

evidence | *note*

In response to an enquiry from Dr Gerald Lip on behalf of the SBSP (Scottish Breast Screening Programme) Clinical Directors Group.

Number 77 April 2018

In asymptomatic women attending for breast screening, what is the clinical and cost effectiveness of digital breast tomosynthesis (DBT) in addition to full-field digital mammography or synthetic 2D (FFDM or S2D) images, compared to FFDM alone?

Questions addressed

The aim of this evidence note is to establish whether the existing evidence supports the addition of digital breast tomosynthesis (DBT) to the breast screening pathway. The specific questions addressed are:

- In asymptomatic women attending for breast screening, what is the evidence that DBT and full-field digital mammography or synthetic 2D (FFDM or 2SD) images are more accurate in detecting cancers, and reducing unnecessary recall of women for further imaging and testing, compared to FFDM alone?
- What is the impact on longer-term outcomes, such as interval cancer rates and overall breast cancer mortality?
- How does breast density and type/stage of cancer influence the results?

Key points

- Evidence from systematic reviews suggests that DBT plus FFDM is better at detecting invasive cancer (particularly early invasive cancer, stage T1 or N0), compared to FFDM alone. However, there appears to be no benefit of DBT in detecting ductal carcinoma in situ.
- Several studies report statistically significant reductions in false positives and recall rates when DBT plus FFDM is compared to FFDM alone. However, the magnitude of effect was not consistent across studies.
- In women with dense breasts attending for screening, the addition of DBT to FFDM results in increased breast cancer detection, compared to FFDM alone.
- While the evidence of diagnostic accuracy suggests that the addition of DBT to FFDM results in improved cancer detection, further information is required before the use of DBT can be

justified in the screening population. For example, further research is needed to evaluate the cost effectiveness of DBT in the breast screening population, and the potential impact of DBT on longer-term outcomes, such as interval cancer rates and mortality.

What is an evidence note?

Evidence notes are rapid reviews of published secondary clinical and cost-effectiveness evidence on health technologies under consideration by decision makers within NHSScotland. They are intended to provide information quickly to support time-sensitive decisions. Information is available to the topic referrer within a 6-month period and the process of peer review and final publication of the associated advice is usually complete within 6–12 months. Evidence notes are not comprehensive systematic reviews. They are based on the best evidence that Healthcare Improvement Scotland could identify and retrieve within the time available. The evidence notes are subject to peer review. Evidence notes do not make recommendations for NHSScotland, however the Scottish Health Technologies Group (SHTG) produces an Advice Statement to accompany all evidence reviews.

Definitions and acronyms

Interval breast cancer: Cancer diagnosed after a screening appointment at which a woman received a normal result, and before her next screening appointment.

DBT: Digital breast tomosynthesis

FFDM: Full field digital mammography

S2D: Synthetic 2D images, generated from tomosynthesis images.

Literature search

A systematic search of the secondary literature was carried out between 5–8 September 2017 to identify systematic reviews, health technology assessments and other evidence based reports. The Medline, Embase, and Web of Science databases were also searched for systematic reviews and meta-analyses. Results were limited to reviews in the English language.

Key websites were searched for guidelines, policy documents, clinical summaries, economic studies and ongoing trials.

Concepts used in all searches included: tomosynthesis; breast screening; synthetic 2D; 3D mammography, digital breast tomosynthesis (DBT). A full list of resources searched and terms used are available on request.

Given the volume of literature, this evidence note is limited to systematic reviews from 2016 and 2017.

Introduction

Currently in Scotland, women aged between 50 and 70 years are invited to participate in breast screening every 3 years, based on the fundamental principle of informed choice. Women over 70 years are encouraged to arrange an appointment. Over the last decade, most screening programmes have changed from 2D analogue mammography to full field digital mammography (FFDM). FFDM represents the current standard for mammography programmes¹.

Screening with mammography can help identify breast cancer earlier, which is associated with reductions in mortality. At a population level, the estimated effect of the UK breast screening programmes is a 20% reduction in breast cancer mortality in women invited for screening².

As with any screening programme, the population health benefits (for example, reduced breast cancer mortality) are balanced against potential harm to participants. Some possible harms from breast screening are over-diagnosis, radiation exposure, and participant anxiety. This issue was evaluated in the independent Marmot review in 2013. The review considered over-diagnosis to be the main harm from breast cancer screening. Putting together the figures for benefit and over-diagnosis, the review estimated that for 10,000 UK women invited to screening from age 50 for 20 years, about 681 cancers would be found of which 129 would represent over-diagnosis, and 43 deaths from breast cancer would be prevented. The report concluded that UK breast screening programmes confer significant benefit and should continue². Any change to the breast screening programme would require a consideration of these benefits and harms, to ensure that the current balance is maintained or, if possible, improved.

This evidence note explores the use of DBT in addition to FFDM (or S2D) compared to FFDM alone, by considering the research questions detailed above. In the conclusion section, wider consideration is given to how the addition of DBT to the existing breast screening pathway may influence the balance between population benefits and participant harms.

Health technology description

DBT is an advanced form of breast imaging which provides three-dimensional information of the breast. In a conventional mammogram, two X-ray images are taken of the breast (top-to-bottom and from angled side-to-side) while the breast is compressed between a clear plastic paddle and an imaging detector. The purpose of compressing the breast is to reduce overlapping of the breast tissue. However, given that there is an inevitable degree of overlapping tissue when a 2D mammogram is read, abnormal tissue can be hidden or normal tissue may appear abnormal. In DBT, the X-ray tube moves in an arc over the compressed breast capturing multiple images, which can be synthesised into a set of 3D images by a computer³.

