	85%
Radpour et al.	2011	Control, 30 (plasma, discovery); cancer, 36 (plasma, discovery), 20 (serum, validation)	Plasma, serum	10 genes	92%	90%
Sharma et al.	2011	Control, 30; cancer, 100	Serum	Single	Not provided	Not provided
Silva et al.	1999	Control, 17; cancer, 35	Plasma	Single	Not provided	Not provided
Skvortsova et al.	2006	Control, 10; benign, 15; cancer, 20	Plasma	3 genes	95%	100%
Yamamoto et al.	2012	Control, 87; cancer, 159	Serum	3 genes	78%	93%

#### **Retrospective studies**

One retrospective study assessing ctDNA testing in the early detection of breast cancer was identified that was not included in the above systematic review (Kloten et al., 2013). This study assessed the potential of ctDNA measurement in seven tumour-suppressor genes for use in breast cancer screening. Using serum samples from 138 breast cancer patients and 135 healthy adult controls, they found that by using three genes they were able to detect breast cancer with a sensitivity of 67% and specificity of 69%. It was noted that these sensitivity and specificity levels were not sufficient to out-perform general results found for FFDM; however, the authors suggest that this gene panel may be useful for screening women with dense breasts due to the reduced accuracy of digital mammograms in this population.

#### microRNAs

# **Prospective studies**

To prospectively assess the ability of microRNAs to successfully detect breast cancer in the general population, Godfrey et al. (2013) obtained serum samples from 205 women who subsequently developed breast cancer, and 205 women who remained cancer-free. These participants were sourced from the Sister Study cohort, which comprises 50,884 self-selected women from the US or Puerto Rico who had never had breast cancer but who had a full or half-sister who had breast cancer at time of enrolment. This puts the sample at higher risk of developing breast cancer, meaning that the generalisability of results to average-risk women may be affected. The serum samples used for the women who developed breast cancer were taken within 18 months prior to their diagnosis, and controls were matched on criteria including ethnicity, age, and similar date of blood draw.

Results showed statistically significant (although relatively small) differences in the levels of 21 microRNAs between the women who developed breast cancer and the controls, with differences ranging from 4-19% (p<.05). Importantly, previous reports of primary breast tumours had examined seven of these 21 microRNAs, with results from all reports matching the results found in Godfrey et al.'s study. No measures of specificity or sensitivity were reported in this study, however the authors felt that the results questioned the utility of microRNA serum testing for screening purposes due to the relatively small differences between groups (and thereby the impact this would potentially have on sensitivity).

# **Retrospective studies**

One retrospective study provided slightly more promising indications of the utility of microRNA in breast cancer screening for asymptomatic women. Ng et al. (2013) used plasma samples from 70 female breast cancer patients and 100 healthy controls (60 female and 40 male) to validate the use of microRNA measurement in the detection of breast cancer. Levels of four specific microRNAs were assessed in the plasma samples. Results found that there was no significant difference in the level of microRNA markers between female and male controls. Levels of three of the markers were significantly higher, and one significantly lower, in breast cancer patients compared with controls (p<.0001).

Logistic regression was used to identify the two markers that were the best combination for breast cancer detection. Receiver operating characteristic (ROC) curve analysis for these combined markers showed an AUC of 0.931 (95% CI 0.886 – 0.977), indicating a strong ability for predicting breast cancer. A sensitivity and specificity of 83% (95% CI 72% - 91%) and 93% (95% CI 81% - 98%), respectively, were found in discriminating breast cancer from healthy controls. Furthermore, the positive predictive value was 88% and the negative predictive value

was 92%. Perhaps most importantly for future potential of these markers in screening, the results showed that these two markers were specific to breast cancer and did not elevate in other cancers included in the study. The authors suggested these results make a very promising and specific breast cancer screening test based on microRNAs possible, however further research is needed to validate this finding.

# CTCs

# Systematic and/or literature reviews

One literature review was identified relating to the measurement of CTCs in breast cancer. This review only included primary studies quantifying the performance of CTC testing in the diagnosis or treatment of breast cancer (Mostert et al., 2009). The review noted that there are currently specificity and sensitivity issues with CTC testing (as described in section 1.4.1), which was likely to affect its appropriateness and accuracy for breast cancer screening. It highlighted the ongoing efforts to identify an optimal assay of CTC markers that would result in an improvement in sensitivity and specificity to levels that were adequate for prospective studies.

# 1.4.3. What cost and safety findings have been reported?

No information on the cost and safety of blood testing was identified for any of the biomarkers that could potentially be used for breast cancer screening; however, there was wide-spread recognition that blood testing is relatively inexpensive, potentially able to be incorporated into annual general health checks, and relatively non-invasive, and could therefore arguably be more acceptable to screening patients than other modalities currently used (Volik et al., 2016; Cree, 2011).

# 1.4.4. Does this technology reduce deaths due to breast cancer through early detection?

Current research is not sufficient to be able to identify whether blood biomarker testing is able to reduce deaths due to breast cancer through early detection of cancer in asymptomatic women.

None of the articles reviewed for this horizon scan specifically talked whether blood testing can detect smaller cancers. Research for blood testing is still at a concept-testing stage, which means that the current research typically uses a sample of breast cancer patients and compares their blood profiles to healthy controls. This means that in the current research, the trial of blood testing most likely involves patients with relatively large tumours (but again, this is uncertain). This is a key gap in the current evidence base which will develop further.

# **1.4.5.** Has this test been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

Blood biomarker testing has not been incorporated into any national screening programs.

# **1.4.6.** Has a national position statement been published about this innovation, and if so, what is the position? Is there consensus in position statements?

No national position statements on the use of blood biomarker testing in breast cancer screening were identified in the literature search.

# 2. SALIVA TESTING

# Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

The use of saliva testing for the early detection of breast cancer in asymptomatic women was not covered in the 2009 report.

# 2.1. Saliva testing as a breast cancer screening test

The use of saliva testing is increasingly becoming a research focus for the early detection of breast cancer in asymptomatic women. There are a number of advantages to using saliva for screening, including that sampling is easy, cheap and non-invasive, and can be collected by individuals with relatively little training (Liu & Duan, 2012; Pfaffe, Cooper-White, Beyerlein, Kostner, & Punyadeera, 2011). Additionally, due to the continual production of saliva in humans, saliva samples are able to give a picture of body condition and health at the exact point of sample collection (Streckfus, Brown, & Bull, 2010).

Despite the promise of saliva testing for use in the early detection of breast cancer in asymptomatic women, there is currently no saliva-based test that is sufficiently sensitive or specific for use in routine screening (Sugimoto, Wong, Hirayama, Soga, & Tomita, 2010); however, research is still being carried out to further develop saliva testing so that its full clinical potential can be reached. This section provides an overview of our current understanding of the potential for saliva testing to be used in the early detection of breast cancer in asymptomatic women.

# 2.2. Summary of key findings

- The use of saliva testing in breast cancer screening of asymptomatic women is still early stages, with much of the research focused on identifying promising biomarkers that are specific to breast cancer, and which can detect breast cancer with sufficient sensitivity and specificity to support their use in further clinical testing (including prospective studies and RCTs). Current research is limited to retrospective studies comparing samples from breast cancer patients with healthy controls.
- There is currently no indication of the timeframe in which the full clinical potential of saliva testing for breast cancer detection will be realised; however, preliminary results from retrospective studies are promising, with studies routinely finding sensitivity and specificity rates over 80%, or even over 90%, for a range of saliva biomarkers.
- Saliva testing has not been incorporated into any national screening programs, nor are there any national position statements that have been released on their use in breast cancer screening for asymptomatic women.
- Current research is not sufficiently advanced to identify whether saliva biomarker testing is able to reduce deaths due to breast cancer through early detection.
# 2.3. Literature search results (number of studies returned)

From the literature search, a total of 11 articles related to saliva testing were identified. Abstract contents were then reviewed and three of these articles were excluded; one study because it related to the development of a saliva testing device rather than the validation of saliva testing on humans, and two studies because they were not specific to breast cancer. Examination of the reference lists for both included and excluded studies led to the identification of one further included study.

A total of nine articles were then reviewed to answer the key research questions in relation to the use of saliva testing for breast cancer screening in asymptomatic women.

#### Systematic and/or literature reviews

Three general literature reviews: Liu & Duan, 2012; Pfaffe et al., 2011; Streckfus et al., 2010

RCTs

None identified.

#### **Prospective studies**

None identified.

#### **Retrospective studies**

Six retrospective studies: Al-Muhtaseb, 2014; Hernandez-Arteaga et al., 2017; Sugimoto et al., 2010; Takayama et al., 2016; Zhang et al., 2010; Zhong, Cheng, Lu, Duan, & Wang, 2016.

#### **Grey literature**

None identified.

# 2.4. Study findings and discussion

#### 2.4.1. What stage of development or trial are the new tests at?

Research into the use of saliva testing for the early detection of breast cancer is still in a relatively early stage. Human saliva comprises a number of different proteins and peptides, and so one of the important first steps in realising the potential of saliva as a screening test is to discover saliva biomarkers that are specific to breast cancer (Liu & Duan, 2012). Problematically, some studies have found that some of the saliva biomarkers explored are not specific to breast cancer and are confounded by the presence of other cancers (Hernandez-Arteaga et al., 2017; Sugimoto et al., 2010).

The detection of appropriate salivary biomarkers is currently the predominant focus of research in this area, with researchers using studies of proteomes (groups of proteins) to identify biomarkers that are promising in terms of their specificity and sensitivity in detecting breast cancer. As such, all articles identified used retrospective designs, which limits the strength of the conclusions that can be drawn from this research. Prospective or RCT trials conducted in the field would provide more reliable evidence about the performance of saliva testing in the early detection of breast cancer for asymptomatic women.

# 2.4.2. What are their considered potential clinical value in five years? In 10 years?

Because of the relative infancy of research in the use of saliva testing for the early detection of breast cancer, there is no indication of the timeframes in which the clinical potential of saliva testing will be realised. However, there is a growing body of research assessing the potential of saliva testing in breast cancer screening that utilises retrospective designs with samples of breast cancer patients and healthy controls. This research can provide an indication of the potential sensitivity and specificity of saliva testing in the early detection of breast cancer in asymptomatic women.

In order to replace the current standard in breast cancer screening – FFDM – saliva biomarker testing needs to be able to achieve a sensitivity of at least 70% and a specificity of at least 85% (Kloten et al., 2013) as well as be fast to administer, cost-effective, safe and acceptable to women and health practitioners.

#### Systematic and/or literature reviews

The three general reviews did not contain any specific information about the specificity and sensitivity found in studies discussed, so the individual studies were identified in reference lists and reviewed separately below.

#### **Retrospective studies**

The most recent article identified measured sialic acid levels in saliva samples obtained from 100 female breast cancer patients and 106 healthy female controls (Hernandez-Arteaga et al., 2017). They found that levels of sialic acid were significantly higher in breast cancer patients compared with controls (p<.05), with testing able to differentiate between the two samples with a sensitivity of 94% and a specificity of 98%. Although these results are promising, it is important to note that the age profiles of the two samples were markedly different, with the control group being younger on average (M = 28.6 years) than the breast cancer patients (M = 51.8 years). The effect of this age difference was not specifically assessed in the study, although the authors note that previous research indicates that levels of sialic acid do not differ by age to such an extent that they would expect it to affect their results.

High rates of specificity and sensitivity were also found in four other studies. The first, Takayama et al. (2016), measured the concentration of 10 polyamines extracted from saliva samples obtained from 191 breast cancer patients and 61 healthy controls. They found that concentration of eight of the polyamines was significantly higher, and significantly lower in two of the polyamines, for breast cancer patients compared with controls. Additionally, measurement of the concentration of eight of the polyamines was able to detect breast cancer with a sensitivity and sensitivity of 88% each. Although lower than the sensitivity and specificity identified in the previous study, these results still indicate that saliva testing may be more specific and sensitive than FFDM in the detection of breast cancer.

In the second study, Zhong, Cheng, Lu, Duan, & Wang (2016) found high rates of specificity and sensitivity in a range of different saliva biomarkers. They collected saliva samples from 30 female breast cancer patients and 25 healthy female controls, and measured levels of 18 different biomarkers. They found differences in the profiles of each of these biomarkers between

the two samples, with sensitivity ranging from 48-96% and specificity ranging from 54-100% across the biomarkers. They did not provide an indication of the specificity and sensitivity obtained from measuring multiple biomarkers at once, however the high rates found for individual biomarkers are promising. It is likely that combining different biomarkers into the one test would improve overall sensitivity and specificity.

The third study, Zhang et al. (2010), used saliva samples from 30 breast cancer patients and 63 matched healthy controls to assess the cancer detection performance of eight mRNA biomarkers and one protein biomarker found in saliva. They found statistically significant differences in the profiles of these biomarkers between the patients and controls and were able to differentiate between the samples with 83% sensitivity and 97% specificity (p<.05). As a further demonstration of the promise of saliva testing for screening purposes, none of the confounders assessed (including age, ethnicity, menopause, and smoking status) affected the profiles of the biomarkers measured.

The discriminative ability of saliva biomarkers for breast cancer detection was also demonstrated by another retrospective study (Sugimoto et al., 2010). This study measured 14 metabolites extracted from saliva obtained from 30 breast cancer patients and 87 healthy controls. Although sensitivity and specificity were not separately reported, they found an Area Under the Curve (AUC) value of 0.881, demonstrating the ability to differentiate between controls and patients to a high degree of accuracy. Despite finding a significant difference in the median age of the two samples (57 for patients and 43 for controls), the researchers concluded there was no correlation between metabolite levels and age. Although these results are promising, the researchers also found that the 14 metabolites measured were not specific to breast cancer, with testing not able to accurately differentiate between breast cancer patients and patients with oral cancer or pancreatic cancer. This casts doubt on the ability of these metabolites to specifically detect breast cancer in general populations.

One additional study did not provide sensitivity and specificity measures for the use of saliva testing in early breast cancer detection; however, the authors were able to demonstrate differences in saliva profiles between breast cancer patients and healthy individuals. Al-Muhtaseb (2014) measured levels of saliva proteins collected from 40 female breast cancer patients and 40 female healthy controls aged between 50 and 70 years. They were able to show that the levels of saliva protein were significantly lower in breast cancer patients than in the healthy controls (p<.05), demonstrating the potential for saliva protein measurement as a cancer detection test.

Overall, results are promising in terms of the potential sensitivity and specificity of saliva testing for the early detection of breast cancer in asymptomatic women, however all available research is retrospective in design and uses relatively small sample sizes. This limits our ability to reliably determine the specificity and sensitivity of salvia testing in clinical settings.

# 2.4.3. What cost and safety findings have been reported?

There is currently no information on the cost and safety of saliva testing for breast cancer screening; however, much of the literature notes that saliva testing is relatively inexpensive, non-invasive, and safer than other screening tests due to there being no need for needles, radiation or biopsies for assessment purposes (Liu & Duan, 2012; Streckfus et al., 2010). Costs would also be reduced in that the training needed to take samples is relatively low (Pfaffe et al., 2011).

#### 2.4.4. Does this technology reduce deaths due to breast cancer through early detection?

Current research is not sufficient to be able to identify whether saliva testing is able to reduce deaths due to breast cancer through early detection of cancer in asymptomatic women.

None of the articles reviewed for this horizon scan specifically talked about breast cancer tumour size and its impact on the accuracy of saliva testing. Research is still at a concept-testing stage, which means that the current research typically uses a sample of breast cancer patients and compares their saliva profiles to healthy controls. Limited information about tumour size at detection is a limitation of the current evidence base. In addition, saliva profiles are correlated with the size of oral cancer tumours, which indicates that there is potential for smaller tumours to make an impression on saliva profiles, but this is not specific to breast cancer.

# 2.4.5. Has this test been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

Saliva testing has not been incorporated into any national screening programs.

# 2.4.6. Has a national position statement been published about this innovation, and if so, what is the position? Is there consensus in position statements?

No national position statements on the use of saliva testing in breast cancer screening were identified in the literature search.

### **IMAGING MODALITIES**

This section covers:

- Automated whole breast ultrasound
- Contrast enhanced mammography
- Digital breast tomosynthesis
- Ductoscopy
- Magnetic resonance imaging
- Microwave imaging
- Molecular breast imaging (scintimammography)
- Spectroscopy
- Thermography, and
- Tomography.

# 3. AUTOMATED WHOLE BREAST ULTRASOUND

#### Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

*Safety*: No adverse effects associated with the use of ABUS were identified. Ultrasound does not use ionising radiation, contributing to the safety of the procedure.

*Effectiveness*: Overall, the review of three studies was inconclusive regarding the effectiveness of ultrasound compared with mammography. Results suggested that ultrasound could be associated with lower specificity in the early detection of breast cancer for asymptomatic women; however, there was some indication that using ultrasound as an adjunct to mammography may have some benefit.

*Cost*: Two studies estimated the cost per ultrasound examination to range from  $\in$ 22-62. The cost per cancer detected was substantially higher for ultrasound (ranging from  $\in$ 14,618 – 25,847) compared with mammography screening ( $\in$ 5,000).

# **3.1.** Ultrasonography as a breast screening tool

Ultrasonography is used in the assessment and diagnosis of breast cancer and has traditionally been performed by a physician moving a hand-held device (called a transducer) over the breast, often referred to as hand-held ultrasonography (HHUS). In HHUS systems, the operator manually directs the transducer to the parts of the body being examined, which requires a high level of training and skill and creates a level of operator dependency. Ultrasonography uses highfrequency soundwaves that 'echo' as they pass through various types of tissue. These echoes are used to create an image called a sonogram, which depicts the internal structures inside the body.

Ultrasonography is a popular imaging technique because it is comfortable for patients, widely available at a relatively low cost, and does not involve the use of ionising radiation or contrasting agents (Geisel et al., 2018). Despite the appeal of ultrasonography to patients, concerns have been raised about the use of HHUS for screening given its lower rates of specificity compared with FFDM (Geisel et al., 2018). Furthermore, the application of HHUS for the early detection of breast cancer has been limited by a lack of technologists or physicians with the level of experience required to perform HHUS examinations (Geisel et al., 2018).

ABUS was developed in response to the limitations of HHUS. In comparison with HHUS, ABUS systems can produce highly reproducible images without the need for highly trained clinicians to perform the examination (Geisel et al., 2018). This is because in ABUS, interpreting the captured images is separated from the process of capturing the images.

There are currently two types of ABUS systems available. The first uses an automated arm to move a handheld transducer, creating 2D images of the breast; a technician guides this automated arm throughout the examination to ensure sufficient contact between the transducer and the breast. The second ABUS system uses a high-frequency, large-format transducer to obtain 3D volumetric images of the breast. Studies have shown ABUS to have equal or greater lesion detectability compared with HHUS (Gilbert & Selamoglu, 2017). Furthermore, the use of ABUS systems requires less training and is less operator dependent. This has made ABUS systems an attractive and promising alternative to traditional HHUS systems, with much of the

current research in this area focusing on improving its effectiveness in screening settings for women with dense breasts.

# **3.2.** Summary of key findings

- ABUS is already in clinical use for breast cancer screening in the United States and Canada. There is no evidence to suggest that ABUS has been implemented into a national screening program at this stage.
- The potential clinical value of ABUS in five to 10 years is not explicitly discussed in the literature, although findings have indicated its value as an adjunct screening tool for women with denser breasts.
- Many studies have found that using ABUS and FFDM leads to significantly higher cancer detection rates for women with denser breasts than using digital mammography alone.
- ABUS is likely to be highly acceptable to women as no safety issues have been identified and the examination process involves less discomfort than other modalities, including digital mammography.
- There is no evidence that ABUS reduces deaths due to breast cancer through early detection.
- ABUS has not been incorporated into any national breast screening programs, nor are there any national position statements that have been released on its use in breast cancer screening for asymptomatic women.

# **3.3.** Literature search results (number of studies returned)

From the literature search a total of 130 abstracts of peer reviewed articles related to breast ultrasonography were identified. The abstracts were reviewed, and 111 articles were excluded because they related to HHUS only, or the use of ultrasonography in a symptomatic population of for diagnostic or treatment purposes. A total of 19 articles were reviewed to answer the research questions relating to the use of automated breast ultrasonography as a breast cancer screening tool.

#### Systematic and/or literature reviews

Eight articles were identified that were either systematic or literature reviews (Berg & Mendelson, 2014; Brem, Lenihan, Lieberman, & Torrente, 2015; Geisel, Raghu, & Hooley, 2018; Gilbert & Selamoglu, 2017; Kornecki, 2011; Le-Petross & Shetty, 2011; Melnikow et al., 2016; and Youk & Kim, 2010).

#### **Prospective studies**

Eight prospective studies were identified (Berg et al., 2015; Corsetti et al., 2011; Giuliano & Giuliano, 2013; Kim, Kim, Moon, Yoon & Kim, 2016; Kelly et al., 2010; Kelly & Richwald, 2011; Ohuchi et al., 2016; and Wilczek, Wilczek, Rasouliyan, & Leifland, 2016)

#### Retrospective studies

Three retrospective studies were identified (Chang, Moon, Cho, Park & Kim, 2011; Hooley, Scoutt, & Philpotts, 2013; and Skaane et al., 2015)

# 3.4. Study findings and discussion

#### 3.4.1. What stage of development or trial is this innovation at?

#### Systematic and/or literature reviews

Six systematic reviews were identified that discussed the current use of ABUS in breast cancer screening in asymptomatic women. Some of these reviews referred to studies of HHUS, the results of which are of relevance to ultrasonography in general.

In their review, Le-Petross & Shetty (2011) referred to six cohort studies which found that the use of ultrasound as a screening technique identified primarily invasive cancers in 0.32% of women. The mean tumour size was 9.9mm, and 90% of cancers were node negative. Biopsy rate was high at 2.3% to 4.7%, with positive predictive value of 8.4% to 13.7% for those biopsied because of an abnormal finding on the ultrasound examination. They observed, however, that the added benefit of using this imaging modality to screen for breast cancer was lower in women aged 50-69 years.

The Le-Petross and Shetty review also noted that ultrasound was able to identify small nonpalpable masses while undeterred by presence of dense breast tissue, which is an inherent limitation of mammography. However, unlike mammography, DCIS is not usually identified by ultrasound. Their review (as well as the review by Kornecki, 2011) noted that no study had advocated for ultrasound to be used as the only modality to screen for breast cancer. This was said to be due to the low yield of ultrasound alone detected breast cancers.

Melnikow et al. (2016) looked at breast cancer detection outcomes for supplementary ABUS and HHUS in 13 studies. The cancer detection rate ranged from 1.9 to 15.2 per 1,000 screening examinations for ABUS (3 studies) and 0.4 to 18.9 per 1,000 for HHUS, with comparable recall rates ranging from 2 to 14%. While there was little to distinguish between the two ultrasound techniques in terms of cancer detection rates, they noted that the sensitivity rate for ABUS (68.0%) was lower than HHUS (which ranged from 80.0 to 100.0%), although specificity was higher for ABUS (92.0%) than HHUS (72.0 to 95.0%).

While the previous study demonstrated equal or greater lesion detectability with ABUS than handheld imaging, Gilbert & Selamoglu (2017) raised concerns regarding false positives and high recall rates, which were higher when ultrasound was used as a supplementary tool to mammography compared with mammography alone. In one of the studies they referred to, recall rates with ultrasound alone were 20.9% in the prevalent round although they dropped to 10.7% in subsequent rounds, compared to mammography recall rates of 11.5% and 9.4%. Increased recall rates were observed in three other studies. They also noted a reader study of 185 cases, including 52 cancer cases that compared the use of mammography alone to mammography combined with ABUS. This study found that while using mammography combined with ABUS. This study found that while using mammography combined with ABUS. As the performance of the readers in this study was variable, it was stressed that training was essential (Gilbert & Selamoglu, 2017).

In their literature review, Brem, Lenihan, Lieberman, & Torrente (2015) referred to several studies that also concluded that ABUS improved breast cancer detection, noting increases in sensitivity for ABUS (plus FFDM) and FFDM alone ranging from 26.7% to 41%. The most prevalent additional cancers detected with ABUS were invasive node-negative cancers (93.3%).

Commenting on reader performance, Berg & Mendelson, (2014) concluded that to meet future demand this screening technique would need to be done by trained technologists rather than physicians. They observed that while ultrasonography has been used for decades to assess masses detected by mammography, there has been a lack of specific training to ensure that it is being used appropriately and effectively. Training is essential to ensure comparable detection rates are achieved with HHUS and ABUS. They referred to promising evidence suggesting the use of ABUS as an adjunct to FFDM resulted in significantly higher cancer detection rates for women with denser breasts (with acceptably higher recall rates). However, more general population studies and large-scale multi-centre RCTs were needed to confirm the efficacy of ABUS in reducing mortality rates through the early detection of cancer.

#### **Prospective trials**

Ohuchi et al. (2016) examined the efficacy of adjunctive ultrasonography. As part of the Japan Strategic Anti-cancer Randomised Trial (J-Start), they screened 79,998 asymptomatic women aged between 40-49 years old, who were randomly assigned in two groups, one undergoing FFDM plus ABUS (the intervention group) and the other FFDM alone (the control group). They found that using ABUS as an adjunct to FFDM improved sensitivity from 77% to 95% (*p*=.0004), although specificity decreased to 87.7% from 91.4% (*p*<.0001). More cancers were detected in the intervention group than in the control group (184 [0.50%] *vs* 117 [0.32%], *p*=0.0003) and were more frequently stage 0 and I (144 [71.3%] *vs* 79 [52.0%], *p*=0.0194). Eighteen (0.05%) interval cancers were detected in the intervention group (*p*=0.034). Their results demonstrated the usefulness of ultrasonography in the early detection of breast cancer.

One of the largest clinical trials of screening ultrasound to date was the American College of Radiology Imaging Network trial (ACRIN 6666). In the study, 2,662 asymptomatic women with high risk of cancer were randomly assigned to two groups, one receiving ultrasound and FFDM screening and the other FFDM screening alone (see Le-Petross & Shetty, 2011). They found that cancer detection was comparable for both ultrasound and FFDM (52.3 v 53.2%, p=.90), with ultrasonography detected cancers more likely invasive (91.4% for ultrasound compared to 69.5% for FFDM). Invasive cancers detected by ultrasound were more frequently node-negative, 64.2% for compared to 43.9% (p=.003). The study observed higher recall and biopsy rates and lower PPV of biopsy for ABUS (Berg et al., 2015).

In their study, Kim, Kim, Moon, Yoon & Kim (2016) also observed that false positives were more common with ultrasound, with attendant discomfort, emotional stress and medical cost.

#### **Retrospective studies**

Current ABUS units, which are equipped with high-frequency broadband transducers, have the advantages of reproductivity, utility (in terms of surveying large areas of the breast), and reduced operator dependence compared to hand-held HHUS devices (Hooley, Scoutt & Philpotts, 2013, and Chang, Moon, Cho, Park & Kim, 2011). However, it is unclear whether ABUS systems are more effective than HHUS devices at detecting cancer in asymptomatic women. For instance, Chang and colleagues retrospectively assessed radiologist performance with respect to the

detection of breast cancer initially detected by HHUS using 3D breast volume data obtained by ABUS. The sensitivities of the radiologists ranged from 57.1% to 78.6%, with false-positives rates ranging from 8.3% to 20.8%. They found that of the HHUS-detected cancers, only 57.1-78.6% were identified with ABUS, which they attributed to the lack of experience that the radiologists had with the devices. In their view, a substantial level of experience and training was necessary to accurately interpret 3D breast volume data obtained by ABUS.

### 3.4.2. What is its considered potential clinical value in five years? In 10 years?

#### Systematic and/or literature reviews

The potential clinical value of ABUS in five to 10 years is not explicitly discussed in the literature. Studies have tended to focus on women with dense breasts, where the cancer detection rate when ABUS is combined with mammography is significantly greater than when mammography is used alone. The results of these studies are summarised in *section 3.4.3*.

# 3.4.3. Does this innovation show high sensitivity and specificity for women with dense breasts and women who have had breast surgery/augmentation compared to digital mammography?

#### Systematic and/or literature reviews

Gilbert & Selamoglu (2017) noted that many studies have demonstrated that ultrasound is a good screening tool for women with dense breast tissue. They referred to one study that detected 4.2 per 1,000 more cancers in women at increased risk and with dense breasts using ultrasound and mammography than mammography alone. Other studies referred to reported increases in cancer detection ranging from 3.4 to 4 per 1,000 examinations (Gilbert & Selamoglu, 2017).

In their paper, Youk & Kim (2010) agreed that ultrasound was an attractive adjunct to mammography in breast cancer screening for women with dense breasts because it was relatively inexpensive, required no contrast-medium injection, was well tolerated by patients, and was widely available for equipment as compared with MRI. Nonetheless, it had many limitations as a screening tool; including operator-dependence, the shortage of skilled operators, the inability of ultrasound alone to detect microcalcifications, and substantially higher false-positive rates than mammography.

#### Prospective studies

Five prospective studies have examined whether ABUS is a good screening tool for women with dense breast tissue.

Wilczek et al. (2016) evaluated the impact of ABUS when added to FFDM on breast cancer detection and recall rates in 1,668 asymptomatic women aged between 40 and 74 years with dense breast tissue examined at a high-volume breast cancer screening mammography centre. The authors found a statistically significant increase in the cancer detection rate when ABUS was used as an adjunct to FFDM; FFDM and ABUS detected a total of 6.6 cancers per 1000 compared to 4.2 per 1000 screening examinations for FFDM alone (p<.001). That said, the recall rate was also significantly increased with the addition of ABUS. FFDM alone resulted in a recall rate of 13.8 per 1000 screening examinations with a false positive rate of 70%, compared to 22.8 per 1000 with a false positive rate of 71% for FFDM + ABUS (p=.004). For FFDM alone the PPV<sub>1</sub>

percentage was 30% and the PPV<sub>2</sub> and PPV<sub>3</sub> percentage increased to 64%. For FFDM and ABUS, the PPV<sub>1</sub> percentage was 29% and the PPV<sub>2</sub> and PPV<sub>3</sub> percentage increased to 48%. The PPV rates were not significantly different between FFDM alone and FFDM and ABUS at the p = .05 level. Wilczek and colleagues concluded that using ABUS as an adjunct to FFDM 'significantly improved invasive breast cancer detection rate with an acceptable recall increase'.

A study conducted by the Huntington Memorial Hospital in California, USA, compared the use of FFDM alone with FFDM and ABUS in 4,419 women at elevated risk of breast cancer and/or dense breasts (Kelly et al., 2010). There was a statistically significant increase in the cancer detection rate when ABUS was used as an adjunct to FFDM. The sensitivity increased from 40% for FFDM alone to 81% for FFDM and ABUS. Furthermore, the number of detected invasive cancers, measuring 10mm or less, tripled from 7 to 21 when ABUS supplemented mammography. Although this study found an increase in recall rates for FFDM and ABUS (9.6%) compared with FFDM alone (4.2%), specificity of recalls was significantly higher for FFDM and ABUS (98.7%) compared with FDDM alone (95.2%; p<.05). Overall, Kelly et al. (2010) concluded that the combination of FFDM and ABUS produced significant improvements in cancer detection compared with FFDM alone. In their view, additional detection and the smaller size of invasive cancers justified the technology's expense for women with dense breasts.

A further two studies also reported promising results for the combination of mammography and ABUS. The Somo Insight Study of 15,318 women conducted between 2009 and 2011 sought to determine the improvement in early cancer detection by using ABUS with mammography compared with mammography alone in asymptomatic women with dense breasts (Brem et al., 2015). Following screening, 112 women were diagnosed with cancer; 82 of these women were diagnosed through mammography alone and a further 30 were diagnosed following ABUS examinations. This meant that adding ABUS to mammography led to an additional 1.9 cancers being detected per 1000 women screened (p<.001). Of the 82 cancers detected with mammography and ABUS, 17 were also detected by mammography alone, while ABUS alone detected 30. These additional cancers detected were clinically significant, with 62% of the cancers detected with MaBUS (p=.001). Sensitivity for mammography + ABUS increased by 27% compared with mammography alone, while the recall rate increased by 284.9 per 1000 women screened (p < .001) (Brem et al., 2015).

The second study by Giuliano & Giuliano (2013) assessed the efficacy of FFDM compared with FFDM and ABUS using a sample of 3,418 asymptomatic women with denser breasts. They found that ABUS had a significantly greater cancer detection rate of 12.3 per 1000 women screened compared with FFDM which had a cancer detection rate of 4.6 per 1000. The sensitivity (76%) and specificity (98%) of FFDM alone was lower than the sensitivity (98%) and specificity (100%) of FFDM + ABUS; however, only the difference in specificity was significant at the p<.05 level. The PPV was also significantly lower for FFDM alone (20%) compared to FFDM + ABUS (81%; p<.05). Based on these results, Giuliano and Giuliano concluded that the cost-benefit of supplementing FFDM with ABUS for women with denser breasts may be justified.

Surveillance of intervals cancers provides a measure of breast screening efficacy. Corsetti et al. (2011) explored the interval cancer rates of women with dense breasts who underwent adjunct ultrasound screening. Their study reported on a retrospective cohort of 8865 women who had 19,728 screening examinations (2001-2006): women with non-dense (BIRADS 1,2) breasts received mammography screening only, while women with dense (BIRADS 3, 4) breasts also received ultrasound. The underlying cancer rates (cancers observed within 1-year from

screening) were 6.3/1,000 screens in the non-dense breast group and 8.3/1,000 in the dense breast group. Screening sensitivity was 83.5% for mammography alone in non-dense breasts relative to 86.7% for mammography plus ultrasound in dense breasts. They observed that including ultrasound as an adjunct screening in women with dense breasts brought the interval cancer rate to similar levels as that for non-dense breasts. In their view, this suggested additional cancer detection by ultrasound was likely to improve screening benefit in dense breasts.

Overall, studies have demonstrated that using ABUS as an adjunct to FFDM could result in a significant increase in the cancer detection rates for women with denser breasts. While the use of FFDM and ABUS can also lead to an increase in recall rates compared with FFDM alone, most studies suggested that rates of false positives for these recalls were equal to or lower than the rates of false positives for FFDM recalls. However, the studies considered that the increase in recall rates was acceptable given the increased cancer detection rates for FFDM and ABUS compared with FFDM alone.

#### **Retrospective studies**

A retrospective study by Skaane et al. (2015) tested how successful five radiologists were in distinguishing between cancer findings (n = 38) and normal or benign findings (n = 76) in screening results from 90 women, using both ABUS and FFDM. The sample included 48 cases of women with denser breasts (BIRADS 3 and 4), however results were not stratified by breast density. The radiologists first made assessments using ABUS only, and then used digital mammography plus ABUS. The average area under the curve (AUC) value for the five readers was significantly higher for FFDM and ABUS (0.823) compared with ABUS alone (0.730; p<.05), indicating that the ability to accurately detect cancer findings was significantly higher with FFDM and ABUS. The consistency of ratings between readers was also higher for FFDM and ABUS compared with ABUS alone, with kappa values ranging from 0.14 to 0.44 for the former and 0.07 to 0.34 for the latter. The average interpretation time for a normal bilateral ABUS examination was 9 minutes. Because of the greater levels of performance for FFDM and ABUS, Skaane and colleagues said that combined reading should be the standard if ABUS is implemented for women with dense breasts.

#### 3.4.4. What cost and safety findings have been reported?

No safety risks regarding ABUS have been identified.

In terms of cost, Brem et al. (2010) gave a figure of \$300 per ultrasound screening but did not compare this to other imaging modalities or specify whether this was the cost of using ABUS of HHUS.

The cost effectiveness of this imaging modality will be affected by reader performance. Wilczek et al. (2016) noted that ABUS has a standardised imaging procedure that can be performed by medical personnel after a short training and without the need for radiologists. This reduces the amount of time and training required to perform the scan compared with HHUS systems. The time required to perform HHUS has been estimated at 19 mins, whereas for ABUS it is estimated to take 5-7 mins (Kelly & Richwald, 2011). Therefore, ABUS is likely to have a strong impact on workflow and therefore practitioner cost, especially in practices or programs managing high volumes of ultrasound images. Youk & Kim (2010) also noted that ABUS was time-consuming

compared to mammography alone. They noted that a breast radiologist would read about 50 studies per hour, which made ultrasound less efficient.

# **3.4.5.** Is there evidence that this innovation is more acceptable to women (in general and by ethnic group) compared to digital mammography?

No specific research on the acceptability of ABUS compared with FFDM was identified; however, ABUS is likely to be appealing to women due to its accessibility, relatively low cost, good patient tolerance – the examination process involves less discomfort than FFDM (because of the lack of breast compression) – and lack of ionising radiation (Gilbert & Selamoglu, 2017).

# 3.4.6. Does this technology reduce deaths due to breast cancer through early detection?

There is no evidence yet that ABUS reduces deaths due to breast cancer through early detection. Le-Petross & Shetty (2011) said that because the incidence of cancers seen on ultrasound is low, to prove mortality rate reduction, a large cohort would have to be studied in a randomised blinded controlled clinical trial. These studies are unlikely to be conducted anytime soon.

# **3.4.7.** Has this innovation been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

There is no evidence to suggest that ABUS has been implemented into a national screening program at this stage. However, ABUS has been in clinical use in the United States for several years and Health Canada has approved three ABUS systems, which have primarily been marketed for imaging dense breast tissue. Two of these have only limited approval and are intended only for use as an adjunct to mammography, rather than as a primary screening modality.

In response to the limitations of using FFDM to screen dense breast, more than 30 states in the United States have passed breast density notification laws (Geisel et al. (2018)). These laws require radiologists to inform patients of their breast density and provide information about available alternative screening options, including ABUS (Geisel et al., 2018). The public education campaign 'Are you Dense?' has also increased the public's awareness of ultrasonography as an adjunct to FFDM (Gilbert & Selamoglu, 2017).

# **3.4.8.** Has a national position statement been published about this innovation, and if so, what is the position? Is there a consensus in position statements?

The only published national position statement identified was that of New Zealand's National Screening Unit, published in September 2014. It concluded: *The National Screening Unit (NSU)* does not support the use of ultrasound as a primary screening tool or the routine use of ultrasound as an adjunct screening tool in the BreastScreen Aotearoa (BSA) programme, as presently there is insufficient evidence to do so.<sup>1</sup>

<sup>&</sup>lt;sup>1</sup> <u>https://www.nsu.govt.nz/system/files/page/position-statement-addressing-the-use-of-ultrasound-as-a-breast-cancer-tool.pdf</u>

# 4. CONTRAST ENHANCED MAMMOGRAPHY

#### Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

Contrast-enhanced mammography was not discussed in the 2009 Bulletin on new and emerging technologies for breast cancer detection.

# 4.1. What is contrast-enhanced mammography?

In the literature, contrast-enhanced mammography (CEM) is commonly referred to as contrast enhanced spectral mammography (CESM) or contrast enhanced digital mammography (CEDM) and other variations including temporal CEM and contrast enhanced dual energy spectral mammography, for example. For the purposes of this review CEM is used to describe any of the above variations – unless there is a clear difference in the method of CEM used.

CEM is a breast imaging modality that typically involves the following key steps (Bhimani et al., 2017; Gilbert & Selamoglu, 2017; Lewis et al., 2017):

- 1. Patient assessment, which includes a questionnaire to ensure the modality is appropriate for the patient those who are pregnant (due to radiation exposure), have allergies to the contrast agent, or have poor renal function are unable to undergo CEM.
- 2. Insertion of an intravenous (IV) line into the patient's arm.
- 3. Administration of a standard iodinated contrast agent by IV (concentration between 300-370mg/ml, at approximately 1.5ml/kg of body weight) using a power injector.
- 4. After at least 90 seconds, the patient is positioned for two standard mammography views (craniocaudal and mediolateral oblique) of each breast.
- 5. Rather than a standard single energy mammogram, the CEM technology acquires dualenergy image pairs in each projection. Since there is less than one second between the low-energy and high energy images, the imaging time is the same as that needed for a standard mammogram – additional projections may be obtained since optimally enhanced images can typically be obtained up to seven to 10 minutes following injection
- 6. The contrast agent, which blocks x-rays, causes cancer to show up as a white area on the mammogram.
- 7. Post-processing: Following the CEM, the contrast-enhanced subtraction images are produced using a weighted-logarithmic subtraction of the low-energy image from the high-energy image. Because the difference in iodine absorption between the images is larger than the difference in tissue absorption, this dual energy subtraction technique has the effect of increasing the visibility of the iodine while almost eliminating the visibility of background tissue.
- 8. Image interpretation: The resulting images are then reviewed and interpreted by a radiologist. Low-energy images, which are almost identical to standard unenhanced digital mammograms, are also used in the interpretation.

# 4.2. Summary of key findings

- The use of CEM in population-based breast cancer screening of asymptomatic women is seen to have some potential in clinical studies, and there are current clinical trials related to CEM, including with a focus on screening.
- There is currently no indication of the timeframe in which the full clinical potential of CEM for breast cancer screening will be realised; however, results from prospective clinical studies show that CEM reports higher percentages of sensitivity, specificity, positive predictive value and negative predictive value than conventional mammography.
- Current research is not sufficient to be able to identify whether CEM as a breast imaging tool is able to reduce deaths due to breast cancer through early detection.
- CEM has not been incorporated into any national breast screening programs, nor are there any national position statements that have been released on its use in breast cancer screening for asymptomatic women.

# 4.3. Literature search results (number of studies returned)

A total of 38 articles related to CEM were identified from the literature search. There were no population-based randomised controlled trials of CEM as a screening modality. Articles were excluded because they clearly indicated studying CEM in a symptomatic population, for diagnostic or treatment purposes, or were studying developing technologies that may have an application within CEM, such as developing contrast agents (eg, Karunamuni et al., 2016). Other articles were excluded because they could not be located or were dissertations. Additional articles were sourced in bibliographic searches. A total of 20 articles were then reviewed for CEM, most of which were from cohort studies.

#### Systematic and/or literature reviews

Three articles were identified that were either systematic or literature reviews (Gilbert & Selamoglu, 2017; Helvie, 2010; Lewis et al., 2017). One study gave a broad overview and discussion of the implementation of CEM in a clinical setting over five years, since 2012, and incorporated a literature review (Bhimani et al., 2017).

#### RCTs

None identified.

#### **Prospective studies**

Seven prospective cohort studies on CEM were included in this review. These included women who participated in population-based breast screening programs with indeterminate or suspicious findings, so were recalled for further examination by CEM (Chou et al., 2015; Dromain et al., 2011; ElSaid, Farouk, Shetat, Khalifa, & Nada, 2015; Jochelson et al., 2013; Kariyappa, Gnanaprakasam, Anand, Krishnaswami, & Ramachandran, 2016; Luczynska et al., 2014; Phillips et al., 2017).

#### Retrospective studies

Retrospective cohort studies involved reviewing CEM images obtained previously during screening follow-up (Houben et al., 2017; Lalji et al., 2015, 2016; Lewis et al., 2017; Patel, Gray, & Pockaj, 2017; Sogani et al., 2017; Tardivel et al., 2016; Yagil et al., 2016).

#### Grey literature

A search of the Australian New Zealand Clinical Trials Registry (ANZCTR) for "contrast enhanced mammography" yielded three results. One was related to CEM for assessing patients with dense breasts:

• 'Efficacy of contrast enhanced spectral mammography versus standard of care imaging test (DBT and ultrasound) in women with mammographically dense breast tissue recalled for investigation of abnormalities detected on routine screening mammograms'. The trial is reported to be in recruitment stages at the Royal Perth Hospital in Western Australia and involves 60 women between the ages of 18 and 65 years with dense breast tissue and non-calcified lesions detected on screening mammography<sup>2</sup>.

A search of the U.S. National Library of Medicine <u>ClinicalTrials.gov</u> website for "breast cancer" and "contrast enhanced mammography" identified 38 related clinical trials, with three studies related to CEM for breast cancer screening purposes, mostly in comparison to other imaging modalities:

- 'Comparison of whole breast screening ultrasound and contrast enhanced mammography for supplemental breast cancer screening' at the Memorial Sloan Kettering Cancer Center, New York, New York, United States. The study started in 2014, is due to be completed in 2018, and involves 800 women age 30 years and older.
- 'Comparison of contrast enhanced mammography to breast MRI in screening patients at increased risk for breast cancer' at the Memorial Sloan Kettering Cancer Center, New York, New York, United States. The study started in 2012, is due to be completed in 2018, and involves 1000 women age 21 years and older.
- 'Comparison of contrast-enhanced spectral mammography (CESM) to MRI in screening high risk women for breast cancer' at the Beth Israel Deaconess Medical Center, Boston, Massachusetts, United States. The study started in 2014, is due to be completed in 2019, and involves 220 women age 30 years and older.

# 4.4. Study findings and discussion

#### What stage of development or trial is this innovation at?

CEM as a modality for population-based breast cancer screening in clinical settings has not been realised yet, though there is evidence to suggest that it may be useful as an adjunct to digital mammography for screening purposes.

<sup>&</sup>lt;sup>2</sup> Available at:

https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=369664&isReview=true

An early review of literature from 2010 suggested that the application of dual-energy CEM was clinically feasible, though the number of patients studied at that time was limited so clinical utility was yet to be established (Helvie, 2010). Bi-lateral dual-energy CEM was determined feasible in one very small study of 10 patients that ranged in age from 28 to 64 years (mean age of 51.6 years) who had been newly diagnosed with breast cancer (Jochelson et al., 2013). No adverse events were observed in the patients in this study, and the quality of the resulting CEM images were deemed satisfactory (Jochelson et al., 2013).

Chou (2015) indicated that the prototype CEM technology they employed was at an early clinical stage, and noted that their mammography unit and imaging processing software underwent minor improvements during their study. They suggested that more technological advances and upgrades were needed for clinical application (Chou et al., 2015). However, Lalji et al. (2016) suggested that the introduction of CEM into clinical practice was safe and feasible, and noted it had been adopted for clinical use in some settings, but as a follow-up to screening, an assessment tool, or as diagnostic workup for breast cancer (eg, Bhimani et al., 2017; Lewis et al., 2017).

Sensitivity of CEM was reported to be in the range of 90 to 100% (ElSaid et al., 2015; Helvie, 2010; Lalji et al., 2016; Lewis et al., 2017; Luczynska et al., 2014; McGuire et al., 2017; Tardivel et al., 2016) while specificity was reported to be between 41% and 74% (ElSaid et al., 2015; Lalji et al., 2016; Luczynska et al., 2014; Tardivel et al., 2016). Furthermore, Tardivel and colleagues reported positive predictive values (PPV) for CEM as 92% and negative predictive values (NPV) for CEM as 81% (Tardivel et al., 2016). Luczynska and colleagues reported PPV values at 77% and NPV values at 100%, compared to 68% and 47% for mammography respectively (Luczynska et al., 2014).

Table 1 reports sensitivity, specificity, PPV and NPV increases between digital mammography and CESM for a recent study of 199 patients recalled from a Dutch breast screening program (Lalji et al., 2016):

Technique	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)
CESM	96.9	69.7	58.2	98.2
FFDM	93	35.9	38.7	92.6

CEM Table 1. Sensitivity, specificity, PPV and NPV for CEM compared with FFDM to detect breast cancer (Lalji et al., 2016)

CEM as an adjunct imaging modality with standard mammography was found to have higher sensitivity (93%) than mammography alone (78%) with no loss in specificity, and actual lesion size closer to histological size with CEM (Dromain et al., 2011).

Low energy CEM images have been assessed as not inferior to full field digital mammography images (Lalji et al., 2015). CEM Table 2 (overleaf) outlines sensitivity and specificity findings from Lewis's (2017) review compared to full-field digital mammography:

Study	No. of lesions	CEM Sensitivity (%)	CEM Specificity (%)	FFDM Sensitivity (%)	FFDM Specificity (%)
Fallenberg et al. (2014)	80	100	-	97	-
Badr et al. (2014)	37	95	85	-	-
Luczynska et al. (2014)	173	100	41	91	15
Jochelson et al. (2013)	52	96	-	81	-
Dromain et al. (2011)	142	93	63	78	58

CEM Table 2. Summary of sensitivity and specificity of CEM compared with FFDM to detect breast cancer (Lewis et al., 2017)

Examples of research to further refine CEM as a breast imaging tool included evaluating Background Parenchymal Enhancement (BPE) as a risk factor for breast cancer using CEM (Sogani et al., 2017), and a quantitative study of photon-counting detectors to support an accurate diagnosis of breast cancer, that assessed the feasibility of CEM to characterise neoangiogenically induced vascular changes in suspicious lesions (Ding & Molloi, 2017). Early research has demonstrated CAD CEM as a feasible complementary tool to potentially reduce false-positive findings during image reading (Patel et al., 2018).

CEM does not appear to have been adopted into clinical use as a population-based screening strategy for breast cancer, though a number of studies report on it's potential, and the need for further large-scale studies to determine whether CEM's applicability as a screening modality is viable (Jochelson et al., 2013; Kariyappa et al., 2016; Lalji et al., 2016; Lewis et al., 2017). Several clinical trials are in the pipeline for CEM, including related to screening, as reported above.

# What is its considered potential clinical value in five years? In 10 years?

The potential clinical value of CEM in five to 10 years is not explicitly discussed in the literature, though, as shown above, findings have indicated CEM's value in terms of higher sensitivity and specificity in reading/interpretation compared with digital mammography.

CEM has been described as a useful adjunct modality for patients recalled from screening mammography (Bhimani et al., 2017; Lalji et al., 2016). It may also improve clinical performance of mammography (Luczynska et al., 2014), and is seen to be crucial for accurate assessment (Lewis et al., 2017). CEM can also provide added value in cancer staging and assisting with unclear findings on conventional imaging (Tardivel et al., 2016). CEM is better at identifying lesions than mammography (Houben et al., 2017; Jochelson et al., 2013; Kariyappa et al., 2016). In mammographically visible lesions, CEM characterised the lesion, affirmed the finding and better demonstrated response to treatment (Kariyappa et al., 2016). Furthermore, CEM provides more accurate prediction of cancer size than digital mammography and is a useful adjunct to preoperative planning (McGuire et al., 2017).

While diagnosis is not the focus of this review, CEM has been assessed as comparable in diagnostic accuracy to contrast-enhanced MRI in patients with suspicious breast lesions (Chou et al., 2015).

#### What cost and safety findings have been reported?

In terms of cost, no studies were able to provide clear findings around the true cost of CEM compared to mammography, or other breast imaging modalities. A number of articles inferred that CEM had lower, or potentially lower costs than MRI (Chou et al., 2015; ElSaid et al., 2015; Helvie, 2010; Lewis et al., 2017) and molecular breast imaging (Lewis et al., 2017). CEM was also noted to be a faster procedure than MRI (Chou et al., 2015; ElSaid et al., 2017; Tardivel et al., 2016) with similar results as MRI (Chou et al., 2015; Jochelson et al., 2013).

With regards to safety, the main risks noted with CEM were potential severe allergy with the administered contrast solution (Bhimani et al., 2017; Helvie, 2010; Houben et al., 2017; Lewis et al., 2017), abnormalities in renal function (Bhimani et al., 2017; Lewis et al., 2017), and higher radiation exposure, especially if used as an adjunct with other radiation based imaging procedures (Houben et al., 2017), and/or for pregnant women (Chou et al., 2015; Lewis et al., 2017). CESM was described as having a higher radiation dose than standard mammography – 54% greater than the standard 2.65mGy (Gilbert & Selamoglu, 2017).

These risks were potential barriers for CEM to be implemented as a routine population-based screening procedure.

#### Does this innovation show high sensitivity and specificity for women with dense breasts and women who have had breast surgery/augmentation compared to digital mammography?

At this stage, CEM shows some promise as a breast cancer screening tool for women with dense breasts, including improved sensitivity (Gilbert & Selamoglu, 2017).

One study demonstrated CEM's usefulness in identifying lesions in mammographically dense oedematous breasts with sensitivity of 95%, specificity of 73%, positive predictive value (PPV) of 88% and negative predictive value (NPV) of 88% (ElSaid et al., 2015). Four articles discussed the potential for CEM to assist with further evaluation of patients with dense breast tissue due to CEM, reporting similar sensitivity with dense breasts to MRI (Bhimani et al., 2017; Dromain et al., 2011; Helvie, 2010; Jochelson et al., 2013), greater than digital mammography.

Sensitivity and specificity of CEM for women who have had breast surgery or augmentation was not discussed in the literature reviewed.

# Is there evidence that this innovation is more acceptable to women (in general and by ethnic group) compared to digital mammography?

With regards to acceptability, some studies inferred that CEM was more tolerated (Bhimani et al., 2017; Jochelson et al., 2013) or well accepted by patients because they were able to have a complete assessment without remaining questionable findings on the same day (ElSaid et al., 2015) though these inferences were not discussed explicitly in relation to digital mammography. CEM was also discussed as more appealing to patients who could not tolerate MRI (Helvie, 2010; Tardivel et al., 2016) or MBI (Lewis et al., 2017). There was no evidence that CEM was more acceptable to women compared to digital mammography.

However, one article reported on a survey of patient acceptability of CEM as part of a prospective study assessing performance characteristics of CEM related to breast MRI in high risk patients (Phillips et al., 2017). No ethnicity information was reported for these patients, but they were 30 years old or older, at high-risk for breast cancer<sup>3</sup>, and receiving screening breast MRI as part of their standard of care. Out of the 38 patients who had undergone CEM that completed a survey, 79% said that they preferred it over MRI, if the exams had equal sensitivity (Phillips et al., 2017). Furthermore, 89% of those patients said that they would be comfortable with receiving contrast as part of an annual screening test (Phillips et al., 2017). However, more research into patient preferences (Houben et al., 2017), particularly in comparison to mammography, was required.

#### Does this technology reduce deaths due to breast cancer through early detection?

There is no evidence to suggest that contrast enhanced mammography reduces deaths due to breast cancer through early detection, and further large studies are warranted to determine this.

# Has this innovation been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

There is no evidence to suggest that contrast enhanced mammography has been implemented into a national screening program.

# Has a national position statement been published about this innovation, and if so, what is the position? Is there consensus in position statements?

One position paper was located for contrast enhanced mammography for the American College of Radiology Society of Breast Imaging. Their position is:

"Based on its ability to image neovascularity in a fashion similar to MRI, CEDM is a promising technique for depicting cancers that are not visible on standard unenhanced mammography. It is approved for clinical use and is performed on commercial systems. Results of clinical studies show it to be significantly more sensitive and specific than mammography alone and to have sensitivity and specificity comparable to contrast-enhanced breast MRI. Current and proposed uses include additional evaluation of symptomatic patients or patients with abnormal screening examinations, assessing local extent of newly diagnosed breast cancers, problem solving, monitoring of neoadjuvant chemotherapy and high-risk screening"<sup>4</sup>

There are no national position statements published about CEM as a screening test for the early detection of breast cancer in asymptomatic women.

<sup>&</sup>lt;sup>3</sup> High-risk patients were defined as patients with one of the following characteristics:  $\geq 20\%$  lifetime risk of breast cancer, BRCA mutation or other hereditary germ line mutation carrier, history of chest wall radiation, history of LCIS, history of breast cancer diagnosed at  $\leq 40$  years of age, history of breast cancer with 1st or 2nd degree relative with breast cancer and either patient or relative diagnosed at  $\leq 50$  years of age, history of breast cancer with mammographically occult lesions, or history of breast cancer for which a medical oncologist feels that breast MRI screening is important (Phillips et al., 2017). <sup>4</sup> Available at <u>https://www.sbi-</u>

online.org/RESOURCES/WhitePapers/TabId/595/ArtMID/1617/ArticleID/601/Contrast-Enhanced-Digital-Mammography.aspx

# 5. DIGITAL BREAST TOMOSYNTHESIS

#### Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

DBT was not discussed in the 2009 Bulletin on new and emerging technologies for breast cancer detection.

# 5.1. What is digital breast tomosynthesis?

DBT (also known as breast tomosynthesis, mammographic tomosynthesis or three dimensional/3D mammography) is an imaging technology that can be used to detect breast cancer and assess suspicious areas. DBT records between 11 and 25 low-dose images of a compressed breast depending on the imaging system used<sup>5</sup>. These images are reconstructed in 1mm (or more) parallel slices to form a three-dimensional image of the breast. Radiologists (or other readers) then analyse these images to determine the presence of suspected abnormalities or to further investigate an area identified as suspicious on a mammogram. The thin cross-sectional images created by DBT minimise the masking effects of breast tissue overlap, which can improve margin visibility for soft tissue tumours and increase lesion conspicuity. This potentially increases screening sensitivity and specificity (especially for women with more dense/more non-fatty breasts) as abnormalities are easier to see.

Radiation dose varies depending on whether DBT is used alone, with integrated synthesised two-dimensional mammogram (s2DM<sup>6</sup>) image acquisition or is used as an adjunct to FFDM. We know that FFDM + DBT requires a higher radiation dose than FFDM alone to acquire images during a screening examination. Concerns about radiation dose plus the longer image acquisition and interpretation time required with FFDM + DBT means that this screening strategy could be potentially unacceptable to women and practitioners on the grounds of radiation dose. DBT + s2DM developed in response to this concern. DBT's use (both in clinical and research settings) is evolving as is the evidence base underpinning its potential use as a screening tool grows.

<sup>&</sup>lt;sup>5</sup> Hologic's Dimensions system takes 15 projections taken over approximately 4 seconds. Other CE mark or FDAapproved systems use 9 or 25 projections taken over 3 to 25 seconds (Sechopoulos, 2013).

<sup>&</sup>lt;sup>6</sup> s2DM is a two-dimensional mammogram that is generated from a DBT source data. These reconstructed images are similar to those captured in the mediolateral (MLO) and craniocaudal (CC) views used in a standard FFDM screening examination (Freer et al., 2017).

# 5.2. Summary of key findings

- The Department of Health commissioned Allen + Clarke to undertake a literature review on the use of DBT in screening. The results of this literature review are included in this report for completeness. Further detail on the methodology is contained in the primary literature review document.
- There is strong evidence that both FFDM + DBT and DBT + s2DM provide superior performance for improved cancer detection and that DBT may be a more sensitive test than FFDM alone. The magnitude of improvement may be affected by reading strategy (eg, double reading approaches result in higher detection rates compared to single reading).
- There is emerging evidence that, used as an adjunct screen to FFDM, DBT can reduce recall rates and false positives results compared to FFDM alone; however, some inconsistent results between large prospective trials are reported (which could reflect the already low rates of recall seen in some population-based screening programs in which the trials are embedded). Further research investigating the comparative performance will help to unpick areas of uncertainty including the impact of double/single reading strategies and the impact of access to previous DBT images.
- Further research may also help determine which combination of approaches (FFDM + DBT, two-view DBT + s2DM, DBTMLO + FFDM, or some form of DBT alone) achieves the best balance between radiation dose, sensitivity and specificity.
- While DBT improves breast cancer detection with (potentially) lower rates of recall than FFDM alone, there is insufficient evidence about the long-term mortality benefit to support the use of DBT alone as a primary screening test. Few studies identified for this literature review investigated DBT alone compared to FFDM alone, although literature exploring different ways to integrate DBT into screening continues to develop quickly.
- Current research is not sufficient to be able to identify whether DBT as a breast imaging tool is able to reduce deaths due to breast cancer through early detection, but it does increase cancer detection.
- DBT has not yet been incorporated into any national screening programs; however, fully-paired trials embedded in population-based screening programs have been completed or are underway. In addition, DBT is available for private breast screening in several jurisdictions including Australia and New Zealand.

# 5.3. Literature search results (number of studies returned)

Information presented in Chapter 5.2 is based on a literature review on DBT which was also completed by *Allen + Clarke* in 2018. A summary of the returned literature is provided here. Further information on the methodology can be found in *Allen + Clarke*'s literature review.

*Allen + Clarke*'s review found 85 relevant articles including two systematics reviews, 12 narrative literature reviews and 42 studies. Primary studies already incorporated into high-quality systematic or literature reviews were not assessed unless additional material not described in the systematic or literature review was included.

#### Systematic and/or literature reviews

Two systematic reviews were identified: Coop et al. (2016) and Hodgson et al. (2016). We also referred to the following narrative literature reviews: Houssami (2017); Poplack (2017); Skaane (2017); Gilbert et al. (2016); Houssami (2015); Svahn et al. (2015A); Vedantham et al. (2015); and Houssami and Skaane (2013).

# RCTs

We identified one RCT: Maxwell et al. (2017). We identified prospective fully-paired trials:

- Oslo Tomosynthesis in Screening Trial (Skaane et al., 2014; Skaane et al., 2013B)
- STORM trial (Bernardi & Houssami, 2017; Bernardi et al., 2016; Bernardi et al., 2014; Caumo et al., 2014; Ciatto et al., 2013; Bernardi et al., 2012), and
- Malmö trial (Lång et al., 2016A, 2016B)

# Other studies

There were 56 other studies including: Aujero et al. (2017); Freer et al. (2017); Hunter et al. (2017); Liberatore et al. (2017); Miller et al. (2017); Pan et al. (2017); Powell et al. (2017); Rafferty et al. (2017); Rodriguez-Ruiz et al. (2017); Carbonaro et al. (2016); Conant et al. (2016); Kalra et al. (2016); McDonald et al. (2016); Sharpe et al. (2016); Wang et al. (2016); Zuckerman et al. (2016); Abdullah Suhaimi et al. (2015); Bonafede et al. (2015); Durand et al. (2015); Lourenco et al. (2015); McDonald et al. (2015); Shin et al. (2015); Sumkin et al. (2015); Dang et al. (2014); Destounis et al. (2014); Friedewald et al. (2014); Greenburg et al. (2014); Lee et al. (2014); McCarthy et al. (2014); Rose et al. (2014); Zuley et al. (2014); Haas et al. (2013); Sechopoulos (2013); Zuley et al. (2013); Gur et al., (2012), Olgar et al. (2012), Wallis et al. (2012); Svane et al. (2011); Gennaro et al. (2010); Svahn et al. (2010)

# Practical or technical evaluations and modelled analysis

Four papers practical or technical evaluations were used: Bonsall et al. (2016); Strudley et al. (2015); Mungutroy et al. (2014) and Strudley et al. (2014).

# 5.4. Summary of Allen + Clarke's literature review on DBT

#### What stage of development or trial is this innovation at?

DBT (using Hologic's Dimensions system) was first approved for use in breast cancer screening by the FDA in 2011. Since then, other systems capable of performing DBT have also been approved (including Siemens Mammomat and GE's SenoClaire). DBT is not widely used as a primary screening tool within any national breast-screening program although some European jurisdictions including France and Monaco include DBT as a screening option (Liberatore et al., 2017). DBT for screening is available to women in a range of private settings across different jurisdictions including in Australia and New Zealand.

#### What is its considered potential clinical value in five years? In 10 years?

#### **Cancer detection**

There is strong evidence that cancer detection rates (CDR) increase when using DBT compared to FFDM alone. Increases were reported in a range of studies (including large prospective trials) for different combinations of screening strategy including FFDM + DBT, DBT + s2DM, and DBT<sub>ML0</sub> compared to DM<sub>CC</sub> or FFDM alone. The direction of effect is consistent across study design, setting and location. There is variance in magnitude of effect.

Pooled analysis based on data from the STORM and OTS trials (Coop et al., 2016; Hodgson et al., 2016) reported a statistically significant increase of 2.43 cancers detected per 1000 screening examinations compared to FFDM alone. Adjunct screening with DBT also increases invasive cancer detection compared to FFDM alone. The incremental CDR and invasive CDR are similar for FFDM + DBT and DBT + s2DM: either detected significantly more cancers than FFDM alone. The Malmö trial also reported a significant increase in CDR: 2.6 cancers detected per 1000 screening examinations using  $DBT_{MLO}$  compared to FFDM alone (Lång et al., 2016A).

DBT + s2DM also performed better, detecting 8.8 cancers per 1000 screening examinations compared to 6.3 with FFDM alone (Houssami, 2017). The evidence of comparative CDR performance for DBT + s2DM compared to FFDM + DBT was inconsistent but the CDR is similar in both screening strategies (in both smaller and larger studies). DBT + s2DM is a promising screening strategy, especially as it significantly reduces radiation dose.

Data from retrospective studies showed a similar effect (that is, increases in CDR when DBT is used) but the increases were smaller than those reported from the prospective trials. Statistically significant CDR results from retrospective studies ranged from 1.6 to 1.9 cancers detected per 1000 screening examinations. Reasons for the lower CDR in the retrospective studies could relate to differences in reading strategy (eg, double reading compared to a single reader approach), participant selection, under-powering or other study design limitations.

#### Types of cancer detected

Earlier studies showed that FFDM + DBT's performance did not appear to be superior for the detection of DCIS because of reduced visibility of microcalcifications (Houssami et al., 2016). Later studies are reporting no differences in the types of cancers detected by either FFDM + DBT or FFDM alone (Lång et al., 2016A). Further research is needed to determine DBT's ability to detect microcalcifications.

There was very limited data about the long-term mortality benefits, treatment morbidity or quality of life improvements associated with FFDM + DBT as a screening strategy. Almost no data exists on results for incident screening compared to prevalent screening, mortality benefit or surrogate indictors of this. Reliable data on interval cancer rate is also scarce.

#### Recall rates and false positive recall rates

The literature is not settled about the association between DBT and recall rates.

Some results show that overall recall rates can be reduced when using FFDM + DBT compared with FFDM alone (often, smaller retrospective studies in screening programs offering annual screening and with higher program recall rates). Other prospective trials reported increased recall with double reading (either by two radiologists or through an arbitration process) but

reduced false positive rate. This may reflect that the overall false positive recall rates within screening programs where prospective trials are embedded are generally low anyway.

There is less literature exploring associations between DBT + s2DM and the rate of false positive recalls although research generally favours a reduction in false positives with DBT + s2DM compared with both FFDM + DBT and FFDM alone. Like overall recall rates, there is some variance in the direction of effect. Results from a large prospective trial (STORM 2) showed a false positive recall rate for DBT + s2DM that was significantly greater for FFDM + DBT and FFDM alone. It is possible that the results from the STORM-2 trial relate to early experiences of incorporating s2DM into real-world screening practice for the first time without previous experience with s2DM images relative to FFDM. Secondary analysis from the STORM 2 trial indicated that false positive recall rates for FFDM + DBT and DBT + s2DM significantly reduced compared to those for FFDM. These results reflect developing and increasing knowledge in the use of FFDM + DBT, with some interpretation issues still present for s2DM. Interim results from the false positive recall rate of screens using DBT over the first 1.5 years, which also indicates that false positive recall could be associated with a learning curve in interpretation.

Information from the smaller retrospective studies (most of which used single reading strategies) reported that recall rate was reduced with the addition of DBT to FFDM. Other trial data reported both reduced recall rate and reduced false positives with the addition or use of DBT. Differences in overall program false positive recall rates, reading strategy and arbitration protocols used to determine which women to recall from screening may account for some of the inconsistency. Increasing reader experience, knowledge of DBT and interpreting 3D images and availability of prior DBT images may also further decrease recall rates.

#### PPV

Overall results on PPV<sub>1</sub> indicated that FFDM + DBT accurately detected proportionally more women recalled from screening who had breast cancer compared to FFDM alone. DBT + s2DM also showed promise of increased accuracy. Screening based on DBT + s2DM screening may correctly identify between one and three more women with diagnosable breast cancer for every 100 women recalled, compared with recalls based on FFDM + DBT screening. Results on PPVs for biopsy recommended and biopsy performed indicate that FFDM + DBT was also more accurate the FFDM alone when used as a basis for recommending or performing biopsies. PPV<sub>2-3</sub> results for DBT + s2DM are also promising but present more varied effect size than results for FFDM + DBT.

#### What cost and safety findings have been reported?

Radiation dose varies with the image acquisition process used (DBT or FFDM or combination mode), the number of and type of views, the use of automatic exposure control, positioning, breast size and composition, and by DBT system used.

Much of the published evidence about the sensitivity and specificity of DBT is based on dual acquisition protocols (i.e., FFDM + DBT). Using average breast thickness, the radiation dose required to acquire acceptable images with FFDM + DBT is approximately double that of FFDM alone (2.98mGy compared to 1.49mGy). This 'double dose' is still within the dose limits set for overseas quality and safety standards but is higher than the per view dose limit set for the BSA program. The radiation dose for DBT compared with FFDM is lower:  $DBT_{MLO}$  has about 70% of

the mean glandular dose (MGD) compared to FFDM alone. Two-view DBT results in a similar MGD compared with FFDM. Other possible single view combinations also result in lower MGD.

Synthesised acquisition of 2D images halves the effective dose of combined FFDM + DBT, making it comparable to FFDM alone but with the improved detection rates associated with DBT. Initial studies indicate that the quality of images reconstructed from s2DM is acceptable, but further evidence is required to ensure that they can be used to accurately interpret microcalcifications.

Having FFDM + DBT as the preferred screening strategy will have implications for accumulative dose if separate acquisitions are used for 2D and 3D images, if the screening interval is annual not biennial, or if women participate in mammography-based screening from their early 40s.

The literature is dominated by studies that used modelled analyses to discuss the effect of DBT implementation on insurance programs in the United States. Modelled analyses showed that FFDM + DBT demonstrated economic favourability when considering clinical benefits like cancer detection and recall rates. Unfortunately, due to differences in health sector policy and service delivery, the modelled analyses may have limited applicability to the BSA program. No modelled analyses focused on the implementation of DBT have been conducted elsewhere.

Implementation of DBT would require capital upgrade costs (eg, new equipment requirements), increased capacity for data storage or transmission, training and additional time for radiologists to read DBT images. Incremental costs could be offset by health system savings associated with increased cancer detection and reduced rates of recall (eg, the costs associated with recalling women unnecessarily, additional unnecessary biopsies and further assessment in cases where breast cancer is not present). At the time of this review, no detailed cost analysis had been reported. Cost is still something that needs to be balanced against potential benefits.

#### Does this innovation show high sensitivity and specificity for women with dense breasts and women who have had breast surgery/augmentation compared to digital mammography?

Findings for CDR stratified by breast density present results that may be surprising, given that DBT improves conspicuity and should, in theory, provide more quality images of more dense breasts. While results for FFDM + DBT show increased CDR for all women, data from prospective trials does not demonstrate a significant increase in CDR when comparing women with more dense breasts to those with less dense breasts. The use of breast density classification systems can result in unreliable allocation because density classification can be affected by factors like hormone levels, genetic factors, parity, use of oestrogen, place in menstrual cycle, use of tamoxifen, weight and (importantly) inter/intra reader variability. It is possible for women to be classified as having non-dense breasts (BIRADS 2) in one mammogram but be reclassified to having more dense breasts to those that could be determined to be more fatty but still have significant areas of mammographic density. This could account for the smaller-than-expected increase in CDR between women with more dense or less dense breasts. Research which reports CDR, recall rates and false positives by "extremely dense" (BIRADS 4) and "almost entirely fatty" (BIRADS 1) could result in clearer (and possibly truer) results on CDR differences.

The age at which screening participation begins and screening interval may also influence sensitivity, specificity and overall lifetime radiation dose received. Studies reporting age stratification used different age bands (i.e., 10-year bands or groups like over 60 years/under 60

years). This impacts on our ability to draw useful conclusions about the possible relationships between age, and clinical outcomes or performance metrics associated with the use of DBT.

# Is there evidence that this innovation is more acceptable to women (in general and by ethnic group) compared to digital mammography?

There was little evidence describing the acceptability to women or health practitioners on the use of DBT as a screening tool. To consider dimensions of acceptability, we sought evidence exploring women's choice, compression time, and mental health outcomes.

No data was presented on women's overall choice to have either FFDM alone or DBT. However, there is anecdotal commentary that women may choose either FFDM + DBT or FFDM when given a choice. For example, 88% of study participants in Rose et al. (2013) consented to have FFDM + DBT (other study participants chose to have FFDM). When enrolling in Freer et al.'s study, women who are more informed about DBT either chose to have DBT because they were aware of its cancer detection benefits or chose not to have it because they were aware of the increased radiation dose associated with dual image acquisition (Freer et al., 2017). While not robust indicators of acceptability to women, these examples demonstrate that, if well-explained or if women are well-informed, they may choose DBT over FFDM.

We infer that women may appreciate the lower compression that can be used to acquire acceptable images with  $DBT_{MLO}$  compared to FFDM, but the literature is not settled on the best balance between reduced compression, image quality and women's preference. Some early studies (cited in Coop et al., 2016) indicated that reducing compression from 4cm to 6cm did not adversely affect image quality. Sechopoulos (2013) reported that women preferred reduced compression (citing Fornvik et al., 2010), but the three participating radiologists did not as image quality was poorer in this study. Data from the STORM trial found that compression time is slightly longer for FFDM + DBT compared to DBT alone: 4m 3s (range = 3m 53s to 4m 18s) compared to  $3m \ 13s$  (range =  $3m \ 0 \ s$  to  $3m \ 26s$ ; *p*=.01) (Bernardi et al., 2012). This may increase overall pain/discomfort associated with this test but no information about women's views on compression or reader feedback about image quality were presented. In the Malmö trial, Lång et al. (2016A) performed the DBT with reduced compression compared to FFDM to determine if reduced compression would compromise acceptable image acquisition and cancer detection. They reported that reduced compression of up to 50% was achieved in 90% of cases (with larger breast requiring more compression). Lång et al. reported women's positive feedback about the reduced compression but did not collect specific data on this outcome.

Two practical evaluations for the National Health Service (for the Hologic and GE DBT systems) reported that most radiologists considered compression times to be acceptable (although for the Hologic system 4/10 participating radiologists rated DBT compression time to be "worse" than FFDM alone with the remaining six noting that it was the same as FFDM: the dimensions of "worse" are not explained) (Bonsall et al., 2016; Mungutroy et al., 2014). Women's comfort was rated average to excellent but, for GE SenoClaire, radiologists reported that they had received no feedback from women to indicate that the system was more uncomfortable that FFDM alone. None of the assessments of women's comfort appear to be validated by women themselves.

Abdullah Suhaimi et al. (2015) reported on a study of 130 Malaysian women's anxiety during participation in a FFDM + DBT screening examination. Using a validated questionnaire (State-Trait Anxiety Inventory' Form Y-1), two study radiologists reported a reduction in women's pain and anxiety with reduced compression (38.5 newtons compared to 93.0 newtons for standard

compression). They found that the mean anxiety score decreased with reduced compression (from 57.15 to 47.23; p<.001). The mean pain during procedure score reduced from 2.13 to 0.69 (p<.001). The authors noted that image quality (as reported by the two participating radiologists) was not compromised. In addition, women's anxiety at participating in a screening program or receiving screening results may be reduced if CDR is improved (with DBT) and if false positive results are reduced and unnecessary recalls are avoided.

#### Does this technology reduce deaths due to breast cancer through early detection?

Because long-term, adequately powered studies comparing DBT (alone or as an adjunct screening test) to FFDM are limited, there is limited data about DBT's impact on interval cancer rate, mortality benefits and improvement in treatment morbidity compared to FFDM. Proxy measures (such tumour size at detection) are available but this does not provide a sense of the long-term mortality reduction benefits which DBT may offer compared to FFDM alone. Our review of the <u>U.S. National Library of Medicine clinical trials database</u> identified large active or recruiting studies investigating the role of DBT in population-based screening for asymptomatic women. Ongoing or upcoming trials include the:

- Malmö Breast Tomosynthesis Screening Trial (a single site study in Sweden)<sup>7</sup>
- Tomosynthesis Trial in Bergen (a single site study in Norway)<sup>8</sup>
- Tomosynthesis Mammographic Imaging Screening Trial (a multi-armed, multicentre study in Canada)<sup>9</sup>
- PROSPECTS trial in the United Kingdom<sup>10</sup>
- TOSYMA study in Germany<sup>11</sup>, and
- large Italian studies<sup>12</sup>.

Future research may provide more information that can be used to better assess mortality benefit.

Has this innovation been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

DBT is not used as a primary test for average-risk women in the BreastScreen Australia (BSA) program; however, access to DBT imaging for screening purposes is offered through some private radiology clinic settings. Outside the BSA program, DBT is increasingly used for the assessment of both screen-detected abnormalities and symptomatic breast cancers.

<sup>&</sup>lt;sup>7</sup> Study outcomes include interval cancer rate.

<sup>&</sup>lt;sup>8</sup> Study outcomes focus on cancer detection, interval cancer rates, PPV, recall rates, prognostic and predictive tumour characteristics, radiation dose, interpretation time, and cost-effectiveness.

<sup>&</sup>lt;sup>9</sup> Study outcomes include cancer detection, recall rates, interval cancers, prevalence of breast cancer subtypes, clinical characteristics of cancers, radiation dose, observer performance studies, BIRADS imaging features, breast cancer mortality, quality monitoring outcomes, PPV, health care utilisation, false positives/true negatives, biopsy rates and biomarker correlation.

 $<sup>^{10}</sup>$  Study outcomes include the cost-effectiveness of breast cancer screening using FFDM + DBT compared to DBT + s2DM.

<sup>&</sup>lt;sup>11</sup> Study outcomes include cancer detection rates including by cancer type and category, interval cancer rate, recall rate and PPV.

<sup>&</sup>lt;sup>12</sup> Study outcomes include interval cancer, recall, PPV, biopsy rates, cancer detection and self-reported pain/discomfort during mammography.

Has a national position statement been published about this innovation, and if so, what is the position? Is there consensus in position statements?

In 2014, the Community Care and Population Health Principal Committee of the Australian Health Ministers' Advisory Council endorsed BSA's position statement on DBT. This position statement was based on a literature review completed in 2009 (Department of Health and Ageing, 2009) and other papers published between 2009-13. The BSA position statement on DBT says that it:

"has the potential to decrease the number of women who are recalled for further tests (reduce recall rates) and possibly increase the detection of breast cancer (improve sensitivity)."; however, the balance between "relative harms and benefits to well women of radiation dose, and the cost, efficiency and effectiveness of using this technology are as yet unclear".

The Standing Committee on Screening concluded that FFDM remained the most effective population screening technology for breast cancer.

Seven countries and one region (including Brazil, Europe, France, Italy, Japan, New Zealand, the United Kingdom and the USA) also have current position papers describing views on the role of DBT (either as a stand-alone or adjunct screening tool) in breast cancer screening. The International Agency on Research into Cancer (IARC) has released a position statement which authors from 16 countries<sup>13</sup> contributed to. Except for Brazil, all the position statements report the same conclusion: existing evidence favours FFDM + DBT compared to FFDM alone for key screening outcomes like CDR and recall rates. DBT is a promising technology that will have some role in the future of screening programs; however, concern remains around the increased radiation dose associated with dual acquisition (which may be addressed by s2DM), the lack of evidence on long-term performance clinical outcomes like interval cancer rates and the impact on longer-term cancer mortality reduction. Currently, all jurisdictions (except Brazil) recommend that further evidence from prospective trials and RCTs be acquired and used to inform decisions about integration into national screening programs.

<sup>&</sup>lt;sup>13</sup> Contributing authors to the Working Group represented Australia, Canada, Chile, Finland, France, Italy, Italy, The Netherlands, New Zealand, Nicaragua, and the USA.

# 6. DUCTOSCOPY

#### Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

The Bulletin advised that "it is unlikely that ductoscopy would ever be used as a routine tool for the diagnosis of breast cancer in asymptomatic women as it is an invasive and labour-intensive procedure" (National Horizon Scanning Unit, 2009, p.52).

# 6.1. What is ductoscopy?

Ductoscopy, commonly referred to as mammary ductoscopy, is a medical diagnostic procedure for viewing and collecting epithelial cells and other internal features of the milk ducts of the breast. While it may not be applicable in routine population-based screening programs, it has been described as an emerging diagnostic procedure for the following indications – nipple discharge, those at high-risk of developing breast cancer, or those with breast cancer undergoing lumpectomy (National Horizon Scanning Unit, 2009).

Ductoscopy is a procedure rationalised by an assumption that the majority of benign and malignant lesions of the breast develop from the epithelium of the terminal duct-lobular unit (Tang, Twelves, Isacke, & Gui, 2011). The advantage of mammary ductoscopy is that it enables direct access to the ductal system via the nipple, where direct visualisation of the duct epithelium allows the clinician/operator to precisely locate any present intraductal lesions (Tang et al., 2011). It also means that ductal anatomy can be mapped in relation to any lesions and can therefore act as a guide for a surgeon, should a lesion need to be removed (Tang et al., 2011).

Ductoscopy can involve flexible or rigid scopes, with diameters ranging from 0.7 to 1.2mm which magnify tissues up to 60 times their normal size to produce high-quality images (Tang et al., 2011). The working channel in the microendoscope allows accessories such as hooks and rhomboids to be passed through the scope for localization and to gather specimens (Tang et al., 2011).

The general procedure of ductoscopy as a non-surgical (i.e., diagnostic) procedure includes the application of a local anaesthetic by topical gel to the nipple, or by periareolar infiltration or infusion down the cannulated nipple duct which facilitates relaxation of the major duct muscle sphincters (Tang et al., 2011). Clinicians/operators of the microendoscope should have a good understanding of the breast's three dimensional anatomy to prevent loss of orientation and ensure that all possible ducts are explored, and not re-explored (Tang et al., 2011). Markers (for example, colour dyes, marking wires, clips), transillumination, and lacrimal probes can be used, or sheaths from the ductoscope (when disposable scopes are used) can be sutured into place, to allow pathologists to precisely identify lesions and undertake accurate assessment (Tang et al., 2011).

Benefits of the intraductal screening approach are discussed in terms of the procedure being a unique opportunity to sample the breast duct fluid in direct contact with epithelial cells that potentially undergo early malignant change, with the opportunity for direct visualisation and biopsy where necessary (Tang et al., 2011). Furthermore, detection of breast cancer may be possible in ductoscopy several years before the lesion is clinically palpable or visible mammographically – offering a promising medium for biomarker studies (Tang et al., 2011).

There are a number of limitations to mammary ductoscopy, including pain, inflammation and infection (though these are uncommon) (Tang et al., 2011). Ductoscopy can fail as a result of lumen occlusion from scarring and sclerosis, and access to peripheral lesions may be limited by the scope length, which is said to be around 6cm – one study reported that 37.3% of patients undergoing ductoscopy had lesions more than 5cm from the nipple (Tang et al., 2011). Perforation of breast ducts by microendoscope was also discussed as a limitation (Tang et al., 2011).

# 6.2. Summary of key findings

- The role of ductoscopy in breast cancer screening of asymptomatic populations is not clear as most of the research focuses on its role in the assessment and diagnosis of DCIS. For example, recent research focuses on the use of ductoscopy with symptomatic patients that present with nipple discharge.
- There is no indication that ductoscopy will achieve clinical potential as a screening tool. There is little evidence to suggest ductoscopy is acceptable to women as a screening modality.
- Current research is not sufficient to be able to identify whether ductoscopy as a screening procedure can reduce deaths due to breast cancer through early detection.
- Ductoscopy has not been incorporated into any national breast screening programs, nor are there any national position statements that have been released on their use in breast cancer screening for asymptomatic women.

# 6.3. Literature search results (number of studies returned)

Five articles were returned from the literature search that related to ductoscopy. None of the articles returned related to asymptomatic patients, therefore, we will only give a brief overview of the technology and its current application as an assessment and diagnostic tool.

#### Systematic and/or literature reviews

Two articles were systematic reviews. One was included in this review (Tang et al., 2011), another could not be located.

RCTs

None identified.

#### **Prospective studies**

Three articles were prospective studies of symptomatic patients (eg, with nipple discharge). One was excluded as it related to diagnostic values of breast tumour biomarkers which was deemed out of scope for this topic, and another looked at the diagnostic accuracy of shear wave elastography, out of scope for the horizon scan. One article was included as it was a more recent study (Yang et al., 2014) than the review article described above (Tang et al., 2011). However, this prospective study reports on

the use of breast ductoscopy in evaluating nipple discharge, not screening for breast cancer (Yang et al., 2014).

Retrospective studies None identified. Grey literature None identified.

6.4. Study findings and discussion

# 6.4.1. What stage of development or trial is this innovation at?

Research so far has not focused on trialling the application of ductoscopy for population-based breast screening of asymptomatic women.

Intraluminal microendoscopic technology has been developing and has improved over the past 20 years (Tang et al., 2011). Earlier limitations of this technology were poor optical resolution, and access restriction of large-calibre scopes, which have now been overcome – and working channels within microendoscopes have been developed to give clinicians/operators the ability to perform biopsy of lesions and therapeutic procedures (Tang et al., 2011). Furthermore, earlier techniques such as air insufflation for distention of mammary ducts has been superseded by saline infusion for superior image quality (Tang et al., 2011).

There is a lack of prospective randomised trials of ductoscopy with a breast screening application (Tang et al., 2011). While technology appears to have improved for microendoscopy of breast ducts, it is difficult to say whether it will have a breast screening application for asymptomatic populations in the future.

# 6.4.2. What is its considered potential clinical value in five years? In 10 years?

The potential for clinical use of ductoscopy as a breast cancer screening modality in asymptomatic populations is unlikely to be realised within the next five to 10 years, due to the lack of any prospective studies or randomised control trial studies in this population (Tang et al., 2011).

# 6.4.3. Does this innovation show high sensitivity and specificity?

The intraductal approach of mammary ductoscopy is not discussed as a useful procedure for women with dense breasts or for women who have had breast surgery or augmentation (Tang et al., 2011; Yang et al., 2014), though it was suggested to have a potential role in screening individuals with a high risk for breast cancer, described as 'uniquely motivated', such as those who have a family history of breast cancer or are predisposed to breast cancer through a genetic mutation (Tang et al., 2011).

# 6.4.4. Is there evidence that this innovation is more acceptable to women (in general and by ethnic group) compared to digital mammography?

There is little evidence to suggest that ductoscopy is an acceptable procedure for women undergoing breast screening.

Acceptability of ductoscopy by women was not discussed by Tang and colleagues but their article discussed some uncommon potential complications and risks of ductoscopy, including pain, inflammation and infection and perforation of the duct which "may be a transient cause of post-procedure discomfort" (Tang et al., 2011, p.1714). Acceptability of ductoscopy was not discussed by Yang et al. (2014).

# 6.4.5. What cost and safety findings have been reported?

There were no findings reported with regards to cost of ductoscopy, though in terms of safety, the lack of radiation from the procedure is reported as a benefit in Tang and colleagues' review (2011).

In a study with 419 female patients with nipple discharge enrolled, ductoscopic investigation was successfully completed in 405 patients (96.7%). Fourteen patients (3.3%) were unable to have the ductoscopy device introduced due to nipple inversion, pain, or narrow canal. After ductoscopic investigation, no complication was reported for any of the 405 patients, thus, it was deemed a safe modality for patients with nipple discharge (Yang et al., 2014).

# 6.4.6. Does this technology reduce deaths due to breast cancer through early detection?

Based on this review, there is no evidence to suggest that ductoscopy reduces deaths due to breast cancer through early detection.

# 6.4.7. Has this innovation been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

Ductoscopy has not been implemented into any national screening programs for breast cancer.

# 6.4.8. Has a national position statement been published about this innovation, and if so, what is the position? Is there a consensus in position statements?

No national position statements have been published about ductoscopy as a breast cancer screening tool.

# 7. MAGNETIC RESONANCE IMAGING

Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

The use of Magnetic Resonance Imaging was not discussed in the 2009 ANZHSN report.

# 7.1. Introduction

Magnetic resonance imaging (MRI) is a non-invasive medical test that uses magnetic fields to produce detailed cross-sectional images of tissue structures. MRI creates images of the breast by measuring changes in the movement of protons in fat and water with the application of changing magnetic fields and by utilizing the differences in tissue relaxation characteristics. The contrast between different types of breast tissues (fat, glandular tissue, lesions, etc.) depends on the mobility and the magnetic environment of the hydrogen atoms in water and fat, which contribute to the signal intensity (brightness) of the breast image. In the breast, this results in images showing predominantly parenchyma and fat, and lesions, if they are present.

Contrast-enhanced MRI generally requires the use of a gadolinium-based contrast agent that is administered intravenously to provide better detail and detection of breast cancers and lesions. The use of contrast-enhanced MRI for breast cancer detection is based on the concept of tumour angiogenesis or neo-vascularity. That is, the blood vessels in cancerous tumours have increased permeability, which leads to prompt take up and release of gadolinium within the first one to two minutes after administration, leading to a pattern of rapid enhancement and washout on contrast-enhanced MRI. This dynamic rapid enhancement pattern helps to distinguish breast cancers from benign lesions. Because parenchymal tissue also enhances, but generally more slowly than malignant lesions, and because contrast can wash out rapidly from some tumours, the images must be looked at an earliest point of time after contrast injection (typically 1 to 3 minutes).

The technique of dynamic contrast-enhanced magnetic resonance imaging (DCE MRI), in which multiphase MRI scans are taken following the intravenous injection of a contrast agent, has been widely used in clinical practice, particularly as a screening modality for young women with a high risk for familial breast cancer and women with dense breasts. This technique is used because invasive cancer in the breast will generally show a more than 70% increase in signal intensity over baseline within the first 60-90 seconds because of large vessels in the tumour. This marked increase in signal intensity is then followed by a wash-out phase resulting from increased background parenchymal enhancement. DCE MRI ensures that appropriate images are captured during the relevant time slot for the best differentiation between breast cancer and normal surrounding tissue.

During a contrast-enhanced MRI exam, a patient lies face down and the breasts are positioned into two holes of a "coil" as this improves the quality of the image. Scanning takes approximately 30 to 45 minutes, and the MRI images are then interpreted by a radiologist. To reduce the influence of normal hormonal changes in the breast which can interfere with the accurate interpretation of the MRI examination, screening with MRI is best performed from day 7-10 of the menstrual cycle. MRI does not use X-rays, so it does not involve any radiation exposure.

Breast MRI is commonly used for diagnostic purposes to evaluate abnormalities detected on mammography or ultrasound, assess the status of breast implants, perform presurgical breast

cancer staging, evaluate palpable breast lumps or other breast-related symptoms, guide breast biopsies, and monitor chemotherapy. It also has an application in a screening environment.

In high-risk populations, screening with both MRI and mammography annually improves the sensitivity of screening but decreases specificity relative to screening with mammography alone.

Research in high-risk woman has indicated that MRI is limited in its ability to identify noninvasive breast cancer (eg, DCIS) and should therefore be used as an adjunct to, rather than a replacement for, mammography. The benefits and harms of adjunct screening breast MRI among women at less than high risk for breast cancer, however, are unclear.

# 7.2. Summary of key findings

- The availability of MRI as a supplementary examination to mammography offers a clear clinical benefit in some situations, particularly women with a high risk for breast cancer, especially those with dense breast tissue. Compared to mammography, MRI is less specific but more sensitive to detect small tumours in subjects with high breast cancer risk. MRI is a non-invasive technique that gives extremely clear, detailed images of soft-tissue structures that other imaging techniques cannot achieve. Unlike mammography, MRI does not expose the tissue to ionising radiation and the contrasting agent used in MRI is less likely to produce an allergic reaction that may occur during the use of iodine-based substances in other imaging modalities.
- MRI has not been recommended for the general population due to high false-positive rates (which can lead to over-diagnosis with attendant cost and anxiety), high cost, time consumption, lack of adequate number of units, the need for experienced radiologists and lack of clinical utility. Some cancers, such as DCIS, are better detected by mammography than by MRI.
- MRI is expensive, and although not painful the patient must remain still during the examination, which was an issue for claustrophobic women. Data has also emerged indicating there can be accumulation of gadolinium in patients who have undergone multiple contrast-enhanced MRI studies.
- MRI has not been incorporated into any national breast screening programs. The American Cancer Society and the European Society of Breast Cancer Specialists have both released statements on the use of MRI in breast cancer screening.

# 7.3. Literature search results (number of studies returned)

From the literature search a total of 253 abstracts of peer reviewed articles related to MRI were identified. Abstract contents were then reviewed, and 151 articles were excluded because they indicated studying MRI in a symptomatic population, for diagnostic or treatment purposes, or were studying developing technologies that may have an application within MRI, such as developing contrast agents. 36 other articles were subsequently excluded because they could not be located or were dissertations. A total of 66 articles were reviewed to answer the key research questions relating to the use of MRI as a breast cancer screening tool.
#### Systematic and/or literature reviews

22 papers were either systematic or literature reviews (Bick, 2015; Chhor & Mercado, 2017; Cott Chubiz et al., 2013; Fischer et al., 2012; Gilbert & Selamoglu, 2017; Greenwood et al., 2013; Health Quality Ontario, 2016; Heller & Moy, 2016; Heywang-Köbrunner, Hacker, & Sedlacek, 2013; Kanal & Tweedle, 2015; Kaniklidis, 2015; Lee et al., 2010; Lehman, 2010; Le-Petross & Shetty, 2011; Mainiero et al., 2016; McLaughlin, Mittendorf, Bleicher, McCready, & King, 2014; Mehnati & Tirtash, 2015; Melnikow et al., 2016; Morrow, Waters & Morris, 2011; O'Flynn, Ledger & de Souza, 2015; Partridge, Nissan, Rahbar, Kitsch, & Sigmund, 2017; Porembka, Seiler, & Sharma, 2016; Rahbar & Patridge, 2016; Runge, 2013; Salem, Kamal, Mansour, Salah and Wessam, 2013; Sardanelli et al., 2010; Sathya & Geetha, 2013; Sutcliffe & Otto, 2013; Wellings, Vassiliades & Abdalla et al., 2016; and Zhang & Ren, 2017).

#### RCTs

None identified

#### **Prospective studies**

19 prospective studies were identified (Badan et al., 2016; Berg et al., 2012; Brandzel et al., 2017; Brédart et al., 2012; Evans et al., 2016; Evans et al., 2014; King et al., 2013; Kuhl et al., 2014; Lehman et al., 2016; Lourenco, Donegan, Khalil, & Mainiero, 2014; McDonald et al., 2016; Møller et al., 2013; Moschetta, Telegrafo, Rella, Stabile Ianora, & Angelelli, 2016; Ng et al., 2013; Passaperuma et al., 2012; Rijnsburger et al., 2010; Saadatman et al., 2013; Sardanelli et al., 2011; and Tieu et al., 2014).

#### **Retrospective studies**

27 retrospective studies were identified (Ahern et al., 2014; Chiarelli et al., 2014; de Bock et al., 2013; Destounis, Arieno, & Morgan, 2016; Freitas et al., 2013; Friedlander, Roth & Gavenonis, 2011; Giess, Poole, Chikarmane, Sippo & Birdwell, 2015; Grimm et al., 2015; Gubern-Merida et al., 2015; Gweon et al., 2014; Harvey et al., 2016; Jain, Jain, Hyzy, & Werth, 2017; Lowry et al., 2012; Moftah et al., 2014; Niell et al., 2014; Pederson et al., 2015; and Santoro, Podo & Sardanelli, 2014).

#### **Grey literature**

A search of the U.S. National Library of Medicine <u>ClinicalTrials.gov</u> website for "breast cancer screening" and "magnetic resonance imaging" identified 19 related clinical trials, with six studies related to MRI for breast cancer screening purposes:

- 'Evaluation of Gadolinium Deposits in Healthy Women Participating in a High Risk Screening Program for Early Breast Cancer Detection' sponsored by the Medical University of Vienna, Austria. The study, which is due to be completed in April 2018, seeks to determine whether patients at high risk to develop breast cancer having received at least 6 cumulative dosages of macrocyclic Gd-based contrast media in the context of breast cancer screening by means of contrast-enhanced MRI.
- 'Contrast-enhanced MR Imaging as a Breast Cancer Screening in Women at Intermediate Risk (MRIB)' sponsored by the IST Istituto Nazionale per la Ricerca

sul Cancro, Genoa, Italy. The randomised control trial, which was due to be completed in April 2016, seeks to evaluate the performance of MRI, in terms of sensitivity, specificity, and predictive value, in the screening of women at intermediate risk of breast cancer.

- 'Initial Evaluation of Ultra FAST Breast Magnetic Resonance in Breast Cancer Screening: Comparative Study With Mammography and Ultrasound' sponsored by Brugmann University Hospital, Belgium. The study, which is due to be completed in January 2020, seeks to evaluate the performance of FAST breast magnetic resonance in normal screening population.
- 'MRI for Detecting Cancer in Women Who Are at High Risk of Developing Breast Cancer' sponsored by the University of Pennsylvania, United States. This study seeks to explore the effectiveness of MRI scans in women who are at high risk for developing breast cancer.
- 'Breast Cancer Screening With MRI in Women Aged 50-75 Years With Extremely Dense Breast Tissue: the DENSE Trial' sponsored by UMC Utrecht, the Netherlands. This study, which is due to be completed in December 2019, seeks to determine the cost-effectiveness of biennial screening with mammography and MRI compared to mammography alone in women aged 50-75 years and who show > 75% mammographic density.
- 'Feasibility of Short Breast MRI (SBMRI) for Screening Patients at High Risk for Breast Cancer' sponsored by University of Texas MD Anderson Cancer Center, Houston, Texas, United States. This study, which is due to be completed in 2020, seeks to determine if a short breast MRI scan (about 10 minutes) can be used for breast cancer screening in high risk people and to learn if it is as effective as a routine breast MRI scan (30-60 minutes).

### 7.4. Study findings and discussion

#### 7.4.1. What stage of development or trial is this innovation at?

Mammography is considered the standard test for breast cancer screening throughout the world, and has been shown to decrease breast cancer mortality But the sensitivity of mammography is lower in young women, women with dense breast tissue, and women who carry the genetic BRCA mutations, and this has led to a search for alternative methods of screening in women at a high-risk of breast cancer (Morrow, Waters, & Morris, 2011). MRI reveals at least 10 additional cancers per 1,000 women screened after mammography and is a useful adjunct screening strategy for women who are at high risk for breast cancer (although the biological significance of these cancers is unknown). The following section looks at those papers that examine the current state of use of MRI as a breast cancer screening tool.

#### Sensitivity

#### Systematic and literature reviews

Five systematic reviews discussed the rate of sensitivity. In a recent literature review of 11 papers, Zhang & Ren (2017) found a pooled sensitivity of 92.0% for asymptomatic women undergoing MRI breast screening. This was 17% higher than the average sensitivity rate for mammography. These figures compared favourably to a study by Heywang-Köbrunner, Hacker, & Sedlacek. (2013) who found sensitivities ranging from 90 to 91%. Similar results were reported in the High Breast Cancer Risk Italian 1 study, in which 18 screening centres enrolled 501 women and performed 1,592 studies (3.2/women) (Sardanelli et al., 2011).

Most cancers detected by MRI screening were between 10 and 20mm in size and were invasive cancers (Heywang-Köbrunner et al., 2013, Le-Petross & Shetty, 2011). Despite the small size of the tumours, 12 to 26% of these patients had a node-positive disease at the time of detection. Le-Petross & Shetty (2011) considered that this raised the question of whether an annual screening interval was adequate, although they acknowledged that the node-disease may be simply related to the aggressiveness of the detected breast cancer.

However, the ability of MRI to detect DCIS was not so clear (Heywang-Köbrunner et al., 2013, and Morrow et al., 2011), with detection rates being reported as low as 40%. The literature suggests that MRI should not be used to exclude malignancy in cases with suspicious microcalcifications. Instead, mammography should remain the preferred modality for the detection of DCIS, particularly given its superiority in the detection of microcalcifications.

#### **Retrospective observational studies**

In their study of 135 asymptomatic women who underwent MRI screening as part of a mammogram follow-up, Bick et al. (2015) reported a sensitivity of 100%, which they attributed to two factors. Many cases of DCIS associated with microcalcifications did not enter their study because they would have been referred straight to biopsy, reducing the potential for false-negatives. The other factor was that the MRI was read in full knowledge of the prior screening mammogram. This would have improved the sensitivity of their results.