In the current pathway of care, women with abnormalities on their mammograms are recalled for additional imaging, clinical examination, ultrasound and potentially biopsy. The use of DBT in addition to FFDM may increase clinical confidence about diagnosis, compared to FFDM alone, thus reducing the number of women who are recalled for further investigation. There are potential issues regarding adding DBT to FFDM as a screening tool, including increased radiation dose, increased costs, over-diagnosis, increased quality assurance requirements, increased data storage requirements and increased scan and read times⁴. However, as the technology evolves, and is refined, some of these potential issues (for example, increased scan time) may become obsolete.

One review reported an increase in radiologist read time for DBT compared to FFDM ranging from 47% to 135%⁵. Though read times may reduce with experience, this is an important consideration given the existing demands on the radiology workforce. Depending on the DBT system, data storage requirements for DBT images can be 10–50 times greater than for standard 2D mammograms, though it is not clear to what extent this would hinder future implementation (Dr G. Lip Clinical Director, North East Scotland Breast Screening Programme. Personal Communication, November 2017).

The radiation dose from DBT is similar to FFDM. When DBT is used as an adjunct to FFDM, this results in a double radiation dose. Alternatively, the DBT data set can be used to generate a synthetic 2D (S2D)

image, avoiding the need for two separate radiation doses. When DBT is used to generate S2D images, the total radiation dose is similar or slightly higher than FFDM¹. However, it should be noted that an NHS England position paper on the use of tomosynthesis states that at present there is insufficient evidence to support the use of the synthetic 2D image generated from a tomosynthesis acquisition as a replacement for a standard digital mammogram⁶.

Five DBT systems have received a CE mark, permitting distribution in the European Union (EU)⁵. These systems have numerous differences in their technical specifications, which would need to be considered if DBT was to be implemented in new settings. Public Health England has published practical evaluations for the GE SenoClaire DBT system and the Hologic [Selenia] Dimensions system as equipment reports⁷. However, no studies compare the different DBT systems, and the effect of these differences on patient-oriented outcomes is unknown⁵.

DBT is not currently used in the NHSScotland breast screening programme. All of the six screening centres in Scotland use the Hologic [Selenia] Dimensions mammography system. These systems have DBT capability, but this has been disabled on all but two systems. In the two units which have DBT enabled, one (in Dundee) is about to be used in a research project, and the other (in Glasgow) is occasionally used as an add-on for specific assessment cases (it cannot be used to replace existing imaging for assessment as it is not linked to the Picture Archiving and Communication System, PACS). Enabling DBT capability on the Hologic [Selenia] Dimensions systems requires a software upgrade and purchase of a license. Dundee also uses the Siemens Mammomat Inspiration System, which has DBT available, but it is only used for symptomatic patients (not for screening or assessment). In addition, Inverness have a Fuji Amulet unit, which cannot be upgraded to have DBT capability (Information provided by peer reviewers. Personal Communication, January 2018).

Epidemiology

The incidence of breast cancer is highest in developed countries, with an age-adjusted incidence rate of 80 per 100,000 in the EU. It is the second most common cause of cancer death in women in developed countries¹.

There are many different types of breast cancer. The first major division is between in situ and invasive carcinoma.

- In situ carcinoma is 'pre-invasive' carcinoma that has not yet invaded the breast tissue. These in situ cancer cells grow inside of the pre-existing breast lobules or ducts.
- Invasive cancers have cancer cells that infiltrate outside of the breast lobules and ducts to grow into the breast connective tissue.

Approximately 80% of breast carcinomas are invasive ductal carcinoma, followed by invasive lobular carcinomas which account for approximately 10–15% of cases. Invasive ductal carcinomas and invasive lobular carcinomas have distinct pathologic features⁸.

There are several risk factors for breast cancer. However, the one that is most relevant to this evidence note is breast density. Denser breasts have more glandular tissue relative to fat. High mammographic breast tissue density is associated with increased breast cancer risk, and is also associated with an increased risk of interval breast cancer (cancer diagnosed after a screening appointment at which a woman received a normal result, and before her next screening appointment) in women who have mammography screening⁹.

With regards to interval cancer rates in the NHS breast screening programme, one study reported that in the periods 0<12 months, 12<24 months and 24<36 months after a negative screen, overall interval cancer rates were 0.55, 1.13 and 1.22 per 1000 women screened, respectively¹⁰.

Clinical effectiveness

Seven systematic reviews were identified from 2016 and 2017^{1, 5, 9, 11-14}. These were considered to sufficiently capture the existing evidence base, and so primary studies and older reviews have not been included. Within the systematic reviews, reference is made to four large prospective studies, which together encompass 37,085 participants. The prospective studies are summarised in Table 1. It should be noted that the four studies do not represent the entire evidence base, but are the largest prospective studies currently available.