Several studies have examined the rate of sensitivity of asymptomatic women falling within different population groups based on their risk of contracting breast cancer.

#### MRI in women with a high lifetime risk of breast cancer

Rahbar & Partridge, (2016) cited multiple studies that found that MRI had the highest sensitivity of any imaging modality for breast cancer detection in asymptomatic high-risk women – defined as women with a lifetime risk of breast cancer greater than 20-25%. This includes individuals with a known BRCA1 and BRCA2 mutation and their first-degree relatives, women who underwent therapeutic chest radiation in their youth and those with specific genetic mutations known to increase the risk of breast cancer such as Li-Fraumeni syndrome (TP53 gene mutations) and Cowden and Bannayan-Riley-Ruvalcaba syndromes (PTEN gene mutations).

Most studies evaluating the sensitivity of MRI screening have focused on this population subset (high risk women). In the studies Heller & Moy (2016) reviewed, MRI sensitivity (77.0%-92.6%) was consistently greater than for mammography (32.6%-50.0%) for the screening of high-risk women. The Wellings, Vassiliades & Abdalla (2016) review reported similar rates of sensitivity (77% to 100% for MRI compared to 33% to 56% for mammography). Sardanelli et al., (2011) observed that MRI outperformed mammography and ultrasound, whether used separately or in

combination, for screening of women at high risk of cancer, whether younger or older than 50 years. Morrow, Waters, & Morris, (2011) observed similar results finding that across all the papers they reviewed sensitivity was only 32% for mammography, while MRI sensitivity was 75%. Combining the two procedures increased sensitivity to 84%.

*BRCA 1 and BRCA 2:* Heller & Moy (2016) commented on the sensitivity of MRI screening on BRCA mutation carriers. One of the trials they cited reported a 91% sensitivity for cancer detection in all women undergoing MRI screening versus 31% with mammograms, but a 100% sensitivity in known mutation carriers versus 25% sensitivity with mammography.

Comparing BRCA1 and BRCA2 mutation carriers, Heller & Moy (2016) noted that the MRI was better at detecting cancers in BRCA1 carriers than detecting cancers in BRCA2 carriers, which were better detected using mammography. In their view, this result supported calls to tailor screening regimens to individual mutation types.

It is important to note that the results from these studies cannot be directly extrapolated to a lower-risk population, as MRI's sensitivity and specificity has been found to differ based on risk status (Health Quality Ontario, 2016). Additionally, women at high risk for breast cancer exhibit altered tumour histology (microscopic anatomy) and morphology (tissue structure) on mammography and have a higher rate of interval cancers than those at less than high risk. The benefit of MRI screening in populations other than high risk needs to be examined separately (Health Quality Ontario, 2016).

*Chest Irradiation*: Screening breast MRI is recommended for women who underwent therapeutic chest radiation (eg, for Hodgkin lymphoma (HL)) during their youth because up to 25% of women so treated develop breast cancer by ages 40-45 (Heller & Moy, 2016). In a prospective study on 148 women treated with chest radiation for HL during their youth and were 8 years beyond treatment, Ng et al., (2013) found no difference in sensitivity between MRI (68.0%) versus mammography (67.0%). However, when the modalities were combined, sensitivity increased to 94.0%. Similar results were reported in two other studies (Tieu et al., 2014, and Freitas et al., 2013).

#### MRI use for women at moderate risk

Women who have a family or personal history of breast cancer, those with a history of lobular carcinoma in situ (LCIS) and those with heterogeneously dense breasts are at moderate risk of developing breast cancer (lifetime risk of between 15 to 20%).

Heller & Moy, (2016) cited four studies that examined the screening of asymptomatic women with a family or personal history of breast cancer (Destounis, Arieno, & Morgan, 2016; Lehman et al., 2016; Gweon et al., 2014; and Giess, Poole, Chikarmane, Sippo, & Birdwell, 2015). These studies found no difference in the rate of sensitivity between women with a family or personal history of breast cancer and those with no family history or genetic risk of cancer. Given the high rate of false-positive MRI screenings (with attendant cost and anxiety), Sutcliffe & Otto, (2013) felt referring clinicians needed to be cautious before referring women with a family or personal history of breast cancer for MRI screening. They recommended that, where possible, such patients should be referred to a cancer genetics counsellor to assess whether they have a high-risk of developing breast cancer (as opposed to merely a moderate risk). Lee et al., (2010) made a similar recommendation.

Friedlander, Roth & Gavenonis (2011) found that screening MRI identifies breast cancer in asymptomatic women diagnosed with LCIS at a rate like that shown in high-risk population groups. While this study excluded women with an abnormality depicted on other imaging modalities, it did not indicate if all women received mammography (Health Quality Ontario, 2016). This limited the study's applicability to a screening context.

King et al., (2013) analysed 776 cases of asymptomatic women diagnosed with LCIS to compare rates of cancer detection with or without MRI. Analysing results from1, 2 and 3 years post diagnosis screening, they found that MRI failed to demonstrate increased cancer detection rates among women (p = 0.23, 0.26 and 0.13 respectively). They concluded that a diagnosis of LCIS remains a significant risk factor for breast cancer, the routine use of MRI does not result in increased cancer detection rates (short-term) which illustrated the need of defining optimal screening strategies for high-risk patients based on tumour biology rather than numerical risk.

It is noted, however, that details regarding the use of mammography and MRI screening were not described in this paper, and insufficient information was provided to determine how and when MRI was performed relative to mammography screening (Health Quality Ontario, 2016). The cancer detection rate was also substantially higher than that reported by McLaughlin, Mittendorf, Bleicher, McCready & King (2014) who identified three other studies that evaluated the role of MRI in patients with LCIS that collectively documented an incremental annual increase in the cancer detection rate of around 4%.

Again, Sutcliffe & Otto, (2013) stressed the need for caution and suggested that women with LCIS should only be referred for MRI screening on a case-by-case basis.

#### MRI use for women at low-risk

Although the use of MRI to screen asymptomatic women with low risk of breast cancer (lifetime risk less than 15%) is not encouraged, Heller & Moy (2016) identified one study that observed a cancer detection rate of 11 per 1000 screening examinations in this population group. They felt that the development of abbreviated MRI screening examinations (which are described in detail below) could enable screening MRI to be expanded to a wider pool of women.

Heywang-Köbrunner et al. (2013) could not identify any studies that examined the sensitivity of MRI screening in low-risk women. Given the low yearly incidence of breast cancer in the general population (<3/1000), the low specificity of MRI and the associated higher screening costs, they said MRI screening of women at low risk was not sensible.

#### Interval cancer rate

Only one study discussed interval cancer rate of MRI. Le-Petross & Shetty (2011) noted that despite the addition of MRI to mammography screening in asymptomatic high-risk women, some published trials still report cancers developing between the screening intervals (i.e., interval cancer) at a rate of 2% to 9%. They noted that the screening regimens had one-year intervals, with MRI and mammography performed within 90 days of one another. They said that staggering examinations at 6-month intervals could decrease the incidence of interval cancers. This will be discussed later in this section.

#### Positive predictive value (PPV)

#### PPV1 (cancers diagnosed per the number of women recalled from screening)

Overall results on PPV<sub>1</sub> indicate that, on average, MRI accurately detected proportionally more women recalled from screening who had breast cancer compared to FFDM alone. Three studies reported that on average, recalls based on MRI screening are correctly identifying an additional five women with diagnosable breast cancer for every 100 women recalled, compared with FFDM alone (Chiarelli et al., 2014, Niell et al., 2014 and Friedlander et al., 2011).

#### PPV<sub>2</sub> (cancers diagnosed per the number of biopsies recommended)

Only two studies reported on the rate of PPV2. The rate of cancers diagnosed from the number of biopsies that were recommended varied from 18.5% (Friedlander et al., 2011) to 24% (Niell et al., 2014). No comparison was given in either study to the rate of PPV2 for FFDM alone.

#### PPV<sub>3</sub> (cancers diagnosed per the number of biopsies performed)

The rate of PPV<sub>3</sub> was reported in three studies (Chiarelli et al., 2014, Niell et al., 2014, and Friedlander et al. 2011), who found that the rate of cancers diagnosed after biopsy following MRI screening (ranging from 23.8% to 26.1%) was higher than the rate following mammography screening (16.9%).

#### Specificity

From the 2 mega-analyses they reviewed, Heywang-Köbrunner et al. (2013) found that the rate of specificity for MRI as a screening tool was only 72% to 75%.

Four studies compared the rate of specificity between MRI and mammography. Zhang & Ren (2017) found an average 70.0% for women undergoing MRI breast screening returned a negative test result for the disease, which was like that of mammography (71.0%). Wellings et al. (2016) also found a lower rate of specificity for MRI (79% to 97%) that that of mammography (91% to 99%). Le-Petross & Shetty (2011) observed similar rates (MRI ranged from 79% to 95% and mammography ranged from 93% to 99%). In contrast to these studies, Morrow et al. (2011) reported that the rate of specificity for mammography (98.5%) was only marginally higher than the specificity of MRI (96.1%).

Heywang-Köbrunner et al. (2013) did not support the use of MRI in general population screening. In their view, its low specificity would lead to additional biopsies that ultimately yield benign pathology (with attendant cost and anxiety) or short-term follow-up in a significant proportion of women. The high cost of using MRI also worked against its use as a general population screening technique.

#### False positive recall rate

False positive recall rate is a significant concern as women who are recalled for further investigation often experience high levels of anxiety, along with the inconvenience and expense of attending a further appointment that bring no health benefit to the woman.

MRI is associated with a higher false positive rate because of its higher sensitivity but lower specificity, which might lead to additional scans and biopsies with their attendant risks of increased cost and anxiety. Le-Petross & Shetty (2011) referred to one study that compared

biopsy rate following MRI and mammogram screening on 195 women. The study found a biopsy rate of 8.5% for MRI but only 2.2% for mammography screening.

Examining the effectiveness of the Ontario Breast Screening Program (which screened 2,207 asymptomatic women with high risk of breast cancer), Chiarelli et al. (2014) found that the recall rate was significantly higher for MRI (15.1%) compared with mammography (6.4%), which was consistent with other prospective studies.

#### When to Screen/Screening Intervals/When to Stop Screening

In terms of when to start screening with MRI, the Society for Breast Imaging and the American College of Radiology recommend MRI starting at age 30 or 10 years earlier than the age of diagnosis of the youngest family member with breast cancer but not before age 25.

Once screening commences, debate exists about whether screening mammogram and MRI should be completed together or alternated every 6 months. Cott Chubiz et al. (2013) and Lowry et al., (2012) both proposed alternating screening mammography and breast MRI every 6 months starting at age 30. The advantage of alternating screening is to decrease the incidence of interval cancers as well as to offer the patient the psychological reassurance of being observed every 6 months. The proposed screening regimen also provided the most favourable balance of clinical benefit, radiation exposure and cost-effectiveness compared to annual screening. McLaughlin et al., (2014) noted, however, that in the MRI screening trials employed in these studies, the interval cancer rate was less than 3% when annual mammogram and MRI screenings were performed together. It was unlikely that an alternating strategy could further improve this rate.

Heller & Moy, (2016) reported a lack of consensus on when and if to stop MRI screening; some national guidelines recommend a re-evaluation of or end to MRI screening at age 50 (UK/ Netherlands) based on breast density (UK), arising in part from the assumption that older women will have lower density breasts and will therefore be more likely to have mammographically visible cancers.

### 7.4.2. What are their considered potential clinical value in five years? In 10 years?

The potential clinical value of MRI in five to 10 years is not explicitly discussed, though studies have explored various technological advances in MRI to assess if these innovations could lead to improved rates of sensitivity and specificity.

For instance, Le-Petross & Shetty (2011) observed that most of the published studies have looked at breast screening MRIs performed with a 1.5 tesla magnetic system. In the United States, many private practices and academic centres routinely use magnetic systems with higher field strengths such as 3.0 tesla. It has been reported that 3-T, when compared to 1.5-T offers improved spatial resolution and temporal resolution for dynamic studies. It also results in a 3-fold increase in cancer detection (Lourenco, Donegan, Khalil & Mainiero, 2014). Both authors hope that the introduction of these higher field magnetic systems will lead to improved rates of sensitivity and specificity, and a reduction in the rate of false-positives.

Rahbar & Partridge (2016) and Heywang-Köbrunner et al. (2013) identified several functional MRI approaches at early stages of development that are showing promise for lesion differentiation and patient outcomes. Their usefulness as screening techniques is not clear.

Diffusion weighted imaging (DWI) is an MRI technique that provides information about the random motion of protons in tissues. Restricted diffusion of water molecules is observed in tissue with increased cellular density or increased fibrosis, which often occurs in malignant tumours. It is visualized as hyperintense signal on diffusion-weighted images or as low signal on calculated images of the apparent diffusion coefficients. Preliminary studies have shown higher AUC in benign lesions and normal breast tissues than most malignancies. Although sensitivity and specificity were only 84% and 79% respectively, this technique has potential value as a screening tool when consideration is given to the fast imaging acquisition time and lack of reliance on an injection of intravenous contrast (Porembka, Seiler & Sharma, 2016, Rahbar & Partridge, 2016, Partridge, Nissan, Rahbar, Kitsch & Sigmund, 2017, and Heywang-Köbrunner et al., 2013). Several challenges have been reported that currently prevent DWI from being implemented in a widespread breast cancer screening protocol, such as the lack of standard protocols for diffusion weighted image acquisition. While this technique has the greatest potential for future use, further investigation is warranted (Porembka et al., 2016, Partridge et al., 2017 and Rahbar & Partridge, 2016).

Three studies have looked at improving the detection of DCIS (which is detected by microcalcifications on mammograms) as they saw its low detection rate as one of the main weaknesses of MRI screening compared to FFDM. They noted that most studies that investigated the use of MRI for DCIS detection were performed with high temporal resolution and relatively low spatial resolution. Badan et al. (2016) showed that MRI was able to detect the most prevalent morphologies and kinetic characteristics of DCIS. However, the study was not randomised (n = 25 patients with suspicious microcalcifications) so the usefulness of their findings is unclear. An earlier systematic review found that sensitivity of MRI in the assessment of DCIS reached 89.0%, compared to only 55.0% in mammography (Greenwood et al., 2013). It was felt that additional MR imaging tools, such as DWI and 3-T imaging, may help increase the specificity and sensitivity of MRI in the detection of DCIS (Greenwood et al., 2013 and Lehman, 2010). Another field of future research includes sodium imaging as it produces different signals in malignant tumours (Rahbar & Partridge, 2016 and Heywang-Köbrunner et al., 2013).

Improvements are also being made to the contrast agent used in MRI screening. One study compared the diagnostic performance of the standard MR contrast agent (gadopentetate dimegumine) with a new contrast agent (gadobenate dimegluine) that had higher relaxivity. They found that the new contrast agent had higher sensitivity and specificity (99.0% and 92.4%) than the standard contrast agent (93.0% and 83.8%).

The use of computer-assisted mass classifier models to automatically analyse lesion features in MRI images to differentiate between normal and abnormal breast tissue has also been examined. Sathya & Geetha (2013) developed the artificial bee colony algorithm to optimise a neural network performing classifications on the regions of interest. The algorithm achieved sensitivity and specificity rates of 92% and 89% respectively. The experimental results demonstrated the usefulness of these systems, particularly for reducing the number of negative biopsies. Other studies have demonstrated the overall accuracy of computer-assisted detection (see Moftah et al., 2014 and Gubern-Merida et al., 2015).

#### 7.4.3. What cost and safety findings have been reported?

MRI is more costly than mammography and may also lead to larger numbers of false positive examinations and unnecessary biopsies which in turn contribute to increased medical expenditure (Berg et al., 2012).

As part of their study on the cost-effectiveness of alternative strategies for integrating MRI into breast cancer screening for women at high risk, Ahern et al. (2014) compared the screening-related costs between bilateral mammography and MRI and found that the latter was approximately five times more expensive. This figure compared favourably to earlier studies that found cost differences of three and four times (de Bock et al., 2013; Cott Chubiz et al., 2013 and Saadatman et al., 2013). Kaniklidis (2015) noted that an MRI screening session in the United States costs anywhere between \$277 and \$965, which was still substantially more expensive than mammography which averaged \$115 to \$135.

Seven studies have looked at ways to shorten the standard MRI screening protocol to save examination time and costs without reducing the rate of sensitivity and specificity. Fischer et al. (2012) noted that the maximum contrast enhancement in cancers occurred within the first three minutes after intravenous application, while normal parenchyma had a slow increasing uptake of the contrast material. This meant that the relevant time slot for the best differentiation between breast cancer and normal surrounding tissue was the first minutes after contrast injection. After this period, the wash-out seen in malignant tumours, as well as the increasing enhancement of normal parenchyma, made the depiction of breast cancer increasingly difficult. They developed a shortened MRI screening protocol for breast cancer detection.

They found that if there was no ambiguous in-breast enhancement an examination could be completed within 4-5 minutes after contrast administration. In their study, three-quarters of women did not display ambiguous findings and were only examined using the shortened MRI protocol (the remainder were required to undergo further imaging, including T2 imaging). They concluded this approach would lead to a higher examination rate and reduction of the cost per MRI.

Kuhl et al. (2014) developed their own abbreviated protocol (AP) for MRI screening, which achieved diagnostic accuracy and cancer yield equivalent to those of regular breast MRI protocols. It also was associated with reduced time for image acquisition (i.e., MRI system time) and image interpretation (i.e., radiologist reading time). In a study of 606 screening patients, they obtained an MRI acquisition time of 3 minutes and reading time of less than 30 seconds, while achieving a sensitivity and specificity of 90.9% and 94.3%, which were not significantly different than the full diagnostic protocol (FDP). They predicted the AP had the potential to identify 18.2 additional cancers per 1,000 screening examinations. Gilbert & Selamoglu (2017) praised the use of the abbreviated breast MRI protocol noting that it was less expensive and more feasible in terms of machine time. However, they acknowledged that it would not be cost-effective for women who were only at immediate risk of developing breast cancer.

Chhor & Mercado (2017) identified several other studies that had investigated the use of an abbreviated breast MRI protocol for screening purposes (see table 1). The AP resulted in shorter acquisition time compared with that for FDP. Among the three studies reviewed, the average time to perform breast MRI with FDP was approximately 25:01 minutes. The average time to perform a study with AP was 11:08 minutes, with an average time savings of approximately 13:53 minutes. In the Grim Harvey and Moschetta studies, the average time to interpret images with AP was 3-4 times faster than with FDP. However, the Grimm study reported no difference

in interpretation time. Chhor & Mercado offered two explanations for this result: (i) the additional sequences for FDP were not useful for clinical interpretation, so the readers spent less time reviewing the images and (ii) readers spent more time reviewing the AP images because they did not have available the images obtained with the additional standard sequences. Nonetheless, in their view these studies showed that comparable screening results can be achieved with abbreviated MRI protocols and conventional full diagnostic protocols.

Results	Grimm et al. (2015)	Harvey et al. (2016)	Moschetta et al. (2016)
FDP acquisition time (min)	20:00	23:20	16:00
AP acquisition time (min)	11:13	4:40	10:00
FDP average interpretation time(s)	177	385.8	360
AP average interpretation time(s)	178.8	93	120
FDP Number of cancers	12	7	69
AP Number of cancers	11	7	67
FDP sensitivity	95	-	92
AP sensitivity	86, 89	-	89
FDP specificity	52	-	92
AP specificity	52, 45	-	91

MRI Table 1: Results from studies investigating use of abbreviated breast MRI protocol

Jain, Jain, Hyzy, & Werth (2017) explored whether the AP also had an impact on recall rates. Using the AP to read 591 scans, they found that the recall rates for AP (6.6%) and FDP (5.8%) were not statistically significant (p = 0.63). They concluded that AP had the potential to replace FDP in the screening setting so long as there is the ability to recall patients with any abnormality for further evaluation.

In terms of safety, MRI has been found to be a safe imaging method that did not involve ionising radiation (Sutcliffe & Otto, 2013). The contrasting agent used in the technique is less likely to produce an allergic reaction that may occur when iodine-based substances are used for other imaging modalities. However, not all patients can tolerate a contrast agent and it may pose a risk for women with kidney disease. There is a small chance that a patient may develop a skin infection at the site of injection (Mehnati & Tirtash, 2015). In their study of women's experiences with breast cancer screening, Brandzel et al. (2017) also observed that a few women had experienced adverse reactions to the contrast agent used in the procedure.

While MRI screening is not painful, the patient must lie still in an enclosed machine, which may be a problem for claustrophobic patients. This position can be difficult to keep for some patients with neck problems, obesity or breathing issues (Mehnati & Tirtash, 2015). Also, MRI cannot be performed in women who have certain metallic implants (such as pacemakers) as these may be affected by the strong magnet of the MRI. Pregnant women are not encouraged to undergo MRI (Mehnati & Tirtash, 2015).

There are also growing health concerns related to the long-term use of gadolinium contrast agents used in MRI screening. Researchers have noticed an accumulation of residual gadolinium in the brain of patients who have undergone multiple contrast-enhanced MRI studies. While the clinical significance of this observation is unclear, there is a need for caution (Kanal & Tweedle, 2015)

# 7.4.4. Does this technology reduce deaths due to breast cancer through early detection?

Three systematic reviews have reported on MRI and the rate of the mortality. Morrow et al. (2011) identified no prospective randomised trials of breast cancer screening in general or highrisk populations using MRI with survival as an endpoint. Their review supported other studies that found the sensitivity of MRI was better than that of mammography for the detection of invasive cancers and the occurrence of fewer interval cancers. Nonetheless, they were careful not to claim that MRI screening had a greater survival advantage. As noted elsewhere, MRI was better at detecting breast cancer in women who carry BRCA mutations. However, the cancers that occur in BRCA1 mutation carriers commonly have a prognostic phenotype, lacking oestrogen and progesterone receptors, so the effect of early detection on survival was unclear.

Heller & Moy (2016) agreed that there was little data on the impact of MRI screening on mortality in other high-risk groups or moderate-risk groups. However they referenced a Evans et al. (2014) study that evaluated 5- and 10-year survival of high-risk women by comparing MRI and mammographic surveillance groups to mammography only or no enhanced surveillance groups. While there was no difference in 10-year survival between the MRI and mammography and mammography-only groups, survival was significantly higher in the MRI-screened group (95.3 %) compared to no intensive screening (73.7 %). The authors concluded that there appears to be benefit from screening with MRI, particularly in *BRCA2* carriers.

Santoro, Podo & Sardanelli (2014) also commented on the Evans trial, noting that the higher survival rate among women screened with MRI could certainly be due to the earlier stage of MRI-detected breast cancers but could also arise from more up-to-date therapies applied to *BRCA* mutation carriers diagnosed with cancer. In a later observational study, Evans et al. (2016) evaluated *BRCA2* mutation carriers with mammography alone (53 women), MRI and mammography (34 women), or no intensive screening at all prior to cancer development (274 women). Ten-year survival was 85.5 % in the mammography group, 100 % in the MRI and mammography group, and 74.6 % in the control group.

Three other observational studies have been identified. Møller et al. (2013) reported on the 5and 10-year survival rate of *BRCA1* mutation carriers. As part of a Norwegian screening program, 802 women with a *BRCA1* mutation were examined annually with MRI in addition to mammography and followed for a mean of 4.2 years. As of December 2011, 68 of 802 women in the screening program were diagnosed with DCIS or invasive breast cancer (8.5 %). Sixty-three of the cancers were invasive and five were in situ. Ten of the 68 patients died of cancer in the follow-up period. The 5-year breast cancer-specific survival for women with cancer was 75% and the 10-year survival was 69%. Since these rates were less than anticipated, the authors considered that the benefit of annual MRI surveillance on reducing breast cancer mortality in *BRCA1* mutation carriers had yet to be proven.

In the Passaperuma et al. (2012) study, which had an eight-year follow-up, there was only one death out of the 54 *BRCA* mutation carriers diagnosed with breast cancer. This equated to a 0.5% mortality rate. An earlier Dutch MRI screening trial found that four of the 42 *BRCA* mutation

carriers diagnosed with invasive breast cancer died of the disease, and one developed metastatic disease. The study reported an 84.0% distant free-survival at six years and an annual mortality rate at 1.2% (Rijnsburger et al., 2010).

In their systematic review, Wellings, Vassiliades, & Abdalla (2016) claimed that if MRI could detect cancers at an earlier stage, it could be inferred that there would be an equivalent decrease in mortality. However, they acknowledged that support in terms of decreased mortality would need to come from recurrence and survival data from more observational studies.

# 7.4.5. Does this innovation show higher sensitivity and specificity for women with dense breast and women who have breast surgery/augmentation compared to single human view?

Increased mammographic density is known as an independent risk factor for breast cancer. Heller & Moy (2016) cited various studies demonstrating MRI's increased sensitivity compared to mammography in subsets of women with dense breasts. This was also observed by Runge (2013), who considered that this population should be considered for routine evaluation.

In their systematic study of supplementary breast screening techniques, Melnikow et al. (2016) referred to three studies that found the sensitivity of MRI screening on women with dense breasts ranging from 75.0% to 100.0%. Specificity also varied, ranging from 78.0% to 93.0%. It should be noted that it was not clear whether the papers referred to in this study examined women who had no apparent risk factors other than having dense breasts.

One retrospective study by Berg WA et al. (2012) comprised 612 participants with heterogeneously or extremely dense breasts who had at least one other risk factor (such as a familiar risk). All the participants had been previously screened as part of an annual mammography screening program. The study reported a sensitivity of 31.3% for mammography alone, which increased to 100% by adding MRI imaging, and yielded 14.7 additional cancers/1000 screening exams. However, the increase in sensitivity was accompanied by a reduction in specificity for mammography (92.1%) and for MRI (70.6%), which resulted in higher false positives and a recall rate of 20%.

O'Flynn, Ledger, & de Souza (2015) raised concerns about the lack of papers that compared performance of different screening modalities in women with dense breasts but with no other risk factors. The cancer yield in women in the general population with dense breasts who are otherwise at low or intermediate risk will not be as high as that in high-risk women: therefore, the costs and time required to scan and interpret large number of images would be high for a small cancer detection. While they acknowledged the potential of MRI to increase screening sensitivity in women with dense breasts, they said there was very little clinical evidence justifying its use in this population cohort.

One study has looked at whether a different screening technique could be employed to improve the detection of breast cancer in asymptomatic women with dense breasts. McDonald et al. (2016) evaluated the performance of incorporating DWI into MRI screening to detect mammographically occult breast cancer (invasive carcinoma and DCIS) in high-risk women with dense breasts. They observed that mammographically occult breast cancers exhibited higher signal intensity on DWI than normal screening methods. They found sensitivity and specificity rates of 45% and 91% respectively. While sensitivity was lower than expected they attributed this to the study design employed. They concluded that DWI had potential as a rapid supplemental screening tool.

Only one paper discussed women who have had breast surgery/augmentation. The paper noted that despite the imaging challenges posed by patients with breast implants, MRI could delineate cancer obscured by the implant. MRI was also the method of choice in the evaluation of the augmented breast and detection of problems over time, with very high accuracy in spotting implant rupture and in cancer detection (Salem, Kamal, Mansour, Salah, & Wessam, 2013).

# 7.4.6. Is there evidence that this innovation is more acceptable to women compared to digital mammography?

Three studies discussed the experience and acceptability of MRI in women who underwent breast screening.

Brandzel et al. (2017) conducted six focus groups attended by 41 women (aged 38-75 years) who had undergone MRI screening as part of their breast cancer treatment. Many women experienced discomfort during the procedure, because of the position they had to hold or claustrophobia, and anxiety related to the examination, primarily because they feared subsequent cancer detection.

In their study of 320 high-risk patients who underwent 757 screening procedures over time, Pederson et al. (2015) also observed that the patients often experienced significant anxiety when called back for additional studies or biopsy procedures. However, MRI screening seemed to be acceptable by most of their test subjects as evidenced by the high rates of participation in recommended screening. They noted that withdrawal from MRI screening programs because of claustrophobia or stress was rare, and that women at elevated risk accepted the need for additional tests, including those that resulted in negative findings, as part of intensified surveillance.

In an earlier study, Brédart et al. (2012) invited 1561 women to complete a questionnaire within 2 days after breast screening (compliance >91%). Of these women, 329 underwent MRI screening. These women reported significant MRI discomfort related to duration (35%), immobility (38%), prone position (21%), noise (65%) and panic feelings (6%). They reported that MRI was associated with a more favourable examination psychological experience than mammography ( $p \le .001$ ), especially in women younger than 50. This finding allowed them to conclude that perception of care and experience with MRI was more positive than with standard imaging procedures.

# 7.4.7. Has the technology been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

We have not identified any evidence that MRI has been incorporated into a national screening program, however, it is being used as part of the Ontario Breast Screening Program in Canada. The program is a province-wide, organised, publicly funded screening program for breast cancer. Average-risk women between the ages of 50 and 74 years are offered screening with mammography every two years, whereas women at higher than average (but not high) risk are offered screening with mammography annually. Screening with both mammography and MRI is currently indicated for women aged 30 to 69 years at high risk for breast cancer (Chiarelli et al., 2014).

# 7.4.8. Has a national position statement been published about the technology, and if so, what is the position? Is there consensus in position statements?

In 2007, the American Cancer Society recommended MRI screening in addition to mammography for women at high (20–25% or greater) lifetime risk of developing breast cancer, individuals with a known BRCA1 or BRCA2 mutation, or those with a first-degree relative with a BRCA mutation. MRI screening is also recommended for women with a clinical history of radiation to the chest between age 10 and 30 and those with specific genetic mutations known to increase the risk of breast cancer such as Li-Fraumeni syndrome (TP53 gene mutations) and Cowden and Bannayan-Riley Ruvalcaba syndromes (PTEN gene mutations), or their first degree relatives (see Heller & Moy, 2016 and Mainiero et al., 2016).

In 2010, the European Society of Breast Cancer Specialists (EUSOMA) published a paper evaluating the available evidence regarding clinical value of and indications for breast MRI. This paper reported the results of all the cohort studies investigating the diagnostic performance of different imaging modalities in the surveillance of high-risk women. They recommended that women with a family history suggesting an inherited predisposition to breast cancer should have their risk assessed by an appropriately trained professional group (eg, genetic counselling). If found to be at high risk (20-30% lifetime risk or greater), these women should be given oral and written information regarding their risk and the risks and benefits of mammography and MRI screening or alternative risk-reducing interventions. If these women accept to be screened by MRI, they should be informed about screening intervals and logistics. This should be determined based on regional or national considerations reflecting an area-specific cumulative risk in the general population, resource availability and practical feasibility. They recommended that annual MRI screening should be available starting at age 30 (Salem et al., 2013; Sardanelli et al., 2010).

### 8. MICROWAVE IMAGING

#### Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

The Bulletin on new and emerging technologies for breast cancer detection published in 2009 advised that microwave imaging may not be of clinical value within a five-year timeframe (from the date of publication). Radar-based microwave imaging was discussed as being in early stages of development, with patient studies planned for 2009 (National Horizon Scanning Unit, 2009).

# 8.1. How does microwave imaging work as a breast imaging tool?

Microwave imaging (MI) as a tool for the early detection of breast cancer in asymptomatic women is based on the contrast of electrical properties in the microwave spectrum of healthy tissue and malignant tumours in the breast. Microwave energy propagation, reflection and attenuation is sensitive to water, and due to the different dielectrical properties of normal breast tissue and tumours (i.e., cancer lesions contain more water than healthy tissue), microwave images show contrasts and allow for detection of abnormalities in the breast (Modiri, Goudreau, Rahimi, & Kiasaleh, 2017). Initial experiments found significant differences in the dielectric properties of normal breast tissue and tumours, whereas large scale experiments found that the differences were much smaller but still detectable (Kwon & Lee, 2016).

MI systems transmit microwave signals into the breast, and then measure the scattered signals reflected from the breast tissue (including tumours, if present) through reconstruction of the breast in an image. While the resolution becomes higher as the MI frequency increases, loss of penetration subsequently increases, making obtaining a clear image difficult. To account for this, a limit of upper frequency of the band is needed for appropriate penetration into the breast tissue – an ultra-wideband signal works in this manner (Kwon & Lee, 2016). Since the penetration loss of healthy fat tissue is less than 4dB/cm with microwave signal (centred at 6GHz), it is possible for low power signals to reach antenna on the other side of the breast.

In order to reduce false positives in MI of the breast, it is important to account for spatial and temporal changes in breast tissue, that is, biological variations due to hormonal changes or weight gain/loss, and inconsistencies in patient positioning or system noise (Modiri et al., 2017).

Wang (2017) details MI approaches that can be grouped as 'passive' and 'active'. 'Passive' MI uses radiometry to measure the temperature differences between normal and malignant tissues – but is not further discussed in the review, presumably because passive MI is not as advanced as active MI techniques but this is unclear (Wang, 2017). Microwave tomography<sup>14</sup> and radarbased microwave imaging are 'active' approaches that both use the scattering of microwave signals. The application of microwave tomography and radarbased microwave imaging as they relate to breast cancer screening are further described below.

#### Microwave tomography

With microwave tomography, an inverse scattering method is used to obtain an image through creating a complete map of the electrical properties of the breast tissue. The inverse scattering method uses scattering signals including diffraction from objects, and in doing so creates a map

<sup>&</sup>lt;sup>14</sup> Note: other forms of tomography are discussed in a separate section of this Horizon scan.

of permittivity and conductivity (Kwon & Lee, 2016; Modiri et al., 2017). The calculation required for the inverse method to obtain the breast image can take a lot of time (a total image acquisition time of 10 to 15 minutes and an image processing time of more than 10 hours) (Kwon & Lee, 2016). MI also requires an (undisclosed) heavy computational load (Wang, 2017).

Recent techniques to improve the accuracy and specificity of microwave tomography include using magnetic nanoparticles as a contrast agent, with compressive sensing techniques to show the magnetic contrast within the breast (Wang, 2017). Summarising results from other studies, Wang et al. noted that microwave tomography demonstrated that similar quality breast images can be obtained via a compressive sensing-based microwave tomography with 12 sensors, and via microwave tomography with 70 sensors. The compressive sensing-based system significantly reduced operation cost and data collection time; however, the authors did not further describe this (Wang, 2017).

#### Radar-based microwave imaging

Radar-based MI is a technique that specifies the location of one or multiple strongly scattering objects (tumours) within tissue (of the breast) (Modiri et al., 2017) through reconstructing the image using the reflected wave (Kwon & Lee, 2016). Wang describes five classifications for radar-based MI, including confocal microwave imaging (CMI), tissue sensing adaptive radar (TSAR), microwave imaging via space time (MIST), multi-static adaptive (MSA) MI, and holographic microwave imaging (HMI) (Wang, 2017). MI Table 1 (below) compares some of these techniques from prototype/experimental stages.

Technique	Ability to detect	Benefits	Limitations	Author
2D CMI	2mm tumours	Ability to generate high resolution images	Limited ability to discriminate against artefacts and noise	Hagness et al. (1999)
3D CMI	6mm tumours	Ability to generate high resolution images	Limited ability to discriminate against artefacts and noise	Hagness et al. (1999)
TSAR	4mm lesions	Can detect and localise lesions	Large reflections by skin, expensive electronics for real- time imaging	Fear et al. (2003)
MIST	4mm tumours	Significant performance due to ultra- wideband	Caused skin-breast artefacts in the images	Bond et al. (2003)
Near-field indirect HMI	Not reported	Real-time images at lower cost	More validation required on theory/proof of concept	Smith et al. (2014)
Far-field HMI	Not reported	No matching medium required, complex permittivity of object not required to calculate to generate an image – reduced imaging	Not discussed	Wang et al. (2014)

MI Table 1: Radar based microwave imaging techniques (Wang, 2017)

Technique	Ability to detect	Benefits	Limitations	Author
		construction time.		

Radar-based MI can also be done through frequency domain analysis. Kwon and Lee noted that there were several frequency domain measurement systems reported in the literature; however, this approach required costly equipment (eg, a vector network analyser (VNA)) and took a significant amount of time, which was a barrier to clinical use as patient movement could create artefacts in the imaging process (Kwon & Lee, 2016). Time domain measurement was another radar-based MI approach for breast imagining discussed by Kwon and Lee, though the only apparent clinical study has only been used for breast health monitoring with 13 healthy volunteers (Kwon & Lee, 2016). Barriers to using this approach were noted, including dependence on a high-precision pulse generator and a very high-speed oscilloscope which may require complicated switching circuits, as well as requiring a very fast sampling clock whereby a "small jitter" may blur resulting images (Kwon & Lee, 2016).

# 8.2. Summary of key findings

- The use of microwave imaging in breast cancer screening of asymptomatic women is still in the early stages. Much of the research focus is on identifying feasible microwave imaging systems that demonstrate sufficient sensitivity and detectability to support their use in further clinical testing. Research is most advanced in the use of ultrawideband frequency systems.
- There is currently no indication of the timeframe in which the full clinical potential of microwave imaging for breast cancer detection will be realised. Preliminary results from clinical studies indicates that this technology has the potential for clinical use.
- Current research is not able to identify whether microwave imaging as a breast imaging tool is able to reduce deaths due to breast cancer through early detection.
- Microwave imaging has not been incorporated into any national breast screening programs, nor are there any national position statements that have been released on their use in breast cancer screening for asymptomatic women.

# 8.3. Literature search results (number of studies returned)

Sixty-five studies relating to microwave imaging of the breast were returned from the literature search. Most of the studies were experimental, based on testing components of MI systems or testing system prototypes for imaging the breast. Most involved testing on non-humans (eg, breast phantoms). These studies were excluded after PICOT criteria were applied. We included two systematic reviews. A further review and one article were identified in a grey literature search.

### Systematic and/or literature reviews

Three reviews: Kwon & Lee, 2016; Modiri et al., 2017; Wang, 2017

RCTs

None identified.

Prospective studies

None identified.

Retrospective studies

None identified.

Grey literature

One news article: "Microwave breast screening getting closer," 2017.

# 8.4. Study findings and discussion

#### 8.4.1. What stage of development or trial is this innovation at?

There is no evidence to suggest that MI is at a stage where it could be effectively applied as a breast screening method. Three systematic reviews have discussed studies where clinical trials have taken place with symptomatic patient populations, but these studies noted that MI for breast cancer screening (asymptomatic or symptomatic) is at an early stage of development.

#### Systematic and/or literature reviews

MI technology has been developing over the last 30 years (Modiri et al., 2017). It is considered to be a promising new method for early-stage breast cancer detection as it may be able to detect small-sized tumours in the breast (Kwon & Lee, 2016). Clinical trial results have demonstrated that tumours with a size of 1cm diameter could be detected, which show's MI's potential as "an alternative or additional tool to mammography for diagnosing breast cancer" (Wang, 2017). Wang describes 'active' MI (microwave tomography and radar-based microwave imaging) as emerging mammography techniques for diagnosis (Wang, 2017). Furthermore, it is inferred that MI can obtain images quickly compared to other breast imaging methods (Kwon & Lee, 2016). However, the technology has not advanced to the point where a population-based RCT has been conducted for any MI technique with a breast screening application.

MI Table 2 (overleaf), from Modiri and colleague's 2017 systematic review, shows a summary of clinical MI breast imaging studies since 2013. Modiri et al. also summarised six phantom-based MI studies since 2013 (Modiri et al., 2017). Results are discussed in the following section.

MI Table 2: Summary of clinical MI studies since 2013 (Modiri et al., 2017, p.453)

Research group	Number of patients	Technique
Fear et al. (2013) University of Calgary, AB, Canada	8 patients	0.05 – 15 GHz Microwave imaging based on tissue-sensing adaptive radar
Meaney et al. (2013) Dartmouth College, NH	8 patients	0.7 – 1.7 GHz in 200 MHz increments Microwave tomography
Porter et al. (2016) McGill University, QC, Canada	13 patients	2 – 4 GHz Time-domain microwave imaging
Preece et al. (2016) Bristol University, Bristol, UK	86 patients	3 – 8 GHz Ultrawideband radar

As seen in MI Table 2 (above), the most comprehensive recent clinical study by Preece et al., had 86 enrolled patients. The patients were recruited from a breast care clinic for symptomatic women, and mainly had cysts and cancers, and rarely had hematoma, lipoma, or fibradenoma. They were scanned with a prototype MI design (MARIA M4), and the resulting images were compared blindly with ultrasound and mammography in terms of detection rate. The findings reported sensitivity rates from MI that were comparable to digital mammography for the total sample (74% compared with 77%, respectively), but which outperformed digital mammography in dense breast tissue (86% for MI compared with 69% for FFDM; Modiri et al., 2017).

While there are promising findings from early clinical studies of MI of the breast, it is noted that it is not at a clinical trial stage yet, as most of the research available has utilised breast phantoms (materially constructed with fatty and cancerous tissue or numerical), that suitable frequency ranges are not confirmed, and spatial resolution of microwave images of the breast are inadequate (Wang, 2017). Many challenges need to be addressed before MI can be implemented in clinical trials, such as improving algorithms for imaging, and implementation of hardware that is highly sensitive, compact, and cost effective with enough sensors and senor arrays to gain high resolution images. Furthermore, a highly dynamic system needs to be developed to capture the small differences in the scattered field or contrast agents to enhance malignant tissue (Wang, 2017). Additionally, (Modiri et al., (2017) stated that results from MI studies need to be reported in relation to current breast imaging standards, (eg, tumour size, location, metastasis extension) and MI slices need to be correlatable to other imaging modalities for accurate comparison).

A University of Waikato Professor of Engineering, Yifan Chen, is leading a large-scale clinical trial of MI as a breast screening tool in China, in partnership with ET Medical, a leading medical instrument company headquartered in Shenzhen, China. Recent reports about the trial include that:

"In the first stage of the trial, the system has been deployed to examine 11 healthy women with six of them having mammary hyperplasia. These preliminary tests confirm the system's effectiveness and suggest that responses from mammary hyperplasia can be identified successfully".

The trial began in December 2016, and the results were due to be reported in the Final Working Group and Management Committee Meeting of the EU Framework Programme Horizon 2020

COST Action TD1301 "MiMed" in Malta in June 2017 ("Microwave breast screening getting closer," 2017).

# 8.4.2. Does this innovation show high sensitivity and specificity for women with dense breasts and women who have had breast surgery/augmentation compared to digital mammography?

MI has showed some promise for better detection of breast cancer for women with dense breasts compared with digital mammography, but more research is needed to confirm these findings. In a clinical trial of 86 symptomatic patients, Preece et al. (2016) observed that one MI technique may have significant potential for detecting breast cancer for patients with dense breasts as it outperformed digital mammography in this group (Modiri et al., 2017). Kwon & Lee, (2016) stated that MIST methods have suffered from performance degradation when used for investigating dense breasts, however, development of a CMOS chip-based time-domain measurement microwave system shows promise as an imaging method for younger women with dense breasts.

Because Professor Yifan Chen's MI system study would include a large group of Asian women "...who usually have denser breasts than European women which makes it harder to detect tumours using X-ray or ultrasound" the study's findings would provide useful insight on sensitivity and specificity for women with dense breasts ("Microwave breast screening getting closer," 2017).

# 8.4.3. What is its considered potential clinical value in five years? In 10 years?

It is unclear from the literature reviewed for MI what the potential clinical value of any of the discussed techniques may have in the next five to 10 years. The articles reviewed all regarded MI of the breast as having potential as a breast cancer detection or diagnosis tool but did not specify within what period this might be realised.

Professor Yifan Chen hoped that commercialisation of MI breast screening technology could begin in the next one to two years ("Microwave breast screening getting closer," 2017).

# 8.4.4. What cost and safety findings have been reported?

### Systematic and/or literature reviews

MI systems utilise nonionizing radiation which is seen as safer than radiation exposure in digital mammography. There have been concerns about the use of ultrawideband frequency, but this has been assessed as safe up to 15 GHz in one study, as absorbed tissue in their MI setup was below prescribed standard limits (Modiri et al., 2017). Kwon and Lee (2016)discuss the utilisation of a specific absorption rate (SAR) for setting safety standards for the maximum allowed exposure to power to be absorbed in tissue per volume. Standard C95.1-1999 states that in the case of devices operating at 100 kHz to 6 GHz, exposure should not exceed a maximum of 1.6 W/kg for an average of 1g of human tissue. Compared to a mobile phone using the same frequency band, microwave breast cancer imaging should not be hazardous to a patient's health, because there is less exposure than a mobile phone.

With regards to cost, all included reviews described MI equipment/systems as inexpensive (Kwon & Lee, 2016; Modiri et al., 2017; Wang, 2017); however, there was little discussion of

evidence to support this. Wang suggested that as MI systems become more complex (eg, including the use of more sensors) and the use of coupling medium, costs can increase significantly (Wang, 2017). Kwon & Lee (2016) argued that a specific MI system (CMOS chipbased time-domain measurement microwave system) could be implemented at a lower price because it did not require a VNA, oscilloscope, or a pulse generator which could cost between \$10,000 and \$100,000 – a transceiver chip costs approximately \$100, so this imaging system prototype could cost less than \$3,000.

# 8.4.5. Is there evidence that this innovation is more acceptable to women (in general and by ethnic group) compared to digital mammography?

Studies have inferred that MI techniques may be more acceptable to patients undergoing screening as they generally do not require breast compression, as in digital mammography. Two reviews mentioned that an advantage of MI techniques was that breast tissue compression, which can cause discomfort for patients, is generally not required for MI (Kwon & Lee, 2016; Modiri et al., 2017). No commentary has been made around acceptability based on ethnicity or age of patients.

#### 8.4.6. Does this technology reduce deaths due to breast cancer through early detection?

Based on this review, there is no evidence to suggest that MI reduces deaths due to breast cancer through early detection.

# 8.4.7. Has this innovation been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

MI techniques have not been implemented into any national screening programs for breast cancer.

# 8.4.8. Has a national position statement been published about this innovation, and if so, what is the position? Is there a consensus in position statements?

No national position statements have been published about MI as a breast cancer screening tool.

# 9. MOLECULAR BREAST IMAGING

#### Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

*Safety*: No adverse effects associated with the use of molecular breast imaging (MBI) were identified; however, MBI is an invasive screening modality and not suitable for pregnant women due to the use of radiopharmaceuticals. The 2009 report concluded that the trade-off in cancer detection versus invasiveness of the MBI procedure would need to be considered before incorporating MBI into general population screening.

*Effectiveness*: Overall, the literature generally suggested that MBI can detect tumours with sensitivity and specificity equal to or better than mammography, although levels of specificity and sensitivity varied across studies. There was indication from one study that dual-head MBI systems may perform better than single-head systems.

*Cost*: Fees for MBI tests ranged from \$286.80 to \$834.90.

### 9.1. Molecular breast imaging as a breast cancer screening modality

Molecular breast imaging (MBI), also known as breast-specific gamma imaging (BSGI) or breast scintigraphy, is a functional imaging technique that uses a breast-specific gamma camera to create high resolution images with a small field-of-view. The literature is inconsistent in its use of the terms MBI and BSGI, sometimes using them interchangeably and sometimes using BSGI to refer to camera systems using a single-head detector and MBI for those using a dual-head detector. Because of its improved ability to detect smaller lesions, dual-head systems are generally now used for clinical MBI (Berg, 2016). For this reason, we use the term 'MBI' throughout, and will specify where results relate to single-head detectors (where this information is available).

The MBI technique represents an improvement on an earlier nuclear imaging technique called scintimammography, which used a traditional gamma camera rather than a breast-specific gamma camera and therefore had difficulty identifying cancers smaller than 1cm in diameter (Holbrook & Newel, 2015). These improvements in imaging equipment have led to higher resolution images and decreases in radiation dosage, introducing the possibility for MBI to be used in breast cancer screening programs (Adrada, Moseley, & Rauch, 2016).

Like positron emission tomography (PET), MBI uses the interaction between biological cells and administered chemical tracers (in this case the radiopharmaceuticals <sup>99m</sup>Tc sestamibi or <sup>99m</sup>Tc tetrofosmin) to detect cancerous tissue. In cancerous tissue, both delivery and uptake of the radiotracers are higher compared with surrounding healthy tissue. This means that the sensitivity of MBI is not affected by the density of breast tissue, scars from prior surgery, or implants (Sun, Wei, Yang, & Liu, 2013). That said, concerns with levels of radiation dosage compared to mammography are the primary barrier for incorporating MBI into routine screening programs (Hruska & O'Connor, 2013); however, some clinics in the USA are offering supplemental MBI to women with dense breast tissue who do not qualify for, or do not want, MRI (Hruska, 2017).

To conduct an MBI test, the breast is minimally compressed from both top to bottom (CC position) and side to side (MLO position) to limit movement, with the gamma camera placed in

direct contact with the surface of the breast. The radiotracer is then injected intravenously and imaging is performed within 5 minutes after injection (Adrada et al., 2016). Two images of 10 minutes duration are taken of each breast (one in each compression position), meaning the entire examination lasts approximately 40 to 45 minutes (Adrada et al., 2016).

The role of MBI in the early detection of breast cancer has not expanded at a fast rate, in part because of historical problems with image resolution, high levels of radiation dosage, and lack of an evidence base on its use in an asymptomatic population (Wahl, 2016). Current recommendations from the American College of Radiology do not support the use of MBI in routine screening, partially due to these limitations (Holbrook & Newel, 2015). However, recent advances in MBI technology are re-awakening interest in MBI as a screening modality for asymptomatic women, particularly those with more dense breasts (BIRADS 3 or 4). These advancements and their impact on MBI effectiveness are outlined below.

# 9.2. Literature search results (number of studies returned)

From the literature search, a total of 52 articles related to innovations in MBI were identified. Abstract contents were then reviewed and 35 of these articles were excluded for various reasons, including: use of non-human, high-risk or asymptomatic samples; a focus on diagnosis or treatment; or a focus on general cancer detection rather than specifically breast cancer. Sources that were not peer-reviewed, such as theses and dissertations, were also excluded. Three further articles were excluded because the full-text article could not be accessed (McPeak, 2014; Shermis, Redfern, Burns, & Kudrolli, 2017; Zardavas et al., 2017). Articles included in identified literature reviews were not separately reviewed for the current scan.