Table 1: Large prospective DBT + FFDM versus FFDM paired studies in the screening context

References	Study characteristics	Participant characteristics	Main results
STORM (Ciatto <i>et al</i> 2013) Screening with Tomosynthesis OR Mammography	Italy (multicentre) FFDM versus FFDM+DBT Selenia Dimensions Unit Double reader	n=7,292 Median age: 58 (54–63) Rate of dense breast: 16.7%	Cancer detection rate: 8.1/1000 screens in DBT+FFDM versus 5.3/1000 screens in FFDM Potential 17.2% reduction in recall rate using DBT
OTST (Skaane <i>et al</i> 2013) Oslo Tomosynthesis Screening Trial	Norway (single centre) FFDM versus FFDM+DBT Selenia Dimensions Unit Single reader, with arbitration/double	n=12,621 Mean age: 59.3 (50–69) Rate of dense breast: N/A	Cancer detection rate: 8.0/1000 screens in DBT+FFDM versus 6.1/1000 screens in FFDM False positive: 53.1/1000 screens in DBT+FFDM versus 61.1/1000 screens in FFDM Positive predictive value recall: 16.2% in DBT+FFDM versus 6% in FFDM
MBTST (Lang <i>et al</i> 2016) Malmö Breast Tomosynthesis Screening Trial	Sweden (single centre) One view DBT versus two-view digital mammography Mammomat inspiration Double reader, with arbitration	n=7,500 Mean age: 56 (40–76) Rate of dense breast: 42%	Cancer Detection Rate: 8.9/1000 in DBT versus 6.3/1000 in DM Recall rate: 3.8% (95% CI 3.3% to 4.2%) in DBT versus 2.6% (95% CI 2.3% to 3.0%) in DM
STORM-2 (Bernardi <i>et al</i> 2016)	Italy (multicentre) DBT+FFDM/DBT + synthetic 2D versus FFDM Selenia Dimensions Unit Double reader, with arbitration	n=9,672 Median age: 58 (53–63) Rate of dense breast: 26.8%	Cancer Detection Rate: FFDM + DBT = 8.5/1000 (95% CI 6.7 to 10.5) S2D + DBT = 8.8/1000 (95% CI 7.0 to 10.8) FFDM = 6.3/1000 (95% CI 4.8 to 8.1) False positive recalls: FFDM + DBT = 381/9587 (3.97%; 95% CI 3.59 to 4.38)

References	Study characteristics	Participant characteristics	Main results
			S2D + DBT = 427/9587 (4.45%; 95% CI 4.05 to 4.89) FFDM = 328/9587 (3.42%; 95% CI 3.07 to 3.80)

There is overlap between all the studies included in the systematic reviews, but particularly with regards to the studies included in Table 1. The reviews reported on different outcomes from the same studies, and pooled the data in different ways. The different outcomes from the reviews have been described below.

*Yun et al (2017)*¹⁴

The most recent systematic review evaluated the benefit of adding DBT to FFDM compared to FFDM alone for breast cancer detection, focusing on cancer characteristics (histologic characteristics, tumour size, and lymph node status). The review is well reported and appears to be methodologically robust. The authors included studies that compared the diagnostic value of DBT plus FFDM to FFDM alone for routine breast screening, and used pathologic confirmation as the reference standard.

Eleven studies met the review inclusion criteria, including the four prospective studies detailed in Table 1 (STORM, OTST, MBTST and STORM-2). The remaining seven studies were retrospective and encompassed 75,532 participants who received DBT plus FFDM, and 175,825 participants who received FFDM alone. The authors reported that the included studies were of satisfactory quality, but all had a high risk of bias concerning the reference standard test. This was because women who were not recalled did not receive a reference standard and the pathologist might have been aware that biopsies were recommended based on radiologic findings.

The authors calculated pooled estimates separately for overall cancer, invasive cancer, and carcinoma in situ. Subgroups of invasive carcinoma by T-stage, N-stage, histologic grade, and histologic type were also compared between the two screening modalities. The results are summarised in Table 2.

Table 2: Summary of results from Yun et al (2017) review

Cancer characteristic	Pooled risk ratios (95% confidence interval); I ² DBT plus FFDM <i>versus</i> FFDM alone	Explanation
<i>Pooled results from 11 studies</i>		
Overall cancer detection rate	1.290 (1.164 to 1.429); I ² , 0%	Overall cancer detection rate was significantly higher in the DBT plus FFDM group, compared with the FFDM group. Detection of invasive cancer was significantly higher in the DBT plus FFDM group, compared with FFDM alone. However, for carcinoma in situ there was no observed benefit of adding DBT to FFDM.
Invasive cancer	1.327 (1.168 to 1.508); I ² , 7.19%	
Carcinoma in situ	1.198 (0.942 to 1.524); I ² , 29.37%	
<i>T staging subset analysis (pooled results from five studies)</i>		
T1 cancer (≤2cm)	1.388 (1.137 to 1.695); I ² , 0%	The pooled estimates suggest that DBT plus FFDM detects more T1 cancers than FFDM
T2 cancer (>2cm)	1.391 (0.895 to 2.163); I ² , 0%	

		alone; but there was no observed benefit of adding DBT to FFDM for T2 or larger cancers.
<i>N staging subset analysis (pooled results from six studies)</i>		
Invasive cancer without nodal metastases (N0)	1.451 (1.209 to 1.742); I ² , 0%	The pooled estimates suggest that DBT plus FFDM detects more invasive cancers without nodal metastases compared with FFDM alone; but there was no observed benefit of adding DBT to FFDM for invasive cancers with nodal metastases.
Invasive cancer with nodal metastases (≥N1)	1.336 (0.921 to 1.983); I ² , 0%	
<i>Histologic cancer grade subset analysis (pooled results from five studies)</i>		
Grade I	1.812 (1.372 to 2.393); I ² , 0%	Pooled RRs showed a greater cancer detection for DBT plus FFDM compared with FFDM alone for all histologic grades of cancer.
Grade II/III	1.403 (1.174 to 1.676); I ² , 0%	
<i>Histological cancer type subset analysis (pooled results from four studies)</i>		
Invasive ductal carcinoma	1.437 (1.189 to 1.737)	Pooled RRs showed a greater cancer detection for DBT plus FFDM compared with FFDM alone for all histologic types of invasive cancer.
Invasive lobular carcinoma	1.901 (1.213 to 2.979)	

Based on these results, the authors surmised that the benefit of adding DBT to detect cancer in screening populations was associated with T-/N-staging. DBT benefit in detecting carcinoma in situ was uncertain, however it showed significant benefit for detecting invasive cancer - in particular early invasive cancer (stage T1 or N0). The results also suggested that the benefit of invasive cancer detection was in favour of DBT, regardless of histologic cancer grade or type.

The authors conclude that adding DBT to FFDM enabled the detection of invasive breast cancer that might have been missed with FFDM alone. They also suggest that knowing which cancer characteristic DBT detects may allow it to play a complementary role in predicting long-term patient outcomes and facilitate treatment planning. However, they note that proving more cancers are found is probably not sufficient to demonstrate that DBT plus FFDM should replace conventional FFDM alone for breast cancer screening. There is a need to demonstrate that the improved detection is sufficient to tolerate the additional issues of DBT, such as additional scan and read time, and increased radiation dose. They say: 'to establish the net efficacy of screening using DBT plus FFDM for reducing breast cancer mortality, further evaluation should be conducted using surrogate endpoints, such as the change of the interval cancer rate or progression-free survival'¹⁴.

Hodgson et al 2016¹

This systematic review (with meta-analyses) evaluated the performance of DBT (alone or in combination with FFDM or S2D images) for detecting breast cancer compared with FFDM alone when screening asymptomatic women. The reference standard for the positive cases of cancer was histological results confirmed by biopsy or surgical resection. The reference standard for the negative cases was any follow-up period, where reported.

Five studies were eligible for inclusion in the review, including OTST and STORM (Table 1). The remaining three studies were conducted in the USA, and were retrospective in design (only one, Lourenco *et al*, was included in the systematic review by Yun *et al*¹⁴). The retrospective studies included all eligible

participants from a given time-frame. The authors reported the results of the US and European studies separately, because of differences in breast cancer rates, demographics and screening practices.

Two European studies (STORM and OTST)

- Higher cancer detection rates were observed with DBT plus FFDM compared with FFDM alone. The pooled difference per 1,000 screens was 2.43 (95% CI 1.8 to 3.1). Both studies reported a higher invasive cancer detection rate, but did not report a statistically significant difference in non-invasive (in situ) cancer detection.
- The two studies reported different results with respect to false positive rates and recall rates. However, the processes used within each study to decide which women were eligible for recall were different. Both used two independent readers; in the OTST study, cases with a positive interpretation by one reader were sent to a consensus group for a final decision, whereas in the STORM study, women were recalled if either reader had a positive screen.
 - In STORM, lower false positive and recall rates were observed when DBT plus FFDM was compared with FFDM alone. The difference per 1,000 screens for false positives was -9.3 (95% CI -11.8 to -7.2); and for recall rate was -6.6 (95% CI -8.7 to -4.9).
 - In OTST, lower false positive rates using DBT plus FFDM were found pre-arbitration, but higher false positive and recall rates were found post-arbitration.
- STORM presented data on interval cancer rates based on the limited follow-up data available (women were screened every 2 years). The sensitivity for DBT plus FFDM (90.77; 95% CI 80.70 to 96.51) was higher compared with FFDM alone (60; 95% CI 47.1 to 71.96). The authors reported that the specificity for DBT plus FFDM (96.49; 95% CI 96.04 to 96.90) was higher compared to FFDM alone (95.55; 95% CI 95.04 to 96.01). However, while the result is statistically significant, the difference may not be clinically significant.

Three USA studies (Friedwald et al, 2014; Destounis et al, 2014; Lourenco et al, 2014)

- One large multicenter US study showed a higher cancer (invasive and non-invasive) detection rate for DBT plus FFDM, while two smaller US studies did not find statistically significant differences. The largest (Friedwald et al, 2014) reported a statistically significant difference per 1,000 screens in favour of DBT plus FFDM over FFDM alone (1.21 95% CI 0.82 to 1.63). The other two studies reported no statistically significant differences (Destounis et al, 2014 and Lourenco et al, 2014).
- The largest study (Friedwald et al, 2014) reported a statistically significant higher invasive cancer detection rate. The other two studies reported no statistically significant differences.
- None of the studies reported a statistically significant difference in non-invasive cancer detection rates between DBT plus FFDM and FFDM alone.
- Statistically significant reductions in recall and false positive rates were observed in all three studies in favour of DBT plus FFDM, however the magnitude of difference varied between the studies. The results were too heterogeneous to combine in meta-analyses, so the results were reported separately:
 - *False positive rates (difference per 1,000 screens)*
 - Friedwald et al: -17.4 (95% CI -15.6 to -19.2)
 - Lourenco et al: -28.7 (95% CI -35.1 to -22.2)
 - Destounis et al: -74.4 (95% CI -105.6 to -43.1)
 - *Recall rates (difference per 1,000 screens)*
 - Friedwald et al: -16.2 (95% CI -18 to -14.5)
 - Lourenco et al: -29.4 (95% CI -36 to -22.8)
 - Destounis et al: -72.5 (95% CI -104.7 to -40.2)

The authors' overall conclusions were that 'cancer detection rates and invasive cancer detection rates are higher using DBT plus FFDM than with FFDM, but non-invasive cancer detection rates are unchanged'¹. This mirrors the conclusion by Yun *et al*¹⁴.