A total of 13 articles were then reviewed to answer the key research questions in relation to innovations in MBI for breast cancer screening in asymptomatic women.

#### Systematic and/or literature reviews

Twelve literature reviews: (Adrada et al., 2016; Berg, 2016; Even-Sapir, Golan, Menes, Weinstein, & Lerman, 2016; Ho, Jafferjee, Covarrubias, Ghesani, & Handler, 2014; Holbrook & Newel, 2015; Hruska, 2017; Hruska & O'Connor, 2013; Surti, 2013; Tafreshi, Kumar, Morse, & Gatenby, 2010; Wahl, 2016; Weigert & Kieper, 2012; Yang, 2011).

RCTs

None identified

#### **Prospective studies**

One study: (Hruska et al., 2015)

#### **Retrospective studies**

None identified outside of those included in identified literature reviews.

**Grey literature** 

None identified

# 9.3. Summary of key findings

- Results from largely retrospective studies are promising regarding the effectiveness of MBI for screening purposes; however, more large-scale multi-centre prospective studies are required before conclusions can be drawn about the role of MBI in the early detection of breast cancer in asymptomatic women.
- There is currently no indication of the timeframe in which the full clinical potential of MBI for breast cancer detection will be realised. High sensitivity and moderate specificity rates have been found in existing research, with improvements in the detection of sub-centimetre lesions. Rates of cancer detection are improved when MBI is used as a supplement to screening mammography.
- Despite initial issues with radiation dosage, more recent MBI systems are demonstrating good detection ability with reduced effective radiation dosages (approximately 2.4 mSv). One study found that the cost of supplementing screening with MBI is higher per examination than for mammography alone, however combining MBI with mammography leads to a decrease in cost per cancer detected.
- There is growing evidence for the increased efficacy of MBI for the early detection of cancer for women with dense breasts (BIRADS 3 or 4) compared with screening mammography, although these results may underestimate the impact of MBI for women with more dense breasts due to categorisation issues.
- No specific information on the acceptability of MBI to women was identified; however, the average length of time for an MBI examination (40-45 minutes) is substantially longer than the length of time needed for a mammogram.
- MBI has not been incorporated into any national screening programs, nor are there any national position statements that have been released on its use in breast cancer screening for asymptomatic women.
- Current research is not sufficient to be able to identify whether MBI is able to reduce deaths due to breast cancer through early detection.

# 9.4. Study findings and discussion

#### 9.4.1. What stage of development or trial is this innovation at?

Although earlier literature raised a number of concerns that prevented the adoption of MBI into routine screening programs, more recent research is promising in terms of advancements in MBI effectiveness and safety; however, more large-scale multi-centre prospective studies are required before conclusions can be drawn about the role of MBI in the early detection of breast cancer in asymptomatic women (Hruska, 2017). A large focus of current research is on reducing the effective radiation dosage of MBI whilst retaining sufficient image quality to ensure high rates of sensitivity and specificity for the early detection of cancer in asymptomatic women.

#### 9.4.2. What is its considered potential clinical value in five years? In 10 years?

#### Systematic and/or literature reviews

No specific timeframes are mentioned in the literature regarding the clinical potential of MBI; however, with increases in the number of industry partners manufacturing and selling MBI equipment, improvements in overall MBI design and techniques are expected during the next five years (Hruska, 2017). Current research regarding the effectiveness of MBI systems developed in recent years is summarised below.

Results from retrospective studies are promising in terms of the effectiveness of MBI in breast cancer detection. Three studies identified in Holbrook & Newel's (2015) literature review reported relatively high rates of sensitivity and moderate specificity, ranging from 91% to 96% and 60% to 77%, respectively. They also report on a meta-analysis conducted on studies using symptomatic samples, which found MBI calculated a sensitivity of 95% and a specificity of 80%; however, it is important to note that the use of symptomatic samples may artificially inflate examination accuracy.

Together, these results suggest that recent advancements in MBI technology have improved the potential for MBI to be successfully used in the early detection of cancer, particularly given previous improvements in its ability to detect relatively small lesions (National Horizon Scanning Unit, 2009). That said, sensitivity is still found to be lower with lesions of smaller sizes. One study identified in Berg's (2016) literature review found that although overall sensitivity for MBI in detecting 357 known lesions was 87%, this reduced to 71% for lesions 5mm or smaller and 76% for lesions 6 to 10mm in diameter; sensitivity was 93% for lesions 11 to 15mm in diameter, and 98% for lesions 16mm or larger. Reduced sensitivity for relatively smaller lesions was also found in a study of 67 patients identified in Even-Sapir et al.'s (2016) literature review, which found sensitivity of 90% for MBI in overall lesion detection, but 60% sensitivity in lesions smaller than 1cm. Additionally, one retrospective study identified in Wahl's (2016) review found that the sensitivity of MBI for detecting invasive lobular breast cancer (86%).

Overall, these findings suggest that there are subsets of cancers – relatively small lesions and lobular cancer – that may remain undetected by MBI. That said, the evidence in this area is not clear-cut, with an additional study identified in Even-Sapir et al.'s (2016) review finding high rates of sensitivity (89%) for lesions smaller than 1cm, using a sample of 139 patients. Furthermore, there are several studies indicating that slight alterations to the MBI examination process or image interpretation may improve overall effectiveness. For instance, there is some indication that routinely taking two sets of images in an MBI examination would improve cancer detection. One study identified in Holbrook & Newel's (2015) literature review found that the specificity of MBI significantly increased from 83 to 95% (p = .008) when a second set of images was taken one hour after administering the radiotracers. The authors of this study concluded that the time delay allowed for improved differentiation between benign and malignant lesions. However, an additional study identified in the review believed that producing a second set of images would not be acceptable for patients due to increased duration of breast compression.

Another possible improvement that could increase MBI effectiveness is the use of a semiquantitative approach to image interpretation. In one study identified in the Holbrook & Newel (2015) review, lesion-to-background uptake was calculated and used in combination with visual analysis to identify potential cancers. Using a ratio of 1.5 or greater led to an increase in the specificity of MBI from 82% to 92%, with this improved specificity being higher than that obtained for mammography (82%).

Because more research is required to establish whether screening using MBI is able to reduce mortality due to early cancer detection, much of the research looking at the effectiveness of MBI focuses on the improved accuracy of screening when using MBI as a supplement to mammography (Hruska, 2017). One study identified in Holbrook & Newel's (2015) literature review used MBI to assess 94 women who had normal findings on screening mammography. This study found that 16 of these women had abnormal MBI findings, with two of these abnormal findings being confirmed as cancers. Two other studies identified in Hruska's (2017) review found that supplementary MBI screening resulted in an incremental cancer detection rate ranging from 13.1 to 16.5 cancers per 1,000 women screened, compared with mammography alone. Although these results are promising, more information is needed before it can be determined whether MBI screening could be used as an alternative to mammography, rather than a supplement (Hruska, 2017).

#### 9.4.3. What cost and safety findings have been reported?

#### Systematic and/or literature reviews

High doses of radiation remain a concern for MBI, with most research using effective dosages in the range of 6.3 to 9.4 mSv (Adrada et al., 2016; Berg, 2016). Additionally, because the radiotracers are systematically distributed throughout the whole body, this radiation exposure is not limited to the breast but instead places most of the radiation burden on the upper and lower intestines (Adrada et al., 2016; Holbrook & Newel, 2015).

That said, recent advances in MBI technology have led to a decrease in the administered dose of radiotracers required for imaging, and therefore the radiation dosage. The literature review conducted by Adrada et al. (2016) identified a 2010 study which found that improved MBI systems could produce images of sufficient quality with an effective radiation dose of 2.4 mSv. This result was replicated in an additional study identified in Berg's (2016) review that found that an improved MBI system was able to maintain cancer detection performance using an effective dosage of approximately 2.4 mSv. Although promising, this effective dosage is still higher than the 2.0 mSv (or 2.0 mGy) mean glandular dose guideline recommended for FFDM by the BreastScreen Australia National Accreditation Standards (BreastScreen Australia, 2015); lower dosage still needs to be achieved before MBI can be incorporated into standard screening procedures (Adrada et al., 2016). This is a large focus of current research, with one team currently trialling an MBI system using an effective dose of approximately 1.3 mSv (Berg, 2016).

#### **Prospective studies**

The cost of MBI performed as an adjunct to mammography screening was estimated in one prospective study of 1,585 women with dense breasts (BIRADS 3 or 4; Hruska et al., 2015). In total, combined screening mammography plus MBI led to 279 diagnostic workups (18% of the sample), 67 biopsies (4% of the sample), and yielded 19 malignancies (PPV<sub>3</sub> = 28%). This was compared with five malignancies identified through mammography alone (PPV<sub>3</sub> = 25%). The addition of MBI increased the cost of screening per patient from \$176 for mammography alone to \$571 for supplementary MBI. However, because the number of malignancies detected was higher for the MBI and mammography combination, the cost per cancer detected was lower overall for the combination (\$47,597) compared with mammography alone (\$55,851). This

finding suggests that the improved rate of detection using MBI as a supplement to mammography may compensate for the higher direct costs.

# 9.4.4. Does this innovation show high sensitivity and specificity for women with dense breasts and women who have had breast surgery/augmentation compared to digital mammography?

#### Systematic and/or literature reviews

Existing research suggests that MBI is a highly promising imaging modality for women with mammographically dense breasts (Hruska & O'Connor, 2013). Two studies identified in Holbrook & Newel's (2015) literature review found that MBI was significantly more sensitive for women with dense breasts (with dense breasts defined using the BIRADS 3 or 4 categories) than mammography. The sensitivity of MBI in these studies ranged from 82% to 83%, whereas the sensitivity of mammography ranged from 27% to 44%. No significant differences in specificity were identified. Interestingly, one of these studies found that combining MBI with mammography produced even higher sensitivity for women with dense breasts (91% compared with 82% for MBI alone and 27% for mammography alone).

An additional study identified in Holbrook & Newel's (2015) review found that MBI was able to detect known cancers with equal sensitivity for women with (95%) and without (97%) dense breasts (defined as BIRADS 3 and 4). Furthermore, in 20 cases of cancer that were undetected by mammography, MBI was able to detect 100% of cancers in less dense breasts and 93% in more dense breasts. This increase for women with dense breasts needs to be offset against the higher rates of radiation used in MBI. This is a particular concern for women with dense breasts due to their already increased risk of malignancy (Holbrook & Newel, 2015).

Results in this area are promising, with one study identified in Adrada et al.'s (2016) literature review able to demonstrate high levels of sensitivity (81%) and specificity (94%) using single-head MBI for screening 1,585 women with dense breasts (defined as BIRADS 3 or 4), using a radiation dosage of 2.4 mSv. This result was replicated at the same dosage level in an additional study identified in the review, which found an incremental detection rate of 7.7% for single-headed MBI compared with mammography, using a community-based sample of 1,696 women with dense breasts. This dose is higher than the effective dose for mammography and DBT (1.2 mSv), however the increase in cancer detection rate potentially offsets the slightly higher radiation dose for this population (Adrada et al., 2016).

It is also important to note that the results found above may underestimate the improvement gained by supplementing FFDM with MBI. In addition to this horizon scan, *Allen + Clarke* has also prepared a literature on breast density and the use of mammography for breast cancer screening for asymptomatic women. We note that the use of BIRADS to assess breast density can result in unreliable allocation between BIRADS categories 2 and 3. This is because density classification can be affected by factors like hormone levels and weight and inter/intra reader variability. It is possible for women to be classified as having non-dense breasts (2) in one mammogram but be reclassified to having more dense breasts in the next mammogram (and vice versa). This creates a level of unreliability that could account for smaller-than-expected incremental increases in CDR between women with more dense or less dense breasts. It may be that density classifications which report CDR, recall and false positives by 25th percentile (very dense) and 75th percentile (very fatty) could result in clearer (and possibly truer) incremental differences in CDR by density.

# 9.4.5. Is there evidence that this innovation is more acceptable to women (in general and by ethnic group) compared to digital mammography?

#### Systematic and/or literature reviews

No specific information on the acceptability of MBI to women was identified; however, the average length of time for an MBI examination (40-45 minutes) is substantially longer than the length of time needed for mammography, which may be undesirable to some individuals (Holbrook & Newel, 2015).

#### 9.4.6. Does this technology reduce deaths due to breast cancer through early detection?

Current research is not sufficient to be able to identify whether MBI is able to reduce deaths due to breast cancer through the early detection of cancer in asymptomatic women (Holbrook & Newel, 2015).

# 9.4.7. Has this innovation been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

Molecular breast imaging has not been incorporated into any national screening programs.

9.4.8. Has a national position statement been published about this innovation, and if so, what is the position? Is there consensus in position statements?

No national position statements on the use of MBI in breast cancer screening were identified in the literature search.

### **10. SPECTROSCOPY**

#### Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

The Bulletin on new and emerging technologies for breast cancer detection published in 2009 discussed early research and development of surface-enhanced laser desorption ionisation timeof-flight mass spectroscopy (SELDI-TOF) as a tool for analysing breath testing samples as an emerging biomarker test for breast cancer. "SELDI-TOF is capable of identifying proteins over a large size range and is sensitive down to the femtomolar level" (National Horizon Scanning Unit, 2009, p.60).

# **10.1.** What is spectroscopy?

Spectroscopy is the study of the interaction between matter and electromagnetic radiation. Spectroscopic techniques in biomedical applications (including in application for breast imaging) are recognised by a number of terminologies including, for example, Raman spectroscopy (RS; Jermyn et al., 2016), optical mammography, diffuse optical imaging, near-infrared spectroscopy, optical transillumination spectroscopy (Akbari Sari, Mobinizadeh, & Azadbakht, 2013), radiofrequency spectroscopy, and magnetic resonance spectroscopy (which is used in conjunction with magnetic resonance imaging, typically for diagnostic and treatment purposes).

The literature regarding spectroscopy's potential application in breast cancer screening is limited at this time. As such, this horizon scan only focuses on the terminology/techniques outlined in those articles – though we acknowledge that other spectroscopic technologies are in development. While spectroscopic techniques are labelled differently, we have assumed that fundamentally they share a similar basis. For the avoidance of doubt, when we mention specific spectroscopic techniques (eg, RS), we will label it as such, and where we mention "spectroscopic techniques" we discuss the techniques more generally as they are mentioned in the literature. Electrical impedance spectroscopy is discussed in the tomography section of this horizon scanning paper (*NB* it is another name for electrical impedance tomography).

RS techniques and optical mammography are the most comprehensively described in the literature relating to breast cancer screening.

#### Raman spectroscopy

This is an optical technique with the capability of non- or minimally-invasively determining molecular information from biological tissue. Inelastic scattering of light was first observed by Indian Nobel Laureate C V Raman in 1928 but was not reported for biomedical applications until 1970. Subsequent improvements in light sources and signal detection have made possible the emergence of RS techniques across a wide range of applications (Jermyn et al., 2016).

Most light scattering in tissue is elastic (that is, there is no energy exchange between the photons and the molecules – Rayleigh scattering). However, in tissue, typically only one in several million photons undergoes inelastic scattering through energy exchange with the vibrational or rotational modes of molecular bonds. This low probability makes detection of Raman-scattered light challenging, especially on the large background of elastic scattering as well as fluorescence and phosphorescence from the tissue. UV or short-wavelength visible light gives the largest Raman signal from tissue; however, this is more than offset by the high fluorescence background, which falls off markedly in the near-infrared range. Compared with fluorescence from biomolecules, which is characterized by a few, rather broad features (tens of nm wide), Raman spectra from tissue typically have multiple sharp peaks (~few nm wide), allowing the broad fluorescence background to be subtracted. The spectral shift between the incident light and the Raman spectroscopic light is also small, making spectral separation from the dominant elastic-scattering light difficult (Jermyn et al., 2016).

Technological advances have led to increased signal and contrast, enabling shorter and more clinically practical acquisition times. Advances include stable diode laser sources, high-throughput and high-resolution spectrographs and sensitive array detectors. Clinical practicality of RS has also been significantly enhanced using fibre optic light delivery and collection, as well as small and high-efficiency spectral filters to remove background light. In addition, powerful spectral analysis and classification algorithms have been developed and implemented on portable computers (Jermyn et al., 2016).

In oncology, RS has been shown to be highly sensitive to the altered 'molecular signatures' of many different cancers (Jermyn et al., 2016). RS techniques for disease detection include Spontaneous Raman spectroscopy (SpRS), Confocal Raman spectroscopy, Spatially offset Raman spectroscopy (SORS), Coherent Raman spectroscopy (coherent RS) and coherent anti-Stokes Raman spectroscopy (CARS), and Surface enhanced Raman spectroscopy (SERS). Each of these techniques, depending on their biomedical application, differ in terms of light source, acquisition time, spatial resolution and signal/background (Jermyn et al., 2016).

Many advantages have been identified with Raman spectroscopy. It is described as well-suited to clinical translation as it does not use exogenous contrast agents, thereby enabling easier integration into clinical workflow and reduced regulatory barriers. Furthermore, recent advances in instrumentation and spectral analysis have allowed for sub-second acquisition times and high diagnostic performance – the ability to make near real-time measurements is important for in vivo interventional applications. Raman-based imaging and spectroscopy systems can also be made portable and clinically ergonomic, for example, hand-held tools have been developed specifically for in vivo Raman spectroscopy (Jermyn et al., 2016).

However, Jermyn and colleagues have identified several challenges in the clinical implementation of Raman spectroscopic techniques. The Raman signal is typically dominated by elastic light scattering and intrinsic tissue autofluorescence. Techniques such as SERS and Coherent RS (noted above) aim to address this issue with instrumentation changes that can substantially improve the measured signal-to-background ratio. However, use of these techniques may impede clinical translation, as they often come at the cost of increased acquisition time and complex instrumentation cost, for example. Another challenge is that the diagnostically relevant signal is also spread across the spectral bands, which limits the use of standard univariate analysis methods – though sophisticated machine-learning techniques have been developed that can take advantage of the rich spectral information (Jermyn et al., 2016). As with the different types of Raman spectroscopic techniques described above, choosing the most appropriate technology to use often involves trade-offs between field-of-view, spatial resolution, spectral resolution and acquisition time – more research evaluating the different techniques for different applications is needed (Jermyn et al., 2016).

#### Optical mammography

This technique, which may involve slight breast compression, employs near-infrared light – which is usually in the wavelength range of 670-970nm – to non-invasively probe the breast. Over this spectral region, haemoglobin is the dominant contrast agent. "Haemoglobin concentration and its oxygen saturation are two key parameters that may allow not only the detection of cancer, but also the discrimination between malignant and benign breast lesions using optical methods" (Akbari Sari et al., 2013, p.45). There are three optical mammography technologies:

- Continuous Wave (CW) in which light dilution at breast surface is measured
- Time Domain (TD) which involves illuminating the breast with short light pulses
- Frequency Domain (FD) in which system flows continuous light currents in oscillation fields at higher frequencies in the tens to hundreds of MHz into the breast tissue (Akbari Sari et al., 2013).

# **10.2.** Summary of key findings

- The potential use of spectroscopic techniques in breast cancer screening of asymptomatic populations is not currently clear, with recent research focusing on developing and refining technology for clinical testing.
- There is currently no indication of the timeframe in which the full clinical potential of spectroscopic techniques for breast cancer detection of asymptomatic people will be realised; however, results from clinical studies indicates that optical mammography has the potential for clinical use in high risk groups. Research and development into spectroscopy for breast cancer detection is emerging and advancing.
- Current research is not sufficient to be able to identify whether spectroscopy as a screening procedure can reduce deaths due to breast cancer through early detection.
- Spectroscopic techniques have not been incorporated into any national breast screening programs, nor are there any national position statements that have been released on their use in breast cancer screening for asymptomatic women.

# **10.3.** Literature search results (number of studies returned)

A total of 25 articles were returned from the literature search that related to different forms of spectroscopy. None of the articles returned related to screening asymptomatic populations with a screening application, therefore, we will only give a brief overview of the technologies in relation to the research questions, as applicable. All but two articles were excluded as they did not meet the criteria for inclusion (i.e., studies were excluded because they had non-human subjects, were a combined modality with another breast detection/imaging tool, and/or related to diagnosis or treatment of breast cancer).

#### Systematic and/or literature reviews

Two systematic reviews related to RS (with breast and other applications) (Jermyn et al., 2016) and optical mammography, including diffuse optical imaging and near-infrared spectroscopy (Akbari Sari et al., 2013), were included.

RCTs

None identified.

#### **Prospective studies**

None identified.

#### **Retrospective studies**

None identified.

#### **Grey literature**

A search of the U.S. National Library of Medicine <u>ClinicalTrials.gov</u> website for "breast cancer" and "spectroscopy" identified 65 related clinical trials, with one related to spectroscopy for breast cancer screening purposes: 'Light Scattering Spectroscopy for Detection of Breast Cancer' at the USC Norris Comprehensive Cancer Center in Los Angeles, California. The trial involves a light-scattering spectroscopy procedure on healthy people with no evidence of disease, people with Stage II and Stage IIIA, B, and C breast cancer<sup>15</sup>.

# 10.4. Study findings and discussion

#### 10.4.1. What stage of development or trial is this innovation at?

There have been a number of technological advances in spectroscopic techniques generally over the past 20 years (Jermyn et al., 2016), which have been discussed above. Research so far has not focused on trialling the application of spectroscopic techniques for population-based breast screening of asymptomatic women, though there is one clinical trial for the application of light-scattering spectroscopy for breast cancer detection currently recruiting<sup>6</sup>.

In terms of its application for breast cancer, techniques under development include optical coherence tomography, and non-optical methods such as radiofrequency spectroscopy (Jermyn et al., 2016). There have been several studies involving the application of RS to breast cancer diagnosis – but most have used ex vivo human tissue sampling or animal models (Jermyn et al., 2016). Early human studies using Raman spectroscopy, however, show some promise:

"The first in vivo RS [Raman spectroscopy] study, using <1 s integration time, in patients was reported by Haka et al. in 9 patients undergoing partial mastectomy and claimed 100% accuracy in classifying tissue as tumor versus non-tumor. However, the statistical significance of this result

<sup>&</sup>lt;sup>15</sup> Available at:

https://clinicaltrials.gov/ct2/show/NCT01755208?term = spectroscopy&cond = breast + cancer&rank = 6

was limited (Haka et al. 2006). Three years later, the same group published a prospective analysis on ex vivo samples showing 83% sensitivity and 93% specificity (Haka et al. 2009)" (Jermyn et al., 2016).

Akbari et al. (2013) suggested that current clinical application of optical mammography could be as a diagnostic supplement alongside conventional mammography for differentiating benign and malignant tumours. Optical mammography was reported to be able to discriminate breast cancer and non-cancerous breast tissue with 96% sensitivity and 93% specificity.

## 10.4.2. What is its considered potential clinical value in five years? In 10 years?

The potential for clinical use of spectroscopic techniques as a breast cancer screening modality in asymptomatic populations is uncertain within the next five to 10 years, due to the limited number of prospective studies or RCT studies in this population. Akbari Sari et al. (2013) stated that further studies involving larger sample sizes are needed before optical imaging can be used as a reliable tool for predicting breast cancer. Jermyn et al. (2016) noted that as understanding of the molecular basis of disease improves, that information can assist with the design of optimal Raman spectroscopy systems.

Jermyn and colleagues stated in relation to a diagnostic application of Raman spectroscopy that while:

"...there have been dozens of single center studies using Raman technology in human applications, there have thus far been a very limited number of prospective studies, no blinded clinical studies, and no multicenter studies conducted to evaluate clinical impact. Moreover, no companies have obtained regulatory approval (eg, FDA in the USA, CE Mark in Europe) for the use of Raman-based instruments in clinical practice, with the exception of Verisante for skin lesion diagnosis (Canada and Europe)" (Jermyn et al., 2016).

However, Jermyn and colleagues also said that increased portability, convenience and affordability of Raman spectroscopy technology made it increasingly viable for screening applications (Jermyn et al., 2016).

### 10.4.3. What cost and safety findings have been reported?

Raman spectroscopic techniques were described as having complex instrumentation costs in one article (Jermyn et al., 2016) and other spectroscopic techniques were described as cost-effective in the other (Akbari Sari et al., 2013), however, there was no evidence reported to support either of these statements.

In terms of safety, infrared imaging and Raman spectroscopy were described as safe and nonionized techniques (Akbari Sari et al., 2013; Jermyn et al., 2016).

# 10.4.4. Does this innovation show high sensitivity and specificity for women with dense breasts and women who have had breast surgery/augmentation compared to digital mammography?

Spectroscopic techniques were discussed as a potentially useful application for high risk, young women with dense breasts – the technology may serve as a viable, cost effective and nonionized alternative for screening patients who like women with more dense breasts because spectral

analysis can predict radiographic density (Akbari Sari et al., 2013), but there were no sensitivity or specificity results available.

There was no indication whether spectroscopic techniques would have an application for women who have had breast surgery or augmentation.

# **10.4.5.** Is there evidence that this innovation is more acceptable to women (in general and by ethnic group) compared to digital mammography?

There was little evidence to suggest that spectroscopic techniques were an acceptable procedure for women undergoing breast screening, although both reviews discussed these techniques as non-invasive (Akbari Sari et al., 2013; Jermyn et al., 2016).

#### 10.4.6. Does this technology reduce deaths due to breast cancer through early detection?

Based on this review, there was no evidence to suggest that spectroscopy reduced deaths due to breast cancer through early detection.

# **10.4.7.** Has this innovation been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

Spectroscopy has not been implemented into any national screening programs for breast cancer.

# **10.4.8.** Has a national position statement been published about this innovation, and if so, what is the position? Is there a consensus in position statements?

No national position statements have been published about spectroscopy as a breast cancer screening tool.

### **11. THERMOGRAPHY**

#### Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

The studies identified by the 2009 report were, overall, poor quality or reported results that could not be generalised to asymptomatic women. Good sensitivity but low specificity was common across these studies, which suggested an unacceptably high level of false positives in an asymptomatic population. The report noted that thermography was a safe and non-invasive screening tool. It did not draw any conclusions about cost. Overall the report found little evidence to support thermography as a screening tool for the early diagnosis of breast cancer in asymptomatic women.

Cancer Australia has published a public statement against the use of thermography for the early detection of breast cancer, noting that there was insufficient evidence to support the use of thermography as a screening tool (Cancer Australia, 2010).

# 11.1. What is thermography?

Thermography is often referred to as thermal imaging, infrared imaging, thermograms, or infrared thermograms. For simplification, thermography is used to describe any of the above.

Thermography is a breast imaging modality that measures radiation from the surface of the breast to detect localised temperature variations which could be the result of cancerous or precancerous cell expansion (Kandlikar et al., 2017). Infrared cameras with specialised lenses and sensors are used to record thermal images (thermographs). When interpreting thermographs, asymmetry between breasts and the determination of areas of a breast with a high level of blood perfusion are key indicators of breast abnormality (Borchartt, Conci, Lima, Resmini, & Sanchez, 2013).

Thermographs must be taken under carefully controlled environmental conditions to produce the most accurate results (Kandlikar et al., 2017). There are no standard protocols for this, but it generally involves the thermographs being taken in a temperature-controlled room with no uncontrolled heat or light sources (such as sunlight coming through a window). Prior to screening, the patient is often required to avoid using lotions on her chest, remove outer layers of clothing and sit in a temperature-controlled waiting room for >15 minutes.

Three to five images are typically taken of the breasts: a front view, views of each side, and, in some cases, two oblique views (Etehadtavakol & Ng, 2013). These capture the entire chest, underarms and lymph region. Images are usually taken with the subject in either a standing or seated position with their arms raised above their head, on their hips, or on arm rests.

## 11.2. Summary of key findings

- There is not a large body of evidence that supports the use of thermography as a tool for breast screening in asymptomatic women. Most studies use small sample sizes and the results vary significantly. Much of the research is focused on different methods and technologies for obtaining and interpreting thermographs rather than the actual use of thermography for a screening purpose.
- Current research is not sufficient to be able to identify whether thermography is able to reduce deaths due to breast cancer through early detection.
- Thermography has not been incorporated into any national screening programs. Importantly, national position statements discouraging thermography as a breast screening tool have been issued in the United States and Australia; however, these are often based on quite old literature and do not reflect the developments in the evidence that have occurred in more recent years.

# **11.3.** Literature search results (number of studies returned)

From the literature search a total of 47 articles relating to thermography were identified. Abstract contents were reviewed, and 22 articles were excluded for numerous reasons, including: relating to risk assessment or diagnosis rather than cancer detection; involving only non-human subjects or symptomatic participants; or being of inadequate quality. A further six articles were excluded because they could not be located.

A total of 19 documents were reviewed to answer the key research questions in relation to the use of thermography for breast cancer screening in asymptomatic women.

#### Systematic and/or literature reviews

Four reviews: Borchartt et al., 2013; Etehadtavakol & Ng, 2013; Fitzgerald & Berentson-Shaw, 2012; Kandlikar et al., 2017

#### RCTs

None identified

#### **Prospective studies**

Three studies: Rassiwala et al., 2014; Sella et al., 2013; Venkataramani et al., 2015

#### **Retrospective studies**

Nine studies: Acharya, Ng, Tan, & Sree, 2010; Francis, Sasikala, Bharathi, & Jaipurkar, 2014; Francis, Sasikala, & Saranya, 2014; Gerasimova et al., 2013, 2014; Kermani, Samadzadehaghdam, & EtehadTavakol, 2015; Lanisa, Cheok, & Wee, 2014; Madhu, Kakileti, Venkataramani, & Jabbireddy, 2016; Rastghalam & Pourghassem, 2013

#### **Grey literature**

Three position statements: Cancer Australia, 2010; National Screening Unit, 2010; US Food and Drug Administration, 2017
### 11.4. Study findings and discussion

#### 11.4.1. What stage of development or trial is this innovation at?

Thermography has been both investigated and utilised as a breast cancer screening tool for decades. Advancements in thermograph capturing and interpretation technology mean that it continues to be the subject of research and development. The literature indicates that there is no consensus around best practice processes and practices for thermography for breast cancer screening, suggesting that it would require significantly more development and testing before it could be used confidently as a tool for the early identification of breast cancer in asymptomatic women. The two key areas of research highlighted in the literature are technologies for capturing thermographs, and approaches to image processing. These are discussed below.

#### Thermography technologies

In a literature review of thermography technology, Kandlikar et al. (2017) explain that the sensitivity of **infrared cameras** to temperature variations has significantly improved since the 1980s. Modern cameras can achieve a temperature difference of <.02 K compared to 0.3 K in early generation cameras. This means that the thermographs captured by these cameras are of better quality and are easier to analyse than early breast screening thermographs.

**Dynamic infrared thermography** captures images over a given time. Dynamic thermographs are often undertaken following the application of a cold stress to the breasts, with breasts being exposed to cold temperatures of between 5°C and 20 °C for several minutes (Kandlikar et al., 2017). The purpose of dynamic infrared thermography is to detect variations in temperature rhythms produced by cardiogenic and vasomotor frequencies (Gerasimova et al., 2013). Kandlikar et al. (2017) summarised four studies using dynamic infrared thermography with sensitivity ranging from 53% to 97%. As these studies were conducted on symptomatic populations and used a range of analysis software, it is difficult to apply these findings to the breast screening of asymptomatic women.

**Rotational thermography** is a technique that has been developed to overcome the situation where parts of the lower posterior breast are obscured in thermographs because of breast sag. The technique involves the breast being imaged from multiple views using a rotating robotic arm with an affixed infrared camera (Francis, Sasikala, Bharathi, et al., 2014). In a retrospective study of 24 women with healthy breasts and 12 patients with malignant breast cancer, Francis et al. used rotational thermography with images taken before and after a cold stress, using texture analysis (outlined below). They found 83.3% accuracy under the precooled condition, and 70.8% accuracy after cooling. They concluded that while cooling did not improve the effectiveness of the imaging technique, with further research rotational thermography may have potential as an adjunct screening tool.

A prospective study by Sella et al. (2012) screened a sample of 434 women using a high resolution **three-dimensional infrared imaging device** (3DIRI) coupled with multiparametic computer analysis. The sample was divided into two subgroups: a control group of 256 women with negative results from a routine mammogram and 178 women with either breast or DCIS. The study found that 3DIRI had a sensitivity of 86% for detecting breast cancer compared to 75% sensitivity for mammography.

#### Thermography image processing

A key area of thermography research explores different methods for analysing thermographs. These include:

- **Texture analysis** a common method for analysing thermographs involves measuring the smoothness, coarseness, and regularity of pixels in a thermograph to aid in the identification of normal and malignant breast conditions. Acharya et al. (2010) fed texture features to a Support Vector Machine classifier for automatic classification of normal and malignant breast conditions in a study of 50 thermograms, with 25 from women with malignant breast cancer and 25 from subjects with healthy breasts. The study found 85.71% sensitivity and 90.48% specificity.
- Automatic colour segmentation primary breast thermographs are in grey scale, and this can make it difficult for the human eye to identify subtle differences in shade. A range of algorithms have been developed to automatically segment images, and these continue to be refined to effectively separate and distinguish between clusters of pixels (Kermani et al., 2015; Lanisa et al., 2014; and Venkataramani et al., 2015).
- **Wavelet -based multifractal analysis** methods using wavelets have been developed to perform multifractal analysis of cardiogenic and vasomotor frequencies in dynamic infrared thermography (Gerasimova et al., 2013, 2014).
- Automated asymmetry analysis methods have been developed which employ segmentation algorithms to calculate and compare the curvature of breasts in thermographs (Rastghalam & Pourghassem, 2013). In a retrospective study of 28 thermographs from nine different medical clinics, Rastghalam and Pourghassem used an algorithm based on spectral probable features to perform asymmetry analysis. They reported high levels of accuracy but do not include sensitivity or specificity rates. The authors found lower levels of accuracy when interpreting low resolution images. In a small study of 22 women, 11 with breast cancer and 11 with healthy breasts, Francis et al. (2013) tested a curvelet transform based feature extraction method for the automatic detection of breast abnormalities. The method had a sensitivity of 81.82% and specificity of 100%.
- **Extraction of medically interpretable features** Madhu et al. (2016) developed an approach for the automatic extraction of medically interpretable features that are designed to differentiate between malignant and non-malignant breast conditions. The objective of this approach is to increase thermography specificity. In their study of thermographs from 265 subjects, including 78 with cancer and 53 with benign conditions, Madhu et al. achieved 99% specificity and 100% sensitivity.

#### 11.4.2. What is its considered potential clinical value in five years? In 10 years?

The potential clinical value of thermography in five to 10 years is not explicitly discussed in the literature. Specificity and sensitivity of thermography as a breast cancer screening tool varies considerably across the studies identified in the literature. Sensitivity ranged from 53% (Kandlikar et al., 2017) to 99% (Madhu et al., 2016) and specificity was found to be as low as 12% (Fitzgerald & Berentson-Shaw, 2012) and as high as 100% (Madhu et al., 2016; Rastghalam & Pourghassem, 2013). It is likely that this is due to the range of thermography technologies and processes used in the studies and the use of small sample sizes.

With the further development and analysis of thermography technologies and image processing techniques for breast screening asymptomatic populations (outlined above) the potential clinical value of thermography may become clearer and more acceptable. At present, the usefulness and effectiveness of thermography for screening asymptomatic women for the early detection of breast cancer is not well established in the literature, with a lack of large-scale studies and randomised control trials.

#### Systematic reviews/ literature reviews

A systematic review by Fitzgerald and Berentson-Shaw (2012) found that across five studies, the sensitivity of thermography was above 70% for the detection of malignancies, but specificity ranged between 12% and 85%. Studies with low specificity had high numbers of false positive tests. The authors noted that the studies were of an average quality, and that industry funded studies had more favourable results than independent studies. As these studies emphasised breast cancer diagnosis, rather than screening, the results have limited applicability for the purposes of this horizon scan.

In a survey of thermography image processing literature, Borchartt et al. (2012) explained that thermography is a valuable tool for the early detection of breast cancer, but that it was most effective when used in combination with another screening tool such as mammography. They also noted that there needs to be greater standardisation of thermography methods and protocols to support clinical effectiveness, and that this would require significant work.

#### **Prospective studies**

Rassiwala et al. (2014) undertook a prospective study using breast thermography to screen 1008 female subjects between the ages of 20 and 60 years who were asymptomatic for breast cancer. The study found that sensitivity to detect malignancy was 97.6% and specificity was 99.17%. For a study with a significant sample size these are good results, however they should be treated with caution as the study did not provide an overview of the methodology used for interpreting the thermographs and no other large-scale studies have been undertaken to corroborate the findings.

#### 11.4.3. What cost and safety findings have been reported?

The literature emphasised that thermography is a low-cost breast screening tool, however no figures were provided, and no comparisons were made with the costs of mammography.

As there is no contact with the skin surface and no radiation thermography does not pose a safety risk to the patient (Kandlikar et al., 2017).

# 11.4.4. Does this innovation show high sensitivity and specificity for women with dense breasts and women who have had breast surgery/augmentation compared to digital mammography?

Kandlikar et al. (2017) explained that using thermography to screen women with breast implants can be effective as the position of implants at the base of the breast does not significantly affect heat transfer in the breast. The authors reported that this was an advantage over mammography where there is an issue with tumours hiding in the shadows behind implants and failing to be detected. The literature suggested that dense breasts did not negatively affect the sensitivity or specificity of thermography in the detection of breast cancer (Kandlikar et al., 2017).

## **11.4.5.** Is there evidence that this innovation is more acceptable to women (in general and by ethnic group) compared to digital mammography?

The literature did not discuss the acceptability of thermography to women. It suggested however that as there was no contact with the skin and no compression, it was a painless screening tool compared to mammography (Gerasimova et al., 2014).

Kandlikar et al. (2017) noted that cooling stress, which was often applied prior to dynamic infrared thermograms, could lead to patient discomfort.

#### 11.4.6. Does this technology reduce deaths due to breast cancer through early detection?

There was no statistical evidence to suggest that thermography reduced deaths due to breast cancer through early detection.

Etehadtavakol and King (2013) suggested that thermography could detect abnormal physiological changes in the breast that may be cancerous five to eight years earlier than these irregularities could be found in mammograms. Kandlikar et al. (2017) clarified that this was because thermography could detect angiogenesis as this was accompanied by increased metabolic activity due to the increased demand of blood to supply the new vessels (Kandlikar et al., 2017). Both authors referenced studies conducted outside of the date range for this horizon scan, and there does not appear to be a substantial or recent body of research to support these claims. Etehadtavakol and King suggested that this possibility of early detection could result in a reduction of deaths.

# **11.4.7.** Has this innovation been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

There was no evidence to suggest that thermography has been implemented into a national screening program at this stage.

# **11.4.8.** Has a national position statement been published about this innovation, and if so, what is the position? Is there consensus in position statements?

Position statements have been published in Australia, New Zealand, and the United States of America. All three oppose the use of thermography for breast cancer screening.

Cancer Australia first published a statement on the use of thermography to detect breast cancer in 1999, with the most recent revision in 2010. The statement says that "Cancer Australia does not recommend the use of thermography for the early detection of breast cancer". It noted that there was insufficient evidence to support the use of thermography for breast cancer screening, and referred to three studies from the 1980s that found that thermography could not detect small tumours and that less than 50% of breast cancers identified by mammography were identified by thermography (Cancer Australia, 2010).

The New Zealand National Screening Unit, The Cancer Society of New Zealand, The New Zealand Breast Cancer Foundation, and The New Zealand Branch of The Royal Australian New Zealand College of Radiologists released a position statement in 2010 stating that they "do not support the use of thermography as a breast cancer screening or diagnostic tool as there is insufficient evidence to do so" (National Screening Unit, 2010).

The US Food and Drug Administration released an official statement on October 27 2017 which stated that thermography should not be used as a breast screening substitute for mammography (US Food and Drug Administration, 2017). The statement explained that there was no evidence to support thermography as a standalone screening tool, and that it was only approved by the FDA as an adjunct method of screening for breast cancer. The FDA cautioned that the use of thermography as a standalone breast cancer screening tool may result in cancer not being detected at an early stage.

#### **12. TOMOGRAPHY**

#### Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

• *Safety*: Breast CT delivers a radiation dose approximately equivalent to that of two-view mammography, with a more homogenous distribution of radiation dose. Breast CT may be more comfortable for women than mammography.

• *Effectiveness*: Only one study was identified on the effectiveness of breast CT in symptomatic women, which found that breast CT performed equal to or better than mammography in: the visualisation of breast lesions; the differentiation between benign breast lesions and malignancies; the effect of breast density on lesion visualisation; and the visualisation of breast masses. Mammography provided better micro-calcification visualisation. Two false positives and seven true positives identified by mammography were not detected in breast CT, whereas one small satellite breast cancer was identified by breast CT but not by mammography.

• *Cost*: The costs of capital equipment and medical services for breast CT were undetermined, but the professional fee for breast CT may be higher than that for mammography due to longer interpretation time for radiologists.

**Cone-beam breast CT:** Not covered in the 2009 report.

#### Positron emission tomography (PET):

• *Safety*: PET is a non-invasive and safe diagnostic procedure, with the dose of radiation equivalent to that of dedicated breast CT.

• *Effectiveness*: Only studies using symptomatic samples were identified. The dedicated breast PET scan was more effective than whole-body PET at diagnosing breast cancer. The generalisability of efficacy results were poor given the inclusion of symptomatic women only; however, overall PET had issues with low sensitivity and specificity.

• *Cost*: No economic studies on the cost of PET for breast cancer diagnosis were identified.

**Optical coherence tomography (OCT):** OCT is being investigated by a team of researchers at the University of Western Australia and is thought to be approximately five years away from clinical use in breast cancer detection.

#### **Electrical impedance tomography (EIT):**

• *Safety*: EIT is non-invasive and does not use radiation. It appears to be safe.

• *Effectiveness*: Four studies that assessed the use of EIT as a diagnostic tool for breast cancer were identified. Overall, EIT performed poorly at detecting cancers compared with mammography and ultrasound, with a diagnostic accuracy of 69% being found in one study of symptomatic and asymptomatic women.

• *Cost*: Many Australian private companies were offering EIT as a direct-to-market service, with one company offering a routine EIT scan for \$145. The EIT system itself costs less than \$50,000.

### **12.1.** Tomography as a breast cancer screening modality

Five different types of tomography used for breast cancer imaging are reviewed in this chapter: computer tomography (CT); cone-beam breast CT (CBBCT); positron emission tomography (PET); optical coherence tomography (OCT); and electrical impendence tomography (EIT). These techniques all fall within the tomography family because they involve imaging the breast in sections and then combining these sections together to create an overall image. Where they differ is in the use of varying kinds of penetrating waves for the imaging process; whereas CT and CBBCT use x-rays to create images, PET uses positron-emitting chemical compounds, OCT uses near-infrared light, and EIT uses electrical current. These differences have an impact on the safety and effectiveness of the various modalities for breast imaging, which is explored in the current chapter.

A brief description of each of these techniques and their current application to breast cancer screening is provided below followed by a summary of evidence for each type of tomography.

#### 12.1.1. Computer tomography

Computer tomography (CT; also known as dedicated breast CT or breast CT) is a 3D imaging modality that was developed to improve on the relatively low levels of specificity and sensitivity obtained using FFDM (Aminololama-Shakeri, Hargreaves, Boone, & Lindfors, 2016). In a typical breast CT exam, the subject lies face-down on a table with the examined breast hanging suspended (i.e., without compression) through an opening in the table top. The other breast is kept out of the imaging space. Approximately 500 projection images are then taken of the breast, using an x-ray tube and flat panel detector that rotates 360° around the breast. The scan takes less than 20s for one breast, allowing the scan to be completed in one breath-hold. Breast CT can be performed either with (referred to as contrast enhanced CT; CE-CT) or without (referred to as unenhanced CT) an intravenous contrasting agent.

Despite high levels of sensitivity in early trials, dedicated breast CT was suspended as a screening and diagnostic tool in the 1970s due to concerns with the level of radiation used, relatively low specificity, high equipment cost, and lack of spatial resolution (National Horizon Scanning Unit, 2009). Advancements in technology, including flat-panel detectors and improved spatial and temporal resolution, have sparked a renewed interest in breast CT as a screening modality (National Horizon Scanning Unit, 2009).

#### 12.1.2. Cone-beam breast CT

Cone-beam breast CT (CBBCT) is a variant of CT that uses a cone-shaped x-ray beam and 2D detector, as opposed to a fan-shaped x-ray beam and 1D detector. This is a fast procedure that produces high spatial and contrast resolution 3D images of the breast (Wienbeck, Lotz, & Fischer, 2017).

To perform a CBBCT scan, subjects lie face down with the examined breast hanging suspended without compression through an opening in the examination table, with the other breast kept out of the imaging space. The ability to forgo breast compression without compromising diagnostic information is one of the primary advantages of CBBCT (Wienbeck et al., 2017). Before the 360° diagnostic scan is conducted, a low-dose initial scan is used to assess the breast position inside the imaging space. The total scan duration is approximately 10s for each breast, with the entire examination taking approximately 10 minutes per breast (Wienbeck et al., 2017).

When an intravenous contrasting agent is used in conjunction with CBBCT, the scan is referred to as contrast enhanced CBBCT (CE-CBBCT).

#### **12.1.3.** Positron emission tomography

Positron emission tomography (PET) is an imaging modality that uses positron-emitting chemical compounds together with an imaging device to provide information on biochemical changes in the body without the need for sample extraction. Most current clinical PET imaging uses the chemical compound F-18-fluorodeoxyglucose (FDG), which differentiates between different kinds of tissues based on their metabolism of glucose. Tumour cells have increased uptake of FDG compared with healthy tissue (Cintolo, Tchou, & Pryma, 2013). When FDG is used as the imaging compound in a PET scan, the modality is referred to as FDG-PET.

Currently, FDG-PET is not routinely used in breast cancer screening because of previously identified issues with detecting non-invasive cancers and tumours of small sizes, and confounding breast conditions (such as benign and inflammatory lesions) that also increase FDG uptake (Cintolo et al., 2013; and Minamimoto et al., 2015). That said, results from the implementation of an FDG-PET breast cancer screening program performed in Japan between 2006 and 2009 are promising; these results are summarised in *section 12.5.2*.

#### 12.1.4. Optical coherence tomography

Optical coherence tomography (OCT) is a non-invasive and agent-free imaging modality. OCT uses the light-scattering properties of biological tissue to create image contrast, using nearinfrared wavelengths to potentially produce images deeper into the tissue and in higher resolution than other optical imaging modalities (Vakoc, Fukumura, Jain, & Bouma, 2012). OCT also has the potential to detect tumour angiogenesis and lymphangiogenesis, furthering its potential ability to detect breast cancer in early stages (Vakoc et al., 2012).

The development of OCT systems for use in breast cancer screening is still in preclinical stages, with current systems only able to penetrate up to 2mm in most tissue (Jung & Boppart, 2012; and Vakoc et al., 2012). This means that most research into the use of OCT for cancer detection or diagnosis involves the use of either animal models or samples obtained via biopsies. Additionally, most of this research focuses on the use of OCT in treatment and diagnosis, rather than in screening (Jung & Boppart, 2012). As such, no studies using asymptomatic samples were identified in our literature scan, and OCT is not discussed further in the results section below.

#### 12.1.5. Electrical impedance tomography

Electrical impedance tomography (EIT) works by measuring how fast electricity travels through the breast using electrodes placed on or held by the subject, and is therefore a non-invasive and radiation-free imaging technique (Forsyth, Borsic, Halter, Hartov, & Paulsen, 2011). Biological alterations within cancerous cells cause them to have greater electrical conductivity than surrounding breast tissue. Voltage measurement data are used to construct an image of the electrical conductivity distribution, which is then assessed for distortions. Previous research has suggested that EIT may be affected by the hormones released during menopause, which would impact its use for screening in a sub-population of women likely to develop breast cancer; however, this has been refuted in one recent study using a symptomatic sample (Zain & Chelliah, 2014). Research is still ongoing into the development of EIT systems that are able to image the breast with sufficient quality to accurately detect breast cancer (Forsyth et al., 2011). Some of these systems require breast compression for imaging, whereas others use a setup like breast CT, where subjects lay prone on an examination table with the breast hanging suspended through an opening in the table top. Because of the early stages of research in this area, there were no studies identified that used asymptomatic samples. For this reason, EIT is not considered further.

### **12.2.** Literature search results (number of studies returned)

From the literature search, a total of 109 articles related to innovations in tomography were identified. Abstract contents were then reviewed and 100 of these articles were excluded for various reasons, including: use of symptomatic or non-human samples only; focus on diagnosis or treatment; focus on system development rather than clinical effectiveness; reporting on tomography variants not included in the review; or results not specific to breast cancer. Sources that were not peer-reviewed, such as theses and dissertations, were also excluded. One further article was excluded because the full-text article could not be accessed (Ruile et al., 2015). Articles included in identified literature reviews were not separately reviewed for the current scan.

A total of eight articles were then reviewed to answer the key research questions in relation to innovations in DBT for breast cancer screening in asymptomatic women.

#### Systematic and/or literature reviews

**Breast CT:** Four literature reviews (Aminololama-Shakeri et al., 2016; Lindfors, Boone, Newell, & D'Orsi, 2010; O'Connell, Karellas, & Vedantham, 2014; Sarno, Mettivier, & Russo, 2015).