Overall, the review is well-reported and appears to be good quality. The authors state that four of the five included studies were rated as having a low risk of bias, but one was rated as having an unclear risk of bias (Lourenco *et al*, 2014).

Houssami *et al* 2016⁹

This rapid review focused on the evidence relating to DBT screening in women with dense breasts, with the aim of estimating additional breast cancer detection attributable to DBT in comparison with standard 2D mammography. Eight studies were included. Four were prospective trials comparing both modalities (DBT plus FFDM versus FFDM alone) in the same participants (including STORM, STORM-2 and MBTST). The other four were retrospective studies reporting screen-detection measures in different groups of women.

The prospective trials encompassed 10,188 women with dense breasts, receiving predominantly biennial screening (every other year), and pooled results showed a significant increase in breast cancer detection in 3.9/1,000 screens attributable to DBT ($p < 0.001$).

Studies comparing different groups of women screened with DBT and FFDM ($n = 103,230$) or with FFDM alone ($n = 177,814$) yielded a pooled difference in breast cancer detection of 1.4/1,000 screens; representing significantly higher breast cancer detection in DBT-screened women ($p < 0.001$). A pooled difference in recall was reported of -23.3/1,000 screens; representing significantly lower recall in DBT-screened groups ($p < 0.001$).

The authors note that the reported estimates of additional breast cancer detection attributed to DBT differ in magnitude between the prospective and retrospective studies. They suggest that this may be partly explained by the difference in screening frequency: the prospective studies were European, and women were screened every two years; the retrospective studies were from the USA, and women were screened annually.

Other systematic reviews

The remaining four systematic reviews relate to slightly different questions and include additional studies, yet they do not significantly add to the evidence already presented. Therefore, they have not been discussed in detail here. Their main conclusions are summarised below.

Coop *et al* (2016)¹¹: 'Using tomosynthesis with digital mammography increases breast cancer detection, reduces recall rates and increases the positive predictive value of those cases recalled. Invasive cancer detection is significantly improved in tomosynthesis compared to mammography, and has improved success for women with heterogeneous or extremely dense breasts.'

Pozz *et al* (2016)¹³: 'Digital breast tomosynthesis addresses the primary limitations of conventional screening mammography by increasing conspicuity of invasive cancers while concomitantly reducing false-positive results...adding DBT to digital mammograms substantially reduces unnecessary diagnostic services, especially in younger and dense-breast women and allows for earlier, less costly treatment strategies. Further research is needed to evaluate the potential impact of DBT on longer-term outcomes, such as interval cancer rates and mortality.'

Melinkow *et al* (2016)¹²: ‘In most cohort studies, cancer detection rates were somewhat higher with DBT as compared to digital mammography alone, and the proportion of invasive cancers detected was similar or higher than the proportion detected with digital mammography alone...studies are needed...that report on both interval cancers identified by a comprehensive reference standard and longer-term outcomes, including effects of the addition of DBT to digital mammography on the stage distribution of detected cancers, breast cancer recurrence or second (contralateral) breast cancers, and mortality rates.’

ECRI Institute (2016)⁵: The main conclusions (relating to the Hologic Selenia Dimensions DBT system):

- DBT plus FFDM is more sensitive and specific than FFDM alone for breast cancer screening.
- S2D images with DBT is more sensitive than FFDM alone for breast cancer screening.
- The evidence does not clearly indicate how the recall rate of screening using DBT plus FFDM compares with FFDM alone because the studies have conflicting results.
- S2D images with DBT results in a higher recall rate than FFDM.
- The evidence does not clearly indicate how the recall rate or sensitivity of S2D+DBT compares with FFDM+DBT.

S2D images to replace FFDM

As already described, the use of DBT as an adjunct to FFDM results in a double radiation dose. A proposed alternative to FFDM is the use of an S2D image, generated from the DBT dataset. The evidence from the systematic reviews above does not clearly indicate whether S2D images can be used in place of FFDM.

The review by Pozz *et al* included a Health Technology Assessment (HTA) which evaluated this issue¹⁵. Although the HTA did not meet the inclusion criteria for this evidence note - it focused on the use of DBT in people who have been recalled from their initial screen for further imaging and testing and women with a moderate/high risk of developing breast cancer - it has been included here for information as it addresses whether S2D images can be used in place of FFDM. The HTA included 7,060 participants, and readers reviewed (1) 2D or (2) 2D + DBT or (3) S2D + DBT images for each case without access to original screening mammograms or prior examinations. The results are summarised below:

Table 3: Summary of results from Gilbert *et al* (2015)

Test or combination of tests	Sensitivity	Specificity
2D only	87% (95% CI 85% to 89%)	58% (95% CI 56% to 60%)
2D + DBT	89% (95% CI 87% to 91%)	69% (95% CI 67% to 71%)
S2D + DBT	88% (95% CI 86% to 90%)	71% (95% CI 69% to 73%)

The authors concluded that the specificity of DBT + 2D was better than 2D alone, but there was only a marginal improvement in sensitivity. Further, the performance of S2D appeared to be comparable to standard 2D. However, the authors also noted the need for more research into the feasibility of implementing DBT in a screening setting, and further comparisons of standard 2D images with S2D images for different lesion types and breast densities.