**CBBCT:** One literature review (Wienbeck et al., 2017).

PET: One literature review (Cintolo et al., 2013)

RCTs

None identified

**Prospective studies** 

PET: Two prospective studies (Minamimoto et al., 2013, 2015)

**Retrospective studies** 

None identified

**Grey literature** 

None identified

#### 12.3. Computer tomography

#### **12.3.1.** Summary of key findings

- The development of breast CT systems for use in the early detection of breast cancer is still in its infancy, with only retrospective studies identified in the literature reviews.
- There is currently no indication of the timeframe in which the full clinical potential of breast CT for the early detection of breast cancer in asymptomatic women will be realised. Although breast CT scanning can visualise breast lesions and masses as well as or better than mammography, breast CT scanning performs worse in visualising microcalcifications. This may limit the incorporation of breast CT as a primary breast cancer screening modality in screening for asymptomatic women.
- Despite initial issues with radiation dose, more recent breast CT scanning systems are improving, achieving radiation dose levels comparable to conventional mammography. Additionally, cost is unlikely to be a barrier in the use of breast CT for the early detection of breast cancer.
- Although no primary studies were identified, one review of the breast CT literature stated that breast CT could play an important role in breast cancer screening for women with dense breasts.
- Findings suggest that breast CT is significantly more comfortable for women than mammography, primarily due to the lack of breast compression.
- Breast CT has not been incorporated into any national screening programs, nor are there any national position statements that have been released on its use in breast cancer screening for asymptomatic women.
- Current research is not sufficient to be able to identify whether breast CT scanning is able to reduce deaths due to breast cancer through early detection.

#### 12.3.2. Study findings and discussion

#### What stage of development or trial is this innovation at?

In their review of the breast CT literature, Lindfors et al. (2010) noted that the development of breast CT is still in its infancy, with only prototype scanners currently in use for clinical research. Additionally, much of the research is focused on the use of CT for breast cancer diagnosis or treatment. No prospective studies or RCTs using asymptomatic samples were identified in the literature reviews. Further system refinements are still required – including increased spatial resolution, increased frame rates, and increased breast coverage – before breast CT can be incorporated into routine screening for asymptomatic women (Lindfors et al., 2010).

#### What is its considered potential clinical value in five years? In 10 years?

#### Systematic and/or literature reviews

None of the reviews identified provided specific timeframes in which the clinical potential of breast CT is expected to be realised for population screening purposes; however, preliminary results from retrospective studies are promising.

Current research suggested that breast CT produced at least equal visualisation of malignant and benign breast lesions as FFDM, with better visualisation of masses (Aminololama-Shakeri et al., 2016; Lindfors et al., 2010; Sarno et al., 2015). These conclusions were verified in a clinical study identified in Lindfors et al.'s (2010) review of the breast CT literature. In this study, breast CT and mammography was used to compare scans taken from 10 healthy controls and 69 women with suspicious lesions. They found that although there were no significant differences between breast CT and mammography in terms of visibility of malignant versus benign lesions, masses were significantly more visible on breast CT scans (p<.001). Although breast CT was better at visualising masses, mammography was significantly better at visualising microcalcifications in this study (p<.001).

The studies identified in Aminololama-Shakeri et al.'s (2016) review of the breast CT literature replicated these results; however, these studies used symptomatic samples to determine levels of visualisation between breast CT and mammography. Although the results regarding the visualisation of masses were promising, both Aminololama-Shakeri et al. (2016) and Lindfors et al. (2010) concluded that the poor visualisation of microcalcifications compared to FFDM was a significant barrier to the primary use of breast CT in screening assessments.

Lindfors et al. (2010) stated that contrast enhanced CT (CE-CT) provided more promising results, with one study finding that CE-CT was better able to visualise malignant breast lesions than mammography; however, this study used a symptomatic sample, limiting the generalisability of its results to screening assessments. Recent developments in breast CT scanning may also be improving the visualisation of microcalcifications. In a review of the existing literature, Wienbeck et al. (2017) found that previous studies using breast phantoms and mastectomy specimens demonstrated that breast CT could visualise calcifications as small as 0.21 to 0.25mm in diameter.

Overall, three reviews of the literature concluded that breast CT may not become a primary screening modality, but could be useful as an adjunct to mammography for validation and diagnosis (Aminololama-Shakeri et al., 2016; Lindfors et al., 2010; O'Connell et al., 2014).

#### What cost and safety findings have been reported?

#### Systematic and/or literature reviews

Despite initial problems with radiation dose with the first breast CT scanners, Aminololama-Shakeri et al.'s (2016) review identified two studies that reported similar radiation dose levels between breast CT scans and conventional mammography; however, the average levels of dose for breast-CT in these studies (ranging from 5.4 to 8.2 mGy) were still above the 2.0 mGy mean glandular dose guideline recommended for FFDM by the BreastScreen Australia National Accreditation Standards (BreastScreen Australia, 2015). Additionally, two studies reviewed by O'Connell et al. (2014) found that the radiation dose distribution was more uniform in breast CT compared with mammography. It was anticipated that there would be a substantial radiation

dose reduction with further advancements in x-ray detector technology and image reconstruction techniques (O'Connell et al., 2014).

The review conducted by Aminololama-Shakeri et al. (2016) did not specify costs, however the authors noted that the base costs of breast CT should allow for the deployment of breast CT systems to most breast imaging clinics.

Does this innovation show high sensitivity and specificity for women with dense breasts and women who have had breast surgery/augmentation compared to digital mammography?

#### Systematic and/or literature reviews

In their review of the breast CT literature, O'Connell et al. (2014) stated that breast CT could potentially play an important role in the detection of cancer in dense breasts. They did not cite any studies to support this claim, however they note the benefit of 3D images in being able to visualise abnormalities in any plane, and in being able to quantitatively estimate breast density. Improved conspicuity may overcome some of the limitations associated with tissue superimposition in mammography.

Is there evidence that this innovation is more acceptable to women (in general and by ethnic group) compared to digital mammography?

#### Systematic and/or literature reviews

Three clinical trials included in the review conducted by Aminololama-Shakeri et al. (2016) found that women were more comfortable using breast CT than mammography. This was largely due to the lack of breast compression for breast CT. This finding was confirmed in a study of 209 women identified in Lindfors et al.'s (2010) review. This study asked the women to rate the comfort of breast CT compared with mammography on a scale of 1 to 10, with 1 indicating better comfort for mammography and 10 indicating better comfort for breast CT. Results showed a median rating of 10 and a mean of 8.8, indicating greater comfort for the breast CT. Furthermore, overall comfort for the breast CT was rated on a scale from 1 (very uncomfortable) to 10 (extremely comfortable), with a mean of 8.1 and median of 8 being obtained from the women's ratings. The most common issues with the breast CT were difficulties in arching the back sufficiently to push the breast forward and maximise imaging coverage, and discomfort in the neck during scanning.

#### Does this technology reduce deaths due to breast cancer through early detection?

Current research is not sufficient to be able to identify whether breast CT scanning is able to reduce deaths due to breast cancer through early detection of cancer in asymptomatic women.

### Has this innovation been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

Breast CT scanning has not been incorporated into any national screening programs.

Has a national position statement been published about this innovation, and if so, what is the position? Is there consensus in position statements?

No national position statements on the use of breast CT in breast cancer screening were identified in the literature search.

#### **12.4.** Cone-beam breast CT

#### **12.4.1.** Summary of key findings

- The research assessing the use of CBBCT is still in the early stages, with research not yet advancing to clinical tests using asymptomatic samples.
- There is currently no indication of the timeframe in which the full clinical potential of CBBCT for breast cancer detection will be realised. That said, current results using symptomatic samples are promising in terms of its ability to improve on FFDM.
- Current research suggests that there is no statistically significant difference in radiation dose from CBBCT scans compared with FFDM. Findings related to the cost of CBBCT were not reported in the identified literature.
- No studies were identified that assessed the performance of CBBCT scans in asymptomatic women with dense breasts, however results from symptomatic samples suggest that CE-CBBCT may improve scan sensitivity for women with dense breasts compared with FFDM.
- Reported patient comfort is higher for CBBCT than for mammography.
- CBBCT has not been incorporated into any national screening programs, nor are there any national position statements that have been released on its use in breast cancer screening for asymptomatic women.
- Current research is not sufficient to be able to identify whether CBBCT scanning is able to reduce deaths due to breast cancer through early detection.

#### 12.4.2. Study findings and discussion

#### What stage of development or trial is this innovation at?

The research assessing the use of CBBCT is still in the early stages, with research not yet advancing to clinical tests using asymptomatic samples. Instead, much of the current research focuses on the use of CBBCT in the diagnosis and treatment of breast cancer patients.

#### What is its considered potential clinical value in five years? In 10 years?

#### Systematic and/or literature reviews

Based on theoretical considerations and a review of the current literature, Wienbeck et al. (2017) concluded that the early detection of breast cancer using FFDM could be significantly improved by the implementation of CBBCT, particularly for women with dense breasts. However, the authors noted that all existing research on the use of CBBCT to detect breast cancer used symptomatic samples rather than a screening population, which could artificially

inflate the sensitivity and specificity found for CBBCT. They also noted that because of limited experience with CBBCT, it would not currently be advisable to replace FFDM with CBBCT.

#### What cost and safety findings have been reported?

#### Systematic and/or literature reviews

Wienbeck et al.'s (2017) review of the literature found that there was no significant difference in the mean glandular dose (MGD) between non-contrast CBBCT and FFDM in published studies; however, the MGD comparators used were for diagnostic not screening mammography. The mean MGD for CBBCT exams ranged from 8.1 to 10.7 mGy, compared with 6.1 to 16.9 mGy for diagnostic two-view mammography (up to 11 images, which is much higher than the standard four images taken for a screening mammogram), with differences in individual studies not being statistically significant.

That said, two other studies in the review found that the total radiation dose for a CE-CBBCT scan was approximately twice that of a non-contrast CBBCT scan; however, no direct comparison in dosage levels for CE-CBBCT and FFDM was identified in the literature.

Findings related to the cost of CBBCT were not reported in the identified literature.

Does this innovation show high sensitivity and specificity for women with dense breasts and women who have had breast surgery/augmentation compared to digital mammography?

#### Systematic and/or literature reviews

In their review of the CBBCT literature, Wienbeck et al. (2017) identified one study using a symptomatic sample that suggested CE-CBBCT improved the sensitivity for detecting malignant breast lesions for patients with more dense breast tissue (i.e., BIRADS 3-4). In this study, CE-CBBCT had a sensitivity of 99% compared with 78% for FFDM. This increased sensitivity was not found for non-contrast CBBCT however, suggesting that CE-CBBCT could be a better option for women with more dense breasts.

Is there evidence that this innovation is more acceptable to women (in general and by ethnic group) compared to digital mammography?

#### Systematic and/or literature reviews

Because of the lack of breast compression in CBBCT, greater comfort would be expected in CBBCT compared to mammography. Indeed, Wienbeck et al. (2017) cited five studies that found that patient comfort was higher for CBBCT than for mammography, including two clinical studies where approximately 87% of women reported that CBBCT was either equal in comfort or more comfortable than mammography. The neck, the shoulder, and the ribs were reported to be the most prominent areas of discomfort for CBBCT, with refinements in the design of the examination table top leading to greater patient comfort in these areas in one study reviewed.

#### Does this technology reduce deaths due to breast cancer through early detection?

Current research is not sufficient to be able to identify whether CBBCT scanning is able to reduce deaths due to breast cancer through early detection of cancer in asymptomatic women.

## Has this innovation been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

CBBCT scanning has not been incorporated into any national screening programs.

Has a national position statement been published about this innovation, and if so, what is the position? Is there consensus in position statements?

No national position statements on the use of CBBCT in breast cancer screening were identified in the literature search.

#### **12.5.** Positron emission tomography

#### 12.5.1. Summary of key findings

- Research into the use of PET for early breast cancer detection in asymptomatic women has progressed to prospective clinical studies, however further research is still required on the safety and effectiveness of PET as a screening test.
- There are no clear timeframes regarding the clinical potential of PET as a screening test being reached; however, current research suggests that there may still be limitations with the ability of PET to detect small tumours.
- Relatively high levels of radiation continue to be an issue for PET systems, and cost is also identified as being potentially prohibitive to its incorporation into routine screening for asymptomatic women.
- No information was identified on the sensitivity and specificity of PET for asymptomatic women with dense breasts or women who have had breast surgery/augmentation compared with FFDM.
- No information was identified on the acceptability of PET for women compared with FFDM.
- Results from a nationwide FDG-PET cancer screening program in Japan found that FDG-PET had a sensitivity of 84% in detecting breast cancer; this was not significantly different from rates found for mammography. Issues with radiation dosage and the cost of FDG-PET scans were noted as a barrier for incorporation into other screening programs.
- There have been no national position statements released regarding the use of PET for the early detection of breast cancer in asymptomatic women.
- Current research is not sufficient to be able to identify whether PET imaging is able to reduce deaths due to breast cancer through early detection of cancer in asymptomatic women.

#### 12.5.2. Study findings and discussion

#### What stage of development or trial is this innovation at?

Research into the use of PET for the early detection of breast cancer in asymptomatic women is more advanced than for the other forms of tomography covered in this chapter, with results already being reported from its implementation in a nationwide cancer screening program in Japan. That said, there are still relatively few studies assessing the effectiveness and safety of PET for breast cancer screening for asymptomatic women, with research largely focusing on the diagnosis or treatment of cancer, or on the use of PET for other types of cancer. Further research is needed before clear conclusions can be drawn about the effectiveness and safety of PET compared with FFDM.

#### What is its considered potential clinical value in five years? In 10 years?

#### Systematic and/or literature reviews

There were no clear timeframes identified in the literature regarding the clinical potential of PET being reached, however information is growing regarding the efficacy of PET in the early detection of breast cancer.

A literature review conducted by Cintolo et al. (2013) cited two studies that found sensitivity rates of 68 to 83% using FDG PET for breast cancer detection. This variability in sensitivity was potentially due to differences in tumour size, with two further studies identified in the review finding that FDG PET had difficulty in detecting tumours that were less than 8 to 10mm in diameter. Despite this identified difficulty, the review noted the recent developments in PET technology, including the FDG positron emission mammography (PEM), which only images the breast during an examination rather than the whole body. One study included in the review found that FDG PEM had significantly improved sensitivity compared with the whole-body PET (95 versus 87%, respectively), particularly for the detection of lesions less than 2cm in diameter. Comparison with FFDM was not provided in this study.

Overall, Cintolo et al.'s (2013) review of the PET literature concluded that due to limitations with whole-body PET imaging, including safety, cost, and accuracy, it is unlikely that it would be integrated into routine screening for asymptomatic women.

#### **Prospective studies**

Data collected from a nationwide FDG-PET cancer screening program in Japan found that FDG-PET scanning detected breast cancer with similar sensitivity to mammography (Minamimoto et al., 2015). Further studies on this screening program and its outcomes are provided below.

#### What cost and safety findings have been reported?

#### Systematic and/or literature reviews

The review of PET literature conducted by Cintolo et al. (2013) identified high levels of cumulative radiation being a particular concern for PET imaging, as it often scans the entire body rather than just the breast being examined. The authors noted that work is currently ongoing to develop systems that can produce adequate images with low doses of radiation, however they did not provide a timeframe in which these new systems would be ready for clinical use.

Cintolo et al. also noted that FDG PET was a much more expensive choice than other imaging modalities, currently costing around USD\$1,000 per examination. The authors felt this limited its usefulness for breast cancer screening assessments. These concerns with the radiation dosage and cost associated with FDG-PET screening were also reflected in results from a nationwide FDG-PET cancer screening program in Japan (detailed further below; Minamimoto et al., 2013, 2015).

#### Does this innovation show high sensitivity and specificity for women with dense breasts and women who have had breast surgery/augmentation compared to digital mammography?

No information was identified on the sensitivity and specificity of PET for asymptomatic women with dense breasts or women who have had breast surgery/augmentation compared with FFDM.

Is there evidence that this innovation is more acceptable to women (in general and by ethnic group) compared to digital mammography?

No information was identified on the acceptability of PET for women compared with FFDM.

#### Does this technology reduce deaths due to breast cancer through early detection?

Current research is not sufficient to be able to identify whether PET is able to reduce deaths due to breast cancer through early detection of cancer in asymptomatic women.

### Has this innovation been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

A nationwide FDG-PET cancer screening program was conducted in Japan between 2006 and 2009 for asymptomatic individuals (Minamimoto et al., 2013, 2015).

A total of 155,456 individuals (92,739 men, 62,073 women, and 644 gender unidentified) participated in the screening program, with screening performed by 233 facilities. This program aimed to detect any type of cancer at an early stage, rather than breast cancer. For this reason, both whole-body PET scanners (61% of scans) and dedicated PET scanners (39% of scans) were employed in the program. FDG-PETs were also combined with other imaging modalities (such as CT, MRI or ultrasound) in at least one case at 85% of participating facilities.

Overall, 478 suspected cases of breast cancer were detected in the screening program including for five men, however sensitivity was only reported for the 473 women with possible breast cancer (Minamimoto et al., 2015). Sensitivity in detecting breast cancer was 84% for FDG-PET scans, with a PPV of 43%. The sensitivity of FDG-PET was not significantly different from the sensitivity obtained for mammography (78%; p=.34). Specificity was not identified in this screening program.

The authors concluded that FDG-PET cancer screening is not a completely established method, and that one-time FDG-PET screening scans are likely to miss some malignancies. Radiation exposure remained an issue, and the risks and benefits of increased radiation exposure using FDG-PET should be explained to participants. Furthermore, the authors noted the high cost of the screening program, with a cost of \$1,000 per scan in Japan and \$2,000 per scan in the United States. They suggested that maximising cost-effectiveness would require limiting the use of FDG-PET to high risk individuals.

### Has a national position statement been published about this innovation, and if so, what is the position? Is there consensus in position statements?

No national position statements on the use of PET in breast cancer screening were identified in the literature search.

#### **READING STRATEGIES**

This section covers:

- Computer aided detection followed by a chapter on artificial intelligence, and
- Tele-mammography.

#### **13. COMPUTER AIDED DETECTION**

#### Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

The use of computer aided detection methods to support radiologists during the screening process to identify abnormal breast tissue was not discussed in the 2009 ANZHSN report.

#### 13.1. Computer-aided detection in breast cancer screening

FFDM images need to be examined by radiologists to detect small changes that could be indicative of a malignancy. This can result in misdiagnosis due to human error by visual fatigue. The human error rate when reading mammograms is reported to be between 10-30% or missed cancers. Errors may result in overdiagnosis leading to unnecessary biopsies or other procedures and unwarranted health expenditure and anxiety (Azavedo, Zackrisson, Mejàre & Heibert Arnlind, 2012).

To maintain high sensitivity and specificity (i.e., high cancer detection and low false-positive rates), many countries including Australia recommend double reading (i.e., that the breast images are reviewed by two specially trained radiologists). While prospective double reading of screening mammograms has been shown to increase cancer detection rates (Dromain et al., 2012 and Horsch, Hapfelmeier & Elter, 2011), a screening program may not always be able to resource this due to the number of radiologists required.

To offset these resourcing implications, sophisticated computerised methods have been developed to support radiologists to interpret medical images by enhancing the images and providing in-depth analysis of visual details. Two main schemes are available:

- 1. computer-aided detection (CADe), and
- 2. computer-aided diagnosis (CADx).

CADe focuses on finding the location of abnormalities. CADx is targeted to characterisation (i.e., malignancy versus benignity). Although some algorithms address both stages simultaneously, most researchers look at them as separate processes. For the purposes of the horizon scan, studies that focused only on characterisation (i.e., the diagnosis of breast cancer) are excluded.

The programs used in CADe separate suspicious regions that may contain masses from background parenchymal tissues. They then identify and mark areas that the software identifies as abnormal breast tissue. CADe programs are not intended to be the sole method for analysing images. Rather, they are designed to alert the radiologist to possible suspicious areas. A radiologist (or other reader) must interpret and decide how to act upon each mark (accept or dismiss).

#### Standard design of a mammographic CADe system

The standard design of a mammographic CADe system for breast cancer detection has four stages:

- 1. image pre-processing
- 2. segmentation of suspicious mass region
- 3. feature extraction

4. classification.

Pre-processing of mammograms is done to improve the contrast of mammograms, which assists in further stages of the detection pipeline. This step also involves denoising the images, which is necessary as the noise created by random fluctuations contained in radiological images make the detection of small and subtle structures more difficult. Segmenting the breast region from pectoral muscle and surrounding regions is then carried out to make it easier to extract the suspicious tissues from breast segments. After the segmentation process, the desired features (such as texture, shape, gradient and intensity) are extracted from the region of interest. These features are used to classify the suspicious regions as normal or abnormal tissues (benign or malignant).

### **13.2.** Summary of key findings

- CADe, applied as a complementary technology to mammography interpretation, prompts the reader to lesions on the mammogram. The reader then needs to decide whether to recall CADe-prompted findings.
- Non-randomised studies have shown that CADe improves the sensitivity of a single reader, with an incremental cancer detection rate ranging between 1 and 19%. A limitation of CADe is that it also increases screening recall rate through a decrease in specificity. Studies have found that double-reading produces the same results than that for single reading with CADe, without the increased rate of recall. In terms of detecting lesion characteristics, good results have been observed for the detection of breast cancers presenting as microcalcifications. However, the rate of detection for masses and architectural distortions is not so positive.
- To expand the clinical value of CADe in the detection of breast cancer, several studies have focused on improving technical stages of the CADe process. Promising results have been observed, particularly regarding the application of deep learning systems, but further work needs to be done. While CADe could substitute the human second reader in the future, depending on the first reader's experience, without improvements in its effectiveness (eg, a decrease in recall rate), CADe is unlikely to be a cost-effective alternative to double reading for mammography screening.
- Current research is not sufficient to be able to identify whether CADe is able to reduce deaths due to breast cancer through early detection.
- CADe has not been incorporated into any national breast screening programs, nor are there any national position statements that have been released on its use in breast cancer screening for asymptomatic women

### **13.3.** Literature search results (number of studies returned)

From the literature search a total of 145 abstracts of peer reviewed articles were identified that related to CADe. Abstract contents were then reviewed and 78 were excluded because they indicated studying CADe in a symptomatic population, for diagnostic or treatment purposes, or were studying the use of CADe for other imaging modalities. 10 other articles were subsequently excluded because they could not be located or were dissertations. A total of 57 articles were

reviewed to answer the key research questions relating to the use of CADe in breast cancer screening.

#### Systematic and/or literature reviews

7 systematic or literature reviews were identified (Agrawal, Vasta & Singh, 2014; Ayer, Chen & Burnside, 2013; Azavedo, Zackrisson, Mejàre & Heibert Arnlind, 2012; Dromain et al., 2013; Ganesan et al., 2013; Houssami & Ciatto, 2011; and Karssemeijer, 2010).

RCTs

None identified

#### **Prospective studies**

6 prospective studies were identified (Bargalló et al., 2014; Dromain et al., 2012; Fenton et al., 2011; James et al., 2010; Sanchez-Gómez et al., 2011; and Sohns et al., 2010).

#### **Retrospective studies**

44 retrospective studies were identified (Al-Najdawi, Biltawi & Tedmori, 2015; Bargalló et al., 2013; Bessa et al., 2014; Beura, Majhi, Dash and Roy, 2015; Choi & Ro, 2012; Chu, Min, Liu & Lu, 2015; Cole et al., 2014; Cole et al., 2012; de Sampaio, Silva, de Paiva & Gattass, 2015; de Sampaio, Silva, de Paiva, Gattass & Diniz, 2011; Dhahbi, Barhoumi & Zagrouba, 2015; Dheeba, Singh & Selvi, 2014; Dhungel, Carneiro & Bradley, 2017; Dong et al., 2015; Guerriero et al., 2011; Hong & Sohn, 2010; Horsch, Hapfelmeier & Elter, 2011; Hu, Li & Gao, 2011; Hupse et al., 2013; Jen & Yu, 2015; Kendall & Flynn, 2014; Kim et al., 2010; Kooi et al., 2017; Lehman et al., 2015; Lesniak, Hupse, Blanc, Karssemeijer & Szekely, 2012; Liu, Xu, Liu & Feng, 2011; Mac Parthalian, Jensen, Shen & Zwiggelaar, 2010; Mordang et al., 2017; Mordang, Gubern-Merida, den Heeten & Karssemeijer, 2016; Mughal, Sharif & Muhammad, 2017; Neto, Silva, Paiva & Gattass, 2017; Povyakalo, Alberdi, Strigini & Ayton, 2013; Qiu et al., 2015; Raghavendra et al., 2016; Rangayyan, Banik & Desautels, 2010; Rangayyan, Banik & Desautels, 2013; Sadaf, Crystal, Scaranelo & Helbich, 2011; Samulski et al., 2010; Singh & Urooi, 2016; Skaane, Kshirsagar, Hofvind, Jahr & Castellino, 2012; Velikova, Lucas, Samulski & Karssemeijer, 2013; Wang et al., 2016; Wang, Lederman, Tan, Wang & Zheng, 2010; and Wang et al., 2017).

#### **Grey literature**

A search of the U.S. National Library of Medicine <u>ClinicalTrials.gov</u> website for "breast cancer" and "computer aided detection" identified 9 related clinical trials, with two studies related to CADe for breast cancer screening purposes:

- 'Computer-Aided Breast Cancer Detection in Women Undergoing Screening Mammography' sponsored by the Aberdeen Royal Infirmary, Scotland, United Kingdom. The study, which is due to be completed in October 2018, seeks to examine how well computer-aided breast cancer detection works in women undergoing screening mammography.
- 'The Diagnostic Efficacy of Computer-Aided Detection (CAD) in Full-Field Digital Mammography (FFDM)- A Prospective Study' at the National Taiwan University Hospital, Taiwan. The study, which is due to be completed in 2018, seeks to

evaluate whether CADe in FFDM can facilitate the detection rate of breast cancer on mammography compared with FFDM without CADe.

### 13.4. Study findings and discussion

#### 13.4.1. What stage of development or trial is this innovation at?

Most of the literature on computer-aided detection for mammography comprises studies looking at technical aspects, such as improvements to software. As they explore the emerging technological developments in CADe, these studies will be examined in the section looking at the potential clinical value of CADe in future years. The following section looks at those studies that focused on the current use of this reading strategy.

#### Sensitivity

#### **Systematic reviews**

Four systematic reviews discussed the rate of sensitivity. Houssami & Ciatto (2011) and Dromain et al. (2012) reported a moderate increase in sensitivity with the use of CADe. Both reviews referred to studies reporting an increase in the true-positive rate ranging from 1.3% to 19.5% (based on the respective findings of four papers).

In their review, Dromain et al. (2012) noted that increases in the rate of sensitivity were consistent with those for doubling-reading by humans (a finding also reported by Azavedo et al. (2012)). CADe was said to be easier to implement and cheaper to use than double reading. Dromain further suggested that there was a clear benefit in using CADe for less experienced or low volume readers. They referred to papers that observed the use of CADe by junior radiologists improved sensitivity from 61.9% to 84.6%, compared to a slight improvement from 76.9% to 84.6 for experienced radiologists. The use of CAD by low-volume readers led to an increase in recall and cancer detection rate of approximately 19%.

In one of the largest studies comparing the performance of digital screening mammography with or without CADe, Lehman et al. (2015) found no evidence that CADe improved screening mammography performance: in fact, the opposite was observed. They compared the accuracy of FFDM interpreted with (n = 495,818) and without (n = 129,807) CADe from 2003 through 2009 in 323,973 women. Mammography sensitivity was 85.3% with and 87.3% without the use of CADe. Commenting on the difference in their results and those from earlier studies supporting the efficacy of CADe, Lehman et al. (2015) said this was due to inherent biases in the methods previously used to examine the impact of CADe. The bulk of the early studies were laboratory based and measured the ability of CADe programs to mark cancers on selected mammograms. The reported "high sensitivities" of CADe from these studies did not translate to higher cancer detection in clinical practice. Lehman concluded that CADe offered no additional benefit to women in terms of breast cancer detection compared with double reading. Nonetheless, they did consider that CAD had some practical applications such as improving workflow.

Lesion characteristics: Dromain et al. (2012) also looked at different lesion types. In terms of detecting breast cancers presenting as microcalcifications, the level of sensitivity ranged from 86% to 99%. They observed that CADe seemed to increase the efficiency and confidence level of radiologists searching for subtle microcalcifications. The detection of masses was not as reliable,

with sensitivity ranging from 83.0% to 90.0%. This was due to the varying size, shape and appearance of masses, and varying tissue density as such masses often share characteristic features with normal parenchymal tissues. Sensitivity was also found to be less for masses with architectural distortions than for spiculation. The review noted that CADe had limited value in the detection of amorphous calcifications with a sensitivity of only 57.0% being reported.

#### **Retrospective observational studies**

Seven studies have evaluated the impact of CADe in a clinical setting using historical control methodology. The rate of sensitivity in these studies (which are set out in CAD Table 1, below) ranged from 76.8% to 98.8%. The authors of these papers concluded that the use of CADe in the interpretation of mammograms could lead to improvements in early cancer detection.

Study	Year	Base <sup>1</sup>	Quoted number of people in the study	Sensitivity
Neto, Silva, Paiva & Gattass	2017	DDSM	621	97.5%
Wang et al.	2017	DDSM	322	92.26
Beura, Majhi, Dash, and Roy	2015	DDSM	200	98.8%
Dong et al.	2015	DDSM	200	94.0%
de Sampaio et al.	2015	DDSM	1049	93.0%
Hu, Li, & Gao	2011	MIAS	170	91.3%
Liu, Xu, Liu & Feng	2011	DDSM	125	76.8%

CAD Table 1. Summary of sensitivity in studies evaluating the impact of CADe in clinical settings

<sup>1</sup> DDSM: Digital Database for Screening Mammograms; MIAS: Mammographic Image Analysis Society

Three other studies compared single human reading with single reading plus CADe. Fenton et al. (2011) found a non-statistical increase in sensitivity after its introduction at 90 Breast Cancer Surveillance Consortium (BCSC) facilities from 79.7% to 83.1% (3-12 months post-CADe) and 80.1% (over 12 months post CADe) (p=.62). From these results, they concluded that CADe was not associated with improvement in detection rate or prognostic characteristics in invasive breast cancer. In a later study, Cole et al. (2012) found that the CADe systems tested were able to detect significantly more cancers than were found on initial screening by the original radiologist for both digital and film-screen mammograms. Their results showed an increase in sensitivity in digital and film-screen mammograms by 31% and 28% respectively. For each system, sensitivity

was also higher for digital than for film-screen mammography, and the sensitivity of the CADe systems were not influenced by lesion characteristics (eg, histology, lesion size, lesion type) or subject characteristics (age, breast density, menopausal status). In contrast, Sanchez Gómez et al. (2011) found that CADe only increased cancer detection compared to mammography screening by 0.1%.

One study compared single reading done by a radiologist assisted by CADe with double reading. Bargalló et al. (2013) found that cancer detection rate improved by 16.2% using a single reader assisted by CADe. In a separate article on the same study, Bargalló et al. (2014) reported that specialised breast radiologists performed better than general radiologists.

Stand-alone CADe sensitivity assessment is an important part in determining whether CADe should be incorporated into clinical assessments. These studies have shown that CADe improves the sensitivity of a single human reader and has the potential to outperform double reading, suggesting that the use of such technology in the interpretation of FFDM could lead to improved breast cancer detection.

Lesion characteristics: The impact of lesion type on sensitivity was explored by several studies. Retrospectively applying CADe to 127 cases, Sadaf, Crystal, Scaranelo & Helbich (2011) found a statistical difference (p<.05) in CAD detection of breast cancers that appeared as microcalcifications compared to other mammographic features (microcalcifications (100%), masses (88.7%) and architectural distortions (71.4%)). They also found that sensitivity increased as the size of the lesion got bigger (1-10 mm (84.4%), 11-20mm (93.2%) and  $\geq$ 20mm (95.6%)). The overall sensitivity was 90.6%.

Bargalló et al. (2013) also evaluated CADe performance on mammographic features of breast cancer but limited their study to very small (equal to or less than 1 cm) invasive breast cancers. CADe showed an overall sensitivity of 86.7%. Breaking down their results by mammographic features, they found that CADe works better for detecting microcalcifications (100%) than for masses (85.5%). Architectural distortions were the most difficult mammographic pattern to be detected (57.1%).

Other studies also reported findings of 100% sensitivity for microcalcifications (Mohamed et al., 2014), and lower figures for architectural distortion, which was found to be 80.0% in two studies by Rangayyan, Banik & Desautels (2010) and Rangayyan, Banik & Desautels (2013).

These studies indicate that the use of a CADe system will help a radiologist as a second reader to identify microcalcifications but it's ability to detect masses and architectural distortion is less clear.

#### **Prospective trials**

Four prospective trials were identified from the literature search that discussed the rate of sensitivity. In their study, Skaane, Kshirsagar, Hofvind, Jahr & Castellino (2012) examined baseline mammograms that were obtained over a fourteen-month period as part of the Norwegian national screening program. Of the 23,923 women who underwent screening during this period, 208 were found to have cancer, of which 104 were diagnosed at baseline screening and 104 diagnosed as subsequent cancers (44 interval cancers and 60 next screening round cancers). CADe correctly marked 94% of mammograms of the screen-detected cancers. Skaane and colleagues concluded that CADe had the potential to increase cancer detection by up to 16%. They acknowledged, however, that this figure represented a likely scenario of the increase in

cancer detection. The retrospective design with standalone analysis meant that it was not possible to estimate how many correct (true positive) CADe marks would have been overruled by radiologists reading the image.

The two other trials were conducted by Sadaf et al. (2010) and Kim et al. (2010) who reported sensitivities of 91.0% (n = 127) and 96.2% (n = 130) respectively.

Lesion characteristics: James et al. (2010) evaluated the mammographic features of breast cancer that favour lesion detection with single reading and CADe or with double reading. Of the 28,204 women who participated in the study, 227 were found to have breast cancer. Most cancers were picked up by both reading regimens; however, some were detected with one reading regimen and missed with the other. While parenchymal deformities were more likely to be picked up by double reading, more asymmetric densities were identified with single reading and CADe. James and colleagues found no difference in the ability of either reading regimen to identify masses or microcalcifications, although noted that the detection rate for masses (87%) was not as high as that for microcalcifications (100%). They also observed that lesion size had no effect on the performance of either reading regime.

Kim et al. (2010)also reported findings of 100% sensitivity for microcalcifications and 78.3% for cancers manifested as masses.

These trials confirmed the findings from the systematic reviews about the usefulness of CADe systems in breast cancer detection, particularly the detection of microcalcifications.

#### Interval cancer rate

#### **Prospective reviews**

Skaane et al. (2012) examined the impact that CADe could have on the reduction of the rate of interval breast cancer rates. In their study, CADe correctly marked 93% of the subsequent cancers retrospectively categorised as actionable. They concluded that CADe had the potential to reduce the number of interval cancers by up to 20%. However, they recognised that the retrospective design of their study would have caused expectancy (hindsight) bias as the radiologist would be aware of a higher than usual possibility of cancer.

#### Positive predictive value (PPV)

#### PPV1 (verified attributable cancers per number of women recalled from screening)

Only one study reported on PPV<sub>1</sub>. Fenton et al. (2011) reported a decrease in PPV<sub>1</sub> after CADe was introduced at BCSC facilities from 4.3% to 3.7% (3-12 months post-CADe) and 3.6% (over 12 months post CADe).

#### PPV2 (cancers diagnosed per the number of biopsies recommended)

No studies reported on PPV2.

#### PPV3 (cancers diagnosed per the number of biopsies performed)

Only one prospective study reported on PPV<sub>3</sub>. Sanchez Gómez et al. (2011) found that the rate of percutaneous biopsy was unchanged by CADe (20.2%).

#### Specificity

#### **Systematic reviews**

Three systematic reviews discussed the effect that CADe had on specificity. Dromain et al. (2012) reported that CAD-assisted reading was associated with a reduction in specificity compare to mammography. They noted this reduction may cause radiologists to underestimate and disregard the positive features of CADe. Karssemeijer (2010) shared the concern that the reduction in specificity would cause radiologists to lose confidence in CADe. Both authors considered that this highlighted the need to train radiologists properly in the use of CADe. However, Lehman et al. (2015) observed only minor differences in specificity between mammograms interrupted using CADe and those interrupted visually.

#### **Retrospective observational studies**

Analysing the records of 684,956 women who received more than 1.6 million FFDM from 1998 to 2006, Fenton et al. (2011) reported a non-significant decrease in specificity after its introduction at BCSC facilities from 91.9% to 91.5% (3-12 months post-CAD) and 91.4% (over 12 months post CAD) (p=.62). Given its scale, this study offered important insights into the effectiveness of CADe when used in real-world practice. First, the results suggested CADe had limited impact on breast cancer detection. This caused Fenton and colleagues to express concern that CADe offered little or no impact on breast cancer mortality, which may depend on earlier detection of invasive breast cancer. Their findings were consistent with a meta-analysis that suggested CADe caused a modest increase in recall rates with little or no impact on cancer detection rates.

Using screening mammograms from 68 patients, Bargalló et al. (2013) also found that specificity of CADe was low (26% with 1.76 false marks per case). In their view, this also represented a major drawback for the technology.

#### False positive recall rate

False positive recall rates are a significant concern as women who are recalled for further investigation often experience high levels of anxiety, along with the inconvenience and expense of attending a further appointment that bring no health benefit to the woman. Many studies discussed below found that the use of CADe for mammography was accompanied by an increase in recall rate due to the high number of false positive marks generated. This could lead to:

- (a) an increase in the number of women being unnecessarily referred for a clinical followup
- (b) an increase in interpretation time of the mammogram, and
- (c) a loss in confidence in the CADe system, especially when locations are marked that are clearly not suspicious.

For instance, in their systematic review of new imaging methods in breast screening, Houssami & Ciatto (2011) referred to two previous studies that observed that CADe caused additional recall of between 6% to 35% of women. They observed that although CADe systems were useful to avoid perceptual oversights of the radiologist, they were not yet suitable to serve as an independent reader.

This conclusion contrasts with the results found in three other studies. While Bargalló et al. (2014) found a modest increase of 3.3% in recall rate, lower figures were reported by Sanchez Gómez et al. (2011) and Guerriero et al. (2011), who observed recall rates of only 0.4% and 0.5% respectively.

Three studies explored whether algorithms could be designed to decrease the rate of false positive marks. Mordang et al. (2017) focused on obvious false positives generated by microcalcification detection algorithms. They added an additional step in the detection method in which dedicated classifiers learn to recognise the patterns of obvious false-positive subtypes that occur more frequently. The insertion of this step enabled their system to detect between 73% and 83% of false positives, compared to only 68% detected by standard CADe systems. Comparing their method to a conventional approach, they found that the number of false positives decreased significantly (p=.0002).

In an earlier study, Mordang, Gubern-Merida, den Heeten & Karssemeijer (2016) noted that breast arterial calcifications were one of the most frequent false positives marked by CADe systems. They demonstrated that by using dedicated algorithms to detect and remove breast arterial calcifications, the performance of CADe systems could be improved. The number of false positives per case that were marked by the CADe system reduced by 29% on average. In their view, using dedicated algorithms to detect and remove breast arterial calcifications improved the performance of CADe systems.

Lesniak, Hupse, Blanc, Karssemeijer & Szekely (2012) used support vector machine classification to reduce the rate of false positive marks. Using a database of 2,516 film mammography examinations and 73 input features, they were able to train the classifiers and evaluate their performance. They found that using an SVM-based CADe significantly reduced the number of false positive marks, enabling their system to outperform conventional approaches for breast cancer CADe.

#### 13.4.2. What are their considered potential clinical value in five years? In 10 years?

Several studies have focused on technical aspects looking at whether improvements could be made to software associated with specific stages of the mammographic CADe pipeline. The authors of these studies felt that the incorporation of these technical developments into the CADe system would improve the rate of breast cancer detection as well as decease the recall rate associated with CADe. These improvements would increase the clinical value of CADe in the detection of breast cancer. The following section examines these studies and discusses their potential clinical value.

#### Systematic reviews

Horsch et al. (2011) noted that many of the studies that sought to improve the mammographic CADe system were performed on selected materials enriched with cancer cases so their value as a true screening situation is questionable. This raised concerns about the quality of such studies. Analysing 59 publications presenting 106 evaluation studies (none of which fell within the time range of the current horizon scan), they noted that when assessing the performance of the mammographic CADe systems and their components (eg, segmentation or feature extraction algorithms), the selection of test cases (or images) and the choice of evaluation methods was particularly critical. Horsch et al. observed that the results obtained in several studies were affected by the selection of images in the mammography databases which was not standardised and prevented verification by other researchers. A substantial amount of bias was also caused by the researchers' decisions around study design, such as using small, often non-reproducible, sub-selections of test cases. This issue was compounded by the evaluation methods being employed, such as sampling and statistical models for solving the classification task. This led to

overestimation of measures like the correct classification rate. In their view, these studies should be approached with caution.

Kooi et al. (2017) expressed similar concerns about the work on CADe systems noting that many studies only tested small data sets, which were not always shared, and the proposed algorithms were difficult to compare. Another complication was that breast cancer has two main manifestations (micro-calcifications and masses). Separate systems have been developed for each which made comparison difficult to do.

#### **Retrospective observational studies**

#### Pre-processing

The purpose of the pre-processing stage of the CADe pipeline is to enlarge the intensity difference between objects and background to produce reliable representations of breast tissue. The methods used to manipulate mammogram images can be divided into two main categories:

- (1) Histogram Equalisation techniques; and
- (2) Morphological Enhancement techniques.

#### **Histogram Equalisation**

Enhances the contrast of the image by reducing the effect of over brightness or over darkness to improve the visual appearance of images.

**Morphological Enhancement** 

Improves and enhances the entire image with low contrast, only if that image contains a single object or of there is no apparent change between the object and the background.

Al-Najdawi, Biltawi, & Tedmori (2015) tested different combinations of contrast enhancement and noise reduction algorithms on a dataset of 1,300 mammogram images to see whether they could better enhance the visual details for radiologists and facilitate the segmentation process. A combination of image enhancing algorithms that included Contrast-Limited Adaptive Histogram equalisation and Median filtering produced a sensitivity of 94.1% and 81.4% for CC and MLO images respectively.

#### Segmentation

Distinguishing the suspicious region from its surrounding areas is an important step in analysing mammogram images. This stage segments mammogram images into several non-overlapping regions, from which regions of interest can be identified and a suspicious mass located. The suspicious area is an area that is brighter than its surroundings, has almost uniform density, has a regular shape with varying size, and has fuzzy boundaries.

Different techniques to improve the quality of the segmentation process have been developed. Chu, Min, Liu & Lu (2015) created a CADe scheme using morphological enhancement combined with simple linear iterative clustering (SLIC). They found that morphological enhancement was effective in suppressing the interferences of the breast tissue and structural noises surrounding the mass regions. By enhancing the contrast between masses and the surrounding tissue, morphological enhancement enables the clustering method SLIC to achieve a segmentation performance, based on simple features such as intensity and spatial distance. The proposed CADe system achieved a satisfactory sensitivity (ranging from 75.9% to 77.9%) while maintaining a relatively low false positive rate (the overall rate dropped from between 8 to 11 images at initial detection to 1 to 2 images per case).

De Sampaio, Silva, de Paiva, Gattass & Diniz (2011) used cellular neural networks to segment mammography images and generate regions of interest. Their algorithm combined shape

features (eccentricity, circularity, circular density disproportion circular and density) and texture features (Ripley's K function, indexes of Moran and Grey) to describe the region of interest. The extracted features were classified as mass and non-mass using support vector machine. This method produced 80.0% sensitivity, 85.6% specificity, with average rate of false positive per image and false negatives per image of 0.84 and 0.2 respectively.

Hong & Sohn (2010) developed an algorithm to detect and segment regions of interest in mammograms using a topographic approach. They observed that topographic representation was largely invariant to brightness and contrast, and therefore provided a robust and efficient representation of the characterisation of mammographic features. As the algorithm achieved 100% detection rate (with 3.8 false positives per image), they suggested that the proposed method would be an effective prompting tool to assist radiologists in breast screen detection.

Improving the quality of the segmentation process has enabled researchers to observe lower rates of false positives per image. Continued development in this area should further reduce this.

#### **Detection of the pectoral muscle**

Pectoral muscles also appear in the mammogram, which may be misinterpreted as masses due to their high intensity values. To reduce the rate of false positives, algorithms generally involve removing the pectoral muscles or a segmentation step to suppress the pectoral muscle region prior to the detection of symptoms. For instance, Mughal, Sharif, & Muhammad (2017) developed a bi-model processing algorithm for the detection of breast cancer that removed pectoral muscles from the test images. Testing the algorithm on two datasets of 322 images and 400 images respectively using different classifiers, they reported performances of 91.0% to 98.5% sensitivity, 88.0% to 97.0% specificity and 88.5% to 96.9% accuracy. To reduce potential bias, they applied the algorithm to a different dataset of 322 images reporting performances of 88.5% to 98.2% sensitivity, 88.1 to 97.0% specificity and 89.0% to 97.5% accuracy.

Agrawal, Vasta & Singh (2014) developed a different algorithm that automatically detected masses from mammograms using saliency-based segmentation even in the presence of pectoral muscles. Testing the model on 322 mammograms (207 are normal and 115 show signs of breast cancer), their algorithm out-performed seven other commonly-used algorithms in reducing false positives while detecting masses in digital mammograms with improved sensitivity. They reported an Area under Curve (AUC) of 0.891 compared to an AUC of 0.601 to 0.787 for the other algorithms.

Successful detection of the pectoral muscles is crucial for meaningful diagnosis. The improvements that have been made to this aspect of the CADe pipeline will aid the subsequent classification of abnormal tissue mass and improve breast cancer detection.

#### Feature extraction

Texture is a commonly used feature in the analysis and interpretation of images. Texture is characterised by a set of local statistical properties of pixel intensities (Mohamed et al., 2014)). Noting that feature extraction is a key issue in designing a CADe system, many studies have proposed new feature extraction methods.

For instance, Singh & Urooi (2016) developed a system that used polar complex exponential transform (PCET) moments as texture descriptors. Testing the system on 200 database images, they found that the system attained a sensitivity of 98.2% and specificity of 97.2%. This

compared favourably with other studies with reported sensitivities ranging from 90.1% to 100.0%, and specificities ranging from 88.1% to 100.0%.

Several studies have noted the incremental computation cost of detection for large numbers of features and have explored reducing the feature subset. For instance, Dhahbi, Barhoumi & Zagrouba (2015) concluded that the selection of a reduced feature subset increased the accuracy of the CADe system. Exploring a new feature extraction method for mammogram description based on curvelets transform and moment theory, they found that curvelet moments yield an accuracy of 91.27% with 10 features for abnormality detection.

Kendall & Flynn (2014) demonstrated a smaller feature size was even more accurate. Using online datasets, they generated forty-one features which were then tested singly and in combinations of two or three. They achieved sensitivities as high as 98% with a specificity of 66% using a k-nearest neighbour classifier (described below) and sensitivity as high as 100% with a specificity of 64% using a naïve Bayesian classifier. In their view, using a small number of features ensured that the CADe system did not become over-trained.

Choi & Ro (2012) developed a new approach for extracting texture features for the characterisation of mammographic masses. They found that by combining texture patterns extracted from both core and margin regions, the algorithm achieved a sensitivity of 77.9% at 2.0 false-positives per image. While previous studies on the same database reported a sensitivity of 72-81%, FPs was higher ranging from 2.0 to 4.5 per image. The authors concluded that their approach was an effective solution for reducing FP signals in CADe schemes.

Taking a different approach, Kooi et al. (2017) decided against choosing an optimal concise set of features and instead developed a deep learning model in the form of a convolutional neural network to read the mammograms. They found that since CNN learns from data and did not have to rely on domain experts, the system outperformed more conventional CADe systems.

Feature extraction is a key issue in designing a CADe system as the feature set must be compact and discriminative to improve system speed and accuracy. The work carried out in this area has identified efficient and effective ways to extract a reduced set of discriminative features for breast cancer.

#### Classification

Once the features related to masses are extracted and selected, the features are inputted into a classifier to classify the suspicious regions as normal or abnormal tissues (benign or malignant). Classifiers such as Artificial Neural Network (ANN), the k-Nearest Neighbour (k-NN) and Support Vector Machine (SVM) are commonly used in mass classification.

Raghavendra et al. (2016) employed k-NN in their CADe obtaining mean

#### Artificial neural network

An artificial neural network (ANN) is a computational model based on the structure and functions of biological neural networks. Information that flows through the network affects the structure of the ANN because a neural network changes - or learns - based on that input and output.

#### k-Nearest Neighbour

K-Nearest Neighbour (k-NN) is one of the simplest classification methods. It is a non-parametric method used for classification and regression in that no parameters are estimated. Instead the proximity of neighbouring input observations in the training set and their corresponding output values are used to predict the class of objects in the validation data set.

#### Support vector machine

Support vector machine (SVM) are supervised learning models with associated learning algorithms that analyse data used for classification and regression analysis. SVM is a binary linear classifier that predicts for each input data which of two possible classes comprise the input.

accuracy, sensitivity and specificity of 98.7%, 99.3% and 98.3%.

Two studies focused specifically on the use of the ANN classifier in mammography interpretation and decision making. Ayer, Chen & Burnside (2013) noted that ANN had many desirable properties that made it well suited for breast cancer detection. ANN was capable of "learning" complicated patterns from data that is difficult for humans to identify. It could also overcome ambiguous and missing data. There were two different ways ANN was used to aid mammography interpretation: first, by applying the classifier directly to the region of interest and, second, by learning from the features extracted from the pre-processed images.

Ayer looked at a range of studies and identified several benefits and limitations of using ANN for classification. The benefits included the high computational power and practical use of ANN. The limitations were:

- lack of a comprehensive assessment of the discrimination accuracy
- overfitting (which occurs when the network overlearns and mimics the training dataset but performs poorly when presented to an external dataset), and
- complexity issues.

They concluded, however, that these limitations could be carefully overcome for the successful application of ANNs in mammography interpretation.

In their study, Wang, Lederman, Tan, Wang & Zheng (2010) found that this classifier achieved sensitivity of 90.0%, which was encouraging, but its results for specificity (42.0%) was less than ideal.

Four studies were identified that examined the impact of other classification systems. Jen & Yu (2015) developed a two-stage classifier to detect abnormal mammograms. Their algorithm yielded sensitivities of 88% and 86% on two respective datasets. They also obtained a specificity of 84%. These results compared favourably with three other automated mass detection systems, which reported sensitivities ranging from 62% to 95.2%.

The method employed by Dong et al. (2015) for classification used chain code to indicate the regions of interest (ROIs). Its internal structure was enhanced by rough set. The convolution vector fields were used to extract 32 features of the ROIs, which were used in training and classification, where the performance of the classifiers Random Forest, SVM, genetic SVM, PSO, PSO-SVM and decision trees are compared. The best performance was obtained using the classifier Random Forest and produced a sensitivity of 94.8%.

The algorithm created by Al-Najdawi et al. (2015) classified masses into four categories, producing a sensitivity of 96.2% and specificity of 94.4% for mass classification.

Mac Parthalian, Jensen, Shen & Zwiggelaar (2010) developed a new nearest neighbour classifier based on fuzzy-rough sets, which uses the nearest neighbours to construct lower and upper approximations of decision classes and classifies test instances based on their membership to these approximations. Applying their method to two publicly accessible datasets containing 322 and 832 mammograms respectively, they obtained accuracies of 91.4% and 89.2% for the two datasets. These rates were like those observed by Ganesan et al. (2013), who noted from the studies he reviewed that the best results were only around 90%. In his view, this was not sufficient enough for implementation in clinical trials.

#### **Deep learning**

While the studies described above have contributed to the discourse on how CADe systems classify suspicious regions, they all employ conventional computer vision technologies based upon detection of hand-crafted imaging features. The discriminative power of the classifiers described above is therefore limited due to the computational costs of identifying definitive features for subset characterisation and optimisation. Existing systems fail to explicitly represent the working principles and knowledge of human experts: expert radiologists normally compare image parts and different images of the breasts to each other (that is, they interpret potentially suspicious regions of the breasts in the context of all other available image information) (Velikova, Lucas, Samulski, & Karssemeijer, 2013). Researchers have investigated ways in incorporate such principles into CADe systems, using machine learning systems.

One form of machine learning is deep learning, which is a neural network with multiple hidden layers that enhance the recognition accuracy of images, thereby increasing its versatility for capturing representative features. Qiu et al. (2015) developed and tested a short-term risk assessment model based on deep learning method on a training set with 200 cases and a testing set with 70 cases. The deep learning CADe system focused on feature detection and risk prediction. The proposed CAD-based risk model yielded a PPV of 69.2% and an NPV of 74.2%, with a total prediction accuracy of 71.4%. In their view, the use of deep learning technology in risk predicting schemes would lead to improved performance in detecting early abnormal symptoms from the negative mammograms. They acknowledged, however, that this work is still in its infancy.

Dheeba, Singh & Selvi (2014) proposed a methodology for the detection of masses using Participle Swarm Optimised Wavelet Neural Network (PSOWNN). After extracting texture energy measures from the mammograms, the algorithm classified the suspicious regions using a pattern classifier. Applying their method to a database of 216 mammograms, they found that the proposed algorithm produced a sensitivity of 94.2% and specificity of 92.1% for mass classification.

Dhungel, Carneiro & Bradley (2017) used a deep learning classifier that was pre-trained with a regression to hand-crafted feature values and fine-tuned based on the annotations of the breast mass classification dataset. They achieved a sensitivity of 98.0% and specificity of 70.0% for mass detection.

Wang et al. (2016) compared the effectiveness of the deep learning model they developed on different lesion types. They achieved a discriminative accuracy of 87.3% for microcalcifications but only 61.3% for masses. They concluded that for large datasets deep learning was superior to standard methods for the discrimination of microcalcifications.

Machine learning techniques have shown the ability to select complex, previously unimagined features from images. Their application within CADe has the potential to significantly improve the classification of lesions and thereby reduce the rate of false-positives generated by conventional CADe systems.

#### 13.4.3. What cost and safety findings have been reported?

No studies reporting on the safety of CADe were identified.

One study commented directly on the cost of using CADe systems to detect breast cancer. Guerriero et al. (2011) investigated whether the use of single reader with CADe is more cost-effective than double reading. They found that the use of CADe increased costs compared to

double reading because of the cost of the CAD equipment, staff training, and the higher assessment cost associated with CAD are greater than the savings in reading costs. The introduction of single reader with CADe, in place of double reading, would produce an additional cost of £227 and £253 per 1,000 women screened in high and average screening volume areas respectively. In low screening volume areas, the high cost of purchasing and maintaining the equipment would result in an additional cost of £590 per 1,000 women screened. They concluded that without improvements in the effectiveness of CADe (i.e., a decrease in recall rate), this reading strategy is unlikely to be a cost-effective alternative to double reading for mammography screening.

Reader performance: One of the biggest impacts on cost is reader performance. Two systematic reviews discussed the issue of reader performance. Azavedo et al. (2012) observed that the lack of trained radiologists remains a problem even if CADe is used. In their view, new generations of radiologists need to be secured. Besides, being able to discuss uncertain cases with an experienced colleague is essential for both educational purposes and to avoid too many false positives/false negatives. They noted that when working with CADe, a single radiologist would have to make the final decision on whether to recall the woman for further work-up. This decision may depend on a single CADe mark in an area where the radiologist did not react initially. In their view, a radiologist should only use CADe if they were highly experienced.

Karssemeijer (2010) saw reader performance as an opportunity to promote the use of CADe. He noted that radiologists understand the high negative predictive value of CADe. They know the likelihood that a cancer is present becomes lower when CADe does not mark the region and higher when it does, especially if marked in multiple views. He considered that CADe use would lead to better decisions on average. However, Azavedo et al. (2012) noted that mammography screening involves high throughput, which places high demands on smooth screening workflows.

Five studies also looked at the performance of radiologists and other technicians who read mammogram images. These studies found that the main disadvantage of CADe systems was the high rate of false-positive marks. This led in great extent to an elongation of mammogram reading time because the radiologist could be distracted by many CADe marks. The resulting increase to total reading time was observed by Hupse R et al. (2013) who found that the use of CAD lengthened the time that radiologists took to read the images by approximately 10 seconds per image.

Reader performance is effected by the experience of the radiologist. Sohns, Angic, Sossalla, Konietschke & Obenauer, (2010) found that CADe elongates reading time of new and less experienced radiologists. They observed no significant difference between the reading time of radiologists experienced with CADe systems.

In terms of reader performance, Povyakalo, Alberdi, Strigini & Ayton (2013) found that CADe helped readers with less discriminating ability but hindered more experienced readers. They considered that such differential effects were clinically significant and therefore needed to be factored into the future development of CADe systems and protocols.

Cole et al. (2014) examined the performance of radiologists with or without CADe by asking them to rescreen mammograms that they had previously inspected visually. They wanted to know whether increased familiarity with CADe systems changed their decisions. They found that sensitivity and specificity differences between the two-different reading strategies were not significantly different. They concluded that radiologists rarely changed their diagnostic decisions after the addition of CADe. The application of CADe did not have a statistically significant effect on radiologist performance in interpreting digital mammograms.

Reader performance was also looked at by Hupse et al. (2013). They developed software that kept CADe marks and their associated suspiciousness scores hidden unless their location was queried by the radiologist. Comparing their findings with traditional CADe prompts used in current clinical practices, they found that the difference in reader performance was significant (91% for the interactive mode and 75% for commercial systems). In a separate article about the same study, Hupse explained that the interactive use of CADe was more effective than the standard use of CADe because CADe was aimed at the prevention of perceptual oversights. An early study also found an increase in detection performance with interactive CADe (Samulski et al. (2010)).

Taking a slightly different approach to breast cancer screening, Bessa et al. (2014) developed a pre-CADe system that identified normal mammograms instead of detecting suspicious ones. In their view, automatically detecting normal cases and screening them out of the process would alleviate the human effort involved in the CADe process, giving radiologists more time to focus on more ambiguous cases. Applying the system to 17,900 mammograms, they correctly identified 20% of the normal mammograms: a result that is like other more complex approaches.

#### 13.4.4. Does this technology reduce deaths due to breast cancer through early detection?

Two major shortcomings in study design apply to CADe as a reading strategy. One is survival rate, which is the most important outcome in mammography screening. None of the studies investigated the survival rates between a single mammographic reading by one breast radiologist plus CADe or double reading involving two breast radiologists and CADe. The other shortcoming is incomplete follow-up, which may lead to sensitivity being overestimated (Azavedo et al., 2012).

# 13.4.5. Does this innovation show higher sensitivity and specificity for women with dense breasts and women who have breast surgery/augmentation compared to single human view?

#### Systematic reviews

One systematic review looked at the implications of breast density on CADe performance. Dromain et al. (2013) noted that the performance of CADe depended on background breast density. They concluded that while overall breast density did not appear to exert an impact on CADe detection of breast cancer, there was a decrease in sensitivity among cancers that manifested as masses (89.0% for non-dense breasts compared to only 83% for dense breasts). The review also noted that CADe was found to have a greater sensitivity for detecting ILC and DCIS, rather than invasive ductal carcinoma.

#### **Retrospective reviews**

Five retrospective reviews commented on the implications of breast density on CADe performance.

In their study comparing the use of CADe with double reading, James et al. (2010) observed breast density had some bearing on breast cancer detection. Double reading by humans
performed better than single reading with CADe in women with denser background patterns. However, Sadaf et al. (2011) found there were no statistical differences (p>0.1) in CAD detection of cancers in dense breasts (89.8%) versus non-dense breasts (91.2%). Bargalló et al. (2013) also observed that background density did not affect detection. However, they noted that most of their patients had non-dense breast and small lesions were easier to depict in non-dense breasts. Consequently, no definitive conclusions could be drawn on this subject from this study.

The work of de Sampaio, Silva, de Paiva & Gattass (2015) presented a computational method to aid in the detection of masses based on the density of the breast. In the segmentation step, they used a micro genetic algorithm to create a texture proximity mask and select regions suspected of containing lesion. After reducing the number of false positives, the algorithm chose the suspicious regions that generated the best training models. They recorded a sensitivity of 89.1% for dense breasts (which was 4.9% lower than non-dense breasts).

Neto et al. (2017) developed a computational method using particle swarm optimisation to detect masses in dense breasts, obtaining 97.5% sensitivity, 92.3% specificity, 94.8% accuracy and 0.38 false positives per image. In their view, the proposed methodology could assist in detection of masses, providing the radiologist with a second opinion in the early detection of breast cancer in women with dense breasts.

# **13.4.6.** Has the technology been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

CADe has not been incorporated into any national screening programs.

**13.4.7.** Has a national position statement been published about the technology, and if so, what is the position? Is there consensus in position statements?

No national position statements on the use of CADe in breast cancer screening were identified in the literature search.

# **14. ARTIFICIAL INTELLIGENCE**

#### Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

The use of artificial intelligence software to interpret mammograms to predict malignant versus benign tumours, and thereby aid in biopsy decisions, was not discussed in the 2009 ANZHSN report.

# 14.1. What does AI mean in breast cancer screening?

To improve interpretive accuracy, sophisticated computerised systems have been developed to support radiologists during the screening process to interpret medical images by enhancing the images and providing in-depth analysis of visual details. While these systems use AI software to assist feature extraction and classify the suspicious regions as normal or abnormal tissues, CADe still requires human intervention for interpreting the final results (Ganesan et al., 2013). It is therefore likely to be impacted by false-prompting of its human reader (Houssami, Lee, Buist & Tao, 2017).

With recent advances in computer processing capabilities, the rapid growth of digital capture and storage of health data, and cloud-based storage capabilities, researchers have identified the potential of AI to perform image interpretation on its own and, therefore, improve accuracy in clinical medicine (Trister, Buist & Lee, 2017). One rapidly growing field of AI is machine learning (ML), which allows computers to learn without explicit programming through automated extraction and analysis of complex data. ML can build complex statistical models from large datasets to potentially identify new variables and combinations of variables that can predict patient outcomes. Two newer ML techniques are:

- 1. deep learning (a class of learning algorithms that stacks a set of non-linear operations to extract features and explore transformations from training data) and
- 2. reinforcement learning (this technique allows ML to tackle a problem involving sequential decisions and determine the ideal actions within a specific context that maximises rewards and performance)

These ML techniques have shown the ability to select complex, previously unimagined features from images.

Researchers are currently developing AI systems to interpret mammograms (and other imaging technologies) as a stand-alone system and avoid the need for a radiologist to confirm the presence of malignancy. These systems seek to predict malignant vs. benign tumours to aid in biopsy decisions. The technique involves scanning large datasets and cross-checking them with results from mammogram X-rays and clinical reports. The software extracts information from clinical reports and is run through a risk assessment model. The software is developed through machine learning, based on high-risk lesions. The model incorporates patterns among many different data elements that include demographics, family history, past biopsies, and pathology reports.

AI systems are still being refined and are not currently applied in breast screening clinical practice in Australia.

# 14.2. Summary of key findings

- The use of AI in breast cancer screening of asymptomatic women is still in its early stages but is closely linked to developments in CAD. While promising, machine learning is in its infancy with respect to demonstrating its utility in cancer screening.
- There is currently no indication of the timeframe in which the full clinical potential of AI for breast cancer detection will be realised; however, preliminary results for AI are promising.
- Current research is not sufficient to be able to identify whether the use of AI in interpretation is able to reduce deaths due to breast cancer through early detection.
- AI has not been incorporated into any national screening programs, nor are there any national position statements that have been released on their use in breast cancer screening for asymptomatic women.

# 14.3. Literature search results (number of studies returned)

From the literature search a total of 54 abstracts of peer reviewed articles were identified that related to AI software to interpret mammograms. Abstract contents were then reviewed, and 32 articles were excluded because they studied AI for diagnostic or treatment purposes as opposed to a screening modality or were related to the use of AI systems in CADe (which are discussed in the previous section). Eight other articles were subsequently excluded because they could not be located or were dissertations. A total of 14 articles were reviewed to answer the key research questions relating to the use of AI as a breast cancer screening tool.

#### Systematic and/or literature reviews

Four systematic and/or literature reviews were identified (Cabitza, Rasoini & Gensini, 2017; Houssami, Silva, de Paiva & Gattass, 2017; Trister, Buist & Lee, 2017; Ganesan et al., 2013)

#### RCTs

None identified

#### **Prospective studies**

None identified

#### **Retrospective studies**

Ten retrospective studies were identified (Becker et al., 2017; Carneiro & Bradley, 2017; Dhungel, Carneiro & Bradley, 2017; Neto, Silva, Paiva & Gattass, 2017; Teare, Fishman, Benzaquen, Toledano & Elnekave, 2017; Guo et al., 2016; Kallenberg et al., 2016; de Sampaio, Silva, de Paiva & Gattass, 2015; Ertosun & Rubin, 2015; Beck et al., 2011)

#### **Grey literature**

None identified

# 14.4. Study findings and discussion

## 14.4.1. What stage of development or trial is this innovation at?

#### Systematic reviews

Two systematic reviews discussed the current use of AI in breast cancer detection.

Houssami et al. (2017) and Trister et al. (2017) both noted that conditions are ideal for AI solutions to transform medical imaging because of existing capabilities in computer vision, automated digital image feature analysis, large stores of digital images through picture archiving and communication systems, linked electronic medical records, and the binary outcome of imaging-based screening tests like mammography.

Trister et al. (2017) noted that digital images can be analysed at the individual pixel level, converting single mammographic images into millions of individual variables. Given a gold standard to learn from, computer algorithms can cluster these pixel-level variables and identify novel image features associated with clinically relevant breast cancers. This meant, machine learning can combine image-level, patient-level and tumour-level variables to develop complex algorithms that could be better than human breast cancer screening accuracy.

Trister et al. (2017) noted that the ability of machine learning to identify high-risk breast cancers was shown by the C-Path (Computational Pathway) tool, an automated pathologist algorithm. Rather than focusing on known histologic features, C-Path uses unsupervised machine learning and automated image processing to identify thousands of novel imaging features on digital pathology specimens. These novel image-based features were then used to predict patient survival, with performance that surpassed that of community pathologists.

The C-Path system was developed by Beck et al. (2011), who applied the tool to microscopic images from two cohorts of breast cancer patients ( $n_1 = 248$  and  $n_2 = 328$ ). The prognostic score generated by their system was strongly associated with overall survival in both cohorts ( $P \le 0.001$ ). One important lesson from this study was the need for unsupervised learning to identify previously undiscovered imaging predictors, as opposed to simply relying on known associates. Another lesson was that extremely large, high-quality, harmonised, unbiased and generalisable datasets were needed to train and validate algorithms.

Seven retrospective studies explored the application of AI in breast cancer detection. One study evaluated the diagnostic accuracy of multipurpose image analysis software based on deep learning with artificial neural networks (dANNs) for the detection of breast cancer. Using mammograms from 3,228 women, Becker et al. (2017) trained the neural network on 143 control cases from the healthy population and then tested 143 patients with known breast cancer. The performance of the trained neural network was also trained on an external dataset of 35 healthy patients and 35 with breast cancer. Three radiologists (3, 5 and 10 years of experience) evaluated the test data set. In a second step, the neural network was trained with all cases from January to September and tested on cases from October to December 2012. The radiologists also evaluated this second test dataset. In the first step, the AUC of the trained neural network was 0.81 and comparable on the test cases 0.79 (p=.63). One of the radiologists showed almost equal performance (0.83, p=.63) whereas two were slightly better (0.90 and 0.94, p<.016). In the second step, performance of the neural network (0.82) was not significantly different from the human performance (0.77-0.87, p>.016): however, radiologists had lower

sensitivity and greater specificity than the neural network. Their results demonstrated that dANNs were able to detect cancer in mammograms with similar accuracy to radiologists, even in a screening-like cohort with low breast cancer prevalence.

The diagnostic accuracy achieved in the Becker study was comparable to other published deep learning models (Kallenberg et al., 2016 and Ertosun & Rubin, 2015). While the Kallenberg study made use of a specialised dANNs for breast cancer detection, Ertosun compared different generic dANNs. Slightly higher results were obtained by Carneiro & Bradley (2017), Dhungel, Carneiro & Bradley (2017) and Guo et al. (2016), who obtained AUCs of 0.90, 0.90 and 0.97 respectively. In terms of distinguishing between malignant and benign masses, Dhungel et al. (2017) reported a sensitivity of 98.0% and specificity of 70.0%. It should be noted that the Carneiro study was the only one of these five studies not to use a large dataset to train their neural network, which is problematic as only a few high-quality datasets are publicly available. In addition, none of the four studies offered a direct comparison to a radiologist's performance.

In a recent study, Teare, Fishman, Benzaquen, Toledano & Elnekave (2017) explained that CADe systems traditionally used in breast cancer detection deploy conventional computer vision technologies based upon detection of hand-crafted imaging features broadly categorised into masses or microcalcifications. In contrast, machine learning methods were based upon feature discovery within samples of ground truth validated images. While ML had made substantial advances in feature recognition, fundamental differences in data acquisition and content have limited the transferability of ML image algorithms to the domain of radiology.

They presented a ML based mammographic malignancy detection algorithm that applied a novel false-colour enhancement technique and analytics architecture that was capable of discerning malignancy with a sensitivity of 91.0% and a specificity of 80.0%. Their algorithm achieved stand-alone accuracy like that reported for expert radiologists.

In an interesting development, Houssami et al. (2017) mentioned that a crowdsourcing challenge is currently underway that asks participants to develop predictive algorithms that reduce false-positive mammograms, while maintaining or improving cancer detection. The goal of the Digital Mammography DREAM Challenge is to enhance the predictive accuracy of algorithms so that they can be used in routine clinical practice (i.e., to develop an algorithm that matches the performance of expert radiologists). The challenge has brought together over 120 independent teams of data experts from across the globe. Participating teams used hundreds of thousands of de-identified mammograms and clinical trial data to create algorithms that can determine a woman's cancer status in the 12 months following a mammography screening. The first stage of the challenge, where teams were asked to develop their algorithms, has been completed. In Houssami and colleague's opinion, the next stage of the challenge (the 'collaborative' phase) where the best teams work together to improve the final algorithm holds promise for improved accuracy of AI systems. Some papers that were submitted as part of the algorithm development stage are available on-line.<sup>16</sup>

<sup>&</sup>lt;sup>16</sup> See Ribl, Horvath, Unger, Pollner, Csabai, *Detecting and classifying lesions in mammograms with Deep Learning*, available at <u>https://arxiv.org/pdf/1707.08401.pdf</u> (accessed 22 February 2018) and Shen, L., *End-to-end Training for Whole Image Breast Cancer Diagnosis using An All Convolutional Design*, available at <u>https://arxiv.org/ftp/arxiv/papers/1708/1708.09427.pdf</u> (accessed 22 February 2018).

#### 14.4.2. What are their considered potential clinical value in five years? In 10 years?

While several authors discussed the possible future use of the AI systems they are developing, only three talked about the potential clinical value of AI.

Houssami et al. (2017) noted that existing research into the capability of AI in mammography screening has mainly focused on interpretive accuracy with the goal of reducing false positives. In their view, there remains much opportunity to explore the role of AI in breast cancer screening, for instance to improve the detection of more aggressive forms of breast cancer or faster growing lesions. The technology should not be limited to the detection of cancers 'missed' at human screen-reading (with the aim of reducing the frequency of interval cancers).

Trister et al. (2017) noted that machine learning was in its infancy with respect to demonstrating its utility in cancer screening. There is no guarantee that effects to merge clinical, genomic, biological and image data will lead to improved cancer screening outcomes. In reaching this conclusion, they referred to the fact that CADe, which incorporates AI techniques into its systems, has shown no improvements in interpretative accuracy since it was first employed in screening.

Houssami et al. (2017) raised issues relevant to the application of AI in breast cancer screening that warranted consideration. These included the social and ethical concerns and implications inherent in entrusting cancer detection to an AI model, and the possibility of unintended consequences. For instance, Cabitza, Rasoini & Gensini (2017) noted that an over-reliance on the capabilities of automation was starting to be observed which was affecting the ability of clinicians to make informed decisions based on detectable signs, symptoms and available data. They referred to a study of 50 mammogram readers that found a 14% decrease in diagnostic sensitivity when presented with challenging images marked by CADe. They felt that further research was needed to better understand whether the overreliance on AI systems that could out-perform or perform as well as human readers could also cause a subtle loss in self-confidence and affect the willingness of radiologists to provide a definitive interpretation or diagnosis.

Houssami et al. (2017) felt these issues needed to be fully explored to provide an understanding of societal perspectives, and to define an ethical-legal framework for the potential application of AI models in breast cancer screening. This would also help ensure that the purpose for which AI models are developed is acceptable to all stakeholders.

#### 14.4.3. What cost and safety findings have been reported?

No studies commented on the cost effectiveness of using artificial intelligence systems to detect breast cancer, or the safety implications of this technology.

#### 14.4.4. Does this technology reduce deaths due to breast cancer through early detection?

The use of AI software for breast cancer screening is still in its infancy and, accordingly, no studies have been identified that discuss its ability to reduce deaths through early detection.

# 14.4.5. Does this innovation show higher sensitivity and specificity for women with dense breast and women who have breast surgery/augmentation compared to single human view?

No studies have been identified that discussed whether AI operating as a stand-alone system showed higher sensitivity and specificity for women with dense breasts. However, two studies examining the use of machine learning as part of CADe systems commented on the its effectiveness in detecting masses in dense breasts. In these studies, de Sampaio, Silva, de Paiva & Gattass (2015) and Neto, Silva, Paiva & Gattass (2017) reported a sensitivity of 89.1 and 97.5% respectively for women with dense breasts. Neto and colleagues claimed that the computational method they devised worked best on women with dense breasts (97.5% sensitivity and 92.3% specificity). Consequently, the development of future AI systems is likely to have significant application for screening women with dense breasts.

# 14.4.6. Has the technology been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

The use of AI software has not been incorporated into any national screening programs.

14.4.7. Has a national position statement been published about the technology, and if so, what is the position? Is there consensus in position statements?

No national position statements on the use of AI software in breast cancer screening were identified in the literature search.

# 15. TELE-MAMMOGRAPHY

#### Findings from the Australia and New Zealand Horizon Scanning Network's 2009 report: New and emerging technologies for breast cancer detection

The use of tele-mammography for the early detection of breast cancer in asymptomatic women was not covered in the 2009 report.

# **15.1.** Tele-mammography as a reading/interpreting technology

Tele-mammography is the real-time, off-site interpretation of digital mammograms through wireless, fibre-optic or wire-based networks between the site where the mammogram is taken, usually a remote area, and a centre where mammograms are interpreted. This process should not be confused with tele-radiology, where an image acquired in one location is transmitted to another location so that they can be viewed and interpreted for screening or diagnostic purposes by a radiologist at a later point in time.

The primary purpose of tele-mammography is to improve access to mammography for underserved populations (such as those living in isolated rural areas to those in low-income inner-city neighbourhoods) and thereby increase the rate of early breast cancer detection. A digital system can be fitted into a mobile unit for visiting areas where there is a lack of radiologists or mammography units (Bhateja et al., 2014). There are several advantages of this process including the ability for radiologists to cover multiple sites, better consultation with the patient or between physicians, less time to diagnosis and consequently reduced period of patient anxiety. The real-time evaluation of mammogram images allows women from rural areas to obtain their results before they need to travel back to their homes, particularly those living in areas that have more limited communication capabilities.

The process of tele-mammography is based on an essential triad: an image sending station, a transmission network, and a receiving image station that has a high-resolution display monitor that can be used for clinical purposes. Specialised computer programs to transmit digital images to remote locations without losing information content are needed, as is adequate bandwidth to accommodate the timely transmission of images. Tele-mammography also requires the standardisation of image characteristics – such as image resolution, bit-depth and intensity response – so that images may be read at the remote workstation (Patil & Ahmed, 2014).

For optimal effect, images should be displayed simultaneously on workstations at the two sites with on-screen display of dual cursors controlled by the user at each site. This dual display, combined with multi-media videoconference link, facilitates:

- real-time tele-management of screening mammography examinations between the technician at the image acquisition site and the radiologist at the image interpretation site, and
- real-time tele-consultation between physicians at two sites (Patil & Ahmed, 2014).

For tele-mammography to be clinically successful, the interpretation of the transmitted images must be done as rapidly as conventional FFDM (i.e., without reduction in throughput). There are some costs associated with tele-mammography as the rapid transmission of digital images requires the use of high-priced infrastructure (eg, satellite transmission and high volume, highspeed networks). Other issues include the feasibility of sending large data sets, network and storage needs, and protecting patient privacy.

# **15.2.** Summary of key findings

- The use of tele-mammography appears to be at least moderately widespread, and seemingly has a range of benefits, including the ability to provide services to remote areas, and better utilisation of radiologists' time.
- Studies have found that there is no significant difference in screening outcomes between traditional mammography and tele-mammography technologies (such as digital cameras) and visualisation devices (such as LCD screens) are suitable for performing tele-mammography.
- It is not clear whether tele-mammography has been incorporated into any national screening programs, or whether its use has been determined at sub-national level.

# **15.3.** Literature search results (number of studies returned)

From the literature search, a total of 170 abstracts of peer reviewed articles were identified that related to the electronic transmission of breast screen images. Abstract contents were then reviewed, and 166 articles were excluded because they related to tele-radiology, were only tangentially related to tele-mammography, or had little or no application to breast screening. Three other articles were identified through manual searches meaning that a total of seven articles were reviewed to answer the key research questions in relation to the use of tele-mammography for breast cancer screening.

#### Systematic and/or literature reviews

Two studies: Bashshur et al., 2016; McClung et al., 2014

RCTs

None identified

#### **Prospective studies**

None identified

#### **Retrospective studies**

Four studies: Fruehwald-Pallamar et al., 2013; Patil & Ahmed, 2014; Salazar et al., 2016; Salazar et al., 2016 (JACR); Salazar et al., 2014.

#### **Grey literature**

Bhateja et al., 2014

# 15.4. Study findings and discussion

# 15.4.1. What stage of development or trial is the new technology at?

In respect of teleradiology, overall, it was estimated that 40 per cent of radiology practices in the United States performed 'outside readings' in 2007, which accounted for 11% of their total workload and 4% of the total workload of all radiologists (Bashshur et al., 2016). No statistics are presented in the literature about the prevalence of tele-mammography.

There appears to be consensus in the literature that, performed correctly, tele-mammography can produce results that are not significantly different from mammograms read by an on-site radiologist. For example, Fruehwald-Pallamar et al. (2013) found that Contrast Detail Mammography (CDMAM) images were identical before and after transmission, and within the limiting values for CDMAM image technology as specified in European guidelines. They found that uncompressed digital mammograms can be transmitted to different institutions with different workstations, without loss of information. They also found that the transmission process does not significantly influence image quality, lesion detection, or BIRADS assessment.

Similarly, Salazar et al. (2016) noted that computed radiography and FFDM were useful in the implementation of tele-mammography, with several studies reporting no significant differences between film-screen mammography and digital modalities. Fruehwald-Pallamar et al. (2013) argued that state-of-the-art equipment, frequent calibration and inspection, as well as quality assurance by a medical physicist were necessary conditions for obtaining mammograms of good quality.

In addition to these techniques, Salazar et al. (2014) and Salazar et al. (2016) found that lowercost image capturing devices could also produce good results. Salazar et al. (2016) said that digital images from a digital camera and specialised scanner were a good-quality low-cost alternative to computed radiology examinations, even for heterogeneously dense and extremely dense breasts (BIRADS 3-4), and for amorphous calcifications (albeit with better performance observed for a specialised scanner than the digital camera). Overall, Salazar et al. (2014) and Salazar et al. (2016) concluded that that there were no significant differences between the interpretation of film-screen or CR mammography examinations and soft copy examinations produced by a specialised film digitiser or a digital camera. Salazar et al. (2014) also found that consumer-grade colour displays (eg, LCD and LED screens) were suitable for detecting malignant or benign mammograms in tele-mammography screening programs.

When tele-mammography was initially attempted, difficulties arose because of technical limitations, mainly due to limited transmission capabilities and problems with displaying the images. However, Fruehwald-Pallamar et al. (20130 noted that tele-mammography was now viable due to the availability of digital mammography systems and picture archiving and communication systems (PACS) with large storage capacity, and data links with extremely high transmission rates (although network reliability remains an important consideration). In contrast, Bhateja et al. (2014) said that several challenges and tele-mammography constraints continued to exist for the execution of tele-mammography. Firstly, image quality needed to be appropriately maintained during communication, meaning that frequent calibration of tele-mammography equipment may be required. Secondly, in many cases the mammograms could not be verified onsite due to issues such as positioning, artefacts, compression, noises and contrast, and were thus not suitable for direct transmission. Thirdly, there were integration and

interoperability constraints in providing an interface between the mammography units and communication/network hardware for transmission (Bhateja et al., 2014).

Patil & Ahmed (2014) noted that due to the large size of the digital images, a high bandwidth was required for the image to be transmitted. They noted that ISDN (Integrated Services Digital Network) systems, while affordable, did not possess the bandwidth sufficient to transmit digital images, especially in rural locations with weak signal environment. On the other hand, ATM (Asynchronous Transfer Mode) charges were high but suitable for tele-mammography use.

BSA is currently undertaking a trial of tele-mammography through the Remote Radiology Project. This project seeks to contribute to the provision of research into the use of remote radiology as a service delivery model within the BSA program. No publications have come out of this work so far.

#### 15.4.2. What is its considered potential clinical value in five years? In ten years?

The identified literature did not discuss timeframes in which potential clinical value may be realised. However, Salazar et al. (2016) noted that tele-mammography, particularly services that utilised cheaper screening technology, may help to provide widespread screening services in underserved areas.

Bhateja et al. (2014) set out an approach which, they argued, would address many of the limitations in current tele-mammography services. They found that interoperability of the devices involved in tele-mammography services would be improved by incorporating the IEEE 21451-x family of standards (which are standards for smart transducer interfaces for sensors and actuators on networked devices). They also found that adoption of the nonlinear polynomial filters significantly improved the performance of the tele-mammography model in respect of visualisation of suspicious regions in test mammograms.

#### 15.4.3. What cost and safety findings have been reported?

As outlined above, Salazar et al. (2016) noted that a \$450 camera (plus \$400 for a support system and a lightbox) and a \$15,000 specialist digitiser did not produce significantly different screening outcomes than more expensive tele-mammography equipment, suggesting that significant savings could be made, or the reach of tele-mammography services significantly extended, with no significant impacts on safety. In assessing the benefits of teleradiology more generally, the Canada Health Infoway Diagnostics Imaging Benefits Evaluation Report found that PACS had reduced turnaround time of radiology reports by an average of 41%, and increased radiologist productivity by 27% (McClung et al., 2014).

#### 15.4.4. Does this technology reduce deaths due to breast cancer through early detection?

Current research does not explicitly identify whether tele-mammography is able to reduce deaths due to breast cancer through early detection of cancer in asymptomatic women. However, studies such as Salazar et al. (2016) did refer to the potential of tele-mammography to provide widespread screening services in underserved areas.

# 15.4.5. Does this innovation show higher sensitivity and specificity for women with dense breast and women who have breast surgery/augmentation compared to single human view?

No studies have been identified that discussed whether AI operating as a stand-alone system showed higher sensitivity and specificity for women with dense breasts. However, one study was identified that evaluated the reliability of four alternatives for tele-mammography – computed radiology, printed film, a film digitiser and a digital camera – in terms of interpretation agreement when using the BIRADS lexicon. In their study published in the Journal of the American College of Radiology, Salazar et al. (2016) found a very high inter-device agreement on management recommendations among the evaluated devices based on the BIRADS final assessment categories. This suggested the possibility of using less expensive devices in tele-mammographic screening programs, which would enable high-quality screening services to be provided to underserved populations at low cost.

# 15.4.6. Has this technology been implemented into a national screening program? If so, what outcomes have been achieved? What implementation issues arose?

As noted above, it appears from the literature that tele-mammography services are at least moderately widespread. Despite this, it is not clear whether this technology has been implemented as part of national screening programs, or whether it has been implemented as part of sub-national decision-making.

# **15.4.7.** Has a national position statement been published about this innovation, and if so, what is the position? Is there consensus in position statements?

No national position statements on the use of tele-mammography were identified in the literature search. However, broader statements on teleradiology have been issued in several areas, including the United States and Europe. The first American College of Radiology Standard for Teleradiology was issued in 1994, was updated several times before expiry in 2007. In addition, the ACR and the European Society of Radiology recently released white papers on teleradiology practice, with the ACR stipulating that the Food and Drug Administration must approve teleradiological equipment, and that image data integrity must be maintained at all system levels and times for both US and international teleradiology (Bashshur et al., 2016).

### REFERENCES

### BIOMARKERS

## **Blood tests**

Cheuk, I. W. Y., Shin, V. Y., & Kwong, A. (2017). Detection of Methylated Circulating DNA as Noninvasive Biomarkers for Breast Cancer Diagnosis. Journal Of Breast Cancer, 20(1), 12–19. https://doi.org/10.4048/jbc.2017.20.1.12

Cree, I. A. (2011). Improved blood tests for cancer screening: general or specific? BMC Cancer, 11(1), 499. https://doi.org/10.1186/1471-2407-11-499

Cree, I. A., Uttley, L., Buckley Woods, H., Kikuchi, H., Reiman, A., Harnan, S., Whiteman, B.L., Philips, S.T., Messenger, M., Cox, A., Teare, D., Sheils, O., Shaw, J., UK Early Cancer Detection Consortium. (2017). The evidence base for circulating tumour DNA blood-based biomarkers for the early detection of cancer: a systematic mapping review. BMC Cancer, 17(1), 697.

Giordano, L., Gallo, F., Petracci, E., Chiorino, G., & Segnan, N. (2017). The ANDROMEDA prospective cohort study: predictive value of combined criteria to tailor breast cancer screening and new opportunities from circulating markers: study protocol. Bmc Cancer, 17. https://doi.org/10.1186/s12885-017-3784-5

Godfrey, A. C., Xu, Z., Weinberg, C. R., Getts, R. C., Wade, P. A., DeRoo, L. A., Sandler, D.P., Taylor, J. A. (2013). Serum microRNA expression as an early marker for breast cancer risk in prospectively collected samples from the Sister Study cohort. Breast Cancer Research, 15(3), R42. https://doi.org/10.1186/bcr3428

Holdenrieder, S., Pagliaro, L., Morgenstern, D., & Dayyani, F. (2016). Clinically Meaningful Use of Blood Tumor Markers in Oncology. Biomed Research International. https://doi.org/10.1155/2016/9795269

Hyun, K. A., Kim, J., Gwak, H., & Jung, H. I. (2016). Isolation and enrichment of circulating biomarkers for cancer screening, detection, and diagnostics. Analyst, 141(2), 382–392. https://doi.org/10.1039/c5an01762a

Kloten, V., Becker, B., Winner, K., Schrauder, M. G., Fasching, P. A., Anzeneder, T., Veeck, J., Hartmann, A., Knuchel, R., Dahl, E. (2013). Promoter hypermethylation of the tumor-suppressor genes ITIH5, DKK3, and RASSF1A as novel biomarkers for blood-based breast cancer screening. Breast Cancer Research, 15(1), R4. https://doi.org/10.1186/bcr3375

Mostert, B., Sleijfer, S., Foekens, J. A., & Gratama, J. W. (2009). Circulating tumor cells (CTCs): Detection methods and their clinical relevance in breast cancer. Cancer Treatment Reviews, 35(5), 463–474. https://doi.org/10.1016/j.ctrv.2009.03.004

Ng, E. K. O., Li, R., Shin, V. Y., Jin, H. C., Leung, C. P. H., Ma, E. S. K., Pang, R., Chua, D., Kent-Man, C., Law, W.L., Law, S.Y.K., Poon, R.T.P., Kwong, A. (2013). Circulating microRNAs as Specific Biomarkers for Breast Cancer Detection. PLoS One, 8(1). https://doi.org/10.1371/journal.pone.0053141 Volik, S., Alcaide, M., Morin, R. D., & Collins, C. (2016). Cell-free DNA (cfDNA): Clinical Significance and Utility in Cancer Shaped By Emerging Technologies. Molecular Cancer Research, 14(10), 898–908. https://doi.org/10.1158/1541-7786.MCR-16-0

Weigel, M. T., & Dowsett, M. (2010). Current and emerging biomarkers in breast cancer: prognosis and prediction. Endocrine-Related Cancer, 17(4), R245–R262. https://doi.org/10.1677/ERC-10-0136

#### Saliva tests

Al-Muhtaseb, S. (2014). Serum and saliva protein levels in females with breast cancer. Oncology Letters, 8(6), 2752. https://doi.org/10.3892/ol.2014.2535

Hernandez-Arteaga, A., Nava, J. D. Z., Kolosovas-Machuca, E. S., Velazquez-Salazar, J. J., Vinogradova, E., Jose-Yacaman, M., & Navarro-Contreras, H. R. (2017). Diagnosis of breast cancer by analysis of sialic acid concentrations in human saliva by surface-enhanced Raman spectroscopy of silver nanoparticles. Nano Research, 10(11), 3662–3670. https://doi.org/10.1007/s12274-017-1576-5

Kloten, V., Becker, B., Winner, K., Schrauder, M. G., Fasching, P. A., Anzeneder, T., Veeck, J., Hartmann, A., Knuckel, R., Dahl, E. (2013). Promoter hypermethylation of the tumor-suppressor genes ITIH5, DKK3, and RASSF1A as novel biomarkers for blood-based breast cancer screening. Breast Cancer Research, 15(1), R4. https://doi.org/10.1186/bcr3375

Liu, J., & Duan, Y. (2012). Saliva: a potential media for disease diagnostics and monitoring. Oral Oncology, 48(7), 569–577.

Pfaffe, T., Cooper-White, J., Beyerlein, P., Kostner, K., & Punyadeera, C. (2011). Diagnostic Potential of Saliva: Current State and Future Applications. Clinical Chemistry, 57(5), 675–687. https://doi.org/10.1373/clinchem.2010.153767

Streckfus, C. F., Brown, R. E., & Bull, J. M. (2010). Proteomics, morphoproteomics, saliva and breast cancer: An emerging approach to guide the delivery of individualised thermal therapy, thermochemotherapy and monitor therapy response. International Journal of Hyperthermia, 26(7), 649–661. https://doi.org/10.3109/02656736.2010.506470

Sugimoto, M., Wong, D. T., Hirayama, A., Soga, T., & Tomita, M. (2010). Capillary electrophoresis mass spectrometry-based saliva metabolomics identified oral, breast and pancreatic cancer-specific profiles. Metabolomics, 6(1), 78–95. https://doi.org/10.1007/s11306-009-0178-y

Takayama, T., Tsutshi, H., Shimizu, I., Toyama, T., Yoshimoto, N., Endo, Y., Inoue, K., Todoroki, K., Min, J.Z., Mizuno, H., Toyo'oka, T. (2016). Diagnostic approach to breast cancer patients based on target metabolomics in saliva by liquid chromatography with tandem mass spectrometry. CLINICA CHIMICA ACTA, 452, 18–26. https://doi.org/10.1016/j.cca.2015.10.032

Zhang, L., Xiao, H., Karlan, S., Zhou, H., Gross, J., Elashoff, D., Akin, D., Yan, X., Chia, D., Karlan, B., Wong, D. T. (2010). Discovery and preclinical validation of salivary transcriptomic and proteomic biomarkers for the non-invasive detection of breast cancer. PloS One, 5(12), e15573. https://doi.org/10.1371/journal.pone.0015573

Zhong, L. P., Cheng, F., Lu, X. Y., Duan, Y. X., & Wang, X. D. (2016). Untargeted saliva metabonomics study of breast cancer based on ultra performance liquid chromatography

coupled to mass spectrometry with HILIC and RPLC separations. Talanta, 158, 351–360. https://doi.org/10.1016/j.talanta.2016.04.049

# **IMAGING MODALITIES**

# Automated whole breast ultrasound

Berg, W.A., Bandos, A.I., Mendelson, E.B., Lehrer, D., Jong, R.A., Pisano, E.D. (2015). Ultrasound as the Primary Screening Test for Breast Cancer: Analysis From ACRIN 6666. Journal Of The National Cancer Institute, 108(4). https://doi.org/10.1093/jnci/djv367

Berg, W. A., & Mendelson, E. B. (2014). Technologist-performed handheld screening breast US imaging: how is it performed and what are the outcomes to date? Radiology, 272(1), 12–27. https://doi.org/10.1148/radiol.14132628

Brem, R. F., Tabár, L., Duffy, S. W., Inciardi, M. F., Guingrich, J. A., Hashimoto, B. E., Lander, M.R., Lapidus, R.L., Peterson, M.K., Rapelyea, J.A., Roux, S., Schilling, K.J., Shah, B.A., Torrente, J., Wynn, R.T., Miller, D. P. (2015). Assessing improvement in detection of breast cancer with three-dimensional automated breast US in women with dense breast tissue: the SomoInsight Study. Radiology, 274(3), 663.

Chang, J.M., Moon, W.K., Cho, N., Park, J.S., Kim, S.J. (2011) Breast cancers initially detected by hand-held ultrasound: detection performance of radiologists using automated breast ultrasound data. Acta Radiologica, 52(1) 8-14, https://doi.org/10.1258/ar.2010.100179

Corsetti, V., Houssami, N., Ghirardi, M., Ferrari, A., Speziani, M., Bellarosa, S., Remida, G., Gasparotti, C., Galligioni, E. & Ciatto, S. (2011) Evidence of the effect of adjunct ultrasound screening in women with mammography-negative dense breasts: Interval breast cancers at 1year follow-up. European Journal of Cancer, 47(7) 1021-1026, https://doi.org/10.1016/j.ejca.2010.12.002

Hooley, R.J., Scoutt, M.D. & Philpotts, L.F. (2013) Breast ultrasonography: state of the art. Radiology 268(3), 642-659, https://doi.org/10.1148/radiol.13121606

Geisel, J., Raghu, M., & Hooley, R. (2018). The Role of Ultrasound in Breast Cancer Screening: The Case for and Against Ultrasound. Seminars In Ultrasound, CT, And MR, 39(1), 25–34. https://doi.org/10.1053/j.sult.2017.09.006

Gilbert, F. J., & Selamoglu, A. (2017). Personalised screening: is this the way forward? Clinical Radiology. https://doi.org/10.1016/j.crad.2017.11.021

Giuliano, V., & Giuliano, C. (2013). Improved breast cancer detection in asymptomatic women using 3D-automated breast ultrasound in mammographically dense breasts. Clinical Imaging, 37(3), 480–486. https://doi.org/10.1016/j.clinimag.2012.09.018

Kelly, K. M., & Richwald, G. A. (2011). Automated whole-breast ultrasound: advancing the performance of breast cancer screening. Seminars In Ultrasound, CT, And MR, 32(4), 273–280. https://doi.org/10.1053/j.sult.2011.02.004

Kelly KM, Dean J, Comulada WS, Lee SJ, Kelly, K. M., Dean, J., Comulada, W.S., Lee, S.-J. (2010). Breast cancer detection using automated whole breast ultrasound and mammography in

radiographically dense breasts. European Radiology, 20(3), 734–742. https://doi.org/10.1007/s00330-009-1588-y

Kim, S-Y., Kim, M.J., Moon, H.J., Yoon, J.H & Kim, E-K. (2016) Application of the downgrade criteria to supplemental screening ultrasound for women with negative mammography but dense breasts. Medicine, 95(44), https://doi.org/10.1097/MD.00000000005279

Kornecki, A. (2011) Current status of breast ultrasound. Canadian Association Of Radiologists Journal, 62(1), 31-40. https://doi.org/10.1016/j.carj.2010.07.006

Le-Petross, H. T., & Shetty, M. K. (2011). Magnetic resonance imaging and breast ultrasonography as an adjunct to mammographic screening in high-risk patients. Seminars In Ultrasound, CT, And MR, 32(4), 266–272. https://doi.org/10.1053/j.sult.2011.03.005

Melnikow, J., Fenton, J. J., Whitlock, E. P., Miglioretti, D. L., Weyrich, M. S., Thompson, J. H., & Shah, K. (2016). Supplemental Screening for Breast Cancer in Women With Dense Breasts: A Systematic Review for the U.S. Preventive Services Task Force. Annals of Internal Medicine, 164(4), 268–278. https://doi.org/10.7326/M15-1789

Ohuchi, N., Suzuki, A., Sobue, T., Kawai, M., Yamamoto, S., Zheng, Y-F., Shiono, Y.N., Saito, H., Kuriyama, S., Tohno, E., Endo, T., Fukao, A., Tsuji, I., Yamaguchi, T., Ohashi, Y., Fukuda, M., Ishida, T. (2016). Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): a randomised controlled trial. The Lancet. 387, 341-348. https://doi.org/10.1016/S0140-6736(15)00774-6

Skaane, P., Gullien, R., Eben, E. B., Sandhaug, M., Schulz-Wendtland, R., & Stoeblen, F. (2015). Interpretation of automated breast ultrasound (ABUS) with and without knowledge of mammography: a reader performance study. Acta Radiologica, 56(4), 404–412. https://doi.org/10.1177/0284185114528835

Youk, J.H., Kim, E-K. (2010) Supplementary screening sonography in mammographically dense breast: pros and cons. Korean Journal Of Radiology 11(6), 589-593, https://doi.org/10.3348/kjr.2010.11.6.589

Wilczek, B., Wilczek, H. E., Rasouliyan, L., & Leifland, K. (2016). Adding 3D automated breast ultrasound to mammography screening in women with heterogeneously and extremely dense breasts: Report from a hospital-based, high-volume, single-center breast cancer screening program. European Journal of Radiology, 85(9), 1554–1563. https://doi.org/10.1016/j.ejrad.2016.06.004

#### Contrast enhanced mammography

Bhimani, C., Matta, D., Roth, R. G., Liao, L., Tinney, E., Brill, K., & Germaine, P. (2017). Contrastenhanced Spectral Mammography: Technique, Indications, and Clinical Applications. Academic Radiology, 24(1), 84–88. https://doi.org/10.1016/j.acra.2016.08.019

Chou, C. P., Lewin, J. M., Chiang, C. L., Hung, B. H., Yang, T. L., Huang, J. S., Lioa, J.B., Pan, H. B. (2015). Clinical evaluation of contrast-enhanced digital mammography and contrast enhanced tomosynthesis-Comparison to contrast-enhanced breast MRI. European Journal of Radiology, 84(12), 2501–2508. https://doi.org/10.1016/j.ejrad.2015.09.019

Ding, H. J., & Molloi, S. (2017). Quantitative contrast-enhanced spectral mammography based on photon-counting detectors: A feasibility study. Medical Physics, 44(8), 3939–3951. https://doi.org/10.1002/mp.12296

Dromain, C., Thibault, F., Muller, S., Rimareix, F., Delaloge, S., Tardivon, A., & Balleyguier, C. (2011). Dual-energy contrast-enhanced digital mammography: initial clinical results. European Radiology, 21(3), 565–574. https://doi.org/10.1007/s00330-010-1944-y

ElSaid, N. A. E., Farouk, S., Shetat, O. M. M., Khalifa, N. M., & Nada, O. M. (2015). Contrast enhanced digital mammography: Is it useful in detecting lesions in edematous breast? Egyptian Journal of Radiology and Nuclear Medicine, 46, 811–819. https://doi.org/10.1016/j.ejrnm.2015.04.002

Gilbert, F. J., & Selamoglu, A. (2017). Personalised screening: is this the way forward? Clinical Radiology. https://doi.org/10.1016/j.crad.2017.11.021

Helvie, M. A. (2010). Digital mammography imaging: breast tomosynthesis and advanced applications. Radiologic Clinics Of North America, 48(5), 917–929. https://doi.org/10.1016/j.rcl.2010.06.009

Houben, I. P. L., Van de Voorde, P., Jeukens, C., Wildberger, J. E., Kooreman, L. F., Smidt, M. L., & Lobbes, M. B. I. (2017). Contrast-enhanced spectral mammography as work-up tool in patients recalled from breast cancer screening has low risks and might hold clinical benefits. European Journal of Radiology, 94, 31–37. https://doi.org/10.1016/j.ejrad.2017.07.004

Jochelson, M. S., Dershaw, D. D., Sung, J. S., Heerdt, A. S., Thornton, C., Moskowitz, C. S., ... Morris, E. A. (2013). Bilateral Contrast-enhanced Dual-Energy Digital Mammography: Feasibility and Comparison with Conventional Digital Mammography and MR Imaging in Women with Known Breast Carcinoma. Radiology, 266(3), 743–751. https://doi.org/10.1148/radiol.12121084

Kariyappa, K. D., Gnanaprakasam, F., Anand, S., Krishnaswami, M., & Ramachandran, M. (2016). Contrast enhanced dual energy spectral mammogram, an emerging addendum in breast imaging. British Journal of Radiology, 89(1067). https://doi.org/10.1259/bjr.20150609

Karunamuni, R., Naha, P. C., Lau, K. C., Al-Zaki, A., Popov, A. V., Delikatny, E. J., Tsourkas, A., Cormode, D.P., Maidment, A. D. A. (2016). Development of silica-encapsulated silver nanoparticles as contrast agents intended for dual-energy mammography. European Radiology, 26(9), 3301–3309. https://doi.org/10.1007/s00330-015-4152-y

Lalji, U. C., Houben, I. P. L., Prevos, R., Gommers, S., van Goethem, M., Vanwetswinkel, S., ... Lobbes, M. B. I. (2016). Contrast-enhanced spectral mammography in recalls from the Dutch breast cancer screening program: validation of results in a large multireader, multicase study. European Radiology, 26(12), 4371–4379.