More recent studies have supported the conclusion that S2D images can be used in place of FFDM images, on the basis that S2D images are comparable for assessment of breast density^{16, 17}; cancer detection rate/diagnostic accuracy¹⁸⁻²⁰; and BI-RADS categorisation²¹⁻²³. However, these studies do not directly answer the research questions posed, and so the full text has not been obtained and appraised.

Participant and social aspects

The evidence examined for the clinical and cost-effectiveness sections of this evidence review did not discuss participant experience of DBT as compared with current practice in breast screening.

Some articles (for example, on patient information websites) report that DBT is more comfortable than FFDM because the breast is not compressed as much (Sally Greenbrook, Breast Cancer Now. Personal Communication, February 2018). However, this was not mentioned in any of the studies included in the evidence review.

Safety

As already noted, there are unavoidable harms from breast screening (for example, radiation, over-diagnosis and participant anxiety), but the Marmot review in 2013 concluded that UK breast screening programmes overall offer significant benefit and should continue.

An ECRI review states that: 'Adverse events associated with mammography are rare. None of the nine studies included in the evidence base of this report described any adverse events associated with DBT or FFDM'⁵.

According to Public Health England, the mean glandular dose (MGD) for a two-view examination with FFDM is 3mGy²⁴. To put this in context:

- The risk of a radiation-induced cancer for a woman attending FFDM screening (two views) is between 1 in 49,000 to 1 in 98,000 per visit;
- The risk of radiation-induced cancer is between 1 in 7,000 to 1 in 14,000 in women who attend all 7 screening examinations between the ages of 50 to 70;
- It is estimated that about 400 to 800 cancers are detected by the NHS breast screening programme for every radiation-induced cancer²⁴.

The radiation dose from a DBT is similar to FFDM. Therefore, using DBT as an adjunct to FFDM results in a double radiation dose. The use of S2D images in place of FFDM would avoid this double radiation dose.

Cost effectiveness

The reviews by Pozz *et al*¹³ and the ECRI institute⁵ included a total of four studies assessing the cost effectiveness of DBT. Two were conference abstracts and therefore have not been considered here. The remaining two were from the US, and focused on the initial screening population. The generalisability of these to the UK context is limited, and so they provide insufficient evidence to draw conclusions for NHSScotland. For information, they are briefly summarised below:

- Bonafede *et al* 2015: This economic model estimates the financial impact of DBT within a hypothetical US managed care plan with one million members comparing screening with FFDM alone with DBT plus FFDM. Assuming 35.5% of eligible women were screened each year, the authors concluded: 'total annual cost saving to the health plan are \$2.4 million, comprising \$5.5 million saving from avoiding follow-up services and \$1.2 million from earlier detection of breast cancer, both of which more than offset the incremental cost of \$4.2 million by adding DBT to screening.' This represents a saving of \$28.53 per woman screened (details taken from Pozz *et al*¹³ and the ECRI institute⁵).

- Lee *et al* 2015: This cost analysis was focused specifically on women with dense breasts between the ages of 50 and 74, and was conducted within the US health system. The authors used a discrete-event breast cancer simulation model. They concluded that biennial combined DBT added to digital mammography is likely to be cost effective, reporting an incremental cost per quality-adjusted life year gained by adding DBT to conventional mammography of \$53,893 (details taken from Pozz *et al*¹³ and the ECRI institute⁵).

DBT is currently available, but switched off, on most mammography machines in NHSScotland. Enabling DBT capability on the Hologic [Selenia] Dimensions system requires a software upgrade and purchase of a license (£50,000 per unit, with discounts for multi-license purchases). The Siemens Mammomat Inspiration system in Dundee already has DBT capability, and does not require any upgrading. Other costs to consider include, for example, increased servicing costs, cost associated with PACS (Picture Archiving and Communication System) and additional data storage, and the training of staff.

PROSPECTS trial

There is a large multicentre prospective trial underway in England and Wales (PROSPECTS), the primary aim of which is: 'To compare the cost-effectiveness of breast cancer screening using DBT+2DDM [FFDM] or S2D with screening using 2DDM by measuring cancer detection rates, interval cancer rates, size and lymph node status of Grade 2 and 3 invasive cancers in intervention (DBT+2DDM or S2D) and 2DDM (standard care) groups'²⁵. This trial is still in the recruitment phase, with the first round of screening due to complete towards the end of 2018. Interim results should be available before a second round of screening (Dr M. Halling-Brown Head of Scientific Computing, Royal Surrey County Hospital. Personal Communication, June 2017). Final results should be available in 2020. This study should form an important part of the evidence base, giving a clearer picture of the clinical and cost effectiveness of DBT in women attending for breast screening.

Conclusion

Seven systematic reviews from 2016 and 2017 were identified that helped to answer the research questions. Despite the reviews all having a slightly different focus, their conclusions were largely in agreement. The evidence suggests that the overall cancer detection rate is higher with DBT plus FFDM, compared to FFDM alone. While DBT plus FFDM showed a significant benefit for detecting invasive cancer (particularly early invasive cancer, stage T1 or N0), there appeared to be no advantage in detecting carcinoma in situ. It should be noted that increased detection of carcinoma in situ may or may not be a benefit (depending on whether treatment is necessary or whether the cancer is likely to cause an issue within the person's remaining lifetime).