Lalji, U. C., Jeukens, C., Houben, I., Nelemans, P. J., van Engen, R. E., van Wylick, E., Beets-Tan, R.G.H., Wildberger, J.E., Paulis, L.E.,Lobbes, M. B. I. (2015). Evaluation of low-energy contrastenhanced spectral mammography images by comparing them to full-field digital mammography using EUREF image quality criteria. European Radiology, 25(10), 2813–2820. https://doi.org/10.1007/s00330-015-3695-2

Lewis, T. C., Pizzitola, V. J., Giurescu, M. E., Eversman, W. G., Lorans, R., Robinson, K. A., & Patel, B. K. (2017). Contrast-enhanced Digital Mammography: A Single-Institution Experience of the First 208 Cases. Breast Journal, 23(1), 67–76. https://doi.org/10.1111/tbj.12681

Luczynska, E., Heinze-Paluchowska, S., Dyczek, S., Blecharz, P., Rys, J., & Reinfuss, M. (2014). Contrast-Enhanced Spectral Mammography: Comparison with Conventional Mammography and Histopathology in 152 Women. Korean Journal of Radiology, 15(6), 689–696. https://doi.org/10.3348/kjr.2014.15.6.689

McGuire, A., O'Leary, D. P., Livingstone, V., Khan, Z., O'Connell, F., Redmond, H. P., Relihan, N., Corrigan, M., Ryan, M. (2017). Contrast-Enhanced Spectrum Mammography: A Useful Adjunct to Digital Mammography in Predicting Tumor Size. Breast Journal, 23(4), 484–486. https://doi.org/10.1111/tbj.12774

National Horizon Scanning Unit. (2009). New and emerging technologies for breast cancer detection. Canberra: Australia and New Zealand Horizon Scanning Network. Retrieved from http://www.horizonscanning.gov.au/internet/horizon/publishing.nsf/Content/AD1C4F0CFAD 1A5E4CA2575E8001DC431/\$File/ETB\_BreastScreen.pdf

Patel, B. K., Gray, R. J., & Pockaj, B. A. (2017). Potential Cost Savings of Contrast-Enhanced DigitalMammography.AmericanJournalofRoentgenology,208(6),W231–W237.https://doi.org/10.2214/AJR.16.17239

Patel, B. K., Ranjbar, S., Wu, T., Pockaj, B. A., Li, J., Zhang, N., Lobbes, M., Zhang, B., Mitchell, J. R.(2018). Computer-aided diagnosis of contrast-enhanced spectral mammography: A feasibilitystudy.EuropeanJournalofRadiology,98,207–213.https://doi.org/10.1016/j.ejrad.2017.11.024

Phillips, J., Miller, M. M., Mehta, T. S., Fein-Zachary, V., Nathanson, A., Hori, W., Monahan-Earley, R., Slanetz, P. J. (2017). Contrast-enhanced spectral mammography (CESM) versus MRI in the high-risk screening setting: patient preferences and attitudes. Clinical Imaging, 42, 193–197. https://doi.org/10.1016/j.clinimag.2016.12.011

Sogani, J., Morris, E. A., Kaplan, J. B., D'Alessio, D., Goldman, D., Moskowitz, C. S., & Jochelson, M. S. (2017). Comparison of Background Parenchymal Enhancement at Contrast-enhanced Spectral Mammography and Breast MR Imaging. Radiology, 282(1), 63–73. https://doi.org/10.1148/radiol.2016160284

Tardivel, A. M., Balleyguier, C., Dunant, A., Delaloge, S., Mazouni, C., Mathieu, M. C., & Dromain, C. (2016). Added Value of Contrast-Enhanced Spectral Mammography in Postscreening Assessment. Breast Journal, 22(5), 520–528. https://doi.org/10.1111/tbj.12627

Yagil, Y., Shalmon, A., Rundstein, A., Servadio, Y., Halshtok, O., Gotlieb, M., & Sklair-Levy, M. (2016). Challenges in contrast-enhanced spectral mammography interpretation: artefacts lexicon. Clinical Radiology, 71(5), 450–457. https://doi.org/10.1016/j.crad.2016.01.012

#### Digital breast tomosynthesis (from Allen + Clarke's 2018 literature review)

NB The following list is a complete list of references used in Allen + Clarke's literature review on the role of DBT in screening. Not all of these references are specifically referred to in the summarised chapter included in this horizon scan.

Abdullah Suhaimi S.A., Mohamed A., Ahmad M., and Chelliah K.K. 'Effects of Reduced Compression in Digital Breast Tomosynthesis on Pain, Anxiety, and Image Quality'. Malaysian Journal of Medical Sciences 22, no. 6 (2015): 40–46.

Aujero M.P., Gavenonis S.C., Benjamin R., Zhang Z., and Holt J.S. 'Clinical Performance of Synthesized Two-Dimensional Mammography Combined with Tomosynthesis in a Large Screening Population'. Radiology 283, no. 1 (2017): 70–76. https://doi.org/10.1148/radiol.2017162674.

Bernardi D., Belli P., Benelli E., Brancato B., Bucchi L., Calabrese M., Carbonaro L.A., et al. 'Digital Breast Tomosynthesis (DBT): Recommendations from the Italian College of Breast Radiologists (ICBR) by the Italian Society of Medical Radiology (SIRM) and the Italian Group for Mammography Screening (GISMa)'. La Radiologia Medica 122, no. 10 (2017A): 723–30. https://doi.org/10.1007/s11547-017-0769-z.

Bernardi D., Caumo F., Macaskill P., Ciatto S., Pellegrini M., Brunelli S., Tuttobene P., et al. 'Effect of Integrating 3D-Mammography (Digital Breast Tomosynthesis) with 2D-Mammography on Radiologists' True-Positive and False-Positive Detection in a Population Breast Screening Trial'. European Journal of Cancer 50, no. 7 (2014): 1232–38. https://doi.org/10.1016/j.ejca.2014.02.004.

Bernardi, D, S Ciatto, M Pellegrini, V Anesi, S Burlon, E Cauli, M Depaoli, et al. 'Application of Breast Tomosynthesis in Screening: Incremental Effect on Mammography Acquisition and Reading Time.' The British Journal of Radiology 85, no. 1020 (2012): e1174-8. https://doi.org/10.1259/bjr/19385909.

Bernardi, D., and N. Houssami. 'Breast Cancers Detected in Only One of Two Arms of a Tomosynthesis (3D-Mammography) Population Screening Trial (STORM-2)'. Breast 32 (2017B): 98–101. https://doi.org/10.1016/j.breast.2017.01.005.

Bernardi D., Macaskill P., Pellegrini M., Valentini M., Fanto C., Ostillio L., Tuttobene P., Luparia A., and Houssami N. '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'. The Lancet Oncology 17, no. 8 (2016): 1105–13. https://doi.org/10.1016/S1470-2045%2816%2930101-2.

Bonafede, M.M., V.B. Kalra, J.D. Miller, and L.L. Fajardo. 'Value Analysis of Digital Breast Tomosynthesis for Breast Cancer Screening in a Commercially-Insured US Population'. ClinicoEconomics and Outcomes Research 7 (2015): 53–63. https://doi.org/10.2147/CEOR.S76167.

Bonsall J, Akran T, Tsang R, Jones V, Turnbull A, Cornfored E. 2016. Practical evaluation of GE SenoClaire digital breast tomosynthesis system. NHS.

Borelli, C and Oduko J. NHS Breast Screening Programme: current position on use of tomosynthesis 28 March 2016.

BreastScreen Australia Accreditation Review Committee. 2015. BreastScreen Australia National Accreditation Standards.

Carbonaro, L.A., G. Di Leo, P. Clauser, R.M. Trimboli, N. Verardi, M.P. Fedeli, R. Girometti, et al. 'Impact on the Recall Rate of Digital Breast Tomosynthesis as an Adjunct to Digital Mammography in the Screening Setting. A Double Reading Experience and Review of the Literature'. European Journal of Radiology 85, no. 4 (2016): 808–14. https://doi.org/10.1016/j.ejrad.2016.01.004.

Caumo, F., D. Bernardi, S. Ciatto, P. Macaskill, M. Pellegrini, S. Brunelli, P. Tuttobene, et al. 'Incremental Effect from Integrating 3D-Mammography (Tomosynthesis) with 2D- Mammography: Increased Breast Cancer Detection Evident for Screening Centres in a Population-Based Trial'. Breast 23, no. 1 (2013): 76–80. https://doi.org/10.1016/j.breast.2013.11.006.

Ciatto S., Houssami N., Bernardi D., Caumo F., Pellegrini M., Brunelli S., Tuttobene P., et al. 'Integration of 3D Digital Mammography with Tomosynthesis for Population Breast-Cancer Screening (STORM): A Prospective Comparison Study'. The Lancet Oncology 14, no. 7 (2013): 583–89. https://doi.org/10.1016/S1470-2045%2813%2970134-7.

Conant E.F., Beaber E.F., Sprague B.L., Herschorn S.D., Weaver D.L., Onega T., Tosteson A.N.A., et al. 'Breast Cancer Screening Using Tomosynthesis in Combination with Digital Mammography Compared to Digital Mammography Alone: A Cohort Study within the PROSPR Consortium'. Breast Cancer Research and Treatment 156, no. 1 (2016): 109–16. https://doi.org/10.1007/s10549-016-3695-1.

Coop, P, Cowling A, Lawson B. 'Tomosynthesis as a Screening Tool for Breast Cancer: A Systematic Review'. Radiography 22, no. 3 (1 August 2016): e190–95. https://doi.org/10.1016/j.radi.2016.03.002.

Dang, Pragya A., Phoebe E. Freer, Kathryn L. Humphrey, Elkan F. Halpern, and Elizabeth A. Rafferty. 'Addition of Tomosynthesis to Conventional Digital Mammography: Effect on Image Interpretation Time of Screening Examinations'. Radiology 270, no. 1 (1 January 2014): 49–56. https://doi.org/10.1148/radiol.13130765.

Department of Health and Ageing. 2009. Horizon Scanning Technology. Prioritising Summary: Breast tomosynthesis – a breast cancer screening tool. Update 2009, Commonwealth of Australia, 2009

Destounis S., Arieno A., and Morgan R. 'Initial Experience with Combination Digital Brea St Tomosynthesis plus Full Field Digital Mammography or Full Field Digital Mammography Alone in the Screening Environment'. Journal of Clinical Imaging Science 4, no. 1 (2014): 127838. https://doi.org/10.4103/2156-7514.127838.

Durand, Melissa A., Brian M. Haas, Xiaopan Yao, Jaime L. Geisel, Madhavi Raghu, Regina J. Hooley, Laura J. Horvath, and Liane E. Philpotts. 'Early Clinical Experience with Digital Breast Tomosynthesis for Screening Mammography'. Radiology 274, no. 1 (2015): 85–92. https://doi.org/10.1148/radiol.14131319.

Freer P.E., Riegert J., Eisenmenger L., Ose D., Winkler N., Stein M.A., Stoddard G.J., and Hess R. 'Clinical Implementation of Synthesized Mammography with Digital Breast Tomosynthesis in a Routine Clinical Practice'. Breast Cancer Research and Treatment 166, no. 2 (2017): 501–9. https://doi.org/10.1007/s10549-017-4431-1.

Friedewald S.M., Rafferty E.A., Rose S.L., Durand M.A., Plecha D.M., Greenberg J.S., Hayes M.K., et al. 'Breast Cancer Screening Using Tomosynthesis in Combination with Digital Mammography'. JAMA - Journal of the American Medical Association 311, no. 24 (2014): 2499–2507. https://doi.org/10.1001/jama.2014.6095.

Gilbert, Fiona J, Lorraine Tucker, and Ken C Young. 'Digital Breast Tomosynthesis (DBT): A Review of the Evidence for Use as a Screening Tool.' Clinical Radiology 71, no. 2 (2016): 141–50. https://doi.org/10.1016/j.crad.2015.11.008.

Greenberg J.S., Javitt M.C., Katzen J., Michael S., and Holland A.E. 'Clinical Performance Metrics of 3D Digital Breast Tomosynthesis Compared with 2D Digital Mammography for Breast Cancer Screening in Community Practice'. American Journal of Roentgenology 203, no. 3 (2014): 687–93. https://doi.org/10.2214/AJR.14.12642.

Gur, D., M.L. Zuley, M.I. Anello, G.Y. Rathfon, D.M. Chough, M.A. Ganott, C.M. Hakim, L. Wallace, A. Lu, and A.I. Bandos. 'Dose Reduction in Digital Breast Tomosynthesis (DBT) Screening Using Synthetically Reconstructed Projection Images. An Observer Performance Study.' Academic Radiology 19, no. 2 (2012): 166–71. https://doi.org/10.1016/j.acra.2011.10.003.

Haas B.M., Kalra V., Geisel J., Raghu M., Durand M., and Philpotts L.E. 'Comparison of Tomosynthesis plus Digital Mammography and Digital Mammography Alone for Breast Cancer Screening'. Radiology 269, no. 3 (2013): 694–700. https://doi.org/10.1148/radiol.13130307.

Hardesty, Lara A. 'Issues to Consider Before Implementing Digital Breast Tomosynthesis Into a Breast Imaging Practice'. American Journal of Roentgenology 204, no. 3 (25 February 2015): 681–84. https://doi.org/10.2214/AJR.14.13094.

Hodgson, Robert, Sylvia H Heywang-Kobrunner, Susan C Harvey, Mary Edwards, Javed Shaikh, Mick Arber, and Julie Glanville. 'Systematic Review of 3D Mammography for Breast Cancer Screening.' Breast (Edinburgh, Scotland) 27, no. 9213011 (2016): 52–61. https://doi.org/10.1016/j.breast.2016.01.002.

Houssami N. 'Evidence on Synthesized Two-Dimensional Mammography Versus Digital Mammography When Using Tomosynthesis (Three-Dimensional Mammography) for Population Breast Cancer Screening'. Clinical Breast Cancer, no. (Houssami) Sydney School of Public Health, Sydney Medical School, University of Sydney, Sydney, Australia (2017). https://doi.org/10.1016/j.clbc.2017.09.012.

Houssami, Nehmat. 'Digital Breast Tomosynthesis (3D-Mammography) Screening: Data and Implications for Population Screening.' Expert Review of Medical Devices 12, no. 4 (2015): 377–79. https://doi.org/10.1586/17434440.2015.1028362.

Houssami N., Bernardi D., Pellegrini M., Valentini M., Fanto C., Ostillio L., Tuttobene P., Luparia A., and Macaskill P. 'Breast Cancer Detection Using Single-Reading of Breast Tomosynthesis (3D-Mammography) Compared to Double-Reading of 2D-Mammography: Evidence from a Population-Based Trial'. Cancer Epidemiology 47, no. (Houssami, Macaskill) Sydney School of Public Health (A27), Sydney Medical School, University of Sydney, Sydney 2006, Australia (2017): 94–99. https://doi.org/10.1016/j.canep.2017.01.008.

Houssami N., Macaskill P., Bernardi D., Caumo F., Pellegrini M., Brunelli S., Tuttobene P., et al. 'Breast Screening Using 2D-Mammography or Integrating Digital Breast Tomosynthesis (3D-Mammography) for Single-Reading or Double-Reading - Evidence to Guide Future Screening Strategies'. European Journal of Cancer 50, no. 10 (2014): 1799–1807. https://doi.org/10.1016/j.ejca.2014.03.017.

Houssami, Nehmat, Kristina Lang, Daniela Bernardi, Alberto Tagliafico, Sophia Zackrisson, and Per Skaane. 'Digital Breast Tomosynthesis (3D-Mammography) Screening: A Pictorial Review of Screen-Detected Cancers and False Recalls Attributed to Tomosynthesis in Prospective Screening Trials.' Breast (Edinburgh, Scotland) 26, no. 9213011 (2016B): 119–34. https://doi.org/10.1016/j.breast.2016.01.007. Houssami, Nehmat, and Per Skaane. 'Overview of the Evidence on Digital Breast Tomosynthesis in Breast Cancer Detection.' Breast (Edinburgh, Scotland) 22, no. 2 (2013): 101–8. https://doi.org/10.1016/j.breast.2013.01.017.

Houssami, Nehmat, and Robin M Turner. 'Rapid Review: Estimates of Incremental Breast Cancer Detection from Tomosynthesis (3D-Mammography) Screening in Women with Dense Breasts.' Breast (Edinburgh, Scotland) 30, no. 9213011 (2016): 141–45. https://doi.org/10.1016/j.breast.2016.09.008.

Hunter S.A., Morris C., Nelson K., Snyder B.J., and Poulton T.B. 'Digital Breast Tomosynthesis: Cost-Effectiveness of Using Private and Medicare Insurance in Community-Based Health Care Facilities'. American Journal of Roentgenology 208, no. 5 (2017): 1171–75. https://doi.org/10.2214/AJR.16.16987.

Kalra V.B., Wu X., Haas B.M., Forman H.P., and Philpotts L.E. 'Cost-Effectiveness of Tomosynthesis in Annual Screening Mammography'. American Journal of Roentgenology 207, no. 5 (2016): 1152–55. https://doi.org/10.2214/AJR.15.14487.

Lang K., Andersson I., Rosso A., Tingberg A., Timberg P., and Zackrisson S. 'Performance of One-View Breast Tomosynthesis as a Stand-Alone Breast Cancer Screening Modality: Results from the Malmo Breast Tomosynthesis Screening Trial, a Population-Based Study'. European Radiology 26, no. 1 (2016A): 184–90. https://doi.org/10.1007/s00330-015-3803-3.

Lang K., Nergarden M., Andersson I., Rosso A., and Zackrisson S. 'False Positives in Breast Cancer Screening with One-View Breast Tomosynthesis: An Analysis of Findings Leading to Recall, Work-up and Biopsy Rates in the Malmo Breast Tomosynthesis Screening Trial'. European Radiology 26, no. 11 (2016B): 3899–3907. https://doi.org/10.1007/s00330-016-4265-y.

Lauby-Secretan, Béatrice, Chiara Scoccianti, Dana Loomis, Lamia Benbrahim-Tallaa, Véronique Bouvard, Franca Bianchini, and Kurt Straif. 'Breast-Cancer Screening — Viewpoint of the IARC Working Group'. New England Journal of Medicine 372, no. 24 (3 June 2015): 2353–58. https://doi.org/10.1056/NEJMsr1504363.

Lee C.I., Cevik M., Alagoz O., Sprague B.L., Tosteson A.N.A., Miglioretti D.L., Kerlikowske K., et al. 'Comparative Effectiveness of Combined Digital Mammography and Tomosynthesis Screening for Women with Dense Breasts'. Radiology 274, no. 3 (2014): 772–80. https://doi.org/10.1148/radiol.14141237.

Liberatore, Mathieu, Jean-Michel Cucchi, Martine Fighiera, Anne Binet, Marie Christine Missana, Philippe Brunner, Michel Yves Mourou, and Antoine Iannessi. 'Interest of Systematic Tomosynthesis (3D Mammography) with Synthetic 2D Mammography in Breast Cancer Screening.' Hormone Molecular Biology and Clinical Investigation 32, no. 2 (16 December 2017). https://doi.org/10.1515/hmbci-2017-0024.

Lourenco A.P., Barry-Brooks M., Baird G.L., Tuttle A., and Mainiero M.B. 'Changes in Recall Type and Patient Treatment Following Implementation of Screening Digital Breast Tomosynthesis'. Radiology 274, no. 2 (2015): 337–42.

Maxwell, A.J., M. Michell, Y.Y. Lim, S.M. Astley, M. Wilson, E. Hurley, D.G. Evans, et al. 'A Randomised Trial of Screening with Digital Breast Tomosynthesis plus Conventional Digital 2D Mammography versus 2D Mammography Alone in Younger Higher Risk Women'. European Journal of Radiology 94 (2017): 133–39. https://doi.org/10.1016/j.ejrad.2017.06.018.

Mayor S. 2016. 'Reduction in breast cancer deaths is due to treatment not screening' in BMJ 2016; 355; doi: https://doi.org/10.1136/bmj.i5544.

McCarthy A.M., Kontos D., Synnestvedt M., Tan K.S., Heitjan D.F., Schnall M., and Conant E.F. 'Screening Outcomes Following Implementation of Digital Breast Tomosynthesis in a General-Population Screening Program'. Journal of the National Cancer Institute 106, no. 11 (2014): dju316. https://doi.org/10.1093/jnci/dju316.

McDonald, Elizabeth S, Anne Marie McCarthy, Amana L Akhtar, Marie B Synnestvedt, Mitchell Schnall, and Emily F Conant. 'Baseline Screening Mammography: Performance of Full-Field Digital Mammography Versus Digital Breast Tomosynthesis.' AJR. American Journal of Roentgenology 205, no. 5 (2015): 1143–48. https://doi.org/10.2214/AJR.15.14406.

McDonald E.S., Oustimov A., Weinstein S.P., Synnestvedt M.B., Schnall M., and Conant E.F. 'Effectiveness of Digital Breast Tomosynthesis Compared With Digital Mammography: Outcomes Analysis From 3 Years of Breast Cancer Screening'. JAMA Oncology 2, no. 6 (2016): 737–43. https://doi.org/10.1001/jamaoncol.2015.5536.

Miller J.D., Bonafede M.M., Herschorn S.D., Pohlman S.K., Troeger K.A., and Fajardo L.L. 'Value Analysis of Digital Breast Tomosynthesis for Breast Cancer Screening in a US Medicaid Population'. Journal of the American College of Radiology 14, no. 4 (2017): 467. https://doi.org/10.1016/j.jacr.2016.11.019.

Mungutroy EHL, Oduko JM, Cooke JC , Formstone, WJ. 2014. Practical evaluation of Hologic Selenia Dimensions digital breast tomosynthesis system. NHS.

Nelson R. 2017. Make screening mammography personal say the French. https://www.medscape.com/viewarticle/888005#vp\_1. Accessed 18 February 2018.

NHC. (2015). Position statement pamphlet on breast density and screening.

Pan H.-B., Wong K.-F., Yao A., Hsu G.-C., Chou C.-P., Liang H.-L., Huang J.-S., Li H.-J., Wang S.-C., and Yang T.-L. 'Breast Cancer Screening with Digital Breast Tomosynthesis - 4 Year Experience and Comparison with National Data'. Journal of the Chinese Medical Association, no. (Pan) Kaohsiung Veterans General Hospital Tainan Branch, Tainan, Taiwan, ROC (2017). https://doi.org/10.1016/j.jcma.2017.05.013.

Poplack, Steven. 'Breast Tomosynthesis: Clinical Evidence.' Radiologic Clinics of North America 55, no. 3 (2017): 475–92. https://doi.org/10.1016/j.rcl.2016.12.010.

Powell J.L., Hawley J.R., Lipari A.M., Yildiz V.O., Erdal B.S., and Carkaci S. 'Impact of the Addition of Digital Breast Tomosynthesis (DBT) to Standard 2D Digital Screening Mammography on the Rates of Patient Recall, Cancer Detection, and Recommendations for Short-Term Follow-Up'. Academic Radiology 24, no. 3 (2017): 302–7. https://doi.org/10.1016/j.acra.2016.10.001.

Olgar T., Kahn T., and Gosch D. 'Average Glandular Dose in Digital Mammography and Breast Tomosynthesis.' Fortschr Rontgenstr 184, no. 10 (2012): 911-918. https://doi.org/10.1055/s-0032-1312877.

Rafferty E.A., Rose S.L., Miller D.P., Durand M.A., Conant E.F., Copit D.S., Friedewald S.M., et al. 'Effect of Age on Breast Cancer Screening Using Tomosynthesis in Combination with Digital Mammography'. Breast Cancer Research and Treatment 164, no. 3 (2017): 659–66. https://doi.org/10.1007/s10549-017-4299-0.

Rafferty, Elizabeth A., Jeong Mi Park, Liane E. Philpotts, Steven P. Poplack, Jules H. Sumkin, Elkan F. Halpern, and Loren T. Niklason. 'Assessing Radiologist Performance Using Combined Digital Mammography and Breast Tomosynthesis Compared with Digital Mammography Alone: Results of a Multicenter, Multireader Trial'. Radiology 266, no. 1 (1 January 2013): 104–13. https://doi.org/10.1148/radiol.12120674.

Rodriguez-Ruiz A., Gubern-Merida A., Imhof-Tas M., Lardenoije S., Wanders A.J.T., Andersson I., Zackrisson S., et al. 'One-View Digital Breast Tomosynthesis as a Stand-Alone Modality for Breast Cancer Detection: Do We Need More?' European Radiology, no. (Rodriguez-Ruiz, Gubern-Merida, Imhof-Tas, Lardenoije, Karssemeijer, Mann, Sechopoulos) Department of Radiology and Nuclear Medicine, Radboud University Medical Centre, Geert Grooteplein 10, Post 766, Nijmegen 6525 GA, Netherlands (2017): 1–11. https://doi.org/10.1007/s00330-017-5167-3.

Rose S.L., Tidwell A.L., Ice M.F., Nordmann A.S., Sexton R., and Song R. 'A Reader Study Comparing Prospective Tomosynthesis Interpretations with Retrospective Readings of the Corresponding FFDM Examinations'. Academic Radiology 21, no. 9 (2014): 1204–10. https://doi.org/10.1016/j.acra.2014.04.008.

Rose S.L., Tidwell A.L., Bujnoch L.J., Kushwaha A.C., Nordmann A.S., and Sexton Jr. R. 'Implementation of Breast Tomosynthesis in a Routine Screening Practice: An Observational Study'. American Journal of Roentgenology 200, no. 6 (2013): 1401–8. https://doi.org/10.2214/AJR.12.9672.

Rosso A., Lang K., Petersson I.F., and Zackrisson S. 'Factors Affecting Recall Rate and False Positive Fraction in Breast Cancer Screening with Breast Tomosynthesis - A Statistical Approach'. Breast 24, no. 5 (2015): 680–86. https://doi.org/10.1016/j.breast.2015.08.007.

Sardanelli, F., H.S. Aase, M. Álvarez, E. Azavedo, H.J. Baarslag, C. Balleyguier, P.A. Baltzer, et al. 'Position Paper on Screening for Breast Cancer by the European Society of Breast Imaging (EUSOBI) and 30 National Breast Radiology Bodies from Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Israel, Lithuania, Moldova, The Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Spain, Sweden, Switzerland and Turkey'. European Radiology 27, no. 7 (2017): 2737–43. https://doi.org/10.1007/s00330-016-4612-z.

Sechopoulos, Ioannis. 'A Review of Breast Tomosynthesis. Part I. The Image Acquisition Process.' Medical Physics 40, no. 1 (2013): 014301. https://doi.org/10.1118/1.4770279.

Sharpe, R.E., S. Venkataraman, J. Phillips, V. Dialani, V.J. Fein-Zachary, S. Prakash, P.J. Slanetz, and T.S. Mehta. '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 278, no. 3 (2016): 698–706. https://doi.org/10.1148/radiol.2015142036.

Shin, Sung Ui, Jung Min Chang, Min Sun Bae, Su Hyun Lee, Nariya Cho, Mirinae Seo, Won Hwa Kim, and Woo Kyung Moon. 'Comparative Evaluation of Average Glandular Dose and Breast Cancer Detection between Single-View Digital Breast Tomosynthesis (DBT) plus Single-View Digital Mammography (DM) and Two-View DM: Correlation with Breast Thickness and Density'. European Radiology 25, no. 1 (1 January 2015): 1–8. https://doi.org/10.1007/s00330-014-3399-z.

Siu, Albert L. 'Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement.' Annals of Internal Medicine 164, no. 4 (16 February 2016): 279–96. https://doi.org/10.7326/M15-2886.

Skaane, P. 'Breast Cancer Screening with Digital Breast Tomosynthesis'. Breast Cancer 24, no. 1 (2017): 32–41. https://doi.org/10.1007/s12282-016-0699-y.

Skaane P., Bandos A.I., Eben E.B., Jebsen I.N., Krager M., Haakenaasen U., Ekseth U., Izadi M., Hofvind S., and Gullien R. 'Two-View Digital Breast Tomosynthesis Screening with Synthetically Reconstructed Projection Images: Comparison with Digital Breast Tomosynthesis with Full-Field Digital Mammographic Images'. Radiology 271, no. 3 (2014): 655–63. https://doi.org/10.1148/radiol.13131391.

Skaane P., Bandos A.I., Gullien R., Eben E.B., Ekseth U., Haakenaasen U., Izadi M., et al. 'Prospective Trial Comparing Full-Field Digital Mammography (FFDM) versus Combined FFDM and Tomosynthesis in a Population-Based Screening Programme Using Independent Double Reading with Arbitration'. European Radiology 23, no. 8 (2013A): 2061–71. https://doi.org/10.1007/s00330-013-2820-3.

Skaane, Per, Andriy I Bandos, Randi Gullien, Ellen B Eben, Ulrika Ekseth, Unni Haakenaasen, Mina Izadi, et al. 'Comparison of Digital Mammography Alone and Digital Mammography plus Tomosynthesis in a Population-Based Screening Program.' Radiology 267, no. 1 (2013B): 47–56. https://doi.org/10.1148/radiol.12121373.

Starikov A., Drotman M., Hentel K., Katzen J., Min R.J., and Arleo E.K. '2D Mammography, Digital Breast Tomosynthesis, and Ultrasound: Which Should Be Used for the Different Breast Densities in Breast Cancer Screening?' Clinical Imaging 40, no. 1 (2016): 68–71. https://doi.org/10.1016/j.clinimag.2015.10.001.

Strudley CJ, Looney P, Young KC. 2014. Technical evaluation of Hologic Selenia Dimensions digital breast tomosynthesis system. NHS.

Strudley CJ, Warren P, Young KC. 2015. Technical evaluation of Siemens Mammomat Inspirations digital breast tomosynthesis system. NHS.

Sumkin J.H., Ganott M.A., Chough D.M., Catullo V.J., Zuley M.L., Shinde D.D., Hakim C.M., Bandos A.I., and Gur D. 'Recall Rate Reduction with Tomosynthesis During Baseline Screening Examinations: An Assessment From a Prospective Trial'. Academic Radiology 22, no. 12 (2015): 1477–82. https://doi.org/10.1016/j.acra.2015.08.015.

Svahn, T M, N Houssami, I Sechopoulos, and S Mattsson. 'Review of Radiation Dose Estimates inDigital Breast Tomosynthesis Relative to Those in Two-View Full-Field Digital Mammography.'Breast (Edinburgh, Scotland)24, no.2(2015A):93–99.https://doi.org/10.1016/j.breast.2014.12.002.

Svahn, Tony M, Petra Macaskill, and Nehmat Houssami. 'Radiologists' Interpretive Efficiency and Variability in True- and False-Positive Detection When Screen-Reading with Tomosynthesis (3D-Mammography) Relative to Standard Mammography in Population Screening.' Breast (Edinburgh, Scotland) 24, no. 6 (2015B): 687–93. https://doi.org/10.1016/j.breast.2015.08.012.

Uematsu T. 'The Need for Supplemental Breast Cancer Screening Modalities: A Perspective of Population-Based Breast Cancer Screening Programs in Japan'. Breast Cancer 24, no. 1 (2017): 26–31. https://doi.org/10.1007/s12282-016-0707-2.

Urban, L.A.B.D., L.F. Chala, S.P. Bauab, M.B. Schaefer, R.P. Santos, N.M.A. Maranhão, A.L. Kefalas, et al. 'Breast Cancer Screening: Updated Recommendations of the Brazilian College of Radiology and Diagnostic Imaging, Brazilian Breast Disease Society, and Brazilian Federation of Gynecological and Obstetrical Associations'. Radiologia Brasileira 50, no. 4 (2017): 244–49. https://doi.org/10.1590/0100-3984.2017-0069.

Vedantham S., Karellas A., Vijayaraghavan G.R., and Kopans D.B. 'Digital Breast Tomosynthesis: State of the Art1'. Radiology 277, no. 3 (2015): 663–84. https://doi.org/10.1148/radiol.2015141303.

Wang W.-S., Hardesty L., Borgstede J., Takahashi J., and Sams S. 'Breast Cancers Found with Digital Breast Tomosynthesis: A Comparison of Pathology and Histologic Grade'. Breast Journal 22, no. 6 (2016): 651–56. https://doi.org/10.1111/tbj.12649.

Yaffe, Martin J. 'Reducing Radiation Doses for Breast Tomosynthesis?' The Lancet Oncology 17, no. 8 (n.d.): 1027–29. https://doi.org/10.1016/S1470-2045(16)30155-3.

Zuckerman, Samantha P., Emily F. Conant, Brad M. Keller, Andrew D. A. Maidment, Bruno Barufaldi, Susan P. Weinstein, Marie Synnestvedt, and Elizabeth S. McDonald. 'Implementation of Synthesized Two-Dimensional Mammography in a Population-Based Digital Breast Tomosynthesis Screening Program'. Radiology 281, no. 3 (28 July 2016): 730–36. https://doi.org/10.1148/radiol.2016160366.

Zuley, Margarita L., Ben Guo, Victor J. Catullo, Denise M. Chough, Amy E. Kelly, Amy H. Lu, Grace Y. Rathfon, et al. 'Comparison of Two-Dimensional Synthesized Mammograms versus Original Digital Mammograms Alone and in Combination with Tomosynthesis Images'. Radiology 271, no. 3 (2014): 664–71. https://doi.org/10.1148/radiol.13131530.

Zuley M.L., Bandos A.I., Abrams G.S., Cohen C., Hakim C.M., Sumkin J.H., Drescher J., Rockette H.E., and Gur D. 'Time to Diagnosis and Performance Levels during Repeat Interpretations of Digital Breast Tomosynthesis. Preliminary Observations'. Academic Radiology 17, no. 4 (2010): 450–55. https://doi.org/10.1016/j.acra.2009.11.011.

#### Ductoscopy

National Horizon Scanning Unit. (2009). New and emerging technologies for breast cancer detection. Canberra: Australia and New Zealand Horizon Scanning Network. Retrieved from http://www.horizonscanning.gov.au/internet/horizon/publishing.nsf/Content/AD1C4F0CFAD 1A5E4CA2575E8001DC431/\$File/ETB\_BreastScreen.pdf

Tang, S. S. K., Twelves, D. J., Isacke, C. M., & Gui, G. P. H. (2011). Mammary ductoscopy in the current management of breast disease. Surgical Endoscopy and Other Interventional Techniques, 25(6), 1712–1722. https://doi.org/10.1007/s00464-010-1465-4

Yang, X., Li, H., Gou, J., Tan, Q., Wang, L., Lin, X., Feng, X., Jiang, Y., Zhang, S., Lv, Q. (2014). The Role of Breast Ductoscopy in Evaluation of Nipple Discharge: A Chinese Experience of 419 Patients. Breast Journal, 20(4), 388–393. https://doi.org/10.1111/tbj.12275

#### Magnetic resonance imaging

Ahern, C. H., Shih, Y.-C. T., Dong, W., Parmigiani, G., & Shen, Y. (2014). Cost-effectiveness of alternative strategies for integrating MRI into breast cancer screening for women at high risk. British Journal of Cancer, 111(8), 1542–1551. https://doi.org/10.1038/bjc.2014.458

Badan, G. M., Roveda, D. J., Paito, S., de Faria Castro Fleury, E., Maragno, B., Dantas do Amaral Campos, M. S., Pecci Ferreira, C. A., Trocoli Ferreira, F. A. (2016). Ductal carcinoma in situ of the breast: Evaluation of main presentations on magnetic resonance imaging compared with findings on mammogram and histology. REVISTA DA ASSOCIACAO MEDICA BRASILEIRA, 62(5), 421–427. https://doi.org/10.1590/1806-9282.62.05.421

Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG, Bohn-Velez, M., Mahoney, M.C., Evans, W.P., Larsen, L.H., Morton, M.J., Mendelson, E.B., Farria, D.M., Cormack, J.B., Marques, H.S., Adams, A., Yeh, N.M., Gabrielli, G., Berg, W.A., Zhang. (2012). Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. JAMA: Journal of the American Medical Association, 307(13), 1394–1404. https://doi.org/10.1001/jama.2012.388

Bick, U. (2015). Intensified surveillance for early detection of breast cancer in high-risk patients. Breast Care (Basel, Switzerland), 10(1), 13–20. https://doi.org/10.1159/000375390

Brandzel, S., Rosenberg, D.E., Johnson, D., Bush, M., Kerlikowske, K., Onega, T., Henderson, L., Nekhlyudov, L., DeMartin, W., Wernli, K.I. (2017) Patient preference and adherence, 11, 199-204. https://doi.org/10.2147/PPA.S122244

Brédart, A., Kop, J-L., Fall, M., Pelissier, S., Simondi, C., Dolbeault, S., Livartowski, A., Tardivon, A.(2012)PatientEducationAndCounseling,86(3)405-413.http://doi.org/10.1016/j.pec.2011.06.012

Chhor, C. M., & Mercado, C. L. (2017). Abbreviated MRI Protocols: Wave of the Future for Breast Cancer Screening. AJR. American Journal Of Roentgenology, 208(2), 284–289. https://doi.org/10.2214/AJR.16.17205

Chiarelli, A. M., Prummel, M. V., Muradali, D., Majpruz, V., Horgan, M., Carroll, J. C., J. C., Eisen, A., Meschino, W. S., Shumak, R. S., Warner, E., Rabeneck, L. (2014). Effectiveness of screening with annual magnetic resonance imaging and mammography: results of the initial screen from the ontario high risk breast screening program. Journal of Clinical Oncology, 32(21), 2224–2230. https://doi.org/10.1200/JCO.2013.52.8331

Cott Chubiz, J. E., Lee, J. M., Gilmore, M. E., Kong, C. Y., Lowry, K. P., Halpern, E. F., McMahon, P.M., Ryan, P.D., Gazelle, G. S. (2013). Cost-effectiveness of alternating magnetic resonance imaging and digital mammography screening in BRCA1 and BRCA2 gene mutation carriers. Cancer (0008543X), 119(6), 1266–1276. https://doi.org/10.1002/cncr.27864

de Bock, G. H., Vermeulen, K. M., Jansen, L., Oosterwijk, J. C., Siesling, S., Dorrius, M. D., M. D., Feenstra, T., Houssami, N., Greuter, M. J. W. (2013). Which screening strategy should be offered to women with BRCA1 or BRCA2 mutations? A simulation of comparative cost-effectiveness. British Journal of Cancer, 108(8), 1579–1586. https://doi.org/10.1038/bjc.2013.149

Destounis, S., Arieno, A., & Morgan, R. (2016). Personal History of Premenopausal Breast Cancer as a Risk Factor for Referral to Screening Breast MRI. Academic Radiology, 23(3), 353–357. https://doi.org/10.1016/j.acra.2015.11.012

Evans, D. G., Harkness, E. F., Howell, A., Wilson, M., Hurley, E., Holmen, M. M., Tharmaratnam, K. U., Hagen, A. I., Lim, Y., Maxwell, A. J., Moller, P. (2016). Intensive breast screening in BRCA2 mutation carriers is associated with reduced breast cancer specific and all cause mortality. Hereditary Cancer in Clinical Practice, 14(1), 8. https://doi.org/10.1186/s13053-016-0048-3

Evans, D. G., Kesavan, N., Lim, Y. Y., Gadde, S., Hurley, E., Massat, N. J., N. J., Maxwell, A. J., Ingham, S., Eeles, R., Leach, M. O., Howell, A., Duffy, S., Stephen, D. (2014). MRI breast screening in high-risk women: cancer detection and survival analysis. Breast Cancer Research and Treatment, 145(3), 663–672. https://doi.org/10.1007/s10549-014-2931-9

Fischer, U., Korthauer, A., Baum, F., Luftner-Nagel, S., Heyden, D., & Marten-Engelke, K. (2012). Short first-pass MRI of the breast. Acta Radiologica (Stockholm, Sweden: 1987), 53(3), 267–269. https://doi.org/10.1258/ar.2012.110638

Freitas, V., Scaranelo, A., Menezes, R., Kulkarni, S., Hodgson, D., & Crystal, P. (2013). Added cancer yield of breast magnetic resonance imaging screening in women with a prior history of chest radiation therapy. Cancer (0008543X), 119(3), 495–503. https://doi.org/10.1002/cncr.27771

Friedlander, L.C., Roth, S.O., & Gavenonis, S.C. (2011). Results of MR Imaging Screening for Breast Cancer in High-Risk Patients with Lobular Carcinoma in Situ. Radiology, 261(2), 421–427. https://doi.org/10.1148/radiol.11103516

Giess, C.S., Poole, P.S., Chikarmane, S.A., Sippo, D.A., & Birdwell, R.L. (2015). Screening Breast MRI in Patients Previously Treated for Breast Cancer: Diagnostic Yield for Cancer and Abnormal Interpretation Rate. Academic Radiology, 22(11), 1331–1337. https://doi.org/10.1016/j.acra.2015.05.009

Gilbert, F. J., & Selamoglu, A. (2017). Personalised screening: is this the way forward? Clinical Radiology. https://doi.org/10.1016/j.crad.2017.11.021

Greenwood, H. I., Heller, S. L., Kim, S., Sigmund, E. E., Shaylor, S. D., & Moy, L. (2013). Ductal carcinoma in situ of the breasts: review of MR imaging features. Radiographics: A Review Publication Of The Radiological Society Of North America, Inc, 33(6), 1569–1588. https://doi.org/10.1148/rg.336125055

Grimm, L. J., Soo, M. S., Yoon, S., Kim, C., Ghate, S. V., & Johnson, K. S. (2015). Abbreviated Screening Protocol for Breast MRI. Academic Radiology, 22(9), 1157–1162. https://doi.org/10.1016/j.acra.2015.06.004

Gubern-Merida, A., Marti, R., Melendez, J., Hauth, J. L., Mann, R. M., Karssemeijer, N., & Platel, B. (2015). Automated localization of breast cancer in DCE-MRI. Medical Image Analysis, 20(1), 265–274. https://doi.org/10.1016/j.media.2014.12.001

Gweon, H. M., Cho, N., Han, W., Yi, A., Moon, H.-G., Noh, D.-Y., & Moon, W. K. (2014). Breast MR Imaging Screening in Women with a History of Breast Conservation Therapy. Radiology, 272(2), 366–373. https://doi.org/10.1148/radiol.14131893

Harvey, S. C., Di Carlo, P. A., Lee, B., Obadina, E., Sippo, D., & Mullen, L. (2016). An Abbreviated Protocol for High-Risk Screening Breast MRI Saves Time and Resources. Journal of the American College of Radiology, 13(4), 374–380. https://doi.org/10.1016/j.jacr.2015.08.015

Health Quality Ontario. (2016). Magnetic Resonance Imaging as an Adjunct to Mammography for Breast Cancer Screening in Women at Less Than High Risk for Breast Cancer: A Health Technology Assessment. Ontario Health Technology Assessment Series, 16(20), 1–30. Retrieved from

http://ezproxy.massey.ac.nz/login?url=http://search.ebscohost.com/login.aspx?direct=true&A uthType=ip,cookie,url,uid&db=cmedm&AN=27990198&site=ehost-live&scope=site

Heller, S. L., & Moy, L. (2016). Breast MRI Screening: Benefits and Limitations. Current Breast Cancer Reports, 8(4), 248–257. https://doi.org/10.1007/s12609-016-0230-7

Heywang-Köbrunner, S. H., Hacker, A., & Sedlacek, S. (2013). Magnetic resonance imaging: the evolution of breast imaging. Breast (Edinburgh, Scotland), 22 Suppl 2, S77–S82. https://doi.org/10.1016/j.breast.2013.07.014

Jain, M., Jain, A., Hyzy, M. D., & Werth, G. (2017). FAST MRI breast screening revisited. Journal of Medical Imaging & Radiation Oncology, 61(1), 24–28. https://doi.org/10.1111/1754-9485.12502

Kanal, E., & Tweedle, M. F. (2015). Residual or Retained Gadolinium: Practical Implications forRadiologistsandOurPatients.Radiology,275(3),630–634.https://doi.org/10.1148/radiol.2015150805

Kaniklidis, C. (2015). Beyond the mammography debate: a moderate perspective. Current Oncology, 22(3), 220–229. https://doi.org/10.3747/co.22.2585

King, T. A., Muhsen, S., Patil, S., Koslow, S., Oskar, S., Park, A., Morrogh, M., Sakr, R. A., Morrow, M. (2013). Is there a role for routine screening MRI in women with LCIS? Breast Cancer Research and Treatment, 142(2), 445–453. https://doi.org/10.1007/s10549-013-2725-5

Kuhl, C. K., Schrading, S., Strobel, K., Schild, H. H., Hilgers, R.-D., & Bieling, H. B. (2014). Abbreviated Breast Magnetic Resonance Imaging (MRI): First Postcontrast Subtracted Images and Maximum-Intensity Projection-A Novel Approach to Breast Cancer Screening With MRI. Journal of Clinical Oncology, 32(22), 2304–2310. https://doi.org/10.1200/JCO.2013.52.5386

Lee, C. H., Dershaw, D. D., Kopans, D., Evans, P., Monsees, B., Monticciolo, D., Brenner, R. J., Bassett, L., Berg, W., Feig, S., Hendrick, E., Mendelson, E., D'Orsi, C., Sickles, E., Burhenne, L. W. (2010). Breast Cancer Screening With Imaging: Recommendations From the Society of Breast Imaging and the ACR on the Use of Mammography, Breast MRI, Breast Ultrasound, and Other Technologies for the Detection of Clinically Occult Breast Cancer. Journal of the American College of Radiology, 7(1), 18–27. https://doi.org/10.1016/j.jacr.2009.09.022

Lehman, C. D. (2010). Magnetic resonance imaging in the evaluation of ductal carcinoma in situ. Journal Of The National Cancer Institute. Monographs, 2010(41), 150–151. https://doi.org/10.1093/jncimonographs/lgq030

Lehman, C. D., Lee, J. M., DeMartini, W. B., Hippe, D. S., Rendi, M. H., Kalish, G., Porter, P., Gralow, J., Partridge, S. C. (2016). Screening MRI in Women With a Personal History of Breast Cancer. Journal Of The National Cancer Institute, 108(3). https://doi.org/10.1093/jnci/djv349 Le-Petross, H. T., & Shetty, M. K. (2011). Magnetic resonance imaging and breast ultrasonography as an adjunct to mammographic screening in high-risk patients. Seminars In Ultrasound, CT, And MR, 32(4), 266–272. https://doi.org/10.1053/j.sult.2011.03.005

Lourenco, A. P., Donegan, L., Khalil, H., & Mainiero, M. B. (2014). Improving outcomes of screening breast MRI with practice evolution: Initial clinical experience with 3T compared to 1.5T. Journal of Magnetic Resonance Imaging, 39(3), 535–539. https://doi.org/10.1002/jmri.24198

Lowry KP, Lee JM, Kong CY, McMahon PM, Gilmore ME, Cott Chubiz JE, Pisano, E.D., Gatsonis, C., Ryan, P.D., Ozanne, E.M., Gazelle GS. (2012). Annual screening strategies in BRCA1 and BRCA2 gene mutation carriers: a comparative effectiveness analysis. Cancer (0008543X), 118(8), 2021–2030. https://doi.org/10.1002/cncr.26424

Mainiero, M. B., Lourenco, A., Mahoney, M. C., Newell, M. S., Bailey, L., Barke, L. D., O'Orsi, C., Harvey, J. A., Hayes, M. K., Huynh, P. T., Jokich, P. M., Lee, S-J., Lehman, C. D., Mankoff, D. A., Nepute, J. A., Patel, S. B., Reynolds, H. E., Sutherland, M. L., Haffty, B. G. (2016). ACR Appropriateness Criteria Breast Cancer Screening. Journal of the American College of Radiology, 13, R45–R49. https://doi.org/10.1016/j.jacr.2016.09.021

McDonald, E. S., Hammersley, J. A., Chou, S. H. S., Rahbar, H., Scheel, J. R., Lee, C. I., Liu, C. L. Lehman, C. D., Partridge, S. C. (2016). Performance of DWI as a Rapid Unenhanced Technique for Detecting Mammographically Occult Breast Cancer in Elevated-Risk Women With Dense Breasts. American Journal of Roentgenology, 207(1), 205–216. https://doi.org/10.2214/AJR.15.15873

McLaughlin, S., Mittendorf, E. A., Bleicher, R. J., McCready, D. R., & King, T. A. (2014). The 2013 Society of Surgical Oncology Susan G. Komen for the Cure Symposium: MRI in Breast Cancer: Where Are We Now? Annals Of Surgical Oncology, 21(1), 28–36. https://doi.org/10.1245/s10434-013-3307-9

Mehnati, P., & Tirtash, M. J. (2015). Comparative Efficacy of Four Imaging Instruments for Breast Cancer Screening. Asian Pacific Journal Of Cancer Prevention: APJCP, 16(15), 6177–6186. Retrieved from

http://ezproxy.massey.ac.nz/login?url=http://search.ebscohost.com/login.aspx?direct=true&A uthType=ip,cookie,url,uid&db=cmedm&AN=26434814&site=ehost-live&scope=site

Melnikow, J., Fenton, J. J., Whitlock, E. P., Miglioretti, D. L., Weyrich, M. S., Thompson, J. H., & Shah, K. (2016). Supplemental Screening for Breast Cancer in Women With Dense Breasts: A Systematic Review for the U.S. Preventive Services Task Force. Annals of Internal Medicine, 164(4), 268–278. https://doi.org/10.7326/M15-1789

Moftah, H. M., Azar, A. T., Al-Shammari, E. T., Ghali, N. I., Hassanien, A. E., & Shoman, M. (2014). Adaptive k-means clustering algorithm for MR breast image segmentation. Neural Computing & Applications, 24(7–8), 1917–1928.

Møller, P., Stormorken, A., Jonsrud, C., Holmen, M. M., Hagen, A. I., Clark, N., N., Vabø, A., Sun, P., Narod, S. A., Mæhle, L. (2013). Survival of patients with BRCA1-associated breast cancer diagnosed in an MRI-based surveillance program. Breast Cancer Research and Treatment, 139(1), 155–161. https://doi.org/10.1007/s10549-013-2540-z

Morrow, M., Waters, J., & Morris, E. (2011). MRI for breast cancer screening, diagnosis, and treatment. Lancet (London, England), 378(9805), 1804–1811. https://doi.org/10.1016/S0140-6736(11)61350-0

Moschetta, M., Telegrafo, M., Rella, L., Stabile Ianora, A. A., & Angelelli, G. (2016). Abbreviated Combined MR Protocol: A New Faster Strategy for Characterizing Breast Lesions. Clinical Breast Cancer, 16(3), 207–211. https://doi.org/10.1016/j.clbc.2016.02.008

Ng, A. K., Garber, J. E., Diller, L. R., Birdwell, R. L., Feng, Y., Neuberg, D. S., Silver, B., Fisher, D. C., Marcus, K. J., Mauch, P. M. (2013). Prospective study of the efficacy of breast magnetic resonance imaging and mammographic screening in survivors of hodgkin lymphoma. Journal of Clinical Oncology, 31(18), 2282–2288. https://doi.org/10.1200/JC0.2012.46.5732

Niell, B. L., Gavenonis, S. C., Motazedi, T., Chubiz, J. C., Halpern, E. P., Rafferty, E. A., & Lee, J. M. (2014). Auditing a breast MRI practice: performance measures for screening and diagnostic breast MRI. Journal of the American College of Radiology, 11(9), 883–889. https://doi.org/10.1016/j.jacr.2014.02.003

O'Flynn, E. A. M., Ledger, A. E. W., & deSouza, N. M. (2015). Alternative screening for dense breasts: MRI. AJR. American Journal Of Roentgenology, 204(2), W141–W149. https://doi.org/10.2214/AJR.14.13636

Partridge, S. C., Nissan, N., Rahbar, H., Kitsch, A. E., & Sigmund, E. E. (2017). Diffusion-weighted breast MRI: Clinical applications and emerging techniques. Journal of Magnetic Resonance Imaging, 45(2), 337–355. https://doi.org/10.1002/jmri.25479

Passaperuma, K., Warner, E., Causer, P., Hill, K., Messner, S., & Wong, J. (2012). Long-term results of screening with magnetic resonance imaging in women with BRCA mutations. Br J Cancer, 107(1), 24–30. https://doi.org/10.1038/bjc.2012.204

Pederson, H. J., O'Rourke, C., Lyons, J., Patrick, R. J., Crowe Jr., J. P., Grobmyer, S. R., & Crowe, J. P. J. (2015). Time-Related Changes in Yield and Harms of Screening Breast Magnetic Resonance Imaging. Clinical Breast Cancer, 15(5), 348–352. https://doi.org/10.1016/j.clbc.2015.01.006

Porembka, J. H., Seiler, S. J., & Sharma, P. B. (2016). Advanced Breast MRI Techniques: Helpful for Screening Breast Cancer? Current Breast Cancer Reports, 8(4), 236–241. https://doi.org/10.1007/s12609-016-0226-3

Rahbar, H., & Partridge, S. C. (2016). Multiparametric MR Imaging of Breast Cancer. MagneticResonanceImagingClinicsofNorthAmerica,24(1),223-+.https://doi.org/10.1016/j.mric.2015.08.012

Rijnsburger AJ, Obdeijn IM, Kaas R, Tilanus-Linthorst MM, Boetes C, Loo CE, Wasser, M.N., Bergers, E., Kok, T., Muller, S.H., Peters, H., Tollenaar, R.A., Hoogerbrugge, N., Meijer, S., Bartels, C.C., Seynaeve, C., Hooning, M.J., Kriege M., Schmitz, P.I., Oosterwijk JC. (2010). BRCA1-Associated Breast Cancers Present Differently From BRCA2-Associated and Familial Cases: Long-Term Follow-Up of the Dutch MRISC Screening Study. Journal of Clinical Oncology, 28(36), 5265–5273. https://doi.org/10.1200/JCO.2009.27.2294

Runge, V. M. (2013). Current technological advances in magnetic resonance with critical impact for clinical diagnosis and therapy. Investigative Radiology, 48(12), 869–877. https://doi.org/10.1097/01.rli.0000434380.71793.d3 Saadatman, S., Tilanus-Linthorst, M.A.M., Rutgers, E.J.T., Hoogerbrugge, N., Oosterwijk, J.C., Tollenaar, R.A.E.M., Hooning, M., Loo, C.E., Obdeijn, I-M., Heijnsdijk, e.a.M., de Koning, H.J., (2013) Journal Of The National Cancer Institute, 105(17), 1314-1321. https://doi.org/10.1093/jnci/djt203

Salem, D. S., Kamal, R. M., Mansour, S. M., Salah, L. A., & Wessam, R. (2013). Breast imaging in the young: the role of magnetic resonance imaging in breast cancer screening, diagnosis and follow-up. Journal of Thoracic Disease, 5, S9–S18. https://doi.org/10.3978/j.issn.2072-1439.2013.05.02

Santoro, F., Podo, F., & Sardanelli, F. (2014). MRI screening of women with hereditary predisposition to breast cancer: diagnostic performance and survival analysis. Breast Cancer Research and Treatment, 147(3), 685–687. https://doi.org/10.1007/s10549-014-3097-1

Sardanelli, F., Boetes, C., Borisch, B., Decker, T., Federico, M., Gilbert, F. J., Helbich, T., Heywang-Kobrunner, S.H., Kaiser, W.A., Kerin, M.J., Mansel, R.E., Martincich, L., Mauriac, L., Meijers-Heijboer, H., Orecchia, R., Panizza, P., Ponti, A., Purushotham A.D., Regitnig, P., Del Turco, M.R., Thibault, F., Wilson, R. (2010). Magnetic resonance imaging of the breast: Recommendations from the EUSOMA working group. European Journal of Cancer, 46(8), 1296–1316.

Sardanelli, F., Podo, F., Santoro, F., Manoukian, S., Bergonzi, S., Trecate, G., G., Vergnaghi, D., Federico, M., Cortesi, L., Corcione, S., Morassut, S., Di Maggio, C., Cilotti, A., Martincich, L., Calabrese, M., Zuiani, C., Preda, L., Bonanni, B., Carbonaro, L.A., Conteguiacomo, A., Panizza, P., Di Cesare, E., Savarese, A., Crecco, M., Turchetti, D., Tonutti, D., Belli, P., Maschio, A. D. (2011). Multicenter surveillance of women at high genetic breast cancer risk using mammography, ultrasonography, and contrast-enhanced magnetic resonance imaging (the high breast cancer risk italian 1 study): final results. Investigative Radiology, 46(2), 94–105. https://doi.org/10.1097/RLI.0b013e3181f3fcdf

Sathya, D. J., & Geetha, K. (2013). Mass classification in breast DCE-MR images using an artificial neural network trained via a bee colony optimization algorithm. Scienceasia, 39(3), 294–305. https://doi.org/10.2306/scienceasia1513-1874.2013.39.294

Sutcliffe, J. B., 3rd, & Otto, P. M. (2013). Controversies in breast MRI. Current Problems In Diagnostic Radiology, 42(4), 149–163. https://doi.org/10.1067/j.cpradiol.2013.03.001

Tieu, M. T., Cigsar, C., Ahmed, S., Ng, A., Diller, L., Millar, B.-A., Crystal, P., Hodgson, D. C. (2014). Breast cancer detection among young survivors of pediatric Hodgkin lymphoma with screening magnetic resonance imaging. Cancer (0008543X), 120(16), 2507–2513. https://doi.org/10.1002/cncr.28747

Wellings, E., Vassiliades, L., & Abdalla, R. (2016). Breast Cancer Screening for High-Risk Patients of Different Ages and Risk - Which Modality Is Most Effective? Cureus, 8(12), e945–e945. https://doi.org/10.7759/cureus.945

Zhang, Y., & Ren, H. (2017). Meta-analysis of diagnostic accuracy of magnetic resonance imaging and mammography for breast cancer. Journal of Cancer Research & Therapeutics, 13(5), 862–868. https://doi.org/10.4103/jcrt.JCRT\_678\_17

#### Microwave imaging

Kwon, S., & Lee, S. (2016). Recent Advances in Microwave Imaging for Breast Cancer Detection. International Journal of Biomedical Imaging, 25. http://dx.doi.org/10.1155/2016/5054912

Microwave breast screening getting closer. (2017, March 15). University of Waikato. Retrieved from http://www.waikato.ac.nz/news-events/media/2017/microwave-breast-screening-getting-closer

Modiri, A., Goudreau, S., Rahimi, A., & Kiasaleh, K. (2017). Review of breast screening: Toward clinical realization of microwave imaging. Medical Physics, 44(12), e446–e458. https://doi.org/10.1002/mp.12611

National Horizon Scanning Unit. (2009). New and emerging technologies for breast cancer detection. Canberra: Australia and New Zealand Horizon Scanning Network. Retrieved from http://www.horizonscanning.gov.au/internet/horizon/publishing.nsf/Content/AD1C4F0CFAD 1A5E4CA2575E8001DC431/\$File/ETB\_BreastScreen.pdf

Wang, L. (2017). Early Diagnosis of Breast Cancer. Sensors (Basel, Switzerland), 17(7). https://doi.org/10.3390/s17071572

#### Molecular breast imaging

Adrada, B. E., Moseley, T., & Rauch, G. M. (2016). Molecular Breast Imaging: Role as a Screening Modality. Current Breast Cancer Reports, 8(4), 230–235. https://doi.org/10.1007/s12609-016-0225-4

Berg, W. A. (2016). Nuclear Breast Imaging: Clinical Results and Future Directions. Journal Of Nuclear Medicine: Official Publication, Society Of Nuclear Medicine, 57 Suppl 1, 46S–52S. https://doi.org/10.2967/jnumed.115.157891

BreastScreen Australia. (2015). BreastScreen Australia National Accreditation Standards. BreastScreen Australia.

Even-Sapir, E., Golan, O., Menes, T., Weinstein, Y., & Lerman, H. (2016). Breast Imaging Utilizing Dedicated Gamma Camera and (99m)Tc-MIBI: Experience at the Tel Aviv Medical Center and Review of the Literature Breast Imaging. Seminars In Nuclear Medicine, 46(4), 286–293. https://doi.org/10.1053/j.semnuclmed.2016.01.001

Harvey, J. A., & Bovbjerg, V. E. (2004). Quantitative assessment of mammographic breast density: relationship with breast cancer risk. Radiology, 230(1), 29.

Ho, J. M., Jafferjee, N., Covarrubias, G. M., Ghesani, M., & Handler, B. (2014). Dense breasts: a review of reporting legislation and available supplemental screening options. AJR. American Journal Of Roentgenology, 203(2), 449–456. https://doi.org/10.2214/AJR.13.11969

Holbrook, A., & Newel, M. S. (2015). Alternative screening for women with dense breasts: breastspecific gamma imaging (molecular breast imaging). AJR. American Journal Of Roentgenology, 204(2), 252–256. https://doi.org/10.2214/AJR.14.13525

Hruska, C. B. (2017). Molecular Breast Imaging for Screening in Dense Breasts: State of the Art and Future Directions. AJR. American Journal Of Roentgenology, 208(2), 275–283. https://doi.org/10.2214/AJR.16.17131 Hruska, C. B., Conners, A. L., Jones, K. N., O'Connor, M. K., Moriarty, J. P., Boughey, J. C., & Rhodes, D. J. (2015). Diagnostic workup and costs of a single supplemental molecular breast imaging screen of mammographically dense breasts. American Journal of Roentgenology, 204(6), 1345–1353. https://doi.org/10.2214/AJR.14.13306

Hruska, C. B., & O'Connor, M. K. (2013). Nuclear imaging of the breast: Translating achievements in instrumentation into clinical use. Medical Physics, 40(5). https://doi.org/10.1118/1.4802733

McPeak, P. R. (2014). Advancements in molecular breast imaging. Radiologic Technology, 85(5), 523M.

National Horizon Scanning Unit. (2009). New and emerging technologies for breast cancer detection. Canberra: Australia and New Zealand Horizon Scanning Network. Retrieved from http://www.horizonscanning.gov.au/internet/horizon/publishing.nsf/Content/AD1C4F0CFAD 1A5E4CA2575E8001DC431/\$File/ETB\_BreastScreen.pdf

Shermis, R. B., Redfern, R. E., Burns, J., & Kudrolli, H. (2017). Molecular Breast Imaging in Breast Cancer Screening and Problem Solving. RadioGraphics, 37(5), 1309–1606. https://doi.org/10.1148/rg.2017160204

Sun, Y., Wei, W., Yang, H.-W., & Liu, J.-L. (2013). Clinical usefulness of breast-specific gamma imaging as an adjunct modality to mammography for diagnosis of breast cancer: a systemic review and meta-analysis. European Journal of Nuclear Medicine and Molecular Imaging, 40(3), 450–463. https://doi.org/10.1007/s00259-012-2279-5

Surti, S. (2013). Radionuclide methods and instrumentation for breast cancer detection and<br/>diagnosis.SeminarsInNuclearMedicine,43(4),271–280.https://doi.org/10.1053/j.semnuclmed.2013.03.003

Tafreshi, N. K., Kumar, V., Morse, D. L., & Gatenby, R. A. (2010). Molecular and functional imaging of breast cancer. Cancer Control: Journal Of The Moffitt Cancer Center, 17(3), 143–155.

Wahl, R. L. (2016). Quo Vadis: PET and Single-Photon Molecular Breast Imaging. Journal of Nuclear Medicine, 57(Supplement 1), 3S–8S. https://doi.org/10.2967/jnumed.115.159202

Weigert, J. M., & Kieper, D. A. (2012). Current and future roles of molecular breast imaging in the community-based breast center. Imaging in Medicine, 4(4), 383–387. https://doi.org/10.2217/iim.12.36

Yang, W. T. (2011). Emerging techniques and molecular imaging in breast cancer. Seminars In Ultrasound, CT, And MR, 32(4), 288–299. https://doi.org/10.1053/j.sult.2011.03.003

Zardavas, D., Regan, M., Maibach, R., Ruepp, B., Hiltbrunner, A., Blacher, L., Goulioti, T., Gelber, R., Flamen, P., Piccart, M., Malorni, L. (2017). PYTHIA: a phase II study of palbociclib plus fulvestrant versus placebo plus fulvestrant for pretreated patients with ER+/HER2-metastatic breast cancer. Cancer Research. Conference: 39th Annual CTRC-AACR San Antonio Breast Cancer Symposium. United States, 77(4 Supplement 1). https://doi.org/10.1158/1538-7445.SABCS16-P6-11-16

#### Spectroscopy

Akbari Sari, A., Mobinizadeh, M., & Azadbakht, M. (2013). A systematic review of the effects of diffuse optical imaging in breast diseases. Iranian Journal Of Cancer Prevention, 6(1), 44–51.

Jermyn, M., Desroches, J., Aubertin, K., St-Arnaud, K., Madore, W. J., De Montigny, E., Guiot, M.C., Trudel, D., Wilson, B.C., Petrecca, K., Leblond, F. (2016). A review of Raman spectroscopy advances with an emphasis on clinical translation challenges in oncology. Physics in Medicine and Biology, 61(23), R370–R400.

National Horizon Scanning Unit. (2009). New and emerging technologies for breast cancer detection. Canberra: Australia and New Zealand Horizon Scanning Network. Retrieved from http://www.horizonscanning.gov.au/internet/horizon/publishing.nsf/Content/AD1C4F0CFAD 1A5E4CA2575E8001DC431/\$File/ETB\_BreastScreen.pdf

#### Thermography

Acharya, U. R., Ng, E. Y. K., Tan, J. H., & Sree, S. V. (2010). Thermography Based Breast Cancer Detection Using Texture Features and Support Vector Machine. Journal of Medical Systems, 36(3), 1503–1510. https://doi.org/10.1007/s10916-010-9611-z

Borchartt, T. B., Conci, A., Lima, R. C. F., Resmini, R., & Sanchez, A. (2013). Breast thermography from an image processing viewpoint: A survey. Signal Processing, 93(10), 2785–2803. https://doi.org/10.1016/j.sigpro.2012.08.012

Cancer Australia. (2010). Statement on use of thermography to detect breast cancer. Retrieved from https://canceraustralia.gov.au/publications-and-resources/positionstatements/statement-use-thermography-detect-breast-cancer

Etehadtavakol, M., & Ng, E. Y. K. (2013). Breast Thermography as a Potential Non-Contact Method in the Early Detection of Cancer: A Review. Journal of Mechanics in Medicine and Biology, 13(2). https://doi.org/10.1142/S0219519413300019

Fitzgerald, A., & Berentson-Shaw, J. (2012). Thermography as a screening and diagnostic tool: a systematic review. The New Zealand Medical Journal, 125(1351), 80–91.

Francis, S. V., Sasikala, M., Bharathi, G. B., & Jaipurkar, S. D. (2014). Breast cancer detection in rotational thermography images using texture features. Infrared Physics & Technology, 67, 490–496. https://doi.org/10.1016/j.infrared.2014.08.019

Francis, S. V., Sasikala, M., & Saranya, S. (2014). Detection of Breast Abnormality from Thermograms Using Curvelet Transform Based Feature Extraction. Journal of Medical Systems, 38(4). https://doi.org/10.1007/s10916-014-0023-3

Gerasimova, E., Audit, B., Roux, S. G., Khalil, A., Argoul, F., Naimark, O., & Arneodo, A. (2013). Multifractal analysis of dynamic infrared imaging of breast cancer. Epl, 104(6). https://doi.org/10.1209/0295-5075/104/68001

Gerasimova, E., Audit, B., Roux, S. G., Khalil, A., Gileva, O., Argoul, F., Naimark, O., Arneodo, A. (2014). Wavelet-based nultifractal analysis of dynamic infrared thermograms to assist in early breast cancer diagnosis. Frontiers in Physiology, 5. https://doi.org/10.3389/fphys.2014.00176

Kandlikar, S. G., Perez-Raya, I., Raghupathi, P. A., Gonzalez-Hernandez, J. L., Dabydeen, D., Medeiros, L., & Phatak, P. (2017). Infrared imaging technology for breast cancer detection -

Current status, protocols and new directions. International Journal of Heat and Mass Transfer, 108, 2303–2320. https://doi.org/10.1016/j.ijheatmasstransfer.2017.01.086

Kermani, S., Samadzadehaghdam, N., & EtehadTavakol, M. (2015). Automatic color segmentation of breast infrared images using a Gaussian mixture model. Optik, 126(21), 3288–3294. https://doi.org/10.1016/j.ijleo.2015.08.007

Lanisa, N., Cheok, N. S., & Wee, L. K. (2014). Color Morphology and Segmentation of the Breast Thermography Image. In 2014 Ieee Conference on Biomedical Engineering and Sciences (pp. 772–775).

Madhu, H., Kakileti, S. T., Venkataramani, K., & Jabbireddy, S. (2016). Extraction of Medically Interpretable Features for Classification of Malignancy in Breast Thermography. In 2016 38th Annual International Conference of the Ieee Engineering in Medicine and Biology Society (pp. 1062–1065).

National Screening Unit. (2010). Position Statement The use of thermography as a breast cancer screening or diagnostic tool. Retrieved from https://www.nsu.govt.nz/health-professionals/breastscreen-aotearoa/position-statements

Rassiwala, M., Mathur, P., Mathur, R., Farid, K., Shukla, S., Gupta, P. K., & Jain, B. (2014). Evaluation of digital infra-red thermal imaging as an adjunctive screening method for breast carcinoma: A pilot study. International Journal of Surgery, 12(12), 1439–1443. https://doi.org/10.1016/j.ijsu.2014.10.010

Rastghalam, R., & Pourghassem, H. (2013). Breast Cancer Detection Using Spectral Probable Feature on Thermography Images. In 2013 8th Iranian Conference on Machine Vision & Image Processing (pp. 116–120).

Sella, T., Sklair-Levy, M., Cohen, M., Rozin, M., Shapiro-Feinberg, M., Allweis, T. M., ... Izhaky, D. (2013). A novel functional infrared imaging system coupled with multiparametric computerised analysis for risk assessment of breast cancer. European Radiology, 23(5), 1191–1198. https://doi.org/10.1007/s00330-012-2724-7

US Food and Drug Administration. (2017). Breast Cancer Screening: Thermogram No Substitute for Mammogram. Retrieved from https://www.fda.gov/ForConsumers/ConsumerUpdates/ucm257499.htm

Venkataramani, K., Mestha, L. K., Ramachandra, L., Prasad, S. S., Kumar, V., & Raja, P. J. (2015). Semi-automated Breast Cancer Tumor Detection with Thermographic Video Imaging. In 2015 37th Annual International Conference of the Ieee Engineering in Medicine and Biology Society (pp. 2022–2025).

#### Tomography

Aminololama-Shakeri, S., Hargreaves, J. B., Boone, J. M., & Lindfors, K. K. (2016). Dedicated Breast CT: Screening Technique of the Future. Current Breast Cancer Reports, 8(4), 242–247. https://doi.org/10.1007/s12609-016-0227-2

BreastScreen Australia. (2015). BreastScreen Australia National Accreditation Standards. BreastScreen Australia.
Cintolo, J. A., Tchou, J., & Pryma, D. A. (2013). Diagnostic and prognostic application of positron emission tomography in breast imaging: emerging uses and the role of PET in monitoring treatment response. Breast Cancer Research And Treatment, 138(2), 331–346. https://doi.org/10.1007/s10549-013-2451-z

Forsyth, J., Borsic, A., Halter, R. J., Hartov, A., & Paulsen, K. D. (2011). Optical breast shape capture and finite-element mesh generation for electrical impedance tomography. Physiological Measurement, 32(7), 797. https://doi.org/10.1088/0967-3334/32/7/S05

Jung, W., & Boppart, S. A. (2012). Modern Trends in Imaging V: Optical Coherence Tomography for Rapid Tissue Screening and Directed Histological Sectioning. Analytical Cellular Pathology (Amsterdam), 35(3), 129–143. https://doi.org/10.3233/ACP-2011-0047

Lindfors, K. K., Boone, J. M., Newell, M. S., & D'Orsi, C. J. (2010). Dedicated breast computed tomography: the optimal cross-sectional imaging solution? Radiologic Clinics of North America, 48(5), 1043–1054. https://doi.org/10.1016/j.rcl.2010.06.001

Minamimoto, R., Senda, M., Jinnouchi, S., Terauchi, T., Yoshida, T., & Inoue, T. (2015). Detection of Breast Cancer in an FDG-PET Cancer Screening Program: Results of a Nationwide Japanese Survey. Clinical Breast Cancer, 15(2), e139–e146. https://doi.org/10.1016/j.clbc.2014.09.008

Minamimoto, R., Senda M, Jinnouchi S, Terauchi T, Yoshida T, Murano T, Fukuda, H., Iinuma, T., Uno, K., Nishizawa, S., Tsukamoto, E., Iwata, H., Inoue, T., Oguchi, K., Nakashima, R. (2013). The current status of an FDG-PET cancer screening program in Japan, based on a 4-year (2006-2009) nationwide survey. Annals of Nuclear Medicine, 27(1), 46–57. https://doi.org/10.1007/s12149-012-0660-x

National Horizon Scanning Unit. (2009). New and emerging technologies for breast cancer detection. Canberra: Australia and New Zealand Horizon Scanning Network. Retrieved from http://www.horizonscanning.gov.au/internet/horizon/publishing.nsf/Content/AD1C4F0CFAD 1A5E4CA2575E8001DC431/\$File/ETB\_BreastScreen.pdf

O'Connell, A. M., Karellas, A., & Vedantham, S. (2014). The Potential Role of Dedicated 3D Breast CT as a Diagnostic Tool: Review and Early Clinical Examples. The Breast Journal, 20(6), 592–605. https://doi.org/10.1111/tbj.12327

Ruile, G., Dianatliev, A., Kriza, C., Meier, F., Leb, I., Kalender, W. A., & Kolominsky-Rabas, P. (2015). Screening for breast cancer with Breast-CT in a ProHTA simulation. Journal of Comparative Effectiveness Research, 4(6), 553–567. https://doi.org/10.2217/cer.15.42

Sarno, A., Mettivier, G., & Russo, P. (2015). Dedicated breast computed tomography: Basic aspects. Medical Physics, 42(6Part1), 2786–2804. https://doi.org/10.1118/1.4919441

Vakoc, B. J., Fukumura, D., Jain, R. K., & Bouma, B. E. (2012). Cancer imaging by optical coherence tomography: preclinical progress and clinical potential. Nature Reviews. Cancer, 12(5), 363–368. https://doi.org/10.1038/nrc3235

Wienbeck, S., Lotz, J., & Fischer, U. (2017). Review of clinical studies and first clinical experiences with a commercially available cone-beam breast CT in Europe. Clinical Imaging, 42, 50–59. https://doi.org/10.1016/j.clinimag.2016.11.011 Zain, N. M., & Chelliah, K. K. (2014). Breast Imaging Using Electrical Impedance Tomography: Correlation of Quantitative Assessment with Visual Interpretation. Asian Pacific Journal of Cancer Prevention, 15(3), 1327–1331. https://doi.org/10.7314/APJCP.2014.15.3.1327

## **READING STRATEGIES**

## **Computer aided detection**

Agrawal, P., Vasta, M., & Singh, R. (2014). Saliency based mass detection from screening mammograms. Signal Processing, 99, June 2014. https://doi.org/10.1016/j.sigpro.2013.12.010

Al-Najdawi, N., Biltawi, M., & Tedmori, S. (2015). Mammogram image visual enhancement, mass segmentation and classification. Applied Soft Computing, 35, 175–185. https://doi.org/10.1016/j.asoc.2015.06.029

Ayer, T., Chen, Q., & Burnside, E. S. (2013). Artificial neural networks in mammography interpretation and diagnostic decision making. Computational And Mathematical Methods In Medicine, 2013, 832509–832509. https://doi.org/10.1155/2013/832509

Azavedo, E., Zackrisson, S., Mejàre, I., & Heibert Arnlind, M. (2012). Is single reading with computer-aided detection (CAD) as good as double reading in mammography screening? A systematic review. BMC Medical Imaging, 12, 22–22. https://doi.org/10.1186/1471-2342-12-22

Bargalló, X., Santamaría, G., del Amo, M., Arguis, P., Ríos, J., Grau, J., Burrel, M., Cores, E., Velasco, M. (2014). Single reading with computer-aided detection performed by selected radiologists in a breast cancer screening program. European Journal of Radiology, 83(11), 2019–2023. https://doi.org/10.1016/j.ejrad.2014.08.010

Bargalló, X., Velasco, M., Santamaría, G., Amo, M., Arguis, P., & Sánchez Gómez, S. (2013). Role of Computer-Aided Detection in Very Small Screening Detected Invasive Breast Cancers. Journal of Digital Imaging, 26(3), 572–577. https://doi.org/10.1007/s10278-012-9550-y

Bessa, S., Domingues, I., Cardoso, J. S., Passarinho, P., Cardoso, P., Rodrigues, V., & Lage, F. (2014). Normal breast identification in screening mammography: a study on 18 000 images. In 2014 IEEE International Conference on Bioinformatics and Biomedicine.

Beura, S., Majhi, B., Dash, R., and Roy, S., (2015) Classification of Mammogram Using Two-Dimensional Discrete Orthonormal S-Transform for Breast Cancer Detection. Healthcare Technology Letters 2, 2 46–51. https://doi.org/10.1049/htl.2014.0108.

Choi, J. Y., & Ro, Y. M. (2012). Multiresolution local binary pattern texture analysis combined with variable selection for application to false-positive reduction in computer-aided detection of breast masses on mammograms. Physics in Medicine and Biology, 57(21), 7029–7052.

Chu, J. H., Min, H., Liu, L., & Lu, W. (2015). A novel computer aided breast mass detection scheme based on morphological enhancement and SLIC superpixel segmentation. Medical Physics, 42(7), 3859–3869.

Cole, E. B., Zhang, Z., Marques, H. S., Edward Hendrick, R., Yaffe, M. J., & Pisano, E. D. (2014). Impact of computer-aided detection systems on radiologist accuracy with digital mammography.

AJR. American Journal Of Roentgenology, 203(4), 909–916. https://doi.org/10.2214/AJR.12.10187

Cole, E., Zhang, Z., Nishikawa, R. M., Hendrick, R. E., Yaffe, M. J., Padungchaichote W, Kuzmiak, C., Chayakulkheeree, J., Conant, E.F., Fajardo, L.L., Baum, J., Gatsonis, C., Pisano. (2012). Assessing the stand-alone sensitivity of computer-aided detection with cancer cases from the digital mammographic imaging screening trial. American Journal of Roentgenology, 199, W392-401. https://doi.org/10.2214/AJR.11.7255

de Sampaio, W., Silva, A. C., de Paiva, A. C., & Gattass, M. (2015). Detection of masses in mammograms with adaption to breast density using genetic algorithm, phylogenetic trees, LBP and SVM. Expert Systems with Applications, 42(22), 8911–8928. https://doi.org/10.1016/j.eswa.2015.07.046

de Sampaio, W., Silva, A. C., de Paiva, A. C., Gattass, M., & Diniz, E. M. (2011). Detection of masses in mammogram images using CNN, geostatistic functions and SVM. Computers in Biology and Medicine, 41(8), 653–664. https://doi.org/10.1016/j.compbiomed.2011.05.017

Dhahbi, S., Barhoumi, W., & Zagrouba, E. (2015). Breast cancer diagnosis in digitized mammograms using curvelet moments. Computers in Biology and Medicine, 64, 79–90.

Dheeba, J., Singh, N. A., & Selvi, S. T. (2014). Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach. Journal of Biomedical Informatics, 49, 45–52. https://doi.org/10.1016/j.jbi.2014.01.010

Dhungel, N., Carneiro, G., & Bradley, A. P. (2017). A deep learning approach for the analysis of masses in mammograms with minimal user intervention. Medical Image Analysis, 37, 114–128. https://doi.org/10.1016/j.media.2017.01.009

Dong, M., Lu, X., Ma, Y., Guo, Y., Ma, Y., & Wang, K. (2015). An Efficient Approach for Automated Mass Segmentation and Classification in Mammograms. Journal of Digital Imaging, 28(5), 613–625. https://doi.org/10.1007/s10278-015-9778-4

Dromain, C., Boyer, B., Ferré, R., Canale, S., Delaloge, S., & Balleyguier, C. (2013). Computed-aided diagnosis (CAD) in the detection of breast cancer. European Journal Of Radiology, 82(3), 417–423. https://doi.org/10.1016/j.ejrad.2012.03.005

Dromain, C., Thibault, F., Diekmann, F., Fallenberg, E. M., Jong, R. A., Koomen, M., Hendrick R.E., Tardivon, A., Toledano, A. (2012). Dual-energy contrast-enhanced digital mammography: initial clinical results of a multireader, multicase study. Breast Cancer Research, 14(3). https://doi.org/10.1186/bcr3210

Fenton JJ, Abraham L, Taplin SH, Geller BM, Carney PA, D'Orsi C, Elmore, J.G., Barlow, W.E., (2011). Effectiveness of computer-aided detection in community mammography practice. JNCI: Journal of the National Cancer Institute, 103(15), 1152–1161. https://doi.org/10.1093/jnci/djr206

Ganesan, K., Acharya, U. R., Chua, C. K., Min, L. C., Abraham, K. T., & Ng, K.-H. (2013). Computeraided breast cancer detection using mammograms: a review. IEEE Reviews In Biomedical Engineering, 6, 77–98. https://doi.org/10.1109/RBME.2012.2232289

Guerriero, C., Gillan, M. G., Cairns, J., Wallis, M. G., Gilbert, F. J., & Gillan, M. G. C. (2011). Is computer aided detection (CAD) cost effective in screening mammography? A model based on

the CADET II study. BMC Health Services Research, 11(1), 11–11. https://doi.org/10.1186/1472-6963-11-11

Hong, B., & Sohn, B. (2010). Segmentation of Regions of Interest in Mammograms in a Topographic Approach. IEEE Transactions on Information Technology in Biomedicine, 14(1), 129–139. https://doi.org/10.1109/TITB.2009.2033269

Horsch, A., Hapfelmeier, A., & Elter, M. (2011). Needs assessment for next generation computeraided mammography reference image databases and evaluation studies. International Journal Of Computer Assisted Radiology And Surgery, 6(6), 749–767. https://doi.org/10.1007/s11548-011-0553-9

Houssami, N., & Ciatto, S. (2011). The evolving role of new imaging methods in breast screening. Preventive Medicine: An International Journal Devoted to Practice and Theory, 53(3), 123–126. https://doi.org/10.1016/j.ypmed.2011.05.003

Hu K., Li, F., & Gao, X., (2011) Detection of Suspicious Lesions by Adaptive Thresholding Based on Multiresolution Analysis in Mammograms. IEEE Transactions on Instrumentation and Measurement 60(2), 462–72. https://doi.org/10.1109/TIM.2010.2051060.

Hupse R, Samulski M, Lobbes M, den Heeten A, Imhof-Tas MW, Beijerinck D, Pijnappel, R., Boetes, C., Karssemeijer, N., (2013). Standalone computer-aided detection compared to radiologists' performance for the detection of mammographic masses. European Radiology, 23(1), 93–100. https://doi.org/10.1007/s00330-012-2562-7

James JJ, Gilbert FJ, Wallis MG, Gillan MG, Astley SM, Boggis CR, Agbaje, O.F., Brentnall, A.R., Duffy. (2010). Mammographic features of breast cancers at single reading with computer-aided detection and at double reading in a large multicenter prospective trial of computer-aided detection: CADET II. Radiology, 256(2), 379–386. https://doi.org/10.1148/radiol.10091899

Jen, C., & Yu, S. (2015). Automatic detection of abnormal mammograms in mammographic images. Expert Systems with Applications, 42(6), 3048–3055. https://doi.org/10.1016/j.eswa.2014.11.061

Karssemeijer, N. (2010). Computer Aided Detection in Breast Imaging: More Than Perception Aid. In 2010 7th Ieee International Symposium on Biomedical Imaging: From Nano to Macro (pp. 273–273).

Kendall, E. J., & Flynn, M. T. (2014). Automated breast image classification using features from itsdiscretecosinetransform.PlosOne,9(3),e91015-e91015.https://doi.org/10.1371/journal.pone.0091015

Kim S.J., Moon W.K., Kim S.Y., Chang J.M., Kim S.M., Cho N. (2010). Comparison of two software versions of a commercially available computer-aided detection (CAD) system for detecting breast cancer. Acta Radiologica, 51(5), 482–490. https://doi.org/10.3109/02841851003709490

Kooi, T., Litjens, G., van Ginneken, B., Gubern-Merida, A., Sancheza, C. I., Mann, R., den Heeten, A., Karssemeijer, N. (2017). Large scale deep learning for computer aided detection of mammographic lesions. Medical Image Analysis, 35, 303–312. https://doi.org/10.1016/j.media.2016.07.007 Lehman, C. D., Wellman, R. D., Buist, D. S. M., Kerlikowske, K., Tosteson, A. N. A., & Miglioretti, D. L. (2015). Diagnostic Accuracy of Digital Screening Mammography With and Without Computer-Aided Detection. JAMA Internal Medicine, 175(11), 1828–1837. https://doi.org/10.1001/jamainternmed.2015.5231

Lesniak, J. M., Hupse, R., Blanc, R., Karssemeijer, N., & Szekely, G. (2012). Comparative evaluation of support vector machine classification for computer aided detection of breast masses in mammography. Physics in Medicine and Biology, 57(16), NIL\_262-NIL\_274.

Liu, X., Xu, X., Liu, J., & Feng, Z., (2011) A New Automatic Method for Mass Detection in Mammography with False Positives Reduction by Supported Vector Machine. In 2011 4th International Conference on Biomedical Engineering and Informatics (BMEI), 1:33–37, https://doi.org/10.1109/BMEI.2011.6098328.

Mac Parthalian, N., Jensen, R., Shen, Q. A., & Zwiggelaar, R. (2010). Fuzzy-rough approaches for mammographic risk analysis. Intelligent Data Analysis, 14(2), 225–244. https://doi.org/10.3233/IDA-2010-0418

Mordang, J. J., Gubern-Merida, A., Bria, A., Tortorella, F., den Heeten, G., & Karssemeijer, N. (2017). Improving computer-aided detection assistance in breast cancer screening by removal of obviously false-positive findings. Medical Physics, 44(4), 1390–1401. https://doi.org/10.1002/mp.12152

Mordang, J. J., Gubern-Merida, A., den Heeten, G., & Karssemeijer, N. (2016). Reducing false positives of microcalcification detection systems by removal of breast arterial calcifications. Medical Physics, 43(4), 1676–1687. https://doi.org/10.1118/1.4943376

Mughal, B., Sharif, M., & Muhammad, N. (2017). Bi-model processing for early detection of breast tumor in CAD system. European Physical Journal Plus, 132(6).

Neto, O. P. S., Silva, A. C., Paiva, A. C., & Gattass, M. (2017). Automatic mass detection in mammography images using particle swarm optimization and functional diversity indexes. Multimedia Tools and Applications, 76(18), 19263–19289. https://doi.org/10.1007/s11042-017-4710-1

Povyakalo, A., Alberdi, E., Strigini, L., & Ayton, P. (2013). How to discriminate between computeraided and computer-hindered decisions: A case study in mammography. Medical Decision Making, 33(1), 98–107.

Qiu, Y. C., Wang, Y. Z., Yan, S. J., Tan, M., Cheng, S., Liu, H., & Zheng, B. (2015). An Initial Investigation on Developing a New Method to Predict Shor-term Breast Cancer Risk Based on Deep Learning Technology. Medical Imaging 2016: Computer-Aided Diagnosis, 9785. https://doi.org/10.1117/12.2216275

Raghavendra, U., U. R. Acharya, H. Fujita, A. Gudigar, J. H. Tan, and S Chokkadi. (2016). Application of Gabor Wavelet and Locality Sensitive Discriminant Analysis for Automated Identification of Breast Cancer Using Digitized Mammogram Images. Applied Soft Computing 46, 151–61. https://doi.org/10.1016/j.asoc.2016.04.036

Rangayyan, R. M., Banik, S., & Desautels, J. E. L. (2010). Computer-Aided Detection of Architectural Distortion in Prior Mammograms of Interval Cancer. Journal of Digital Imaging, 23(5), 611–631. https://doi.org/10.1007/s10278-009-9257-x

Rangayyan, R. M., Banik, S., & Desautels, J. L. (2013). Detection of Architectural Distortion in Prior Mammograms via Analysis of Oriented Patterns. Journal of Visualized Experiments : JoVE, (78), 50341. https://doi.org/10.3791/50341

Sadaf, A., Crystal, P., Scaranelo, A., & Helbich, T. (2011). Performance of computer-aided detection applied to full-field digital mammography in detection of breast cancers. European Journal of Radiology, 77(3), 457–461. https://doi.org/10.1016/j.ejrad.2009.08.024

Samulski, M., Hupse, R., Boetes, C., Mus, R. D. M., den Heeten, G. J., & Karssemeijer, N. (2010). Using computer-aided detection in mammography as a decision support. European Radiology, 20(10), 2323–2330. https://doi.org/10.1007/s00330-010-1821-8

Sanchez Gómez, S., Torres Tabanera, M., Vega Bolivar, A., Sainz Miranda, M., Baroja Mazo, A., Ruiz Diaz, M., Martinez Miravete, P., Lag Asturiano, E., Munoz Cacho, P., Delgado Macias, T. (2011). Impact of a CAD system in a screen-film mammography screening program: A prospective study. European Journal of Radiology, 80(3), e317–e321. https://doi.org/10.1016/j.ejrad.2010.08.031

Singh, S. P., & Urooi, S. (2016). An Improved CAD System for Breast Cancer Diagnosis Based on Generalized Pseudo-Zernike Moment and Ada-DEWNN Classifier. Journal Of Medical Systems, 40(4). https://doi.org/10.1007/s10916-016-0454-0

Skaane, P., Kshirsagar, A., Hofvind, S., Jahr, G., & Castellino, R. A. (2012). Mammography screening using independent double reading with consensus: Is there a potential benefit for computer-aided detection? Acta Radiologica, 53(3), 241–248. https://doi.org/10.1258/ar.2011.110452

Sohns, C., Angic, B., Sossalla, S., Konietschke, F., & Obenauer, S. (2010). CAD in full-field digital mammography-influence of reader experience and application of CAD on interpretation of time. Clinical Imaging, 34(6), 418–424. https://doi.org/10.1016/j.clinimag.2009.10.039

Velikova, M., Lucas, P., Samulski, M., & Karssemeijer, N. (2013). On the interplay of machine learning and background knowledge in image interpretation by Bayesian networks. Artifical Intelligence in Medicine, 57(1), 73–86.

Wang, J. H., Yang, X., Cai, H. M., Tan, W. C., Jin, C. Z., & Li, L. (2016). Discrimination of Breast Cancer with Microcalcifications on Mammography by Deep Learning. Scientific Reports, 6, 27327–27327.

Wang, X., Lederman, D., Tan, J., Wang, X. H., & Zheng, B. (2010). Computerized detection of breast tissue asymmetry depicted on bilateral mammograms: a preliminary study of breast risk stratification. Academic Radiology, 17(10), 1234–1241. https://doi.org/10.1016/j.acra.2010.05.016

Wang, S. H., Rao, R.V., Chen, P., Zhang, Y.D., Liu, A.J., & Wei, L., (2017) Abnormal Breast Detection in Mammogram Images by Feed-Forward Neural Network Trained by Jaya Algorithm. Fundamenta Informaticae 151(1–4), 191–211.

## Artificial intelligence

Beck, A. H., Sangoi, A. R., Leung, S., Marinelli, R. J., Nielsen, T. O., van de Vijver, M. J., West, R. B., van de Rijn, M., Koller, D. (2011). Systematic Analysis of Breast Cancer Morphology Uncovers

Stromal Features Associated with Survival. Science Translational Medicine, 3(108), 108ra113. https://doi.org/10.1126/scitranslmed.3002564

Becker, A. S., Marcon, M., Ghafoor, S., Wurnig, M. C., Frauenfelder, T., & Boss, A. (2017). Deep Learning in Mammography Diagnostic Accuracy of a Multipurpose Image Analysis Software in the Detection of Breast Cancer. Investigative Radiology, 52(7), 434–440. Retrieved from https://insights.ovid.com/pubmed?pmid=28212138

Cabitza, F., Rasoini, R., & Gensini, G. (2017). Unintended consequences of machine learning in medicine. JAMA, 318(6), 517–518. https://doi.org/10.1001/jama.2017.7797

Carneiro, G. N., & Bradley, A. (2017). Automated Analysis of Unregistered Multi-View Mammograms With Deep Learning. IEEE Transactions on Medical Imaging, 36(11), 2355–2365. https://doi.org/10.1109/TMI.2017.2751523

de Sampaio, W.B., Silva, A. C., de Paiva, A. C., & Gattass, M. (2015). Detection of masses in mammograms with adaption to breast density using genetic algorithm, phylogenetic trees, LBP and SVM. Expert Systems with Applications, 42(22), 8911–8928. https://doi.org/10.1016/j.eswa.2015.07.046

Dhungel, N., Carneiro, G., & Bradley, A. P. (2017). A deep learning approach for the analysis of masses in mammograms with minimal user intervention. Medical Image Analysis, 37, 114–128. https://doi.org/10.1016/j.media.2017.01.009

Ertosun, M. G., & Rubin, D. L. (2015). Probabilistic visual search for masses within mammography images using deep learning. In 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 1310–1315). https://doi.org/10.1109/BIBM.2015.7359868

Ganesan, K., Acharya, U. R., Chua, C. K., Min, L. C., Abraham, K. T., & Ng, K.-H. (2013). Computeraided breast cancer detection using mammograms: a review. IEEE Reviews In Biomedical Engineering, 6, 77–98. https://doi.org/10.1109/RBME.2012.2232289

Guo, Y., Dong, M., Yang, Z., Gao, X., Wang, K., Luo, C., Ma, Y., Zhang, J. (2016). A new method of detecting micro-calcification clusters in mammograms using contourlet transform and non-linking simplified PCNN. Computer Methods and Programs in Biomedicine, 130, 31–45. https://doi.org/10.1016/j.cmpb.2016.02.019

Houssami, N., Lee, C.I., Buist, D.S.M., & Tao, D. (2017). Artificial intelligence for breast cancer screening: Opportunity or hype? Breast (Edinburgh, Scotland), 36, 31–33. https://doi.org/10.1016/j.breast.2017.09.003

Kallenberg, M., Petersen, K., Nielsen, M., Ng, A. Y., Diao, P., Igel, C., Vachon, C. M., Holland, K., Winkel, R. R., Karssemeijer, N., Lillholm, M. (2016). Unsupervised Deep Learning Applied to Breast Density Segmentation and Mammographic Risk Scoring. IEEE Transactions on Medical Imaging, 35(5), 1322–1331. https://doi.org/10.1109/TMI.2016.2532122

Neto, O. P. S., Silva, A. C., Paiva, A. C., & Gattass, M. (2017). Automatic mass detection in mammography images using particle swarm optimization and functional diversity indexes. Multimedia Tools and Applications, 76(18), 19263–19289. https://doi.org/10.1007/s11042-017-4710-1

Teare, P., Fishman, M., Benzaquen, O., Toledano, E., & Elnekave, E. (2017). Malignancy Detection on Mammography Using Dual Deep Convolutional Neural Networks and Genetically Discovered False Color Input Enhancement. Journal of Digital Imaging, 30(4), 499–505.

Trister, A. D., Buist, D. S. M., & Lee, C. I. (2017). Will Machine Learning Tip the Balance in BreastCancerScreening?JAMAOncology,3(11),1463–1464.https://doi.org/10.1001/jamaoncol.2017.0473

## Tele-mammography

Bashshur, R., Krupinski, E., Thrall, J., & Bashshur, N. (2016). The Empirical Foundations of Teleradiology and Related Applications: A Review of the Evidence. Telemedicine and e-health, 22(11), 868-898. https://doi.org/10.1089/tmj.2016.0149

Bhateja, V., Urooj, S., & Misra, M. (2015). Technical Advancements to Mobile Mammography Using Nonlinear Polynomial Filters and IEEE 21451-1 NCAP Information Model. IEEE Sensors Journal, 15(5), 2559–2566. https://doi.org/10.1109/JSEN.2014.2366599

Fruehwald-Pallamar, J., Jantsch, M., Pinker, K., Hofmeister, R., Semturs, F., Piegler, K., Staribacher, D., Weber, M., Helbich, T. H. (2013). Teleradiology with uncompressed digital mammograms: clinical assessment. European Journal of Radiology, 82(3), 412-416. https://doi.org/10.1016/j.ejrad.2012.03.004

McCLung, A., McKibbon, A. & Archer, N. (2014). Novel eHealth Trends In The Field Of Radiology: A Scoping Review (Working Paper No. 49). Ontario: McMaster eBusiness Research Centre. http://hdl.handle.net/11375/17711

Patil, K.K., & Ahmed, S.T. (2014). Digital Telemammography Services for Rural India, Software Components and Design Protocol. 2014 International Conference on Advances in Electronics, Computers and Communications. 10.1109/ICAECC.2014.7002442

Salazar, A. J., Romero, J., Bernal, O., Moreno, A., & Velasco, S. (2016). Reliability of the BI-RADS Final Assessment Categories and Management Recommendations in a Telemammography Context. Journal of the American College of Radiology, 14, 686-692. http://dx.doi.org/10.1016/j.jacr.2016.08.004

Salazar, A. J., Romero, J., Bernal, O., Moreno, A., Velasco, S., & Díaz, X. (2014). Evaluation of Low-Cost Telemammography Screening Configurations: A Comparison with Film-Screen Readings in Vulnerable Areas. Journal of Digital Imaging, 27(5), 679–686. https://doi.org/10.1007/s10278-014-9695-y

Salazar, A. J., Romero, J. A., Bernal, O. A., Moreno, A. P., Velasco, S. C., & Díaz, X. A. (2016). Noninferiority and Equivalence Evaluation of Clinical Performance among Computed Radiography, Film, and Digitized Film for Telemammography Services. International Journal of Telemedicine and Applications, 2016, 1–12. https://doi.org/10.1155/2016/3642960