Several studies reported statistically significant reductions in false positives and recall rates when DBT plus FFDM is compared with FFDM alone. However, this finding was not consistent, and the magnitude of the difference varied between the studies. These differences may be partly explained by variation in processes used to decide which individuals to recall, differences in reader experience, and variation in baseline recall rates between units (the impact may be greater in units with higher baseline recall rates). Therefore, the potential impact of DBT on recall rates in NHSScotland is unclear. In women with dense breasts, the addition of DBT to FFDM also resulted in increased breast cancer detection, compared to FFDM alone. Emerging evidence suggests that S2D images are comparable to FFDM images. The use of S2D in place of FFDM would avoid the double radiation dose of DBT plus FFDM.

While these findings allow us to say with a degree of confidence that DBT plus FFDM detects more cancers than FFDM alone in the asymptomatic screening population, they do not tell us the net value of

screening using DBT plus FFDM on breast cancer mortality. Furthermore, there are potential issues with DBT that would need to be considered, for example the increased scan and read time, and additional data storage requirements.

With any screening programme, an ethical consideration of the balance of population health benefits against participant harms is required. Adding DBT to the screening pathway may mean that cancers are detected earlier and treated sooner, thus reducing overall breast cancer mortality. Furthermore, the improved diagnostic accuracy of DBT may increase clinician confidence about diagnosis, and reduce the number of women facing the stress and anxiety of a recall from routine screening. However, costs at the participant level could be increased radiation dose (if S2D is not used in place of FFDM), and possible over-diagnosis (treatment for a cancer that would never have caused an issue in a person's lifetime). The existing evidence does not allow us to establish whether there would be any improvement of health at a population level, and whether any such improvements would be sufficient to tolerate the potential harms at an individual participant level.

Two economic studies from the US concluded that the use of DBT in addition to FFDM in the screening population is likely to be cost effective, although the generalisability of these studies to the UK context is limited. A large UK-based prospective study is currently underway (PROSPECTS), which aims to compare the cost effectiveness of breast cancer screening using DBT plus FFDM or S2D with screening using FFDM. The results of this study should be published around 2020, and is likely to form an important part of the evidence base.

Identified research gaps

- Further research is needed to evaluate the potential impact of breast screening using DBT on longer-term outcomes, such as interval cancer rates and breast cancer mortality. Some of these gaps should be addressed by the PROSPECTS study, which includes interval cancer rates as an outcome.

Equality and diversity

Healthcare Improvement Scotland is committed to equality and diversity in respect of the nine equality groups defined by age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion, sex, and sexual orientation.

The process for producing evidence notes has been assessed and no adverse impact across any of these groups is expected. The completed equality and diversity checklist is available on www.healthcareimprovementscotland.org

About evidence notes

Evidence Notes are produced to inform a decision at a particular point in time and are therefore not routinely updated. They will however be considered for review if requested by stakeholders, based upon the availability of new published evidence which is likely to materially change the advice given. For further information about the evidence note process see:

www.healthcareimprovementscotland.org/our_work/clinical_cost_effectiveness/shtg/standard_operating_procedures.aspx

To propose a topic for an evidence note, email shtg.hcis@nhs.net

References can be accessed via the internet (where addresses are provided), via the NHS Knowledge Network www.knowledge.scot.nhs.uk, or by contacting your local library and information service.

A glossary of commonly used terms in Health Technology Assessment is available from htaglossary.net.

Acknowledgements

Dr Gerald Lip, Consultant Radiologist, Clinical Director North East Scotland Breast Screening Programme acted as a topic advisor for this evidence note. This involved providing clinical expertise, and reviewing drafts of the document.

Healthcare Improvement Scotland and SHTG invited the following individuals to peer review the draft evidence note:

- Dr Marzi Davies, Consultant Radiologist, Director of West of Scotland Breast Screening Centre.
- Professor Andrew Evans, Professor of Breast Imaging, University of Dundee. Hon Consultant Radiologist, NHS Tayside.
- Sally Greenbrook, Policy Manager, Breast Cancer Now.
- Dr Janet Litherland, Consultant Radiologist, West of Scotland Breast Screening Service.
- Dr Michael J Michell, Consultant Radiologist, King's College Hospital NHS Foundation Trust, London.
- Katherine Schofield, Lead Mammography Physicist, NHS National Services Scotland.
- Dr Yee Ting Sim, Consultant Radiologist, NHS Tayside.
- Patsy Whelehan, Clinical Specialist Radiographer & Senior Research Radiographer, University of Dundee

Declarations of interest were sought from all peer reviewers. All contributions from peer reviewers were considered by the group. However the peer reviewers had no role in authorship or editorial control and the views expressed are those of Healthcare Improvement Scotland.

Healthcare Improvement Scotland development team

- Joanna Kelly, Health Services Researcher/Lead Author
- Paul Herbert, Health Information Scientist
- Charis Miller, Health Information Scientist
- Shonagh Ramsey, Project Officer
- Members of the SHTG evidence review committee

© Healthcare Improvement Scotland 2017

References

1. Hodgson R, Heywang-Kobrunner SH, Harvey SC, Edwards M, Shaikh J, Arber M, *et al.* Systematic review of 3D mammography for breast cancer screening. *Breast*. 2016;27:52-61.
2. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M, *et al.* The benefits and harms of breast cancer screening: an independent review. *Br J Cancer*. 2013;108(11):2205-40.
3. RadiologyInfo.org. For Patients. Breast Tomosynthesis. [cited 24 Nov 2017]; Available from: <https://www.radiologyinfo.org/en/info.cfm?pg=tomosynthesis>.
4. Gilbert FJ, Tucker L, Young KC. Digital breast tomosynthesis (DBT): a review of the evidence for use as a screening tool. *Clin Radiol*. 2016;71(2):141-50.
5. ECRI Institute. Digital Breast Tomosynthesis for Breast Cancer Screening 2016 [cited 04 Sept 2017]; Available from: <https://www.ecri.org/components/Target/Pages/30047.aspx>.
6. Public Health England. NHS Breast Screening Programme: current position on use of tomosynthesis 2016 [cited 04 Sept 2017]; Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/517840/BSP_position_on_tomosynthesis.pdf.
7. Public Health England. Breast Screening: Professional Guidance. [cited 02 February 2018]; Available from: <https://www.gov.uk/government/collections/breast-screening-professional-guidance>.
8. John Hopkins Medicine. Types of Breast Cancers: Histologic examples of in situ & invasive carcinomas of the breast. [cited 26 January 2018]; Available from: <http://pathology.jhu.edu/breast/types.php>.
9. Houssami N, Turner RM. Rapid review: Estimates of incremental breast cancer detection from tomosynthesis (3D-mammography) screening in women with dense breasts. *Breast*. 2016;30:141-5.
10. Bennett RL, Sellars SJ, Moss SM. Interval cancers in the NHS breast cancer screening programme in England, Wales and Northern Ireland. *Br J Cancer*. 2011;104(4):571-7.
11. Coop P, Cowling C, Lawson C. Tomosynthesis as a screening tool for breast cancer: A systematic review. *Radiography*. 2016;22(3):e190-e5.
12. Melnikow J, Fenton JJ, Miglioretti D, Whitlock EP, Weyrich MS. Screening for Breast Cancer With Digital Breast Tomosynthesis. Rockville (MD): Agency for Healthcare Research and Quality (US). 2016.
13. Pozz A, Corte AD, Lakis MA, Jeong H. Digital Breast Tomosynthesis in Addition to Conventional 2DMammography Reduces Recall Rates and is CostEffective. *Asian Pacific journal of cancer prevention : APJCP*. 2016;17(7):3521-6. Epub 2016/08/12.
14. Yun SJ, Ryu CW, Rhee SJ, Ryu JK, Oh JY. Benefit of adding digital breast tomosynthesis to digital mammography for breast cancer screening focused on cancer characteristics: a meta-analysis. *Breast Cancer Res Treat*. 2017;164(3):557-69.
15. Gilbert FJ, Tucker L, Gillan MG, Willsher P, Cooke J, Duncan KA, *et al.* The TOMMY trial: a comparison of TOMosynthesis with digital MammographY in the UK NHS Breast Screening Programme--a multicentre retrospective reading study comparing the diagnostic performance of digital breast tomosynthesis and digital mammography with digital mammography alone. *Health technology assessment (Winchester, England)*. 2015;19(4):i-xxv, 1-136. Epub 2015/01/20.
16. Alshafeiy TI, Wadih A, Nicholson BT, Rochman CM, Peppard HR, Patrie JT, *et al.* Comparison Between Digital and Synthetic 2D Mammograms in Breast Density Interpretation. *AJR Am J Roentgenol*. 2017;209(1):W36-W41.

17. Conant EF, Keller BM, Pantalone L, Gastouniotti A, McDonald ES, Kontos D. Agreement between Breast Percentage Density Estimations from Standard-Dose versus Synthetic Digital Mammograms: Results from a Large Screening Cohort Using Automated Measures. *Radiology*. 2017;283(3):673-80.
18. Aujero MP, Gavenonis SC, Benjamin R, Zhang Z, Holt JS. Clinical Performance of Synthesized Two-dimensional Mammography Combined with Tomosynthesis in a Large Screening Population. *Radiology*. 2017;283(1):70-6.
19. Choi G, Woo OH, Shin HS, Jang S, Cho KR, Seo BK. Comparison of two-dimensional synthesized mammogram (2DSM) and conventional full-field digital mammogram (FFDM) for evaluation of breast cancer. *Clin Imaging*. 2017;43:170-4.
20. Freer PE, Riegert J, Eisenmenger L, Ose D, Winkler N, Stein MA, *et al*. Clinical implementation of synthesized mammography with digital breast tomosynthesis in a routine clinical practice. *Breast cancer research and treatment*. 2017;166(2):501-9.
21. Garayoa J, Chevalier M, Castillo M, Mahillo-Fernandez I, Amallal El Ouahabi N, Estrada C, *et al*. Diagnostic value of the stand-alone synthetic image in digital breast tomosynthesis examinations. *Eur Radiol*. 2018;28(2):565-72.
22. Mariscotti G, Durando M, Houssami N, Fasciano M, Tagliafico A, Bosco D, *et al*. Comparison of synthetic mammography, reconstructed from digital breast tomosynthesis, and digital mammography: evaluation of lesion conspicuity and BI-RADS assessment categories. *Breast cancer research and treatment*. 2017;166(3):765-73.
23. Mumin NA, Rahmat K, Fadzli F, Ramli M, Rozalli FI, Westerhout C, *et al*. Diagnostic efficacy of synthesized 2D images from digital breast tomosynthesis compared to full field digital mammography in Asian population. *Journal of Medical Imaging and Radiation Oncology*. 2017;61(Supplement 1):111.
24. Public Health England. Radiation risk with digital mammography in breast screening 2017 [cited 07 February 2018]; Available from: <https://www.gov.uk/government/publications/breast-screening-radiation-risk-with-digital-mammography/radiation-risk-with-digital-mammography-in-breast-screening>.
25. PROSPECTS trial. [cited 24 November 2017]; Available from: <https://medphys.royalsurrey.nhs.uk/prospects/>.